University College London

Department of Genetics, Evolution, and Environment

Ecological and evolutionary impacts of chromosomal inversions

Carl J Mackintosh

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Declaration of originality

I, Carl Mackintosh, confirm that the work presented in my thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Acknowledgements

I have a lot of people to thank.

I hope I don't forget any of you.

Firstly, to my supervisors. POM, whose guidance has proven invaluable and encouraged me to focus on the bigger picture. Max, whose guidance has proven equally invaluable but encouraged me to focus on the small details. Mike, whose guidance has also proven equally invaluable, who has always been patient with me and generous with his time, and has been instrumental in getting me through this.

To everyone who provided me with advice and feedback, academic or otherwise. To Flo Camus, Tim Connallon, Bhavin Khatri, Filip Ruzicka, those poor, poor, anonymous reviewers, and of course, Ewan Flintham — harbinger of facts, a beacon of knowledge blazing out across a sea of ignorance.

To everyone I've worked with and met over the years. There are far too many of you to list, so consider this a highlight compilation. Here we go: thank you to Sasha (of a certain corner), Harrison "Hoz" Ostridge, (Feral) Jas, Geowizard Enrique, Stefano, Marion, Fiona, James, The Wet Dreams — Stuart, Will, Aaron (efficiency incarnate), Tabs (chaos incarnate), Aidan, Raquel — Flo, Aisha (teabag stealer), Jasmine, Will, Lara, Rob, Gla, Tom A, Filip/Flip/Fil/Filly, Roman, Dr. Functioning Adult Thomas, Iris, Elü, Vitor. To those outside this academic bubble, especially Shamus, Aaron, Hugh, Ben, Ozzy, Jake, and The Pals. To Michael, my partner in moaning. To Raquel (again — I know you'll love that you're here twice), my "PhD-best-friend". To Tom and Fla,

who make delicious banana bread, and to Josh, who once did not. And inevitably, to Ewan “Weaninho” Flintham, a man of agricultural prowess and malleability — thank you for being overgenerous with your time, knowledge, and portmanteaus. It’s been a pleasure to take this journey alongside you.

To Becca, whose presence I am and have been glad for every day. Thank you for being both my and Winnie’s custodian while I’ve been busy — you’ve done an excellent job.

To my family, who have always supported me, but especially to Mum. You’ve always given me the freedom and space to chase whatever it was I wanted to do, and encouraged me all the way. This is for you.

For you, Mum.

Thesis abstract

Inversions are genomic structural variants in which a segment of a chromosome is reversed orientation. One consequence is the near-total suppression of recombination within the inverted region. This region of tight linkage can have profound implications for the evolution of the genome, by enabling sets of alleles to be inherited as a single unit without being in close proximity. This thesis investigates how inversions evolve and influence evolution, from the genome up to the level of populations. One example is the architecture of meiotic drivers — genetic elements that have a transmission advantage over their wild-type counterparts. These systems often comprise multiple loci within an inversion. The first chapter models the spread and demographic consequences of X-linked meiotic drivers in the face of the associated fitness costs. When the costs are such that X-drive can remain polymorphic, the resulting female-biased sex ratio increases the equilibrium population size and persistence time relative to a wild-type population. Inversions may also promote the evolution of local adaptation under gene flow by linking coadapted alleles and preventing the production of hybrid, maladaptive genotypes. Work on this phenomenon often considers the limited case of a ''continent-island" model. We extend this and include the probability of the formation of locally adaptive inversions. Our results contrast with a simple interpretation of existing literature, suggesting strongly selected loci are likely to underpin locally adaptive inversions. Higher mutation load is expected in regions of suppressed recombination, due to the reduced efficacy of purifying selection on deleterious alleles. Furthermore, in the inversion case, the mutation load captured by an inversion is likely to persist throughout future lineages. The relative importance

of each of these phenomena depends on how long the inversion spends at a low frequency. We use simulations to determine the conditions in which the accumulation of deleterious mutations during the early stages of inversion spread is important in deciding the fate of an otherwise adaptive inversion. This is placed in the context of the existing literature, which underestimates the effect of mutation accumulation during the early stages of inversion lineages.

Impact statement

I present work demonstrating the capacity for chromosomal inversions to induce changes at various levels of selection, from the genome to the population. Inversions have been implicated in the existence of a plethora of phenomena, including mimicry, mating strategies, local adaptation, speciation, and the evolution of sex chromosomes. The existence of inversions has been known for much of the history of genetics itself, however their prevalence across genomes was vastly underestimated prior to the emergence of recent sequencing techniques. As such, there has been a surge in interest in understanding how inversions evolve. In addition, population genetic theory has been limited in how far inversion evolution can be understood. Analytic tractability often requires models to comprise one or two loci, or a simplifying assumption of uniformity across many loci. A further assumption often made is that of large population sizes, so that the effects of genetic drift can be ignored. Contrary to this, inversions often comprise many loci of various effects, and the effects of their interference with the process of recombination require the presence of drift to model accurately. Perhaps analogous to advances in genomic methods, advances in computing power have made the quick simulation of inversions possible. This thesis utilises both population genetic theory and simulation to aid our understanding of how inversions evolve.

Work from this thesis has been disseminated to the wider scientific community. Results from Chapter 3 were presented as a poster at Population Genetics Group 53 (Jan 2020), and published in in the journal *Genetics*. I presented a poster based on results

from Chapter 4 at the Congress of the European Society of Evolutionary Biology in Prague (Aug 2022), and gave a talk on the same subject at Population Genetics Group 56 in London (Jan 2023). Further, a manuscript based on this work has been submitted to an academic journal and is available as a preprint.

Contents

List of Figures

List of Tables

General Introduction

The composition and structure of the genome are ever-changing. Mutation constantly introduces genetic variation, from changes at single nucleotides to insertions and deletions of longer sequences of DNA. Selection on this genetic variation is the means by which evolution by natural selection proceeds. A radical source of genetic variation comes in the form of structural variants — changes in the number, order, or location of genes on a chromosome, that often affect gene activity and transmission (Mérot et al. 2020). One particular class of these rearrangements, chromosomal inversions, occur when a segment of chromosome breaks off and is reinserted in the reverse orientation, so that all the genetic content is retained but in an altered order. This happens by means of a double-stranded break in two places along the chromosome or more rarely via intrachromosomal recombination, where ectopic recombination occurs between related and opposite-facing DNA sequences on the same chromosome (Hartwell 2000; Ling and Cordaux 2010). The effects of this are generally deleterious, as inversion breakpoints alter the expression pattern of genes in their vicinity. If the breakpoints lie within the genes themselves, the gene is split and its function more severely disrupted — though adaptive breakpoint effects have also

been documented and could play a significant role in their spread (Villoutreix et al. 2021).

A key property of inversions is that they interfere with recombination in heterokaryotypes (Sturtevant 1917; Roberts 1976). Crossing over and the formation of the chiasma are essential in eukaryotes for proper meiotic segregation (Maynard Smith 1998). Without them, meiosis can be disrupted resulting in gametes without a full chromosomal complement, which lead to aneuploid offspring (having an abnormal number of chromosomes). An inverted chromosome section will not align with the rest of the chromosome during meiosis and therefore tends not to recombine, unless the inverted chromosome forms an inversion loop (Hartwell 2000). This aligns homologous regions of both chromosomes, allowing meiotic recombination between the two to proceed, however this may also prevent the formation of chiasmata (Hale 1986; Coyne et al. 1993). Worse, recombinant gametes produced in this manner often contain duplications and/or deletions that cause genetic imbalance, leading to their loss (Sturtevant and Beadle 1936). Some recombinant gametes can be produced in the rare case of an even number of crossovers within the inversion — usually two (Sturtevant and Beadle 1936). Therefore, inversion heterozygotes produce few viable offspring with crossovers in the inverted region. Due to the lack of recombination, inversions increase linkage between genes located between the breakpoints — in some ways, acting as a form of asex within the process of sexual reproduction.

Recombination, and therefore also the suppression of recombination, has major evolutionary consequences. Genetic variation is key for evolution to proceed, and recombination is responsible for around a quarter of this (Navarro, Betrán, et al. 1997).

Recombination breaks associations between loci, for better or for worse. Selection on one allele is influenced by selection on linked alleles, a concept known as 'selective interference' (Sharp and Otto 2016). So, linkage between alleles tends to weaken the efficacy of selection at an individual locus (Hill and Robertson 1966). For example, without exchanging alleles between chromosomes, distinct beneficial mutations can only appear on the same chromosome if each of them independently arises (Fisher 1930), a fairly unlikely event. Otherwise, these alleles on the same chromosome cannot be joined together and end up in competition with each other (Muller 1932). Recombination also enables purifying selection — the purging of deleterious mutations. Without recombination, genomes accumulate deleterious variation that persists throughout lineages, because there is no way for offspring of relatively fit individuals to lose deleterious mutants that are in physical linkage. Eventually, some mutations will be carried to fixation by genetic drift, resulting in the loss of the fittest class of individual. This process, known as "Muller's Ratchet" (Muller 1964; Felsenstein 1974), continues to occur, steadily decreasing the fitness of individuals in the population. However, recombination introduces variance in the mutation load of progeny. Individuals with high mutation loads are likely to be selected out while those with low mutation loads survive. So, deleterious variation is removed more efficiently from the gene pool while enabling the production of the fittest class of individuals.

On the other hand, recombination is often not immediately adaptive. Recombination load is present when the recombinant genotypes are less fit than the original, by separating beneficial allele combinations (Santos 2009). This could be because there is positive epistasis between a set (or sets) of alleles such that they are fitter than alter-

native allele combinations (Haldane 1957; Wasserman 1968). Alternatively, epistasis need not be invoked if locally adapted allele combinations are subject to maladaptive gene flow, so that recombination can produce maladaptive genotypes (Kirkpatrick and Barton 2006). In these cases, localised recombination suppression in the form of an inversion means that specific allele sets can be maintained without losing the benefits of fitness variation brought about by recombination. Thus recombination can be a double-edged sword — a concept known as Felsenstein's dilemma (Faria, Johannesson, et al. 2019).

Inversions lose some of the benefits of sexual reproduction because of the reduced effective rate of recombination. Recombination will only occur in heterokaryotypes unless there is a double crossover, meaning any such events are more likely to involve the middle part of the inversion. Gene conversion between different arrangements is possible too, and can play a significant role in the evolution of the content of inversions, especially when they aren't too large (Navarro, Betrán, et al. 1997; Berdan et al. 2021). As such, the efficacy of purifying selection is comparatively weak, so that inversions are prone to the accumulation of deleterious variation at low frequencies. However, if an inversion manages to reach a frequency high enough to often be homozygous then recombination can proceed as normal again (Barton and Charlesworth 1998; Charlesworth, Harvey, et al. 2000). Furthermore, alleles captured in the ancestral inverted arrangement are likely to persist throughout the lineage even when recombination becomes more common, because they will usually be homozygous in homokaryotypes. In the presence of recessive deleterious alleles, this results in balancing selection on the inversion through associative overdominance (Ohta 1971).

The interactions between inversions and deleterious variation spawned a body of theory in the 1960s. It was recognised early on that inversions that capture a smaller load than the average collinear region have a selective advantage (Nei, Kojima, and Schaffer 1967). However, this advantage is transient. Eventually, mutation catches up with the inversion such that it carries a load similar to the the standard arrangement as well as its captured load, which is fixed within inversions. So, in the absence of any other selective advantage, inversion selection coefficients degenerate until they are negative, and so will ultimately be lost. This conclusion was called "unrealistic" by Kimura and Ohta 1970, who developed influential methods to analyse the fate of an allele whose fitness decreases through time in a finite population. Applying this method to an inversion, with the same rate of decay derived by Nei et al, shows that inversions can capture mutations and still fix. In finite populations, it is not guaranteed that the transient advantage afforded by capturing a better-than-average background will decay prior to fixation. So, the conditions for inversion fixation are less stringent than previously asserted. More recently, multilocus models and simulations have been able to determine the long-term fate of an inversion in large populations (Connallon and Olito 2021), and the relative importance of deleterious variation between sex chromosomes and autosomes (Connallon, Olito, et al. 2018). Another simulation study considered exclusively recessive deleterious mutations, which revealed how different inverted arrangements can diverge, each with different recessive mutations and giving rise to a balanced lethal system (Berdan et al. 2021).

Adaptation and coadaptation

Inversions have long been a subject of interest to evolutionary biologists, occupying a central role in the development of evolutionary thinking. It was a selected inversion polymorphism in *Drosophila pseudoobscura* that inspired Dobzhansky's school of thought that genetic variation could be maintained by balancing selection, rather than purely by mutation-selection balance (Dobzhansky 1955; Charlesworth 2016). Dobzhansky argued that the generation of linkage disequilibrium was the most important consequence of inversions (rather than the creation of reproductive isolation between populations, Kirkpatrick and Barton 2006), and that inversions consisted of sets of coadapted alleles with positive epistasis (Dobzhansky 1947; Dobzhansky and Dobzhansky 1970). Motivated by Dobzhansky's findings that i) polymorphic inversions tended to be fitter in heterozygotes than either homozygote, and ii) sets of genes differed between the same inversion in different populations as well as between the inversion and the collinear region (Dobzhansky 1947; Dobzhansky 1955), early work focused on scenarios in which there was some explicit heterozygote advantage, or positive epistasis between captured alleles (Haldane 1957; Charlesworth 1974). Otherwise, the deleterious effects of inversions dominated discussions on how inversions might fix, relying on founder effects or genetic drift (Feder, Gejji, et al. 2011).

Inversions can be favoured whenever there is selection for linkage disequilibrium between sets of genes. Positive epistasis is one such example. Linkage disequilibrium is also favoured where there is gene flow between differently adapted populations, such that recombinant offspring are less adapted. Some theory had been developed show-

ing that modifiers reducing recombination between alleles in local adaptation models could spread (Charlesworth and Charlesworth 1979; Lenormand and Otto 2000), but it wasn't linked explicitly to inversions until highly influential work by Kirkpatrick and Barton 2006 (although they note that the spread of inversions in a similar model to theirs had previously been shown in simulations by Trickett and Butlin 1994), who showed inversions capturing locally favoured alleles could be favoured by selection in a model where one population experiences homogeneous maladaptive gene flow. This was significant because all prior explanations for how inversions might be adaptive relied on interactions between genes. Inversion research had fallen out of fashion with the rise of molecular genetics in the years preceding this work (Kirkpatrick 2010). However, in the last 20 or so years, powerful genomic methods and sequencing techniques revealed that inversions are far more common than was previously expected (Kirkpatrick 2010; Wellenreuther and Bernatchez 2018), and associated with a vast array of different phenotypes, sparking a revival in interest in the theory of how inversions evolve.

Empirical studies among this recent work has shown that the maintenance of particular allele combinations within inversions allows for the evolution of sophisticated local adaptations controlled by multiple genes, but inherited as one. Such karyotypes are referred to as "supergenes" when they enable the existence of several discrete phenotypes at polymorphism (Thompson and Jiggins 2014). Often, linkage between these allele combinations is maintained by an inversion (Villoutreix et al. 2021), but supergenes could also evolve in areas of low recombination (eg close to the centromere, Entani et al. 1999), through hemizygosity (Li et al. 2016), where there can be

no recombination because there is no corresponding collinear sequence present for a crossover to occur, or through a combination of multiple structural variants (Kim et al. 2021). Examples include the adaptive colouration of deer mice (Harringmeyer and Hoekstra 2022), "wave" and "crest" ecotypes of the *Littorina* sea snail (Faria, Chaube, et al. 2019; Koch et al. 2021), and dune and prairie ecotypes of *Helianthus* sunflowers (Huang, Andrew, et al. 2020). In each of these cases, inversions protect genotypes that confer locally adaptive phenotypes, where hybridisation between ecotypes could result in unfit offspring. One can also find evidence for inversions maintaining epistatic interactions in mimicry supergenes (e.g. Joron et al. 2011) and different reproductive strategies (e.g. Lamichhaney et al. 2016). Unpicking the exact loci within inversions and their respective contributions to supergene-associated phenotypes is challenging (Tigano and Friesen 2016; Villoutreix et al. 2021), because loci need to be separated in order to examine their individual effects. Many of the supergene candidates whose most significant loci have been characterised are related to meiotic drive systems (Villoutreix et al. 2021) — selfish genetic elements that ensure their own propagation to subsequent generations at the expense of other genes (Burt and Trivers 2006). One such example is the t-haplotype, an autosomal meiotic driver in *Mus musculus* comprising hundreds of genes within four overlapping inversions (Ardlie 1998; Schimenti 2000).

Thesis overview

This thesis aims to deepen our understanding of both inversion evolution and inversion-driven evolution. Inversions are a significant evolutionary force and we

now know, common. As sequencing technology has developed, an understanding of structural variation has become more important to make sense of what we see in genomic data (Mérot et al. 2020). Here, I demonstrate how meiotic drive gene complexes linked by inversions can influence population level traits (Chapter 2), how inversions can aid individuals maintain local adaptation under gene flow (Chapter 3), and how the nature of inversion evolution means they may be more likely to fix in small populations (Chapter 4). Together, these works show how the breadth of the capability of inversions to influence evolution across selection at different scales.

When meiotic drivers are linked to sex chromosomes and are expressed in the heterogametic sex, they can result in population sex ratio bias, with consequences for population viability and productivity (Hamilton 1967). In Chapter 2, I determine conditions for polymorphism of an X-linked male meiotic driver under varying sperm competitive ability, mating behaviours, and direct fitness effects, and then explore the consequences maintenance of polymorphism can have on population size and persistence. Being X-linked has the effect of biasing offspring sex ratios towards females. When X-drivers segregate they increase population productivity, and so can increase equilibrium population sizes, and the persistence time of small populations relative to non-drive populations, all else being equal.

Maladaptive gene flow favours inversions that protect locally adaptive allele combinations (Kirkpatrick and Barton 2006; Charlesworth and Barton 2018). Migration generates linkage disequilibrium without the presence of epistasis or other previously invoked phenomena used to explain inversion fixation. These analytical models consider a single population experiencing homogenous migrant gene flow. While

sufficient as a proof of concept, it is hard to apply to many real world scenarios, in which geographically close populations are likely to exchange migrants. In Chapter 3, I extend analyses of inversions in the "continent-island" model (Kirkpatrick and Barton 2006; Bürger and Akerman 2011; Charlesworth and Barton 2018) to a two-deme model (e.g. Akerman and Bürger 2014). In this setting, heterogeneity in gene flow is an emergent property of the model. Furthermore, I determine the overall likelihood of adaptive inversions arising, given the frequency of locally adaptive haplotypes.

During the early stages of inversion establishment, a combination of the suppression of recombination and low frequency severely weakens purifying selection and strengthens the effects of genetic drift. As such, inversions could be prone to the accumulation of deleterious mutations. Existing work addresses how this impacts the long term fate of inversions, when inversions might spend a long time at low frequency or be maintained at intermediate frequency (Connallon, Olito, et al. 2018; Connallon and Olito 2021; Berdan et al. 2021). However, little attention has been paid to cases where the short-term dynamics are relevant, most notably in finite populations. In Chapter 4, I discuss the current state of the field and use simulations to show where early events can prevent the fixation of otherwise adaptive inversions.

X-linked meiotic drive can boost population size and persistence[1]

Abstract

X-linked meiotic drivers cause X-bearing sperm to be produced in excess by male carriers, leading to female-biased sex ratios. Here, we find general conditions for the spread and fixation of X-linked alleles. Our conditions show that the spread of X-linked alleles depends on sex-specific selection and the way they are transmitted rather than the time spent in each sex. Applying this logic to meiotic drive, we show that polymorphism is heavily dependent on sperm competition induced both by female and male mating behaviour and the degree of compensation to gamete loss in the ejaculate size of drive males. We extend these evolutionary models to investigate the demographic consequences of biased sex ratios. Our results suggest driving X-alleles that invade and reach polymorphism (or fix and do not bias segregation ex-

[1] A manuscript based on this chapter has been published in Genetics, with Andrew Pomiankowski and Michael Scott as co-authors (Mackintosh, Pomiankowski, and Scott 2021).

cessively) will boost population size and persistence time by increasing population productivity, demonstrating the potential for selfish genetic elements to move sex ratios closer to the population-level optimum. However, when the spread of drive causes strong sex ratio bias, it can lead to populations with so few males that females remain unmated, cannot produce offspring and go extinct. This outcome is exacerbated when the male mating rate is low. We suggest that researchers should consider the potential for ecologically beneficial side effects of selfish genetic elements, especially in light of proposals to use meiotic drive for biological control.

Introduction

Meiotic drivers violate Mendel's law of equal segregation by ensuring that they are transmitted to more than half of a carrier's progeny (Burt and Trivers 2006). While beneficial at the chromosome-level, this transmission benefit usually comes at a cost to carrier survival or fecundity (Werren 2011). Meiotic drive has been observed across a wide variety of animal and plant taxa (Sandler, Hiraizumi, and Sandler 1959; Turner and Perkins 1979; Jaenike 1996; Ardlie 1998; Taylor, Saur, and Adams 1999; Fishman and Willis 2005; Tao et al. 2007; Lindholm et al. 2016), particularly in flies and rodents (Helleu, Gérard, and Montchamp-Moreau 2015). Many of the described systems are sex-specific (Úbeda and Haig 2005; Lindholm et al. 2016), arising due to activity in either female (e.g., Fishman and Willis 2005) or male meiosis (e.g., Sandler, Hiraizumi, and Sandler 1959). When meiotic drivers arise on sex chromosomes, they change the relative frequencies of gametes carrying the sex-determining alleles, resulting in

brood sex ratio bias (Burt and Trivers 2006). In particular, where X-linked meiotic drivers bias segregation in males, X-bearing sperm outnumber Y-bearing sperm and the sex ratio among offspring is female-biased. Hamilton 1967 noted that extreme sex ratios caused by X-linked meiotic drivers could lead to population extinction, as eventually the almost entirely female population will go unmated and be unable to produce offspring.

Substantial theoretical work since Hamilton's pioneering study (Hamilton 1967) has investigated the spread of meiotic drive, and the conditions that lead to its polymorphism and prevent population extinction. Polymorphism and population persistence are most directly achieved via suppression systems that evolve at other loci to negate meiotic drive (Hamilton 1967; Charlesworth and Hartl 1978; Frank 1991). In the absence of suppression, fixation of autosomal (Ardlie 1998; Larracuente and Presgraves 2012) or X-linked (Taylor and Jaenike 2002; Taylor and Jaenike 2003; Price, Bretman, et al. 2014) meiotic drive can be prevented by direct fitness costs associated with carrying the driving allele. Meiotic drive systems often occur within inversions that link together the required drive and enhancer loci (Pomiankowski and Hurst 1999). These inversions may also capture deleterious alleles and/or allow deleterious mutations to accumulate through Muller's ratchet, potentially explaining the fitness costs associated with meiotic drivers (Edwards 1961; Curtsinger and Feldman 1980; Dyer, Charlesworth, and Jaenike 2007; Kirkpatrick 2010). Such effects have been demonstrated empirically, with female carriers of X-linked meiotic drive observed to have reduced survival or fecundity, especially when homozygous (Larner et al. 2019; Dyer and Hall 2019; Keais, Lu, and Perlman 2020). However, these fitness costs are not

necessarily sex-specific or recessive (Finnegan, White, et al. 2019).

Meiotic drive can also have deleterious effects by reducing male fertility, most obviously because sperm/spores that do not carry the driving element are rendered dysfunctional or killed (Price, Hodgson, et al. 2008). This effect may be negligible when females mate with a single male, but drive can alter competition between the ejaculates of different males in a polyandrous mating system. Not only do drive-carrying males deliver fewer sperm per ejaculate, but drive-carrying sperm can also perform more poorly in sperm competition with sperm from wild-type males (Price, Hodgson, et al. 2008; Manser, Lindholm, et al. 2017; Dyer and Hall 2019; Manser, König, and Lindholm 2020). Offspring sired by drive males have lower fitness which may favour the evolution of increased sperm competition through female polyandry, an argument for which there is some theoretical and experimental evidence (Price, Hodgson, et al. 2008; Wedell 2013; Price, Bretman, et al. 2014; Holman et al. 2015; Manser, Lindholm, et al. 2017), but see (Sutter et al. 2019). The fertility cost to drive males, and associated selection for female polyandry, becomes less important as male frequency declines leading to lower competition for mates and fertilisation (Taylor and Jaenike 2002; Taylor and Jaenike 2003). In line with this, modelling has shown that polyandry can limit the spread of meiotic drive alleles, but the evolution of polyandry is not sufficient to stop meiotic drive alleles fixing (Holman et al. 2015).

The above models have focused on the evolutionary dynamics of meiotic drive but ignored its demographic consequences. This is surprising as in one of the foundational models of the field, Hamilton 1967 showed that sex-linked drive causes transient population expansion before extinction. Population decline occurs when the sex ratio is

pushed beyond the point where females can find sufficient mates. This model did not include density-dependent population regulation or fertility/viability costs associated with meiotic drive. Nevertheless, it suggests that X-linked meiotic drivers will increase population size when they cause sex ratios to be biased, but not extremely biased. Some subsequent analyses support this hypothesis, but it has not been examined directly. Unckless and Clark 2014 showed that species with X-linked meiotic drivers can have an advantage during interspecific competition, shifting the community competition in their favour (James and Jaenike 1990). Similar effects can occur with other systems that cause female-biased sex ratios. For example, feminisation caused by *Wolbachia* can increase population size until females go unmated due to a lack of males (Hatcher et al. 1999; Dyson and Hurst 2004). Finally, under temperature-dependent sex determination, shifts in climate can bias the sex ratio towards females (West 2009), which is predicted to increase population sizes providing males are not limiting (Boyle et al. 2014).

First, we derive new general analytical expressions for the invasion and maintenance of X chromosome variants. The results define the relative weighting of selection in males/females and maternal/paternal transmission, refining the heuristic that X-linked alleles weight their fitness effects twice as strongly in females because they spend twice as much time in females (Patten 2019; Hitchcock and Gardner 2020). We use these results and a simulation-based model to investigate the interplay between female mating rate (polyandry), male mating rate (limits to the number of females each male can mate with) and male sperm compensation (for losses caused by meiotic drive) in the maintenance of X-drive polymorphism. Having established the evo-

lutionary dynamics, we investigate the demographic consequences of meiotic drive and show that drive can cause population sizes to be larger than wild-type populations, enabling them to persist for longer and with lower intrinsic birth rates.

Materials and Methods

We model a well-mixed population with XY sex-determination where generations are discrete and non-overlapping. There are two types of X chromosome segregating in the population, a standard X chromosome and a drive X_d chromosome. There are three female genotypes XX, X_dX and X_dX_d, and two male genotypes XY and X_dY, which we describe as wild-type and drive males respectively. In XY males, meiosis is fair. The X_d chromosome biases segregation such the ratio of X_d to Y chromosomes among their sperm is $(1+\delta)/2 : (1-\delta)/2$. When $\delta = 0$, meiosis is fair and sex chromosomes are transmitted with equal probability; when $\delta = 1$ drive males produce only X_d sperm.

We assume males (whether drive or wild-type) produce sufficient sperm in an ejaculate to fertilise all a female's eggs. Drive males have reduced ejaculate size because Y-bearing sperm are rendered dysfunctional, reducing their success in sperm competition. The ejaculate size of X_dY drive males is determined by the degree of compensation c ($c \in [0,1]$). When $c = 1/(1+\delta)$, there is no compensation for dysfunctional Y sperm. When $c > 1/(1+\delta)$, drive males produce extra sperm in their ejaculate to compensate for those lost through meiotic drive. In the extreme when $c = 1$, drive male ejaculates contain the same number of viable sperm as wild-type males. Compensation affects the success of drive males in sperm competition which is assumed to

female genotype i	w_i^f	$E_{X,i}$	$E_{X_d,i}$	
$i = XX$	1	1	0	
$i = X_dX$	$1-hs_f$	1/2	1/2	
$i = X_dX_d$	$1-s_f$	0	1	
male genotype j	w_j^m	$S_{X,j}$	$S_{X_d,j}$	$S_{Y,j}$
$j = XY$	1	1/2	0	1/2
$j = X_dY$	$1-s_m$	0	$c(1+\delta)/2$	$c(1-\delta)/2$

Table 1: Relative fitness and transmission parameters for different male and female genotypes

follow a fair raffle (Parker 1990). In this paper, we refer to c in the context of ejaculate size, however it can also be interpreted as the competitive ability of drive male sperm. This could apply to cases where sperm have reduced motility, for example.

We track the genotypes of adults, who experience density dependent competition for resources and mate at random before producing offspring. We assume that fertilization follows sperm competition among the ejaculates of all males a female mates with. The resulting offspring experience selection according to their genotype before they become the adults of the next generation. The fitness of each genotype is given by w_i^f and w_i^m, allelic fitness effects in males and females are given by s_f, $s_m \in [0,1]$ and $h \in [0,1]$ determines dominance in females (Table 1).

Analytical model

The total number of adults in the population is given by $N = \sum_i F_i + \sum_j M_j$, where F_i and M_j represent female and male population densities respectively and $i \in \{XX, XX_d, X_dX_d\}$ and $j \in \{XY, X_dY\}$. We assume that competition for resources among adults linearly reduces the fecundity of females. Specifically, each adult female gives birth to $B_N = b(1-\alpha N)$ offspring, where b is the intrinsic female fecundity in the absence of competition and α is the per-individual competitive effect on fecundity. In the absence of meiotic drive or other genotypic effects on fitness, the population size in the next generation is $N' = (b/2)(1-\alpha N)N$ and the equilibrium population size is $\hat{N} = (b-2)/b\alpha$. This form of density dependence can equally apply to intra-specific competition that reduces female survival probability before reproduction. We consider cases where the strength of density dependence is a function of the birth rate in Appendix .

In this model, we consider various degrees of polyandry determined by a fixed integer λ_f: females mate λ_f times, with a male mate chosen uniformly at random. When each female mates once ($\lambda_f = 1$), the adult female densities of genotype ab in the next generation, summed across matings between all possible female i and male j parents, are given by

$$F'_{ab} = \left(\sum_{i \text{ female}} B_N F_i E_{a,i} \right) \left(\sum_{j \text{ male}} \frac{m_j S_{b,j}}{\sum_k S_{k,j}} \right) w^f_{ab}, \tag{1}$$

and the male densities of gentotype aY are given by

$$M'_{aY} = \left(\sum_{i \text{ female}} B_N F_i E_{a,i} \right) \left(\sum_{j \text{ male}} \frac{m_j S_{Y,j}}{\sum_k S_{k,j}} \right) w^m_{aY}, \tag{2}$$

Term	Definition
M_i	Density of males with genotype i
F_i	Density of females with genotype i
m_i	Within-sex frequency of males with genotype i
λ_f	Number of matings before laying eggs in a females' lifetime
$E_{a.i}$	Proportion of eggs of genotype a produced by a female of genotype i
$S_{b,i}$	Units of sperm of type b produced by a male of genotype i relative to wild-type
N	Total density of males and females
α	Per-adult cost to average female fecundity
b	Intrinsic female fecundity (in the absence of competition)
B_N	Female fecundity in a population of size N
c	Ejaculate size of a drive male compared to a wild-type male
δ	Strength of drive

Table 2: Table of terms

where $E_{a,i}$ is the proportion of eggs with haploid genotype a produced by females with diploid genotype i, $m_j = M_j / \sum_k M_k$ is the frequency of males with genotype j, and $S_{b,j}$ is the proportion of sperm with haploid genotype b contributed by males with genotype j (Table 1). That is, diploid parental genotypes are denoted by subscripts i and j for males and females, while subscripts a and b represent haploid chromosomes inherited maternally and paternally, respectively. As there are no parent-of-origin effects, the sum of F'_{X_dX} and F'_{XX_d} is represented simply as F'_{X_dX}. When each

female mates twice ($\lambda_f = 2$), female densities in the next generations are given by

$$F'_{ab} = \left(\sum_{i \text{ female}} B_N F_i E_{a,i} \right) \left(\sum_{j,k \text{ male}} \frac{m_j m_k \left(S_{b,j} + S_{b,k} \right)}{\sum_l \left(S_{l,j} + S_{l,k} \right)} \right) w^f_{ab}, \tag{3}$$

where there is competition for fertilization of each egg among the sperm contributed by two males, firstly with genotype j and then with genotype k. When each female mates many times (λ_f large), the female densities in the next generation approach

$$F'_{ab} = \left(\sum_{i \text{ female}} B_N F_i E_{a,i} \right) \left(\sum_j m_j \frac{M_j S_{b,j}}{M_{XY} + c M_{X_d Y}} \right) w^f_{ab}, \tag{4}$$

where females effectively sample sperm randomly from the total pool of gametes produced by all males in the population. Recursion equations for male densities follow similarly, replacing $S_{b,i}$ with $S_{Y,i}$ and w^f_{ab} with w^m_{aY} in equations Eq(3) and Eq(4). Full derivations can be found in Appendix 1.

Simulation model

The previous model assumes that male matings are not limiting. Population extinction can only occur when the birth rate is low and/or no males remain. In the simulation model, we allow limitations on the mating rate in both female and male matings which are capped by λ_f and λ_m respectively. When an individual reaches the maximum number of matings they cannot mate again. This constraint precludes the possibility that a small number of males can fertilise a large number of females, which is possible in the analytical model. Under these more realistic conditions, it is possible for a population to become extinct because the sex ratio is female biased and there are insufficient males to sustain the population.

As in the analytical model, adult females experience density-dependent competition for resources. In the absence of any competition, females lay b eggs each. In the case where b is non-integer, females lay a mean of b eggs by laying a minimum of $\lfloor b \rfloor$ eggs with a $100(b - \lfloor b \rfloor)\%$ chance of laying one more. Whether or not a birth occurs depends on the competitive influence of other adults, with birth probability $1 - \alpha N$.

The first generation comprises N_0 wild-type individuals at an equal sex ratio, and the driving X_d chromosome is introduced into the population at a proportion q in Hardy-Weinberg equilibrium. Generations then proceed similarly to the previous model. Adults mate randomly until there are either no females or no males available to mate. Assuming they are able to mate, every individual is picked with equal probability. We track the sperm carried by each female as a 3-tuple (x, y, z), representing the quantity of X, X_d, and Y bearing sperm respectively. When a male mates with a female, he adds to the sperm that the female carries. XY males add $(0.5, 0, 0.5)$, and X_dY males add $(0, c(1 + \delta/2), c(1 - \delta)/2)$. Once mating is complete, each egg is fertilised by a sperm sampled randomly, weighted by the probability distribution (x, y, z) after normalisation. The juveniles then undergo viability selection according to their genotypic fitness, with survival probabilities given in Table 1.

There are three main sources of stochasticity present within the simulation model but not in the analytical model. First, the exact sperm that fertilises an egg is sampled at random. Second, juvenile survival to adulthood and the realisation of births is probabilistic. And finally, mating is at random. These three sources can result in fluctuations in genotype frequencies, which can affect the population sex ratio and population size.

Results

Invasion of a rare X chromosome

We first give general conditions for the spread of a rare X chromosome. A rare X-linked allele increases in frequency if

$$\frac{1}{2}w^{f}_{mat} + \frac{1}{2}w^{m}_{mat} * w^{f}_{pat} > 1, \tag{5}$$

where w^{i}_{j} is the relative fitness of the mutant X chromosome in sex i when inherited maternally ($j = mat$) or paternally ($j = pat$). These relative fitnesses include any transmission biases that arise during gamete production or competition, relative to the transmission of the resident chromosome in the same sex. This is a general expression that covers classical models of sex-specific selection on the X chromosome without sperm competition or meiotic drive (e.g. Curtsinger and Feldman 1980; Rice 1984).

A widespread heuristic posits that X-linked alleles weight female fitness components twice as much as male fitness effects because X chromosomes spend twice as much time in females as in males (Patten 2019; Hitchcock and Gardner 2020). This two-thirds to one third weighting is a linear weak selection approximation of Eq (5), in which all the terms become additive. The more-accurate full expression Eq(5) has two parts, reflecting the two pathways via which a rare X chromosome can increase in frequency in females, which are equally weighted (Figure 1). First, X chromosomes can be inherited from mother to daughter (w^{f}_{mat}). Second, X chromosomes in males are always inherited from the mother and will always then be passed to a daughter

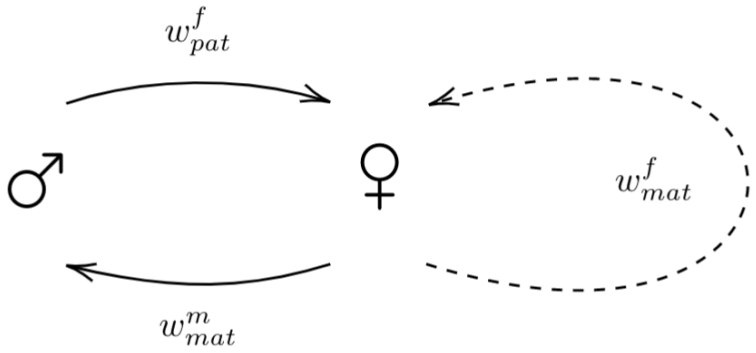


Figure 1: For a rare X chromosome variant to spread in a population, it must increase in frequency in females, which may occur via either of the paths shown. Females transmit X chromosomes (maternally) to either sons or daughters. Sons transmit all X chromosomes (paternally) to females in the next generation

($w^m_{mat} * w^f_{pat}$). If, averaged over these two pathways, the frequency of female carriers increases, then a rare chromosome type will spread in the population. This condition (Eq(5)) shows that the spread of X-linked alleles depends on sex-specific selection and their transmission through the generations rather than the time spent in each sex.[2]

Maintenance of drive polymorphism

We now apply this general condition to a driving X_d chromosome. To remain polymorphic, a rare X_d chromosome must increase in frequency when rare but not fix in the population. That is, $w^f_{mat} = 1 - hs_f$ is the viability of the heterozygous female; $w^m_{mat} = 1 - s_m$ is the viability of the drive male; and $w^f_{pat} = (1+\delta)[c\lambda_f/(c+\lambda_f-1)](1-hs_f)$ is the transmission of meiotic drive alleles through sperm competition and then their viability in female heterozygotes. Combining these terms, the driving X_d chro-

[2]This derivation and interpretation was done by Michael Scott, after the results particular to this paper were obtained.

mosome spreads if

$$\frac{(1-hs_f)}{2}\left(1+\left[\frac{c\lambda_f}{c+(\lambda_f-1)}\right](1+\delta)(1-s_m)\right)-1>0. \quad (6)$$

The success of a rare drive allele in sperm competition is $c/(c+(\lambda_f-1))$, given that a female mates with a single drive male and λ_f-1 wild-type males. Across all matings, the relative success of rare drive alleles during sperm competition is given by the term in square brackets.

Using the same logic, the driving X_d chromosomes will not fix in the population if

$$\frac{(1-hs_f)}{2(1-s_f)}\left(1+\left[\frac{\lambda_f}{1+c(\lambda_f-1)}\right]\frac{1}{(1+\delta)(1-s_m)}\right)-1>0. \quad (7)$$

As X chromosome meiotic drive (X_d) becomes common, the transmission and fitness advantage/disadvantage of X_d chromosomes in males is unchanged (terms involving δ and s_m). The sperm competition term (in square brackets) now reflects the relative competitiveness of sperm from non-drive males.

Importantly, close to fixation, most females are either heterozygous or homozygous for meiotic drive and, unlike Eq(6), Eq(7) depends on these relative female fitnesses. The maintenance of polymorphism (satisfying inequalities in both Eq(6, 7)) occurs when meiotic drive causes low fitness cost in female heterozygotes (hs_f) relative to the cost in female homozygotes (s_f), which allows invasion but prevents fixation. For example, meiotic drive alleles are less likely to reach fixation when the negative fitness effects of drive are recessive ($h=0$, Figure 2).

Sperm competition affects the dynamics of rare X-alleles through a combination of polyandry (λ_f) and any reduction in ejaculate size caused by drive (c) (Figure 2). If

466 females mate with only one male ($\lambda_f = 1$) then sperm competition has no effect. The same holds if drive males produce the same amount of sperm as wild-type males
468 ($c = 1$) (Figure 2). In both cases, the sperm competition term in the square brackets of Eqs(6-7) is equal to 1. At the other extreme, where females mate many times
470 ($\lambda_f \to \infty$) the sperm competition term becomes c - the relative ejaculate size of drive males. If there is also no compensation for Y-bearing sperm killed by meiotic drive
472 alleles ($c = 1/(1+\delta)$), meiotic drive cannot invade (Figure 2B). Between these extremes, increases in polyandry (larger λ_f) and decreases in compensation in drive
474 males (smaller c) hinder both invasion and fixation of meiotic drive alleles (Figure 2). Sperm competition is most important when there is both extensive polyandry and a
476 large reduction in ejaculate size caused by meiotic drive (Figure 2).

Limiting male matings narrows the polymorphism space

478 In the results presented above, we assume that there is no sperm limitation, so even a small number of males is capable of fertilizing a large female population. In this case,
480 extinction by meiotic drive only occurs when there are no males left in the population.

Here, we use the simulation model to consider limitations on the number of matings
482 that a male can perform. First, we compare the proportion of numerical simulations that result in drive polymorphism to the predictions from the analytical model, where
484 there are no limits to male mating. With male mating set at $\lambda_m = 20$ (Figure 3A), the region of polymorphism shrinks (the orange tiles do not completely fill the theoreti-
486 cal polymorphism space). On the upper boundary, this represents conditions where

the polymorphism is unstable because meiotic drive alleles have only a slight advantage and remain at low frequencies where they are exposed to loss by genetic drift. The leftmost boundary is where drive is strong enough to reach a high frequency and the sex ratio is heavily female biased, so many females go unmated due to male mating limitation, and the population can go extinct. When the maximum number of matings per male was reduced to $\lambda_m = 2$ (Figure 3B), just as many lose drive stochastically (on the upper boundary). But the problem of females remaining unmated is exacerbated. More populations go extinct close to the fixation boundary, with fewer simulations resulting in polymorphism. Thus, we predict that population extinction is likely when male mating rates are low and strong meiotic drive alleles reach high frequencies.

Population size in the presence of drive

By creating female biased sex ratios, meiotic drive can influence population size. Figure 4 illustrates two different outcomes when drive spreads (extinction and polymorphism). As a base for comparison, parameter values are chosen under which a wild-type population is stably maintained (Figure 4A). When a driving X allele is introduced into the population it rapidly increases in frequency. This can skew the sex ratio further and further towards females until extinction ensues because there are insufficient males to fertilise all the females (Figure 4B). When the fitness costs of drive in females are higher, drive can be stably maintained. The resulting population is female-biased and larger than it would be in the absence of drive because the higher proportion of females increases the productivity of the population (Figure 4C).

In the absence of meiotic drive ($p = 0$), the population reaches an equilibrium population size ($\hat{N}$) given by the intrinsic birth rate (b) and the density-dependent reduction in female fecundity caused by competition among individuals (α):

$$\hat{N}|_{p=0} = \frac{b-2}{b\alpha}, \tag{8}$$

which is a standard result for logistic population growth with non-overlapping generations (pp.44-46 with $r = b/2$ and $d = b\alpha$ Edelstein-Keshet 1987). The equilibrium population size is larger when the intrinsic birth rate (b) is higher or the competitive effect of other individuals (α) is weaker. For the population to persist, each female must produce at least two offspring ($b_{min}|_{p=0} = 2$).

To derive the population size with meiotic drive, we focus on the case where females mate only once, excluding the effects of sperm competition. First, we define ϕ and ψ as the LHS of Eq(6) and Eq(7) respectively, with $\lambda_f = 1$; ϕ gives the selective advantage of drive alleles when rare and ψ is the advantage of wild-type alleles in a population fixed for drive. If an X chromosome meiotic driver invades (i.e. $\phi > 0$, Eq(6)) and reaches a polymorphic equilibrium (i.e. $\psi > 0$, Eq(7)) then its frequency in females and males is given by

$$\hat{p}_f = \frac{\phi}{\phi + \psi}, \tag{9}$$

$$\hat{p}_m = \frac{(1-s_m)\phi}{(1-s_m)\phi + \psi}. \tag{10}$$

At the polymorphic equilibrium, the sex ratio will be female-biased and this in turn affects the ecological equilibrium population size

$$\hat{N}|_{p=\hat{p}} = \frac{b^* - 2}{b^*\alpha}, \tag{11}$$

where

$$b^* = b(1+\phi p_f/2)\frac{1-p_m}{1-p_f} > b. \tag{12}$$

b^* is the effective birth rate given the change in the sex ratio caused by meiotic drive. The effective birth rate with drive is higher, $b^* > b$, because ϕ and p_f are non-negative and $p_m \leq p_f$ (from Eq(10)). The effective birth rate is increased by a factor equal to the number of females surviving to reproductive age (given the equilibrium frequency of drive) relative to the number of females in a wildtype population (see File S1). As $b^* > b$, the population size with drive is always larger than it would have been without drive (Figure 5A). Drive populations effectively behave like wild-type populations with a higher birth rate, as a result of the sex ratio bias.

A similar outcome holds when a drive allele fixes. The total population size is

$$\hat{N}|_{p=\hat{p}} = \frac{\tilde{b}-2}{\tilde{b}\alpha}, \tag{13}$$

where

$$\tilde{b} = b(1+\delta)(1-s_f). \tag{14}$$

For drive alleles that reach fixation, $\tilde{b} > b$. Again, by biasing the sex ratio towards females, fixed drive increases the population birth rate and thereby increases the overall population size (Figure 5A). However, this result may be most relevant for weak meiotic drivers ($\delta < 1$) because there will be no males in the population when strong meiotic drivers ($\delta \approx 1$) reach fixation.

By increasing population productivity, meiotic drive alleles also help to protect populations from extinction. With strong drive at an intermediate equilibrium frequency, the minimum intrinsic birth rate required for population persistence is $b_{min}|_{p=\hat{p}} =$

538 $2/(1+\hat{p}_f\phi)$, while for weak drive at fixation this is $b_{min}|_{p=\hat{p}} = 2/(1-s_f)(1+\delta)$. Both of these values are less than two, the cut-off value for a population to go extinct in
540 the absence of drive. Populations with drive can persist with a lower average number of offspring per female than those without, because a higher proportion of the
542 population are female. The results of the simulation model align with the analytic model. Whenever a polymorphism is reached, the resulting population size is bigger
544 than in the absence of drive (Figure 5B). The extent of the boost in population size depends on the viability cost associated with drive. As the cost decreases (either h or s_f
546 decreases), the equilibrium frequency of drive increases, the sex ratio becomes more female biased, and the increase in population size becomes larger. Overall, these sim-
548 ulations confirm that meiotic drive can boost population size even when males can only fertilize a limited number of females.

550 Population persistence time

Populations that are relatively small are liable to go extinct within a reasonable time
552 due to demographic stochasticity. To examine the effect of drive on persistence times simulations were run in small populations with a low intrinsic birth rate ($b = 2.4$,
554 $\alpha = 10^{-2.4}$), reflecting for example a small patch in a suboptimal or marginal environment. In these simulations, the mean population size without meiotic drive was
556 $\bar{N} \pm s.d. = 36.3 \pm 12.7$ (consistent with the expected population size from Eq(8), which is $\hat{N} = 41.9$) and the persistence time was mean $\pm s.d. = 1088 \pm 1001$ generations. The
558 approximate alignment of the mean and standard deviation is expected because the persistence times of stochastic logistic growth models are exponentially distributed

560 (Ovaskainen and Meerson 2010).

First, we consider the case where meiotic drive has no fitness costs ($s_f = s_m = 0$) and
562 either spreads to fixation or is lost by drift (Figure 6A). With $\delta = 0$ (i.e. no transmission distortion), the X_d allele is completely neutral and the population persists as if it were
564 wild-type (Figure 6A). For increasingly strong meiotic drivers (increasing δ), the probability of invasion increases and meiotic drive alleles are present at the end of more
566 simulations, causing populations to persist for longer. In this example (Figure 6A), the male mating rate is high ($\lambda_m = 20$), so there are sufficient males to maintain female
568 fecundity and resist extinction, even with strong drive (Figure 6A). However, when drive is very strong ($\delta \geq 0.8$), the sex ratio can become excessively female biased and
570 population extinction becomes more likely.

Population persistence was also evaluated for strong meiotic drivers ($\delta = 1$). For sim-
572 plicity, the dominance coefficient in females was set to $h = 0$, limiting viability reduction to homozygous female carriers (Figure 6B). When drive incurs no or small fitness
574 costs ($s_f < 0.2$), it spreads to fixation and causes rapid extinction through extreme sex ratios. As the cost increases ($0.2 < s_f \leq 0.5$), meiotic drive spreads more slowly and the
576 persistence time increases back towards that found in wild-type populations. Eventually, with higher cost ($s_f > 0.5$), drive does not fix. Here, the sex ratio is skewed towards
578 females but there are sufficient males, leading to longer population persistence than wild-type populations. Where the cost is very high ($s_f > 0.7$), drive is maintained at a
580 low frequency and may itself be stochastically lost. However, the transient presence of drive still increases the overall longevity of the population.

These two examples demonstrate how drive increases population persistence until sex ratio biases are so strong that the males cannot fertilise all the females. The effect of drive on population persistence depends on its frequency and thus the sex ratio bias created. As outlined in our evolutionary analysis above, other parameters affect the frequency of meiotic drive alleles (dominance, male fitness effects, polyandry, ejaculate size compensation) and have corresponding effects on population persistence.

Discussion

This paper sets out a general condition for the spread, polymorphism and fixation of X-linked alleles, Eq(5), which we apply to the study of the evolutionary dynamics of meiotic drive. There are two equally important pathways by which X-alleles spread: either from mother to daughters, or from mother to sons and then into granddaughters (Figure 1). Our condition shows that the success of X-linked alleles depends on sex-specific selection as well as the asymmetric transmission through the sexes. If selection is weak, female fitness effects are twice as influential, as X chromosomes spend twice as much time in females as in males (Patten 2019; Hitchcock and Gardner 2020). But this 2:1 rule does not apply when selection is strong, as is likely to be the case in meiotic drive.

A central finding is that X-linked meiotic drivers generally increase population size. By biasing the sex ratio towards females, meiotic drive effectively boosts the population birth rate which is typically limited by the number of females (Eq(12,14)). This

increases the expected population size beyond the level in wild-type populations (Figures 4 and 5). In small populations at risk of stochastic population extinction, the increase in population size through meiotic drive can dramatically increase population persistence time (Figure 6). This should enable populations to persist in marginal environments where they would otherwise go extinct. The population-level benefit of drive breaks down when males become limiting and are no longer able to mate often enough for females to achieve full fecundity (Figure 6). Previous work (Pominankowski and Hurst 1999; Taylor and Jaenike 2002; Taylor and Jaenike 2003; Dyer and Hall 2019; Larner et al. 2019) has shown that female fitness in drive heterozygotes and homozygotes affects the frequency of meiotic drive alleles. We show the additional dependence on the female (Figure 2) and male mating rate (Figure 3), and how this then impacts the sex ratio bias, population size and persistence time of populations invaded by meiotic drive alleles (Figures 4, 5, 6). We find that the male mating rate (λ_m) is key to determining whether meiotic drive cause population extinction. When males can mate repeatedly, their rarity does not cause sperm limitation amongst females and the distortion in the sex ratio is beneficial to population persistence. Limits on the number of females each male can mate with cause some females to go unmated resulting in population extinction as meiotic drive spreads and skews the sex ratio. This higher likelihood of extinction narrows the space in which meiotic drive is likely to occur as a polymorphism in natural populations (Figure 3).

Most previous work has concentrated on the consequences of female rather than male mating rates, that is polyandry (λ_f), as this is a cause of sperm competition that hinders the spread of meiotic drive alleles (Price, Hurst, and Wedell 2010; Price, Bret-

man, et al. 2014; Holman et al. 2015). Our work shows that this is only the case when ejaculate size is significantly reduced in male meiotic drive carriers (Figure 2). Generally, as compensation increases (i.e. $c \to 1$), so does the likelihood of polymorphism, because drive male success in sperm competition reaches towards that of wild-type males. In the modelling, we consider drive males to have lower fertility because of reductions in ejaculate size (proportional to the strength of drive δ). The same logic applies to other mechanisms that might disadvantage the success of drive males in sperm competition, like slower sperm swimming speeds or reduced sperm longevity (Olds-Clarke and Johnson 1993; Kruger et al. 2019; Rathje et al. 2019).

Although there are few empirically obtained estimates for the fitness costs of X-linked drive, many of them are compatible with polymorphism according to our model. Female viability costs in *Drosophila* are often recessive but strong ($h = 0 - 0.11, s_f = 0.56 - 1$, see Table 1 in (Unckless and Clark 2014) and (Larner et al. 2019; Dyer and Hall 2019)). A counterfactual is the estimate from the stalk-eyed fly *Teleopsis dalmanni* (Finnegan, Nitsche, et al. 2019) which found additivity and weaker viability loss in egg-to-adult viability, though the range on the dominance estimate is large. A limitation of attempts to measure fitness is that they are based on laboratory conditions that may distort the pressures that exist in natural populations. They also typically measure one component of fitness, for example survival over a particular life stage, neglecting others such as reproductive success. Furthermore, we note that these empirical estimates may be biased towards systems with strong meiotic drive ($\delta \approx 1$) because weak meiotic drivers are less easy to detect (Burt and Trivers 2006).

Population persistence is predicted to increase exponentially with population size

(Ovaskainen and Meerson 2010) (Figure 6). Therefore, we predict that populations with meiotic drive are more likely to be observed in marginal habitats where wild-type populations may go extinct. In natural populations, tests of this prediction may be confounded by a range of other factors associated with marginal habitats. For instance the rate of polyandry is likely to be lower in poor quality environments and this will favour the spread of drive (Pinzone and Dyer 2013; Finnegan 2020). A viable first experimental step may be to use lab populations to evaluate whether X-linked meiotic drive can increase population birth rates and/or rescue declining populations from extinction.

A relationship between sex ratios and population size/persistence is also not yet clearly established in species with temperature-dependent sex determination, despite similar predictions (Boyle et al. 2014; Hays et al. 2017). As predicted previously (Hamilton 1967), severely male limited populations should be quickly driven to extinction, which can occur in lab populations (Price, Hurst, and Wedell 2010) and may have been observed in a natural population (Pinzone and Dyer 2013). However, high male mating rates can facilitate population persistence in the face of extremely biased sex ratios. A *Wolbachia* infection in butterflies resulted in a sex ratio of 100 females per male, but these populations persisted perhaps because males can mate more than 50 times in a lifetime (Dyson and Hurst 2004).

The population dynamics of sex ratio distorting elements are thought to be influenced by their propensity to colonise new patches and drive them to extinction, i.e., metapopulation dynamics (Hatcher 2000). When drive is strong and confers little fitness cost in females, new populations cannot be established by drive geno-

types because of the deficit in the numbers of males and resulting weak population growth. This could lead to cycling dynamics where colonisation by non-drive genotypes is needed to establish populations, which can then be invaded by drive genotypes whose spread is followed by extinction (Taylor and Jaenike 2003). These population level costs can decrease the overall frequency of selfish genetic elements across the metapopulation (Boven and Weissing 1999). Our results emphasise the potential for X-linked meiotic drivers to boost population sizes and persistence times, which we expect would increase the proportion of patches expected to have drive. It has also been suggested that individuals carrying selfish genetic elements may show a greater propensity to migrate between populations, increasing their fitness by reaching patches with lower numbers of heterozygotes and less polyandry (Runge and Lindholm 2018). However, the full metapopulation dynamics where local population sizes are affected by drive frequency remains to be investigated.

We generally predict population size to be increased when the sex ratio is biased towards females. Thus we expect our results to hold in species with ZW sex determination when meiotic drive favours W chromosomes (Kern et al. 2015) but not when meiotic drivers favours Y chromosomes or Z chromosomes (Hickey and Craig 1966; Gileva 1987). A general constraint on our conclusions is that they hold for competition models where an increase in birth rate increases population size (Supplementary Information). If the population is limited by the availability of resources regardless of the birth rate, boosts in population size are not expected. Likewise, where males contribute to parental care either through direct care or via control of resources used by females, sex ratio distortion will not have such a profound effect because the expected

change in the number of offspring produced will be reduced and have a lesser effect on population size and persistence (West 2009). A further caveat of these results is that they assume density dependent selection is contributed to equally by both sexes. Where males contribute less than females the sex-ratio skew will have a lesser impact on population size. There may also be cases where increased birth rates cause competition to become increasingly intense and reduce population size. An example is given in the Supplementary Information, where drive counter-intuitively decreases population size by increasing the effective birth rate beyond a critical level (see Figure S1). Although this pattern of density dependence seems likely to be atypical, it points to the need for the biological details of particular species to be taken into account.

Our results are also pertinent to the design of synthetic gene drive systems. Gene drive systems have been proposed as a method of controlling pest populations through altering the sex ratio so that one sex becomes limiting. Many of these proposals are analogous to Y-linked meiotic drive, for example "X-shredders" (Windbichler, Papathanos, and Crisanti 2008; Galizi et al. 2014; Burt and Deredec 2018) that limit the reproductive output of the population by biasing segregation towards Y-bearing sperm. We expect systems that cause male sex ratio bias to be effective. X-drive has also been recently suggested as a tool for biological control (Prowse et al. 2019). As observed in some simulations, as long as males are not limiting, the population may benefit from the introduction of an X-drive that increases the population productivity and carrying capacity (Prowse et al. 2019). That is, less efficient synthetic X-drivers may fix and result in larger populations without causing populations to crash (Prowse et al. 2019); this is analogous to fixation of weak meiotic drive in our model. Another possibility

is that the driving allele does not fix but is maintained at a polymorphic equilibrium by the evolution of suppressors or associated fitness costs, for example. The resulting population will have a female-biased sex ratio, which our results suggest could increase population size and persistence. Thus, we urge caution when considering the use of X-linked gene drive for population control.

At the population level, the optimal sex ratio is likely to be female biased because relatively few males are required for complete fertilization. In some circumstances, such as local mate competition, individual-level and group-level selection can align, and female-biased sex ratios can evolve (West 2009; Hardy and Boulton 2019). Here, we show that selfish genetic elements (specifically, X-linked meiotic drivers) can move populations towards their population-level optimum and benefit population-level traits (such as population size and persistence time), a possibility that has probably been under-emphasised relative to their detrimental effects on populations.

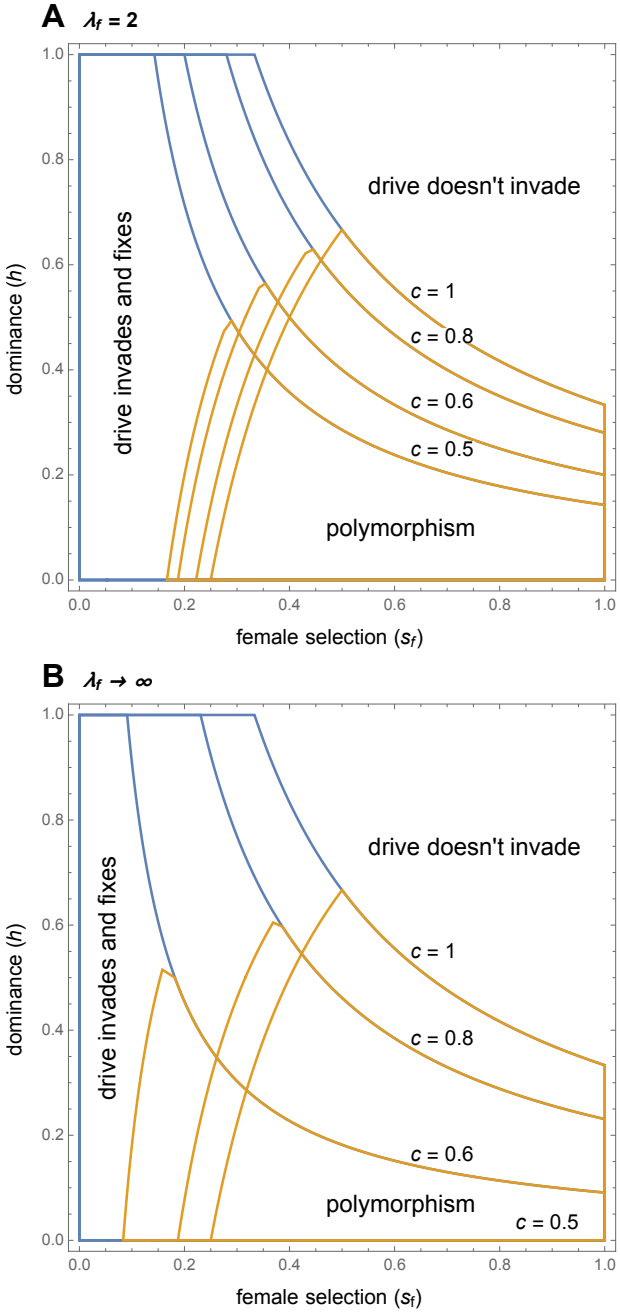


Figure 2: Fitness parameters under which X chromosome meiotic drive invades, reaches polymorphism (orange border), or fixes (blue border), given different levels of sperm compensation (c). Boundaries at $c = 1$ (full compensation) are equivalent to the condition of a single female mating ($\lambda_f = 1$). In A), females mate twice ($\lambda_f = 2$), in B) females mate many times, effectively sampling at random from all male sperm produced. If females mate many times and there is no sperm compensation ($c = 0.5$), then polymorphism is not possible. Other parameter values: no fitness effects in drive males ($s_m = 0$) who only produce X_d-bearing sperm ($\delta = 1$).

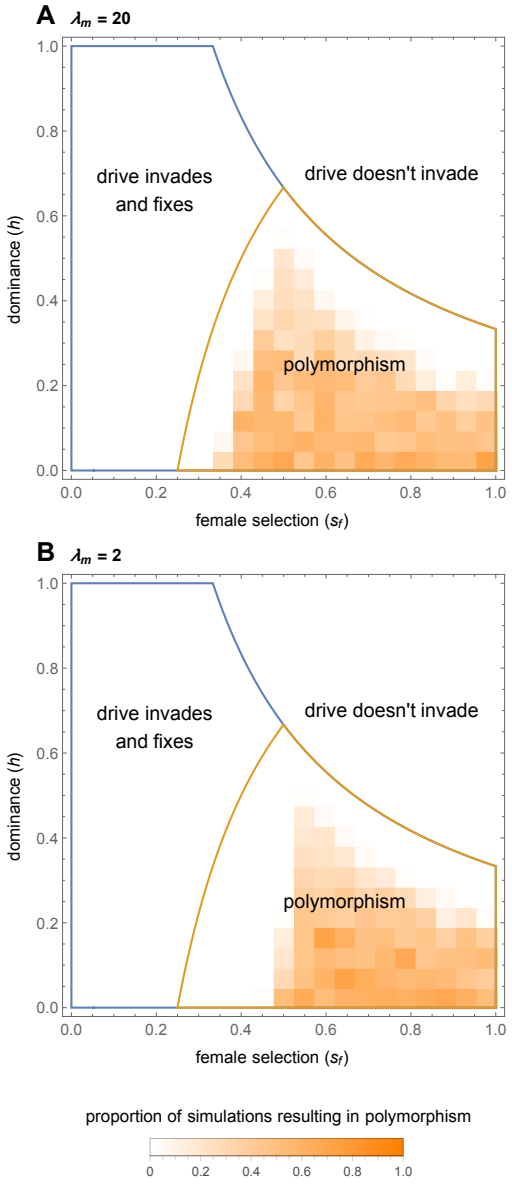


Figure 3: The effect of the male mating rate. Numerical simulation data showing the proportion of times (out of 50 simulations) that a polymorphism was maintained for 2000 generations when A) males mate 20 times ($\lambda_m = 20$) and B) males mate twice ($\lambda_m = 2$). The region of polymorphism is demarcated on the assumption that there are no constraints on male mating (area within orange line). The simulation parameters used were $\delta = 1$, $c = 1$, $\lambda_f = 1$, $q = 0.01$, $N_0 = 200$, $b = 2.4$, $\alpha = 0.001$.

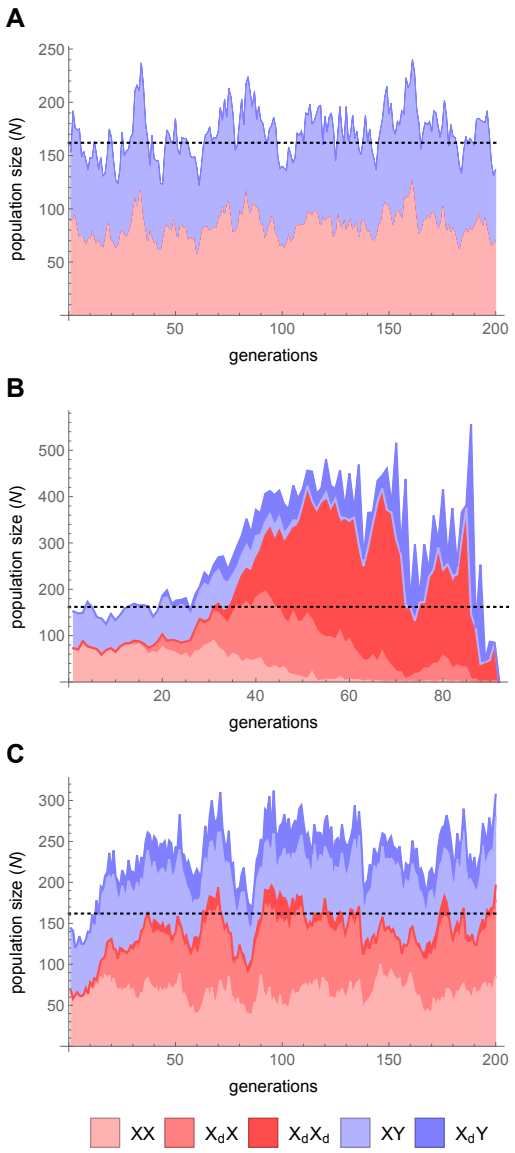


Figure 4: Illustrative examples of population dynamics with and without drive. A) the wild-type population without drive, B) the addition of drive causing rapid population extinction ($h = 0.4, s_f = 0.2$), C) the addition of drive subject to stronger counter selection leading to a population polymorphic for drive ($h = 0.2, s_f = 0.55$). Female genotypes are shown in red, and male in blue. The mean wild-type population size was 161 and is shown by the dotted line (analytical model predicts approximately 167). Other parameters used were $c = 0.75, \delta = 1, b = 2.4, \alpha = 10^{-3}, \lambda_f = 2, \lambda_m = 20, q = 0.01$, and the initial population size was 150.

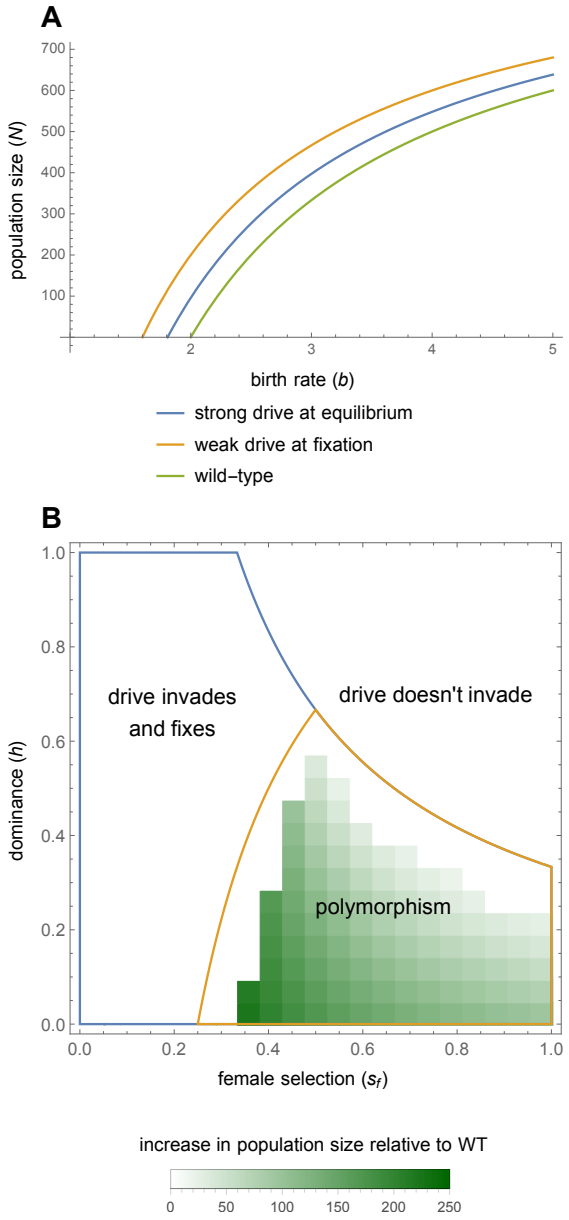


Figure 5: Population size is increased with meiotic drive. A) Two examples where populations with meiotic drive have higher population size and persist with lower intrinsic growth rates ($b < 2$). The first when drive is weak and at fixation ($\delta = 0.25$) and the second when drive is strong and at equilibrium ($\delta = 1$). Other parameter values: $s_f = 0$ for weak drive, $s_f = 0.8$, $h = 0.1$ for strong drive, $s_m = 0$, $c = 1$, $\alpha = 10^{-3}$. B) The average increase in population size compared to a wild-type population without meiotic drive for the data in Figure 3A. The population size for each simulation was taken to be the mean size after a 100 generation burn in period, and the value for each tile in the plot is the mean of those simulations that resulted in polymorphism from a sample of 50.

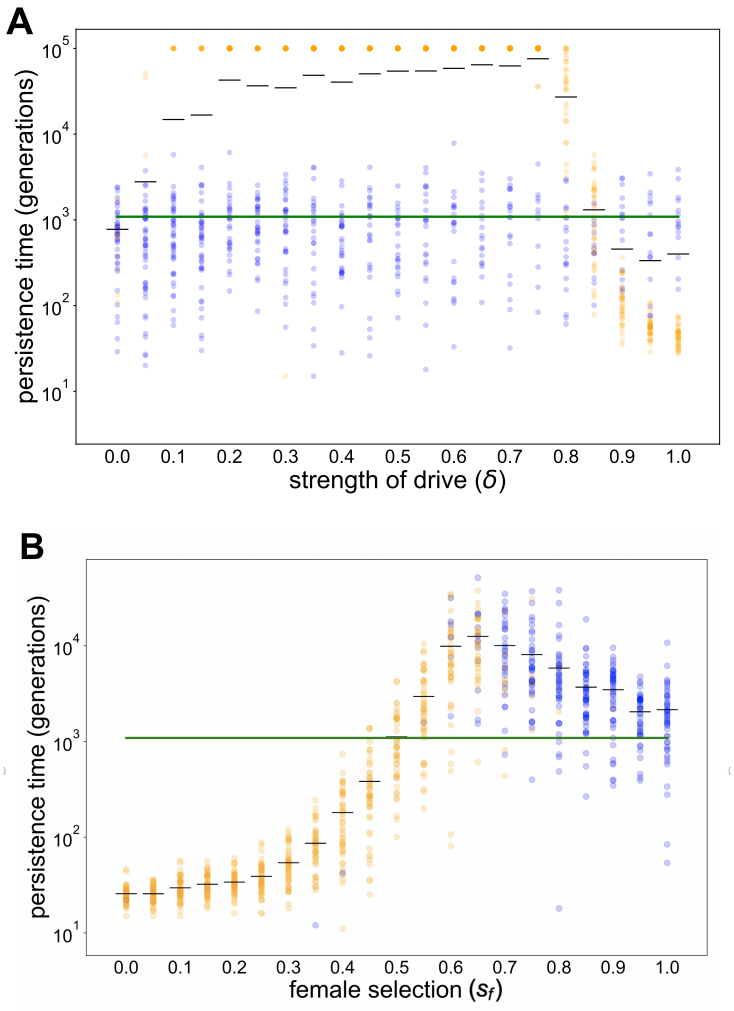


Figure 6: Persistence times for populations as A) the strength of drive increases (δ), and B) the strength of selection in females increases (s_f). Orange points denote populations where drive was present and blue points where drive was absent at the time of extinction or at the maximum simulation duration of 10^5 generations. The green line represents the mean persistence time of wild-type populations without meiotic drive and the black lines show mean persistence times. Populations began with an initial drive frequency of $q = 0.1$. Female adults had a mean birth rate of $b = 2.4$ with a high cost of competition, $\alpha = 10^{-2.4}$. In A) $s_f = 0$, drive acts by killing a fraction of Y sperm with no compensation ($c = 1/(1+\delta)$) and in B) viability costs were in homozygotes only ($h = 0$), males produced only X_d sperm and had full compensation ($\delta = c = 1$). Other parameter values $s_m = 0, \lambda_f = 2, \lambda_m = 20$.

Locally adaptive inversions in structured populations [3]

Abstract

Inversions have been proposed to facilitate local adaptation, by linking together locally coadapted alleles at different loci. Classic prior work addressing this question theoretically has considered the spread of inversions in "continent-island" models in which there is a unidirectional flow of maladapted migrants into the island population. In this setting, inversions are most likely to establish when selection is weak, because stronger local selection more effectively purges maladaptive alleles, thus lessening the advantage of inversions. Here, we show this finding only holds under limited conditions. We study the establishment of inversions in a "two-deme" model, which explicitly considers the dynamics of allele frequencies in both populations linked by bidirectional migration. For symmetric selection and migration, we find

[3] Work based on this chapter is published as a preprint on bioRxiv and has been submitted to an academic journal, with Michael Scott, Max Reuter, and Andrew Pomiankowski as coauthors.

that stronger local selection increases the flow of maladaptive alleles and favours inversions, the opposite of the pattern seen in the asymmetric continent-island model. Furthermore, we show that the strength and symmetry of selection also change the likelihood that an inversion captures an adaptive haplotype in the first place. Considering the combined process of invasion and capture shows that inversions are most likely to be found when locally adaptive loci experience strong selection. In addition, inversions that establish in one deme also protect adaptive allele combinations in the other, leading to differentiation between demes. Stronger selection in either deme once again makes differentiation between populations more likely. In contrast, differentiation is less likely when migration rates are high because adaptive haplotypes become less common. Overall, this analysis of evolutionary dynamics across a structured population shows that established inversions are most likely to have captured strongly selected local adaptation alleles.

Introduction

Chromosomal inversions are a form of structural variant that suppress recombination between loci. Inversions can result in reduced fitness due to the disruption of genes around their breakpoints (Kirkpatrick 2010), or from the capture and accumulation of deleterious alleles due to their lower effective recombination rate (Wasserman 1968; Berdan et al. 2021). Furthermore, inversion heterozygotes may experience reduced fecundity as a result of improper meiosis that results in aneuploid gametes (White 1978). Despite these negative fitness effects, the ubiquity of inversions has led to

several putative explanations for their continued persistence (see reviews Kirkpatrick 2010; Wellenreuther and Bernatchez 2018; Faria, Johannesson, et al. 2019; Huang and Rieseberg 2020; Villoutreix et al. 2021). In particular, inversions could facilitate local adaptation under gene flow by increasing linkage between coadapted alleles and reducing effective migration of maladapted haplotypes (Kirkpatrick and Barton 2006).

Empirical evidence for this hypothesis has since been documented across a wide array of taxa (e.g. Lowry and Willis 2010; Cheng et al. 2012; Ayala, Guerrero, and Kirkpatrick 2013; Lee et al. 2017; Christmas et al. 2019; Faria, Chaube, et al. 2019; Huang, Andrew, et al. 2020; Koch et al. 2021; Hager et al. 2022; Harringmeyer and Hoekstra 2022), and a body of related theoretical work has also developed from the original model, investigating the roles of geography, chromosome type, and inversion length on the fate of adaptive inversions (Feder, Gejji, et al. 2011; Charlesworth and Barton 2018; Connallon, Olito, et al. 2018; Connallon and Olito 2021; Proulx and Teotónio 2022). For simplicity, this work often considers a "continent-island" model, in which inversions are introduced into an "island" population which receives maladapted migrants from a larger "continent" population. In this model, the selective advantage of an adaptive inversion is proportional to the rate of gene flow (Kirkpatrick and Barton 2006), and inversely proportional to the strength of selection on the island (Bürger and Akerman 2011; Charlesworth and Barton 2018). These results rely on the homogeneous maladaptation of migrant alleles which follows from the extreme migration asymmetry assumed between the continent and island populations (Kirkpatrick and Barton 2006). This scenario is unlikely to apply to many empirical systems, where local adaptation occurs in a structured population with greater symmetry and indi-

viduals migrate between similarly sized populations at rates that are similar to and from each population (e.g. Feder, Gejji, et al. 2011, Proulx and Teotónio 2022). With two-way dispersal, selection will interact with migration to determine the overall rate of maladaptive gene flow. However, there has been no thorough analytical dissection of the roles that migration and selection play individually in such a model.

In addition, it is important to consider not only whether an inversion spreads but also how the frequency of adaptive haplotypes affects their probability of being captured by an inversion. This has been briefly discussed before (Kirkpatrick and Barton 2006), and in relative terms when comparing X-linked and autosomal inversions (Connallon, Olito, et al. 2018). But so far models have sidestepped the problem by assuming that either an inversion capturing the coadapted haplotype simply existed or that such an inversion arose during a period of allopatry (Feder, Gejji, et al. 2011). Explicitly modelling the origin of the inversion is important because parameters favourable for the establishment of an adaptive inversion are not necessarily those where adaptive inversions are likely to arise. Assuming an inversion captures a random genotype, the probability of capturing a particular adaptive combination is proportional to its frequency. For example, adaptive inversions are expected to be favoured most when there are high rates of migrant gene flow, so there are fewer fit genotypes to be captured.

Here, we model the fate of locally adaptive chromosomal inversions in a two-locus, two-allele, two-deme model with migration and selection. We consider the case of symmetrical deme sizes and migration, as well as asymmetrical scenarios with the continent-island model as the extreme case. To understand the dynamics of inver-

sions, we determine the probability of an adaptive inversion arising and its subsequent selective advantage in a population in which the locally adaptive alleles have reached their equilibrium frequencies and linkage under migration and selection. By considering the processes of inversion origin and spread in both demes, we determine population structures which favour the evolution of inversions that allow local adaptation under environmentally variable selection.

Methods

We consider a population consisting of two demes linked by bidirectional migration with selection for local adaptation. We first derive analytical expressions for equilibrium allele frequencies at the local adaptation loci and the linkage disequilibrium (LD) between them. This will allow us to assess the frequency of each haplotype and hence the invasion probability of an inversion capturing a locally adapted combination of alleles. We then determine the probability of such an inversion arising and establishing itself in the population.

Model

We model an infinite population of two demes, consisting of haploid, hermaphroditic individuals with discrete non-overlapping generations. The model is equally applicable to the case where there are two sexes at even sex ratio whose genetic determination is unlinked to the adaptive loci under consideration. Selection acts on two loci,

A and B, that have two alleles each, A_i and B_i, where $i \in \{1,2\}$ denotes the deme in which the allele provides a benefit s_i (equal between the two loci). The relative fitness of an individual in deme i is either $(1+s_i)^2$, $(1+s_i)$ or 1, depending on whether it carries two, one or no allele(s) conferring local adaptation to its environment.

The life cycle begins with adults. These individuals reproduce, whereby pairs of parents are sampled according to their relative fitness in their current deme to produce one joint offspring. During reproduction, recombination occurs between the parental chromosomes (and their loci for local adaptation) at rate r. When alleles are held in an inversion, the recombination rate with non-inverted chromosomes drops to zero (double cross-overs and gene conversion are ignored). Migration between demes then occurs such that a proportion m_{kl} of juveniles in deme l are migrants from deme k. After migration, the juveniles in each deme become the adults of the next generation. As the life cycle consist of just two phases, reproduction/selection and dispersal, the order of events within a generation does not affect the results.

At the beginning of a generation, A_iB_j adults in deme k are at proportion p_{ij}^k and have fitness w_{ij}^k. Among the parents sampled for reproduction, the frequencies are $\tilde{p}_{ij}^k = p_{ij}^k(w_{ij}^k/\bar{w}_k)$, where $\bar{w}_k$ is the mean fitness in deme k. $D_k = p_{11}^k p_{22}^k - p_{12}^k p_{21}^k$ is the coefficient of linkage disequilibrium in deme k, and $\tilde{D}_k = \tilde{p}_{11}^k \tilde{p}_{22}^k - \tilde{p}_{12}^k \tilde{p}_{21}^k$ is the linkage disequilibrium after selection, among parents. Among the juveniles of the next generation, the frequency of genotype A_iB_j in deme k after migration, is given by

$$p_{ij}^{k\prime} = (1-m_{kl})(\tilde{p}_{ij}^k - r\tilde{D}_k) + m_{kl}(\tilde{p}_{ij}^l - r\tilde{D}_l) \tag{15}$$

if $i = j$, and

$$p_{ij}^{k\prime} = (1 - m_{kl})(\tilde{p}_{ij}^{k} + r\tilde{D}_k) + m_{kl}(\tilde{p}_{ij}^{l} + r\tilde{D}_l) \tag{16}$$

otherwise.

For analytical tractability, we convert this discrete time system to continuous time by taking the limit when all rate parameters (m, s, r) are small and of the same order. When migration is limited to one direction (i.e., m_{12} or $m_{21} = 0$) or when selection in one environment is very strong ($s_i \gg s_j$), the model approaches the well studied "continent-island" model (hereafter superscript "C-I", e.g., Kirkpatrick and Barton 2006 and Charlesworth and Barton 2018).

Analysis

To use the quasi-linkage equilibrium (QLE) approximation, we first rewrite the genotype frequencies in terms of allele frequencies and LD, and then calculate their equilibria (Kirkpatrick, Johnson, and Barton 2002; Otto and Day 2011). This approximation assumes that recombination between the two loci is sufficiently high compared to migration and selection ($r \gg m_{ij}, s_k$) to allow LD to reach an equilibrium much more quickly than the allele frequencies. This is justified here if we do not consider loci that are already tightly linked. But this is not an interesting case, because inversions then offer minimal advantage from suppressing recombination. To ensure the existence of an equilibrium, migration must also be weak compared to selection (i.e. $\max(m_{12}, m_{21}) < \min(s_1, s_2)$). These values allow the calculation of the equilibrium mean fitness in each deme, and hence the rate of increase of an adaptive inversion.

Using Equations 15 and 16 with $r = 0$, the dynamics of an A_1B_1 inversion are described by the matrix A_{11}, in which the (i, j)-th entry describes an inverted adult experiencing selection in deme i, and whose offspring is located in deme j post-dispersal, given by

$$A_{11} = \begin{pmatrix} \frac{(1-m_{21})(1+s_1)^2}{\hat{w}_1} & \frac{m_{12}(1+s_1)^2}{\hat{w}_1} \\ \frac{m_{21}}{\hat{w}_2} & \frac{(1-m_{12})}{\hat{w}_2} \end{pmatrix}. \tag{17}$$

where $\hat{w}_k$ is the equilibrium mean fitness in deme k, calculated from the allele frequencies at QLE (we use the circumflex symbol ˆ for equilibrium values throughout). A_{11} is a mean matrix, in which the entry a_{kl} describes the expected number of offspring a parent in deme k has that end up in deme l. The rate that a rare A_1B_1 inversion increases in frequency in the whole population is given by the leading eigenvalue of A_{11} (λ_{11}). As the population is at equilibrium the growth rate of a recombining A_1B_1 haplotype is 1, so $\lambda_{11} > 1$ implies a benefit to the inversion that can be ascribed to the absence of recombination.

Capture of locally adaptive alleles

Locally adaptive inversions must have captured locally adaptive haplotypes. The chance of this occurring depends on the frequency of said haplotypes in each population. The probability that an inversion captures coadapted alleles (A_iB_i) and invades is given by

$$\gamma_i = p_1 f_{ii}^1 + p_2 f_{ii}^2, \tag{18}$$

where p_i is the probability of invasion conditional on the inversion arising in deme i. The probabilities p_i can be derived using the theory of branching processes. First,

define the random variable $\mathbf{X}_l^k$ to be the number of offspring that, post-migration, are in deme l from a parent in deme k. Under Wright-Fisher conditions, $\mathbf{X}_l^k$ has a Poisson distribution with mean a_{kl}. Then, let $\mathbf{q} = (q_1, q_2)$, where q_i is the probability an inversion that starts in deme i is ultimately lost. In this case, $\mathbf{q}$ is the unique solution to the pair of equations

$$f^{(k)}(\mathbf{z}) = \mathbb{E}\left[z_1^{\mathbf{X}_1^k} z_2^{\mathbf{X}_2^k}\right] = \mathbf{z}, \tag{19}$$

where $k \in \{1,2\}, \mathbf{z} = (z_1, z_2) \in [0,1]^2$.

After some algebra and using the fact that all the $\mathbf{X}_l^k$ are independent, we can find that the two extinction probabilities are the solution to

$$\begin{aligned} q_1 &= e^{-(a_{11}(1-q_1)+a_{12}(1-q_2))}, \\ q_2 &= e^{-(a_{21}(1-q_1)+a_{22}(1-q_2))}. \end{aligned} \tag{20}$$

The establishment probabilities $\mathbf{p} = (p_1, p_2)$ are thus given by $\mathbf{1} - \mathbf{q}$, and can be found numerically using root-finding algorithms to solve the simultaneous equations. The probability of invasion overall is given by $p_1 + p_2$. This invasion probability is specific to the A_1B_1 haplotype and hence conditional on an inversion capturing this allelic combination. To account for the probability of an inversion actually capturing the A_1B_1 haplotype in the first place, we also need to take into account the frequency of this haplotype in a population at equilibrium. Finally, the probability of any locally adapted inversion establishing when it arises needs to consider both A_1B_1 and A_2B_2 haplotypes, and is given by

$$\Gamma = \gamma_{11} + \gamma_{22}. \tag{21}$$

This is also equal to the probability of an inversion establishing itself overall, because inversions that capture allele combinations that are not advantageous in either deme

(i.e. A_1B_2 or A_2B_1) are never favoured. Invasion probabilities for A_2B_2 haplotypes can be calculated analogously.

Results

Equilibrium allele frequencies and linkage disequilibrium

Define $\alpha_i = m_{ij}/s_j < 1$ and

$$\begin{aligned}\hat{f}_0^1 &= \frac{1}{2}\left(1 - 2\alpha_1 + \sqrt{1+4\alpha_1\alpha_2}\right),\\ \hat{f}_0^2 &= \frac{1}{2}\left(1 + 2\alpha_1 - \sqrt{1+4\alpha_1\alpha_2}\right),\end{aligned} \tag{22}$$

which respectively are the frequencies of adaptive and non-adaptive alleles when there is free recombination between the adaptive loci (i.e. r large or $D = 0$). At equilibrium, the frequencies of the alleles ($\hat{f}_j^i$ for allele j in deme i) are

$$\begin{aligned}\hat{f}_{A_1}^1 = \hat{f}_{B_1}^1 &\approx f_0^1 + \frac{\alpha_1}{r}(f_0^1 - f_0^2)\left(s_1 - \frac{\alpha_2(s_1+s_2)}{\sqrt{1+4\alpha_1\alpha_2}}\right),\\ \hat{f}_{A_1}^2 = \hat{f}_{B_1}^2 &\approx f_0^2 - \frac{\alpha_2}{r}(f_0^1 - f_0^2)\left(s_2 - \frac{\alpha_2(s_1+s_2)}{\sqrt{1+4\alpha_1\alpha_2}}\right),\end{aligned} \tag{23}$$

and the linkage disequilibrium between loci in deme 1 (D_1) is

$$\begin{aligned}\hat{D}_1 &\approx \frac{m_{21}(\hat{f}_0^1 - \hat{f}_0^2)^2}{r}\\ &\approx \frac{m_{21}}{r}\left(\alpha_1 + \alpha_2 - \sqrt{1+4\alpha_1\alpha_2}\right)^2.\end{aligned} \tag{24}$$

Linkage disequilibrium in deme 2 ($\hat{D}_2$) is given by replacing m_{21} with m_{12} and vice versa. These equilibrium values, derived here for haploidy, are in accord with previous results (Akerman and Bürger 2014).

In the case where migration and selection are symmetric, $m_{kl} = m$ and $s_i = s$ (i.e., two populations with exactly opposing local selection pressures exchanging an equal

proportion of migrants), the demes have symmetric allele frequencies ($f^2_{A_1} = \hat{f}^2_{B_1} = 1 - \hat{f}^1_{A_1} = 1 - \hat{f}^1_{B_1}$) and linkage disequilibria ($D_1 = D_2$)

$$\hat{f}^1_{A_1} = \hat{f}^1_{B_1} \approx \frac{1}{2}\left(1 - 2\alpha + \sqrt{1+4\alpha^2} - \frac{m}{r}\left(8\alpha - 2\frac{1+8\alpha^2}{\sqrt{1+4\alpha^2}}\right)\right), \tag{25}$$

$$\hat{D} \approx \frac{m}{r}\left(\sqrt{1+4\alpha^2} - 2\alpha\right)^2. \tag{26}$$

In the other extreme case, where there is unidirectional gene flow from deme 2 ("continent") to deme 1 ("island"), the "continent" genotypes remain fixed to A_2B_2. Setting $s_1 = s$ and $m_{21} = m$

$$\hat{f}_{A_1} = \hat{f}_{B_1} \approx \left(1 - \frac{m}{s}\right)\left(1 + \frac{m}{r}\right), \tag{27}$$

$$\hat{D} \approx \frac{m}{r}\left(1 - \frac{m}{s}\right)^2. \tag{28}$$

With free recombination between the loci, the system is decoupled and the alleles are at migration-selection balance.

Locally adaptive alleles are more abundant in the symmetric scenario (equation 25) than in the continent-island scenario (equation 27). This difference arises because in the symmetric scenario a fraction of locally adapted migrants from a focal deme migrate to and survive in the other deme, only to return back and contribute to the frequency of beneficial alleles in the focal deme. In the continent-island scenario, in contrast, continental migrants can only introduce deleterious alleles into the focal deme.

In both scenarios, linkage disequilibrium is positive, indicating that the adaptive alleles tend to be found together in coadapted haplotypes (A_1B_1 and A_2B_2). This tendency increases with the strength of selection in both models ($\partial\hat{D}/\partial s \geq 0$), because

selection favours the association of coadapted alleles, but decreases with the rate of recombination ($\partial\hat{D}/\partial r \leq 0$) which breaks the coadapted haplotypes apart to create more intermediate haplotypes (A_1B_2 and A_2B_1).

The role of migration is less straightforward and differs between the two scenarios. At small migration rates, selection tends to be stronger relative to migration and demes are enriched for locally adapted haplotypes. Linkage disequilibrium then increases with m because more A_2B_2 combinations are introduced into deme 1 (and more A_1B_1 combinations are introduced into deme 2 in the symmetric scenario). When migration becomes higher, the balance between selection and migration shifts and migration tends to introduce proportionately more maladaptive haplotypes from the other deme, thus degrading the linkage disequilibrium that is built up locally by selection. The rate of migration at which this effect sets in depends on the model. In the continent-island scenario, migration decreases linkage disequilibrium when $m > s/3$. In the symmetric case, migration begins to decrease linkage disequilibrium at a lower point, when $m > s\sqrt{3}/6$, because the presence of A_1B_1 migrants in deme 2 generates more intermediate haplotypes through recombination. These individuals can back-migrate and degrade linkage disequilibrium in deme 1 (with the same process going on in the reverse direction).

Invasion probability of a locally adaptive inversion

Having established the equilibrium composition of populations, we can now consider the fate of a new inversion that captures allele A_1 and B_1, which are locally adaptive in

deme 1. We calculate the rate of increase and probability of fixation of this inversion. We again compare the two extreme models, the continent-island and the symmetric scenarios before examining the full model.

While stronger selection increases the frequency of maladaptive alleles among migrants, it will also remove them more effectively from the focal deme. This effect was not fully captured by our QLE approximation, so we numerically calculate the advantage of a rare inversion while assuming that allele frequencies are at the exact equilibrium calculated to second order in selection and migration (Figure 7). In the continent-island scenario, the genotypic composition of migrants is unaffected by selection. Stronger selection reduces the advantage of an inversion (as found by Bürger and Akerman 2011; Charlesworth and Barton 2018) because the island population becomes better adapted as selection increases, so that recombining adaptive haplotypes results in less fit offspring less often (Figure 7).

In the symmetric scenario, the numerical results confirm that increasingly strong selection favours inversions.This happens because selection reinforces local adaptation and makes migrants more maladapted. However, this advantage plateaus as the strength of selection increases, because adaptive alleles become more common. This decreases the advantage of inversions, as selection alone tends to weed out the maladapted combinations. Unless selection is very strong, the former force dominates, meaning that the selective advantage of inversions is primarily determined by the genotypic composition of migrants. Under very strong selection, the invasion probability under symmetric migration converges on that in the island-continent scenario (Figure 7), because the composition of migrants in each become similar.

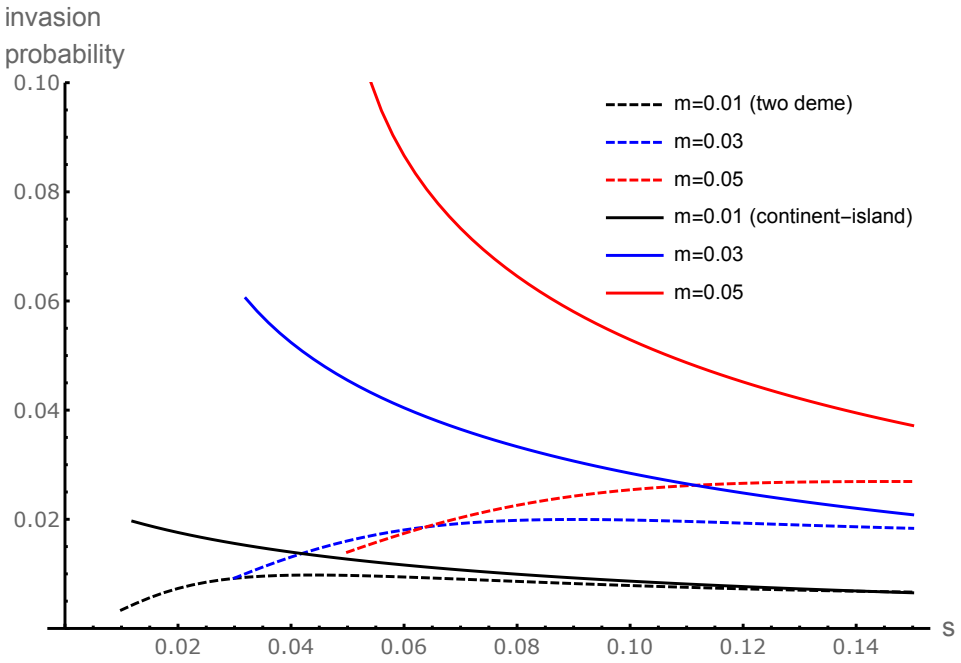


Figure 7: Invasion probabilities approximated to second order in migration and selection for an inversion capturing A_1B_1 in each of the symmetric and continent-island scenarios under various rates of migration.. Data with $s < m$ are excluded as the adaptive alleles may not be at a stable equilibrium. The rate of recombination between the two loci was $r = 0.15$.

Unlike the continent-island model, the two-deme model allows us to include asymmetric local selection and migration (Figure 8). Selection in the focal deme (s_1) increases the degree of local adaptation and inversions therefore have a lesser advantage. This effect is strongest when there are more maladapted migrants entering deme 1 (higher m_{21}, Figure 8A) or when the genotypic composition of migrants is more maladapted (higher s_2, Figure 8C), but has a weaker effect on inversion invasion probability than parameters that change the genotypic composition of migrants (m_{21} and s_2). For a fixed level of migration into deme 1 (m_{21}), the growth rate of the inversion decreases with increasing migration out of deme 1 (m_{12}) because inversions migrate out of the environment in which they are adapted (Figure 8B). Overall, a combination of increased migration from, and selection in, deme 2, are the most important factors

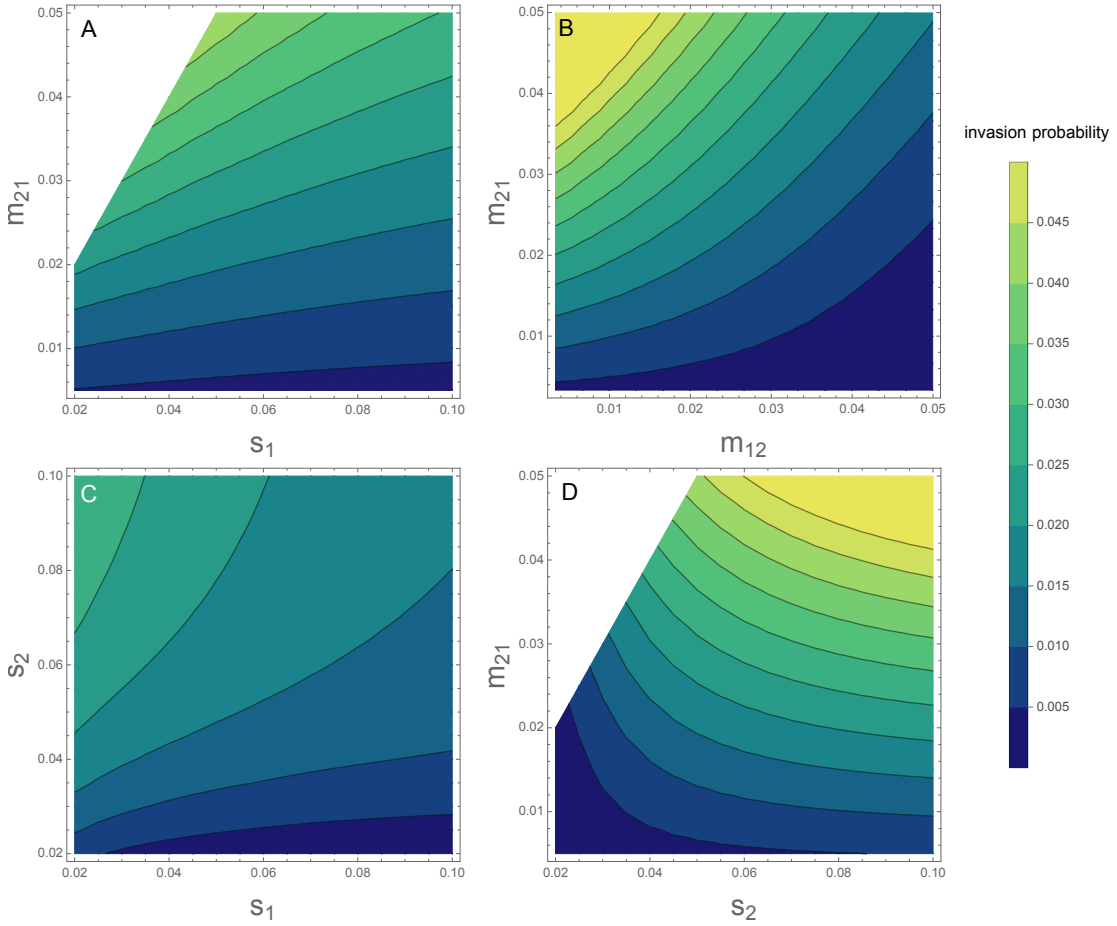


Figure 8: A_1B_1 inversion invasion probabilities calculated to second order in migration and selection terms. Where they do not vary, migration parameters are 0.02 and selection parameters are 0.05. Recombination was set to $r = 0.15$.

in generating the inversion's advantage (Figure 8D) — exactly the two parameters that are most extreme in the continent-island model.

Combined capture and invasion probability of locally adaptive inversions

The analysis above calculates the invasion probability assuming that an inversion captures the A_1B_1 haplotype. It does not take into account the probability that an inversion occurs in an A_1B_1 individual . It seems reasonable to assume that an inver-

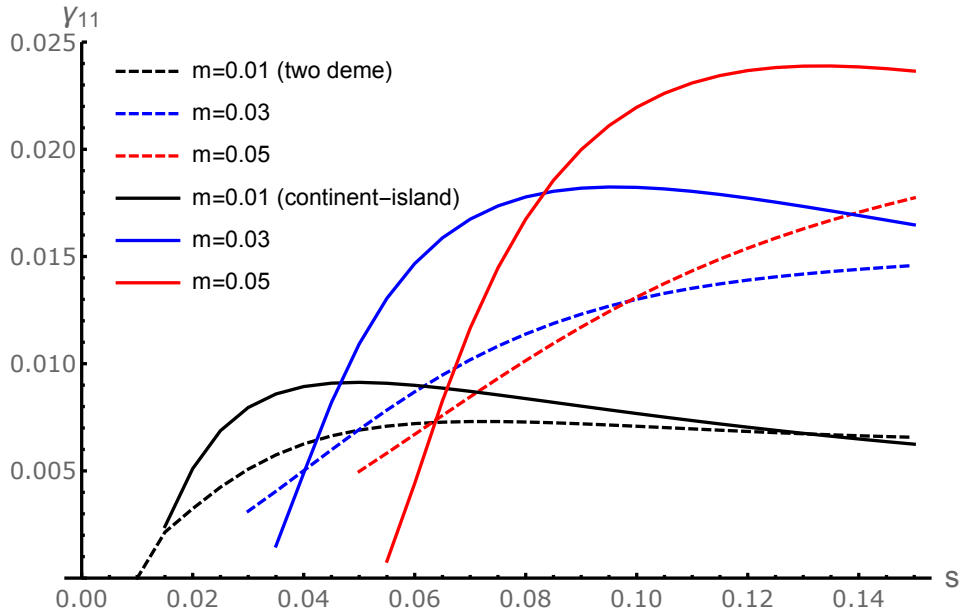


Figure 9: Combined probability of an inversion arising on an A_1B_1 haplotype and then invading (γ_{11}). The invasion probabilities from Figure 7 are adjusted to account for the frequency and relative reproductive value of A_1B_1 in each deme. Equilibria are unstable for $m < s$, $r = 0.15$.

sion captures a random haplotype which means that the invasion probability should reflect the relative frequency of A_1B_1 as well as its reproductive value in each deme. Under this assumption, both the continent-island and two-deme scenarios predict similar patterns of invasion probabilities. As the strength of selection s increases, more locally adaptive genotypes are available to be captured by an inversion (Figure 9). The positive effect of selection on the frequency of locally adapted genotypes (A_1B_1) has a larger positive effect on the combined invasion probability than the negative effect of selection on the inversion's subsequent selective advantage relative to the population (as illustrated in Figure 7). Thus, our results predict that stronger selection is more likely to drive the evolution of locally adaptive inversions. Importantly, this is true for both scenarions and radically alters the prediction for how inversions should contribute to local adaptation in the continent-island scenario (c.f. Figure 7).

We can also see how asymmetric migration or selection affect the combined process of haplotype capture and invasion by inversions. While high migration into deme 1 strongly favours the invasion of existing adaptive inversions (Figure 8B), it also lowers the probability of them arising in the first place, due to the lower frequency of coadapted haplotypes. Thus, adaptive inversions are most likely to form and invade when m_{21} is intermediate, such that the probability of an inversion capturing an adaptive haplotype and the inversion's subsequent selective advantage are both reasonably large (Figure 10A).

Increasing the strength of selection in either deme typically increases the chance that adaptive inversions will arise and spread. Increasing the strength of selection in deme 2 (s_2) increases migration load and therefore the inversion's advantage and increasing selection in deme 1 (s_1) increases the probability of capturing the adaptive haplotype (Figure 10B). Yet, as discussed above, A_1B_1 inversion invasion probabilities decline under very strong selection in deme 1 (very high s_1) by increasing preexisting adaptation. Nevertheless, stronger local selection usually creates a more favourable environment for adaptive inversions to arise and proliferate.

So far, we have only considered the evolution of a specific inversion, adaptive in one deme. This is the only plausible scenario in the continent-island scenario, where only inversions that capture the island-adapted haplotype A_1B_1 are of interest. However, with two demes, divergent local adaptation can occur from either adaptive inversion, both due to the beneficial effects in the favoured deme and due to the protection from deleterious recombination that such an inversion offers to individuals adapted to the other deme. So in this final section we consider the overall probability of local

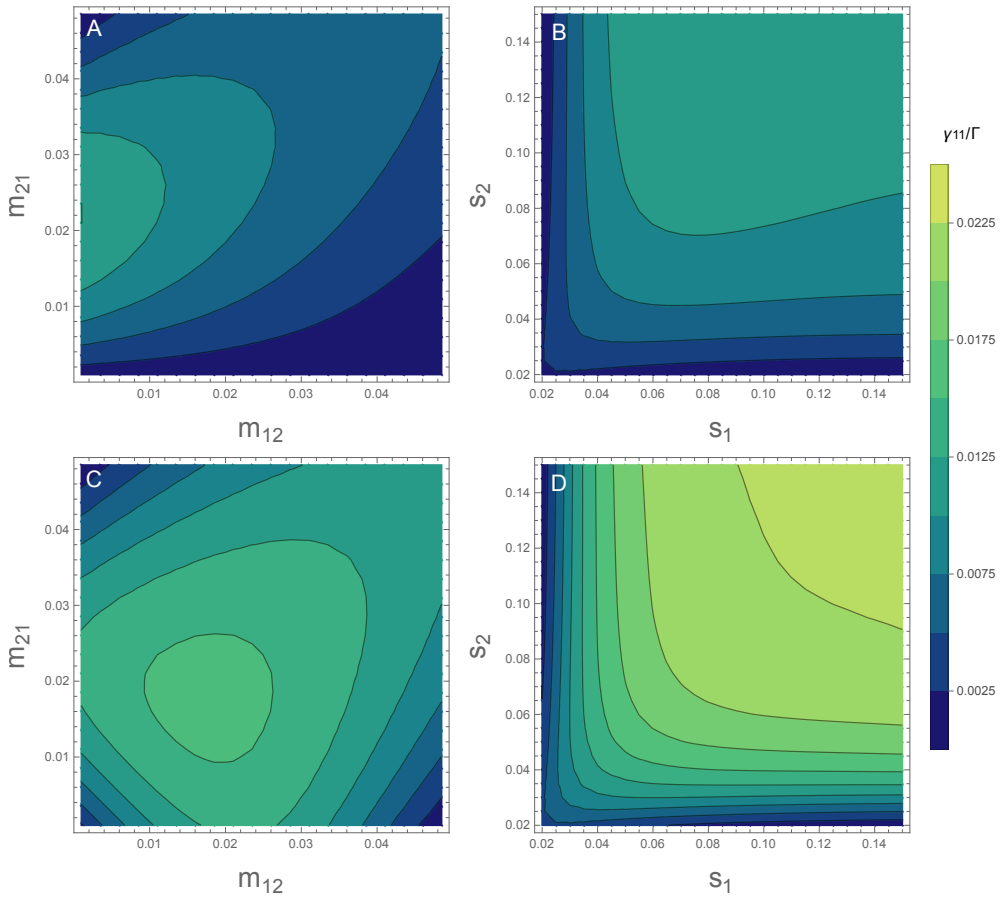


Figure 10: Total establishment probability of an adaptive inversion across the whole population. A, B: Combined probability of an inversion arising on the A_1B_1 haplotype and then invading (γ_{11}) for asymmetric migration (A) or selection (B). C, D: Probability of an inversion capturing either adaptive haplotype (A_1B_1 or A_2B_2) and invading (Γ) for asymmetric migration (C) or selection (D). The continent-island model corresponds to $m_{12} = 0$ (Y axis in A, C) and the symmetric two-deme model corresponds to the $s_1 = s_2$ diagonal in panels B and D. Unless varying along axes, $m_{12} = m_{21} = 0.02$ and $s_1 = s_2 = 0.05$. To ensure stability, we vary parameters in the range where $\max(m_{12}, m_{21}) < \min(s_1, s_2)$, $r = 0.15$.

adaptation through the spread of an inversion that arises anywhere in the population ($\Gamma := \gamma_{11} + \gamma_{22}$; Figure 10C, 10D).

Under symmetric local selection, inversions are most likely to establish when migration is symmetric and intermediate (Figure 10C). Migration rates that are favourable for the establishment of inversion in one deme are not so favourable in the other (γ_{22}

values can be seen by reflecting Figures 10A, 10B across the diagonal) such that symmetric migration rates give the highest overall probability of inversion establishment. Similarly, when migration is symmetric, strong and symmetric local selection is most conducive to the formation and spread of locally adaptive inversions (Figure 10D). Across both demes, this maximises the probability of capturing an adaptive haplotype while maintaining migration load.

Discussion

Here, we have examined the evolution of locally adaptive chromosomal inversions while explicitly modelling selection across a structured population. Inversions can keep locally favoured allele combinations together in the face of maladapted migrants. Therefore, adaptive inversions spread fastest when migrant alleles are homogeneously maladaptive, as assumed in the continent-island scenario that has been well studied (Kirkpatrick and Barton 2006; Charlesworth and Barton 2018). The continent-island scenario represents an extreme, where migrants are fixed in their genetic composition, being purely maladaptive, with the migration rate alone determining selection for the inversion. In comparison, the two-deme model leads to a number of novel insights. By including the dynamics of selection and migration in the source population, we find that inversions capturing alleles experiencing relatively strong selection are more favoured, unlike the condition found when migration is unidirectional in the continent-island scenario (Figure 7). Extending the model to account for the probability that inversions initially capture favourable haplotypes

shows that relatively strong selection is most likely to underlie inversions (Figure 9) and continent-island scenarios aren't necessarily most conducive to inversion evolution (Figure 9). We further examine asymmetric selection pressures across demes, showing that strong selection in either deme generally promotes the establishment of adaptive inversions by either increasing the selective advantage or the probability of capture (Figure 10). Overall, our results suggest that inversions are particularly likely to arise and establish when selection on locally adaptive alleles is strong.

Theories concerning the origins of adaptive inversions can broadly be split into three categories (Schaal, Haller, and Lotterhos 2022): "capture", in which an inversion creates a linkage group of existing adaptive variation and spreads (Kirkpatrick and Barton 2006); "gain", in which an inversion is initially polymorphic (e.g. due to drift, underdominance, or acquisition of a good genetic background), and then accumulates adaptive variation which is subsequently protected from recombination (e.g. Lamichhaney et al. 2016, Samuk et al. 2017); or "generation", in which adaptive variation is created when the inversion occurs through the breakpoint disrupting coding sequence or gene expression (Feder and Nosil 2009; Villoutreix et al. 2021, e.g. Jones et al. 2012). Our work focuses on the "capture" hypothesis in which locally adaptive alleles are already segregating and have reached migration-selection equilibrium and may have already evolved enhanced local fitness. This scenario is the most analytically tractable, and hence we analyse it here. However there is *a priori* no reason why any inversion with "capture" origins could not subsequently gain more adaptive variation at a later date as set-out in the "gain" hypothesis. In a pure "capture" scenario, we show large effect alleles are the most likely to underlie adaptive inversions.

1088 The evolution of the effect size distribution of locally adaptive alleles is likely to favour those that are strongly selected, facilitating the evolution of adaptive inversions. In 1090 the short term, locally adaptive alleles must experience fairly strong selection to be able to resist being swamped by migration (Lenormand 2002; Yeaman 2015). Small 1092 effect alleles can still contribute to local adaptation when they arise in close linkage with large effect alleles, resulting in aggregated regions of adaptation which could be 1094 modelled as a single locus of large effect (Yeaman and Whitlock 2011). Alternatively, they can contribute transiently before being lost (Yeaman 2015). With high gene flow, 1096 and over long timescales, the architecture of local adaptation is expected to evolve towards a few, highly concentrated clusters of small effect alleles linked with large effect 1098 alleles (Yeaman and Whitlock 2011), which are likely to be particularly conducive to inversion establishment.

1100 Migration regimes under which inversions are likely to form and spread are fairly specific because they must satisfy multiple requirements. Firstly, we assume that locally 1102 adaptive alleles are polymorphic, which means they must be able to resist swamping by migration. This condition requires relatively weak migration and is likely to be 1104 a significant constraint on the evolution of local adaptation (Feder, Gejji, et al. 2011). Then, given that locally adaptive alleles are maintained, higher migration rates favour 1106 the spread of inversions because they increase the frequency of the maladaptive alleles and thus the cost of recombination (Figures 7, 8). However, this also has the effect 1108 of reducing the frequency of adaptive haplotypes so that inversions are less likely to capture a full complement of adaptive alleles (Figure 10). The result is that higher 1110 migration rates do not always favour the evolution of inversions. In general, rates

of migration may turn out to restrict the evolution of capture-origin inversions more than previously though.

Schaal, Haller, and Lotterhos 2022 used simulations to study the invasion of inversions capturing variation that influences a polygenic quantitative trait, finding that inversions involved in local adaptation tended to exhibit more of a capture than a gain effect when alleles were unlikely to be swamped. When alleles were prone to swamping by migration, persisting locally adaptive inversions had often gained much more adaptive variation post-capture. Under high rates of gene flow both capture and gain scenarios are plausible, depending on the effect size of the loci captured. Because adaptive alleles can be gained after the inversion arose and spread, recent inversions may offer the best opportunity to test our predictions about the effect size of alleles driving the evolution of locally adaptive inversions. The allelic content of such inversions could depend on how long the populations in question have been diverging, with the expectation that long periods of divergence results in a more concentrated architecture (Yeaman and Whitlock 2011). However, separating the individual trait effects of different loci within the inversion is challenging once they have been linked together. Thus, despite the prevalence of putatively adaptive inversions, mapping of quantitative trait loci has been achieved in only a handful of cases (e.g. Peichel and Marques 2017; Koch et al. 2021; and Poelstra et al. 2014 for an example unrelated to local adaptation) leaving open questions about the number and effect size of loci that underpin inversion selective advantage (Tigano and Friesen 2016).

We only consider the evolution of inversions that link alleles at two relatively nearby loci. It is possible that an inversion could capture more than two loci that affect lo-

cal adaptation. As the number of loci contributing towards adaptation increases, it becomes less likely that an inversion will capture all the adaptive alleles on the same haplotype. Nevertheless, inversions will still spread if they capture more locally adaptive alleles than the population mean. A similar process has been proposed for the evolution of inversions that happen to capture fewer deleterious mutations than average (Nei, Kojima, and Schaffer 1967; Jay et al. 2022; Lenormand and Roze 2022). The relationship between invasion fitness and haplotype frequencies as the number of loci increases remains to be explored, but we expect inversion evolution will continue to depend on a balance between the selective advantage of the captured haplotype and on the probability of capturing a favourable haplotype.

Our model does not include deleterious mutations or breakpoint effects, which can affect the fate of inversions. Low rates of gene flux within inverted arrangements means that deleterious variation captured by the inversion persists for a long time throughout lineages, as purging this variation relies on rare events such as gene conversion and double crossover events. Inversion breakpoints can also disrupt gene function and result in lower individual fitness (White 1978; Kirkpatrick 2010), though this can occasionally be adaptive (e.g. Corbett-Detig 2016). These effects can be incorporated into the model by introducing a fixed cost or benefit. Reduced recombination within inversions severely weakens the efficacy of purifying selection on new mutations (Charlesworth 1996; Betancourt, Welch, and Charlesworth 2009). Mutation accumulation is particularly important while the inversion is at low frequency, because most inverted chromosomes will occur in heterokaryotypes where recombination is suppressed (Navarro, Barbadilla, and Ruiz 2000), though gene conversion

and double crossover events may alleviate this a little (Berdan et al. 2021). We model a haploid population, but in diploids the presence and accumulation of strong recessive mutations within inversion will result in negative frequency-dependent selection which limits inversion frequency and the recombination rate (Nei, Kojima, and Schaffer 1967; Wasserman 1968; Ohta 1971). The generally deleterious effects associated with inversions likely mean that their invasion probabilities are much lower than we obtain here.

In summary, our results emphasise the likelihood that strongly selected loci can contribute to local adaptation in two ways: by increasing the frequency of adaptive haplotypes that can be captured by an inversion, and by increasing the rate of migrant gene flow and thus the potential cost of recombination. High migration rates also increase this recombination load and thus the selective advantage of an inversion, but this also reduces the frequency of adaptive haplotypes. The probability of adaptive inversion formation could be as important as its selective advantage in determining where such inversions are likely to be found.

Mutation accumulation and the establishment of adaptive inversions in finite populations

Abstract

By suppressing recombination, inversions maintain linkage between loci they encompass. This restriction may limit the ability of an inversion to spread through a population by increasing the mutation load associated with them. Recombination is necessary for efficient purifying selection, and any deleterious variation captured by the inversion when it first arises is likely to be present in all future descendants. The presence of deleterious variation has been shown to have a significant impact on whether inversions establish within populations or are ultimately lost. However, some of the theoretical literature is based on early work in which an assumption of free recombination between inversion loci is implicit, leading to results that only apply when the inversion is common. This theory predicts that otherwise neutral in-

versions that capture even a single deleterious mutation will always be lost, because the transient advantage afforded by a relatively mutation-free background decays as mutation-selection balance is reached within inversions. As a result, the possibility such inversions might establish is often discounted. We simulate the fixation of inversions in a finite haploid population under a wide range of parameters. Our results suggest that not only can inversions capturing deleterious variation fix before their transient advantage fully decays, but that mutation accumulation preventing inversion establishment is most significant when inversions are rare, when they undergo Muller's ratchet-style degradation.

Introduction

Through the suppression of recombination, chromosomal inversions can cause segments of a chromosome to be inherited in tandem. Since their discovery in the early 1920s, interest in inversions has grown — especially since the widespread availability of genome-sequencing has revealed their relative abundance and their links to several important evolutionary processes, including local adaptation and mating strategy (Kirkpatrick 2010; Wellenreuther and Bernatchez 2018).

Selection acts both directly on the inversion itself and also on the genetic material it captures, maintained by increased linkage between loci (Berdan et al. 2021). Positive selection on the inversion as a whole can result directly from adaptive breakpoint effects (Corbett-Detig 2016; Villoutreix et al. 2021), or indirectly when the inversion captures a better-than-average mutation load (Nei, Kojima, and Schaffer 1967),

sets of coadapted alleles with positive epistasis (Dobzhansky 1947; Haldane 1957; Charlesworth 1974) or alleles involved in local adaptation (Ch. 3, Kirkpatrick and Barton 2006; Charlesworth and Barton 2018). In general, where there is selection for linkage disequilibrium (LD) between alleles of a set of genes, an inversion that captures this set can be advantageous by preventing the decay of LD (Kirkpatrick and Barton 2006).

While the contents of an inversion are inherited together, they are liable to change through time due to the input of mutations at any of the loci within the inversion. Recombination is suppressed only between standard and inverted arrangements, so that their respective evolutionary trajectories diverge (Sturtevant 1917; Roberts 1976). In particular, the overall rate of recombination in the inverted region is decreased and consequently so too is the efficacy of purifying selection on deleterious alleles (Barton and Charlesworth 1998). This effect is more pronounced when one arrangement, usually the inversion, is rare. The inversion "subpopulation" is also fixed for variation it captures when it first arises, though some could be lost through gene conversion between inverted and standard haplotypes (Navarro, Betrán, et al. 1997). Such deleterious variation cannot be removed through purifying selection when recombination occurs in inversion homozygotes, because they will also be homozygous for the deleterious alleles. When this variation is recessive, it can result in inversions being under balancing selection through associative overdominance (Ohta 1971). Furthermore, this can result in them carrying an inherent load that will fix in the population if the inversion fixes.

The methods used in recent inversion work (e.g. Connallon, Olito, et al. 2018; Con-

nallon and Olito 2021) are based on previous theory that showed that inversions capturing one or more mutations will be unable to fix, unless the inversion has some unique advantage outweighing the cost of the captured mutation load (Nei, Kojima, and Schaffer 1967). This is because the advantage of capturing a good background is transient, and will degenerate in the long term. At this equilibrium, the inversion contents reach mutation-selection balance but are also burdened with the captured mutations, which are present in all inversion copies. The best case scenario is that the inversion captures no mutations, becoming selectively neutral at equilibrium. So, one would expect observable inversions to either be very small, so as to increase the chances of capturing a mutation-free background, or have strong positive selection acting on the inversion itself. However, this assumes that the long-term degeneration of the inversion will always occur before fixation. This problem was noticed by Kimura and Ohta 1970, who extended the model to allow for the possibility that inversions could plausibly fix before degenerating.

Yet even this extended model still does not tell the full story. Implicit in Nei et al.'s derivation is free recombination between loci contained within the inversion. However, this can only have an appreciable effect when the inversion is at a sufficiently high frequency so that recombination can occur between homozygotes. Furthermore, their results assume that the rate of mutation accumulation is constant through time. This is despite the fact that at low frequencies, reduced recombination should qualitatively result in a different manner of degeneration of inverted regions, as new mutations arising can fix or hitchhike as the inversion increases in frequency in a manner similar to Muller's ratchet.

Here, we address these issues in more detail. Using simulations, we show that inversions under directional selection can fix in finite populations even when they capture a mutation load greater than their advantage, contrary to Nei, Kojima, and Schaffer 1967. Transient selective advantages readily drive inversions to fixation. Furthermore, we show that early-stage mutation accumulation can prevent inversions from establishing, and propose that this is a key determinant in deciding inversion fate. These results hold across a range of mutation rates, population sizes, direct inversion advantages, and backgrounds captured. We conclude by proposing a framework for thinking about inversion evolution, and how this might be tackled properly from a theoretical point of view.

Previous Models

Nei and Kimura models

Here, we outline the analytical approach for the case of a finite population as presented in previous work, before discussing its limitations. Nei, Kojima, and Schaffer 1967 derived an expression for the relative fitness of an inversion through time. They model the inversion as a single locus whose selective advantage decreases as it accrues mutations through time. This expression was then used by Kimura and Ohta 1970 in their framework for determining the corresponding fixation probability in a finite population.

For ease of analysis, we modify these approaches to apply to haploids. This is accept-

able here because early events (when recombination would be at its lowest) are the most important and haploid selection coefficients are parallel to heterozygous haploid selection coefficients. Conclusions made here will also apply to the additive case ($h = 1/2$) in diploids. First, consider the frequency (q) of a deleterious mutation at an initially non-mutated locus through time. This can be described by the differential equation

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \mu - uq. \tag{29}$$

Mutations arise at this locus at a rate μ, and where present they are selected against at a rate u. This equation has the solution

$$q(t) = (\mu/u)(1 - \mathrm{e}^{-ut}), \tag{30}$$

with initial condition $q(0) = 0$, from which it can be seen that an average wild-type allele approaches mutation-selection balance (μ/u) at a rate u. Since q is the frequency of mutations at a locus, qu is the average mutation load, i.e. the average reduction in fitness due to alleles segregating at the focal locus. Now consider an inversion that has a direct selective advantage s (due to beneficial effects of rearrangements at the breakpoint, for example), but also captures c deleterious mutations. The fitness of the inversion through time can be approximated as the product of the fitness of all its loci — captured mutations contribute $(1-u)^c$, while wild-type loci degrade as in Equation 30:

$$\begin{aligned} w_t &= (1+s)(1-u)^c \left(1 - \mu\left(1 - \mathrm{e}^{-ut}\right)\right)^{L-c} \\ &\approx (1+s)(1-u)^c \mathrm{e}^{-(L-c)\mu(1-\mathrm{e}^{-ut})} \\ &\approx (1+s)(1-u)^c (1 - (L-c)\mu(1-\mathrm{e}^{-ut})). \end{aligned} \tag{31}$$

At $t = 0$, this inversion has fitness given by the beneficial effect and the captured deleterious alleles $(1+s)(1-u)^c$, whereas the population mean fitness is $(1-u)^{L\mu/u} \approx 1-L\mu$, assuming mutation-selection balance. However, the inversion fitness then decays at a rate u until it reaches a relative fitness $(1+s)(1-u)^c(1-(L-c)\mu)$. At this point, mutation-selection balance is reached at every locus except those where a mutation was captured, where mutations are fixed. While μ determines the rate of input of mutations to a locus and their eventual frequency, it is u alone that determines the rate of decay of fitness. When there is no advantage unique to the inversion, inversions capturing no deleterious mutations will eventually deteriorate until they are as fit as the rest of the population. Further, those capturing even a single mutation will eventually be lost, since $(1-u)^c(1-(L-c)\mu) < 1-L\mu$ for all $c \leq 1$. For inversions to have a long-term advantage, the total effect of s and the capture of c mutations must be net positive. The key disadvantage of an inversion is that whatever deleterious alleles it captures are forever part of its genetic content, they cannot be lost through recombination as all inversion carry the same set of captured alleles.

As Kimura and Ohta 1970 pointed out, this is not especially realistic given that inversions could fix before the long-term dynamics are realised. They use an inversion as an example of an application of a method to find the fixation probability of a new mutant whose selection coefficient changes in time. First, they derive a diffusion equation for the fixation probability of a new mutant whose selective advantage decays over time. Specifically, they model the case where the selection coefficient has an initial value s_0 that declines at a rate k—that is,

$$s_t = s_0 \mathrm{e}^{-kt}. \tag{32}$$

This approach can be used for an inversion by substituting the corresponding rate of decay $k = u$, and initial selection coefficient

$$s_0 = w_0 - 1 \approx s - (c - \bar{c})u, \tag{33}$$

ignoring terms of order us, s^2, u^2. The full detail of the method is omitted, but after adjusting for haploidy, it comes down to solving the following differential equation for $Y = Y(s)$

$$\frac{\mathrm{d}Y}{\mathrm{d}S} = \frac{Y(S-Y)(1-\mathrm{e}^{-Y})}{KS\left(1-(1+Y)\mathrm{e}^{-Y}\right)} \tag{34}$$

where $S = 2N_e s_0$ and $K = 2N_e u$. The invasion probability of the inversion, given that it arises in a single individual, is

$$\gamma = \frac{1-\mathrm{e}^{Y/N}}{1-\mathrm{e}^{-Y}}. \tag{35}$$

It is hard to get any intuitive sense of the fate of inversions from this, as Y has no meaning *per se*, and values for Y need to be obtained numerically.

Nei's method makes implicit assumptions that make it invalid in some scenarios. Firstly, Nei's method considers the inversion as a collection of loci, but implicitly assumes free recombination between them. Selection against mutations dampens their frequency, and with free recombination between loci can act precisely on each locus as they evolve independently. This can only be valid when inversions are at some intermediate frequency, so that homozygotes are common enough for recombination to occur. As a result, it is not applicable to the early stages of inversion evolution, when recombination within inversions is very rare and there is tight linkage between loci.

In truth, a small inversion population will degenerate in a manner similar to Muller's ratchet — mutation accumulation in small, asexual populations. Since inversions

cannot recombine, new mutations will always be carried by inversion descendants. When every individual carries an extra mutation, the fittest class of inversions is lost.

While asexual populations, and so also rare inversions, do ultimately reach mutation-selection balance, they do so at a different rate and in a different way to sexual populations. Inversions are unlikely to degenerate to mutation-selection balance as quickly as in Nei's model because recombination can only introduce existing mutations. Mutations must first arise on an inversion before they can be recombined onto another. Furthermore, selection acts across the entire complement of loci, rather than each individually. Surprisingly, this means that asexual populations have a lower mutation frequency at a given locus. This happens because when loci are linked, the presence of one mutation acts to decrease the frequency of mutations at other loci (Pénisson et al. 2013).

Adjusting Kimura's method to account for this is infeasible. Firstly, the method requires that the selection coefficient changes at a constant rate, however the manner of fitness decay should be dependent on inversion frequency. Secondly, while Nei's model does apply to inversions at intermediate frequency (assuming there are enough mutations segregating within inversions), the application of Kimura's model from this point on may not be valid because it assumes that the inversion starts at low frequency. So, a new approach is needed if we are to faithfully model inversion evolution.

Time-dependence of inversion evolution

The nature of inversion evolution depends on both its frequency and number. So, it is necessary to consider how long the inversion spends at low, then intermediate frequency. The fixation of a new beneficial allele generally consists of three phases. First, at low frequencies there is a stochastic phase in which genetic drift dominates the dynamics of a new mutation (Charlesworth 2020). During this phase, the trajectory can be approximated as that of a neutral mutation (e.g. Kimura and Ohta 1973). The allele will increase in a more or less deterministic fashion only once it frequency exceeds a threshold, given by $1/Ns$ in a haploid population. Finally, there is a second stochastic phase before fixation, where the alternative allele is so rare that frequency dynamics are again dominated by drift.

We can apply this thinking to the inversion as a whole. A Muller's ratchet process will occur when both the number and frequency of inversions in the population is small. The frequency must be low so that recombination is rare, and the inversion population size below which degeneration of the inversion (or of the standard arrangement, when the inversion is close to fixation) is unavoidable is approximately $e^{\mu L/u}$ (Haigh 1978). Large inversions or higher mutation rates increase the inversion number threshold below which they will degenerate. In inversions, this process overlaps with the initial stochastic phase which occurs when the number of inversions is less than $1/s_0$, so more degeneration is likely to occur when inversions have a weak advantage and behave stochastically for longer.

Nei's model can be used once the inversion is at sufficiently high frequency. For the

inversion to decrease in fitness such that it will ultimately be lost, it must be polymorphic for long enough that mutation-selection balance is reached and also so that there has been sufficient mutational input into inversions.

Methods

To address these points, we simulate a Wright-Fisher population consisting of N haploid, hermaphroditic individuals, each of which has a chromosome of G loci. Generations are discrete and non-overlapping. Mutations occur at a rate μ per locus per generation uniformly across the chromosome, each having a deleterious multiplicative fitness cost $u = 0.005$. Recombination occurs with a minimum of one crossover, with further crossovers occurring with probability ρ between each pair of consecutive loci. So, the mutation rate per chromosome per generation is $G\mu$, and the expected number of crossovers is $1 + \rho G$. Inversions are of length L and are introduced into the centre of the chromosome (ie spanning locus $G/2 - L/2$ to locus $G/2 + L/2 - 1$). The inversion directly confers a selective advantage s. Recombination between chromosomes with the same conformation occurs freely, but exchange between the standard and the inverted arrangement is permitted only when there are an even number of crossover points sampled within the inversion. Otherwise, the offspring is assumed inviable and new crossover points are drawn. In general, we use values of $G = 1000, L = 100$ so that the inversion makes up 10% of the chromosome, and $\rho = 5 \times 10^{-4}$ so that each meiosis results in an average of 1.5 crossovers. We calibrate our parameter values such that the chromosome-wide mutation rate is similar to that

of *Drosophila melanogaster*, estimated to be 1.2 deleterious mutations per diploid genome (Haag-Liautard et al. 2007). Since the autosomes are approximately twice the size of the X chromosome, this leads to a chromosome-wide mutation rate estimate of 0.48 mutations per diploid autosome, or 0.24 for a haploid. In our model, with the above chromosome length, 0.24 mutations per chromosome per generation corresponds to a mutation rate of $\mu = 2.4 \times 10^{-4}$.

Simulations were performed using SLiM v4.0 (Haller and Messer 2022). Before introducing the inversion, we simulate a burn-in period of 10^5 generations to reach mutation-selection-drift balance. This process occurs separately for every replicate, and a minimum of 5000 replicates were run for each set of parameter values. After the burn-in period, an inversion occurs in a randomly sampled individual in the population. The simulation continues until the inversion is either lost or fixed, and we record which of these occurred. No other fate was possible because all the modelled alleles experience directional selection only.

We use the term "capture" to describe the deleterious variation that exists within the inversion at the moment it arises (i.e., the deleterious mutations carried by the specific chromosome that gives rise to the inversion). "Accumulation" refers to mutations that arise post-introduction. To differentiate between the separate effects of capture and accumulation we require a control simulation to compare the results of the full simulation to, a scenario in which there is no accumulation effect. The total selection on a new inversion can be broken down into two factors: the direct selective advantage we ascribe to it (s), and the relative fitness of the background it captures relative to the rest of the population. So, if an inversion captures c mutations and the average

number of mutations in the same region across the population is $\bar{c}$, the initial relative fitness of the inversion is

$$w_0 = (1+s)(1-u)^{c-\bar{c}}. \tag{36}$$

While the inversion is segregating, new mutations will occur and reduce the fitness of individual copies below w_0. In contrast, a single locus with selection coefficient w_0 will maintain that coefficient through time. Therefore, any difference between the fixation probabilities of the inversion and the single locus control must be due to new mutations that have appeared post-introduction. Sometimes, we stipulate that an inversion captures a mutation load equal to the population mean in that region. This mean is almost never an integer number of mutations. In these cases, the inversion captures a minimum of $\lfloor \bar{c} \rfloor$ mutations, with a probability $\bar{c} - \lfloor \bar{c} \rfloor$ of capturing one more.

Results

The mutation load captured by an inversion is Poisson distributed with mean μL. The same quantity is proportional to the rate of mutation accumulation within an inversion. So, we expect a proportional increase in μ or L to have the same effect as a similar increase in the other. In order to cover multiple orders of magnitude, it was more practical to vary μ rather than L. In general, conclusions regarding inversion length could also be inferred from μ.

The effects of the mutation rate on invasion probability are multiple. First, the fixation probability of positively selected inversions is negatively impacted by the presence of deleterious mutations around it, because selective interference between the two low-

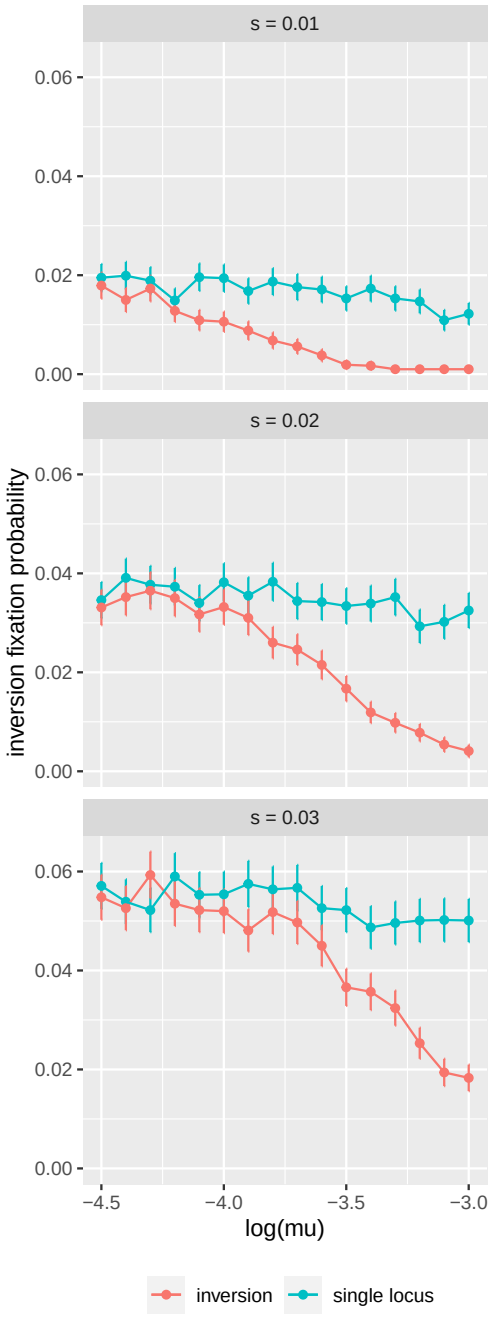


Figure 11: Simulated fixation probabilities for inversions of length 100 which capture the mean mutation load, and a single locus with the same direct selection coefficient. As the mutation rate increases, there is greater mutation accumulation in the inversion which retards the fixation probability compared to a single locus. The effect is more pronounced under weaker selection (top panel). Bars show twice the standard error. Each data point is based on 10^5 simulations. Parameter values are as detailed in the methods, with $N = 1000$.

ers the efficacy of selection. Increasing the mutation rate increases the number and density of deleterious mutations and hence the strength of interference. This effect is independent of inversion length, so has an identical effect on each of the inversion and single locus simulations. This effect is more pronounced when the selective advantage of inversions is small, as weak direct selection on the inversion increases the relative strength of selective interference (Figure 11, blue lines). Since an inversion behaves similarly to a single locus, we can compare the invasion probability of an inversion to that of a single locus of the same selection coefficient (i.e. $L = 1$), which is equally affected by linkage to mutations. We expect both to decrease under stronger selective interference. The single locus comparison allows to determine the effects of new mutations on inversions, as any difference between the two occurs as a result of mutation accumulation.

When an inversion captures the mean mutation load (i.e. when $c = \bar{c}$), its background is neither advantageous nor disadvantageous relative to the population mean. So, in the absence of any further mutation accumulation, it would be expected to have the same fixation probability as a single locus with an identical selection coefficient. Using this idea, we investigate when high mutation rates cause mutation accumulation to the point where they prevent inversion fixation. When mutation rates are sufficiently low, the degree of accumulation is also low and so it has no discernible effect on fixation probability (Figure 11). However, fixation probabilities decrease with higher mutation rates. The rate of mutation at which this effect becomes significant depends on how strongly the inversion is favoured. New deleterious mutations have a higher impact when s is low. There are two possible reasons: fewer mutations are

required to offset a lower selective advantage, and/or the stochastic phase of the inversion's trajectory is longer, increasing the probability of the inversion accumulating mutations in a process similar to Muller's ratchet. In this case, mutation rates can be relatively low and still prevent fixation. Very high mutation rates can cause the inversion to degenerate so quickly that it almost never invades, irrespective of the intrinsic selective advantage it confers. The decline in inversion fitness in these cases must be associated with early mutation events causing Muller's ratchet-style degeneration, because the rate of decay of the transient advantage is a function of u, not μ, which influences only the frequency of mutations at equilibrium.

On average, a new inversion will capture a mutation load equal to the population mean in that region. However, those inversions that capture mutation loads that are relatively low compared to the population mean have an advantage over the standard arrangement and so are more likely to fix. Conversely, inversions that capture poor backgrounds with a greater than average mutation load are less likely to fix. Further, an inversion capturing more mutations than expected can still be favoured as long as the cost of the relative fitness of its background does not exceed the direct benefit s provided by the inversion. The number of mutations an inversion copy can afford to accumulate, and so the importance of mutation accumulation, depends on how close it is to this threshold.

We simulated inversions that captured a fixed number of mutations, and compared them to single locus simulations with selection coefficient s_0 (Figure 12). Increasing the population size lowered the frequency of deleterious mutations at mutation-selection-drift balance, so care must be taken when comparing results from different

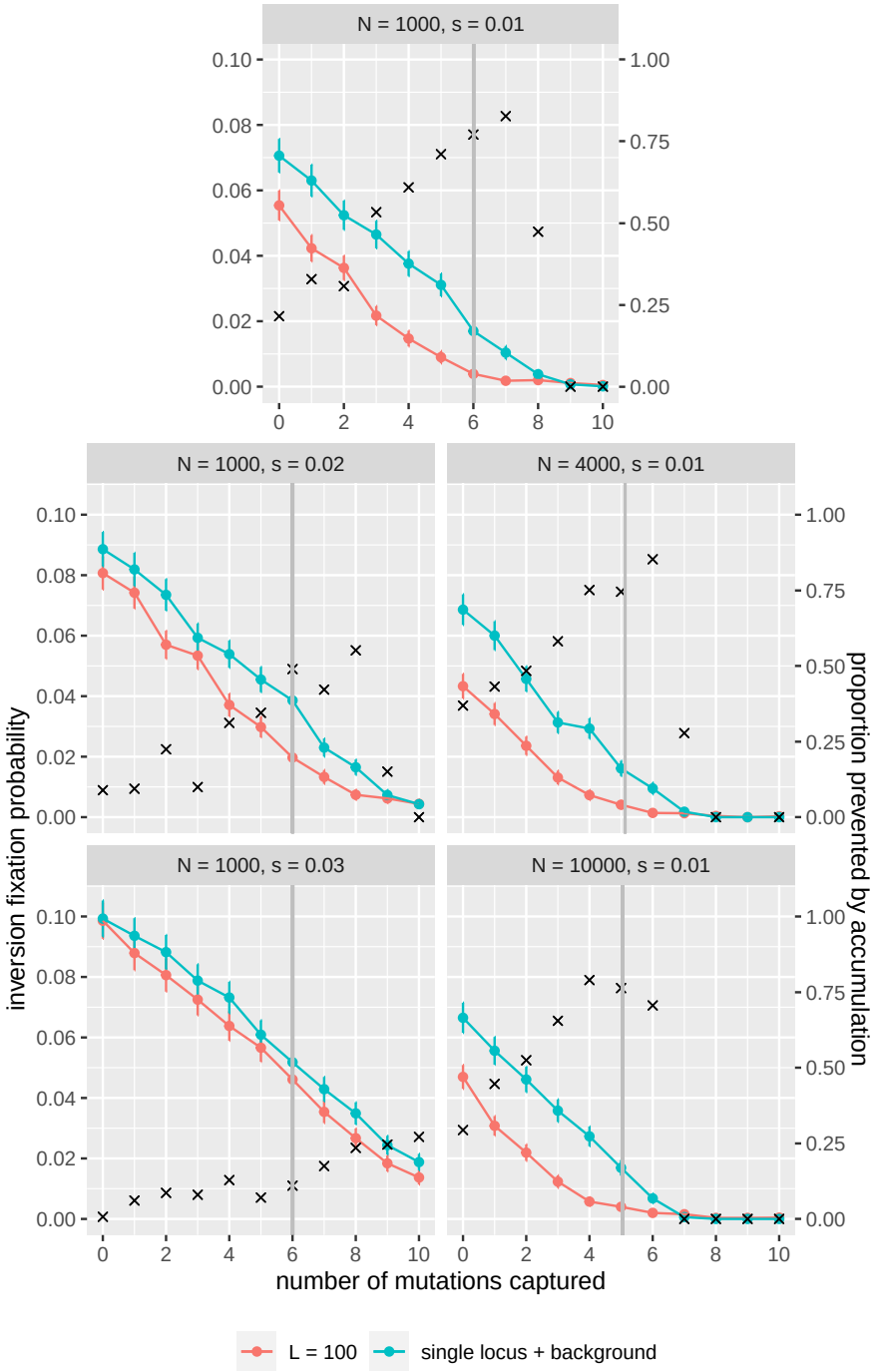


Figure 12: Simulated fixation probabilities for inversions of length 100 which capture a particular load of mutations, and a single locus with selection coefficient s_0 as defined in the Methods. Bars show twice the standard error. Each data point is based on 10^5 simulations. Crosses show the proportion of simulations where non-fixation can be attributed to the presence of mutation accumulation. Vertical grey bars show the population wide mean mutation load in the inversion region at introduction. Parameters values are as detailed in the methods, with $\mu = 2.5 \times 10^{-4}$.

population sizes. For example, the population mean mutation load when $N = 1000$ is
1484 6 mutations, compared to 5 when $N = 10000$. So, inversions that capture 5 mutations are relatively fitter in the smaller population. Taking this into account, alterations
1486 in population size had very little effect on the fixation probability of inversions within the range tested (right hand side of Figure 12). Generally, as s increases the proportion
1488 of fixations prevented by accumulation decreases, as direct selection on the inversion dominates (left hand side of Figure 12).

1490 Accumulation is more likely to cause extinction when inversions capture high mutation loads. There are again two possible explanations for this. These inversions
1492 start closer to the threshold where the cost of the mutation load is greater than the inversion's direct selective advantage (s). As s increases, so too does the dis-
1494 tance from this threshold. Or, the stochastic phase of the inversion is longer as s_0 is smaller, meaning more mutations are accumulated. If many mutations are captured,
1496 then accumulation-associated extinctions are less important because such inversions were unlikely to fix in the first place.

1498 Comparing the simulation results to the probability obtained from the Kimura-Nei method shows their approximation hugely overestimates the importance of popula-
1500 tion size (Figure 13). In the simulations, there was no discernible difference in invasion probability in the range of population sizes from 2000 to 10000. However, the
1502 approximation predicts that inversions should be much less likely to fix in larger populations, because inversions arising as a single copy take longer to fix and degenerate
1504 for longer. This again suggests that a Muller's ratchet process is indeed what prevents inversion fixation rather than long-term degeneration, because it depends on

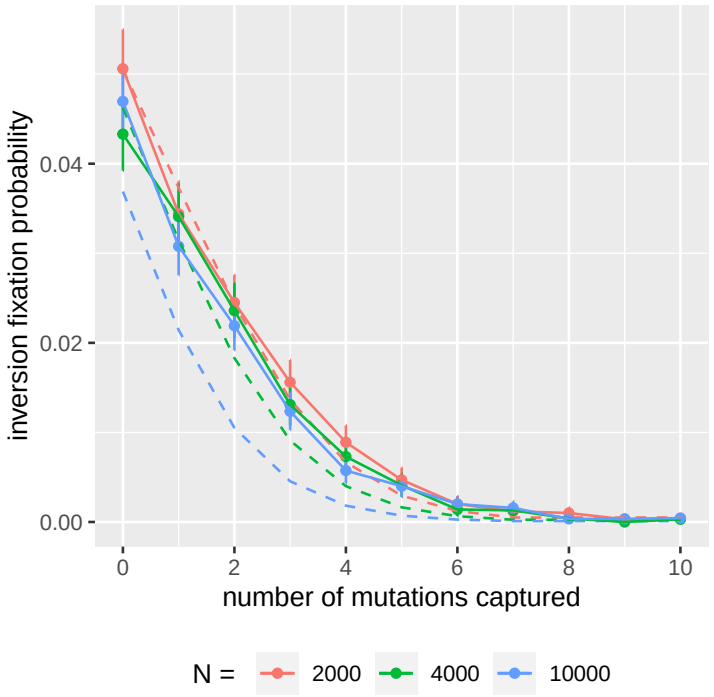


Figure 13: Fixation probabilities from Figure 12 (solid lines) and the corresponding probability derived using the Kimura-Nei method (dashed lines), assuming that population-wide mutation frequencies are at theoretical mutation-selection balance.

1506 the number of inversions rather than their frequency.

Proposed trajectory of directly selected inversions

1508 In summary, we propose the following description of how inversions evolve during each phase of their trajectory:

1510 1. **Mutation capture:** a new inversion arises on a chromosome in the population, which is assumed to be at mutation-selection-drift balance. Any deleterious
1512 variation within the inverted region at this point is fixed within the population of inversions, giving it an innate mutation load. When this mutation load is

lower than the population average in the collinear standard arrangement, the inversion is selectively favoured and may increase in frequency. If the inversion has some other selective advantage independent of deleterious variation, then inversions with an average (or higher) mutation load may also be selected for.

2. **Mutation accumulation:** while the inversion is at low frequency, recombination is rare as recombination can generally only happen within inversion homozygotes, so purifying selection is weak. During this phase, the small inversion population evolves similarly to a small asexual population. In this scenario, Muller's ratchet-style degeneration occurs, increasing the mutation load. For some of this phase, the frequency of the inversion is stochastic and heavily influenced by drift. The time spent in this phase depends on the initial selection coefficient, as the time spent in the drift-dominated phased is inversely related to s_0.

3. **Intermediate frequency:** inversion numbers are sufficiently high that Muller's ratchet degeneration is prevented. Whether they are approaching mutation-selection balance for sexuals or asexuals depends on their frequency, and this can change during the inversion's transit time. The mutation frequencies in the standard arrangement are still at mutation-selection-drift balance, although this is higher than before because their population size is reduced due to the presence of inverted chromosomes. Within inversions, mutations that were captured or accumulated early on are fixed. If inversions spend a long time at intermediate frequency, they approach mutation-selection-drift balance at loci where mutations are not already fixed. But overall these inversions eventually

have a higher mutation load than the standard arrangement, due to their fixed load. Unless there is a source of positive selection the outweighs the fitness cost of this load, the inversion will ultimately be lost.

4. **Fixation:** if the inversion emerges from the mutation accumulation phase fitter than the population mean, and does not spend a long time at intermediate frequency, then it, along with any captured or accumulated mutations, will fix.

Discussion

Nei's model of inversion evolution has been the basis of recent theoretical and simulation studies of inversion evolution (e.g. Connallon, Olito, et al. 2018; Connallon and Olito 2021). In particular, it has been taken for granted that inversions capturing a greater mutation load than any unique advantage they have will ultimately be lost. Here, we show how this relies on the inversion attaining its long-term equilibrium state (though the implausibility of the result was pointed out by Kimura and Ohta 1970). The transient advantage of capturing a better-than-average background generally did not decay before fixation of the inversion could occur. We posit that a Muller's ratchet-style degeneration of inversions occurs while the number of inversions is low, and that this plays a bigger role in determining the inversion's fate than any long-term effects.

The degradation of inversions before fixation requires that they spend sufficiently long at intermediate frequencies. One way this could happen is the case where an in-

version has an intermediate equilibrium frequency. Inversion polymorphism is often maintained by selection on the inversion-linked phenotype, e.g. frequency dependent selection on inversion-linked morphs (e.g. Dagilis and Kirkpatrick 2016; Lamichhaney et al. 2016), or local adaptation inversions under gene flow (e.g. Lee et al. 2017; Stenløkk et al. 2022). However, the vast majority of new inversions are unlikely create supergene-style structures at their time of formation because it relies on pre-existing variation. Alternatively, a new allele could arise on an inversion in a "capture-and-gain" scenario, forming a linked pair of coadapted genes (Schaal, Haller, and Lotterhos 2022) Inversions are often under balancing selection, but this could be a result of selection bias as inversions can only be detected while polymorphic. So, it could be the case that inversions arise and fix relatively often. For example, comparisons between *Drosophila melaongaster* and *Drosophila subobscura* reveal a large number of inverted synteny blocks (Karageorgiou et al. 2019).

This work models haploid individuals rather than diploids. In the diploid case, inversions could carry recessive deleterious alleles such that associative overdominance results (Ohta 1971). If such an allele were fixed among inversions, either by capture or early accumulation, then the inverted arrangement will be under balancing selection and could plausibly degenerate as in the Nei model. Otherwise, recessive mutations could arise on different inversion copies and result in a system of balanced lethals in which the inversion is fixed but various haplotypes segregate (Berdan et al. 2021). Our results are applicable to the codominant case, $h = 1/2$. The early dynamics in a diploid system will depend on hs, though the effect is double-edged when $h < 1/2$. Individual recessive mutations are less likely to contribute towards stochastic loss of the

inversion, but could also slow down or prevent the deterministic increase phase when expressed as homozygotes. In addition, recessive mutations would be at higher frequencies. This case is of particular interest given estimates of $\bar{h} = 0.25$ for deleterious alleles (Manna, Martin, and Lenormand 2011). If $h > 1/2$, then fixation is impeded less, given survival of the initial phase.

Including the probability of capturing a given mutation load, as well as its subsequent probability of fixation, would give a better idea of the nature of inversions likely to be observed (as in Ch. 3 and Connallon and Olito 2021). Inversions capturing relatively mutation-free backgrounds are strongly favoured but rarely arise. The number of mutations carried by an individual is theoretically a Poisson distribution with mean μL, which could easily be incorporated into these results. More accurate would be to record the empirical distribution from the simulations, which would incorporate any additional effects due to selective interference or genetic drift.

We allow recombination to occur within inversions if there are an even number of crossovers. When this occurs in inversions with a high number of mutations it is likely to result in it having fewer, and vice versa. If such events were common, we expect them to depress the invasion probability of invasions capturing good backgrounds, and inflate those that capture poor ones. The probability of a recombination event between an inverted and a standard chromosome is approximately 0.5%, using our model parameters. So, while a purer comparison might use simulations in which inversions completely suppress recombination, it should have little effect here. Berdan et al. 2021 found that gene conversion between arrangements significantly increased the mean time spent segregating, but our results are not compara-

ble because they model only recessive deleterious variation. Then, gene conversion could purge strongly harmful alleles that lower the inversion's equilibrium frequency. In this model, gene conversion between standard and inverted arrangements could remove some of the mutation load fixed in inversions, but further modelling would be required to test how impactful this would be.

A theoretical framework for modelling inversion dynamics is difficult to derive because of the non-linear relationship between mutation accumulation and inversion frequency, and the change in effective rate of recombination. One possibility would be to break the phases down and combine known results about the rate of degeneration due to Muller's ratchet and Nei's model with approximations of the time spent in in each of the stochastic drift and deterministic increase phases (Haigh 1978; Charlesworth 2020). In particular, Muller's ratchet can be considered during the time it takes to reach a population size of at least $e^{\mu L/u}$, which may also include some of the deterministic increase phase. For example, degeneration occurs when there are fewer than approximately 150 inversions under our model parameters. For an inversion capturing the mean mutation load with $s = 0.02$, the stochastic drift phase lasts until it attains a frequency of 0.05, which is expected to take 50 generations. Then the increase from a frequency of 0.05 to 0.15 would take approximately 60 generations (Charlesworth 2020). So, we might expect Muller's ratchet degeneration to occur for 110 generations. From here, the inversion population will start to approach mutation selection balance in a manner dependent on the rate of recombination. In particular, one would need to know at which frequency inversions are common enough to effectively reproduce sexually. Otherwise, branching process models could be used to

model the first phases (Uecker and Hermisson 2011), though this may face tractability issues when inversions can recombine.

We suggest early mutation accumulation probably determines the fate of inversions based on relationships between model parameters and fixation probability, and that long-term mutation-selection balance is seldom attained by inversions. A truer analysis of the effects of transit time in determining the inversion's fate would measure the average mutation load of inversions through time. To do so would require a more detailed analysis of mutation loads over the transit time of an inversion. The simulations also suggest negatively and neutrally selected inversions could also sweep to fixation by capturing a fit background. Although not explicitly modelled, an inversion capturing no deleterious alleles with a selection coefficient of $s = -0.01$ would behave similarly to an inversion capturing 4 mutations with $s = 0.01$ (see Figure 12).

While only inversions under directional selection are considered here, those with intermediate equilibria will undergo the same process of establishment even if they will ultimately be lost. So, these results also relate to the probability that an inversion establishes within the population, because inversion loss was not a result of Nei-style degeneration. Overall, these results suggest that the difficulty inversions experience establishing may be overestimated in the current literature.

General Discussion

The importance of inversions in evolution was theorised from an early stage in the short history of evolutionary biology, before falling out of favour and once again emerging as a hot topic when wide-scale genome sequencing revealed their pervasiveness (Kirkpatrick 2010). The act of suppressing recombination in a confined portion of the genome sounds innocuous, but here we show that it can have far-reaching consequences, influencing evolution at the genome level all the way up to the localised extinction of populations, as demonstrated in this thesis. Understanding the impact of variation in genomic structure is crucial to understand how genomes and individuals evolve. This thesis utilises theoretical population genetics and simulation approaches to model the impacts of inversions on both evolution and ecology. This final section reviews the results contained in this thesis and their importance in aiding our understanding of inversion and inversion-enabled evolution.

Most of the supergene complexes for which we have knowledge of the internal genetic architecture are meiotic drivers. Chapter 2 models the effects of a meiotic drive system, often comprising multiple loci linked by at least one inversion, on X chromosomes. The resultant bias in sex chromosome carried by fertilising sperm causes an

offspring sex ratio bias, which in the extreme can cause population extinction through a lack of males (Hamilton 1967). While substantial theoretical work has since been devoted to studying meiotic drive in the context of mating systems and suppression evolution, little attention has been paid to the demographic consequences of X-linked meiotic drive at polymorphism, despite numerous real life examples of meiotic drivers being maintained at intermediate frequency (e.g. Gileva 1987; James and Jaenike 1990; Fishman and Willis 2005). The key result is that a moderately female biased sex ratio maintained through selection against X-drive can increase the size and persistence time of populations, especially when small. This is because the number of juveniles in a population is limited by the number of reproducing females, and when populations are small they are more susceptible to extinction through stochastic fluctuations in size. This effect occurs within a plausible range of parameters based on experimental data (Finnegan, White, et al. 2019; Meade et al. 2020). So, X-linked meiotic drivers can both increase and decrease population persistence times, depending on the nature of selection against them.

A useful extension of this model would be to add population structure such that patches undergo extinction and recolonisation, with variable patch quality. The hypothesis is that selection will act on a population level such that patches, especially poor quality ones, will tend to have drive populations because wild-type populations will often become extinct. However, X-drive individuals may face difficulties recolonising patches because random sampling could lead to early all-female generations. So, the existence of drive populations may rely on the infiltration of patches colonised by wild-types.

When populations are differently adapted to their environments, dispersal between them can introduce maladaptive alleles into the population, resulting in the production of maladapted offspring. In Chapter 3, we extend the classic model of how inversions can spread when the maintain locally adaptive allele combinations under gene flow. Prior work had considered how this might happen in continent-island models (Kirkpatrick and Barton 2006; Charlesworth and Barton 2018), in which a single population recieves homogenously maladapted migrants from another population, and a two patch model with symmetric migration and selection (Proulx and Teotónio 2022). We analyse a two patch model with the possibility of asymmetric migration and selection, and the role of each model parameter in detail. In contrast to the continent-island model, we found that inversions are most likely to spread in the two-deme model when local adaptation alleles are strongly selected for. Furthermore, when one considers that inversions both have to arise and also capture a locally adaptive haplotype, the establishment of locally adaptive inversions was highest when selection was strong in both models.

Inversions evolve over two scales, because there is selection on both the inversion itself and on its allelic content. In particular, when inversions arise on a background of relatively few deleterious mutations they are afforded a selective advantage over the population mean. Early work examined this in infinite populations, showing that inversions capturing any deleterious variation will ultimately be lost (Nei, Kojima, and Schaffer 1967). This is because the good background captured by the inversion eventually deteriorates to mutation-selection balance like the standard arrangement, but with the addition of extra, fixed mutations. Building on this, (Kimura and Ohta 1970)

showed that in a finite population, fixation can occur before the loss of the inversion's transient advantage. Nei's work has also been used more recently (e.g. Connallon, Olito, et al. 2018; Connallon and Olito 2021) In Chapter 4, we re-analyse Nei's model and show that its implicit assumptions give rise to misleading results. The process of degradation to mutation-selection balance requires recombination to occur within the inversion, which is permissible only when the inversion is at high frequency. Our simulations suggest that the biggest determinant of the fate of an inversion is the extent to which it accumulates mutations early on, in a process similar to Muller's ratchet.

Selection against meiotic drivers is sometimes ascribed to direct fitness costs associated with mutation accumulation within the inversions that maintain meiotic drive gene complexes (Lindholm et al. 2016). We could consider the driver to be a positively selected inversion in the context of the results of Chapter 4. The threshold frequency at which deterministic increase begins is low, and the ascendancy is quick, so we predict that the effects of early stage mutation accumulation should have little effect on strong meiotic drivers. This also suggests that it is unlikely enough degeneration will occur afterwards to prevent fixation. So, unless the first inversion(s) happened to capture strongly deleterious recessive alleles, it is likely that the forces maintaining polymorphism are unrelated to mutation accumulation — at least, initially.

Inversions involved in local adaptation under gene flow are likely to be at some intermediate frequency, either in distinct populations or in a clinal structure. As such, they should reach mutation-selection balance similarly to Nei's model of inversion evolution. From this, we could plausibly infer that selection for the local adaptation alleles

is stronger than selection against both the deleterious mutations load captured by the inversion, and its breakpoint effects (which may also be positive — see (e.g. Lamichhaney et al. 2016; Lee et al. 2017; Li et al. 2016) for an example of breakpoint-induced local adaptation). This imposes a further restriction on the alleles likely to be involved in inversions. Not only must they be strongly selected so that they resist swamping by migration, they must also outweigh any harmful effects of the inversion.

In conclusion, this thesis has investigated the impacts inversions can have at all levels of evolution. Maintaining linkage between alleles can be a powerful force, allowing for the evolution of some of nature's most interesting and extravagant phenotypes. In short, inversions offer a solution to Felsenstein's dilemma. They provide a region in which linkage between coadapted alleles is maintained, while allowing adaptive evolution to proceed in the long term when recombination once again ensues.

Bibliography

Akerman, A. and R. Bürger (2014). "The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model". *Journal of Mathematical Biology* 68.5, pp. 1135–1198.

Ardlie, K. G. (1998). "Putting the brake on drive: meiotic drive of t haplotypes in natural populations of mice". *Trends in Genetics* 14.5, pp. 189–193.

Ayala, D., R. F. Guerrero, and M. Kirkpatrick (2013). "Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito". *Evolution* 67.4, pp. 946–958.

Barton, N. H. and B. Charlesworth (1998). "Why Sex and Recombination?" *Science* 281.5385, pp. 1986–1990.

Berdan, E. L. et al. (2021). "Deleterious mutation accumulation and the long-term fate of chromosomal inversions". *PLOS Genetics* 17.3, e1009411.

Betancourt, A. J., J. J. Welch, and B. Charlesworth (2009). "Reduced effectiveness of selection caused by a lack of recombination". *Current Biology* 19.8, pp. 655–660.

Boven, M. V. and F. J. Weissing (1999). “Segretation distortion in a deme-structured
1758 population: opposing demands of gene, individual and group selection”. *Journal of Evolutionary Biology* 12.1, pp. 80–93.

1760 Boyle, M. et al. (2014). “Under what conditions do climate-driven sex ratios enhance versus diminish population persistence?” *Ecology and Evolution* 4.23, pp. 4522–
1762 4533.

Bürger, R. and A. Akerman (2011). “The effects of linkage and gene flow on local adap-
1764 tation: A two-locus continent–island model”. *Theoretical Population Biology* 80.4, pp. 272–288.

1766 Burt, A. and A. Deredec (2018). “Self-limiting population genetic control with sex-linked genome editors”. *Proceedings of the Royal Society B: Biological Sciences*
1768 285.1883, p. 20180776.

Burt, A. and R. Trivers (2006). *Genes in Conflict*. Harvard University Press.

1770 Charlesworth, B., P. H. Harvey, et al. (2000). “The degeneration of Y chromosomes”. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sci-*
1772 *ences* 355.1403, pp. 1563–1572.

Charlesworth, B. (1974). “Inversion polymorphism in a two-locus genetic system”. *Ge-*
1774 *netics Research* 23.3, pp. 259–280.

Charlesworth, B. (1996). “Background selection and patterns of genetic diversity in
1776 Drosophila melanogaster”. *Genetics Research* 68.2, pp. 131–149.

Charlesworth, B. (2020). “How Long Does It Take to Fix a Favorable Mutation, and
1778 Why Should We Care?” *The American Naturalist* 195.5, pp. 753–771.

Charlesworth, B. and N. H. Barton (2018). “The spread of an inversion with migration and selection”. *Genetics* 208.1, pp. 377–382.

Charlesworth, B. and D. L. Hartl (1978). “Population dyanmics of the segregation distorter polymorphism of *Drosophila melanogaster*”. *Genetics* 89.1, pp. 171–192.

Charlesworth, D. and B. Charlesworth (1979). “Selection on recombination in clines”. *Genetics* 91.3, pp. 581–589.

Charlesworth, D. (2016). “The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations”. *Evolutionary Applications* 9.1, pp. 74–90.

Cheng, C. et al. (2012). “Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach”. *Genetics* 190.4, pp. 1417–1432.

Christmas, M. J. et al. (2019). “Chromosomal inversions associated with environmental adaptation in honeybees”. *Molecular Ecology* 28.6, pp. 1358–1374.

Connallon, T. and C. Olito (2021). “Natural selection and the distribution of chromosomal inversion lengths”. *Molecular Ecology* 31.13.

Connallon, T., C. Olito, et al. (2018). “Local adaptation and the evolution of inversions on sex chromosomes and autosomes”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 373.1757, p. 20170423.

Corbett-Detig, R. B. (2016). “Selection on inversion breakpoints favors proximity to pairing sensitive sites in *Drosophila melanogaster*”. *Genetics* 204.1, pp. 259–265.

Coyne, J. A. et al. (1993). “The fertility effects of pericentric inversions in Drosophila melanogaster.” *Genetics* 134.2, pp. 487–496.

Curtsinger, J. W. and M. W. Feldman (1980). “Experimental and theoretical analysis of the "Sex-Ratio" polymorphism in *Drosophila pseudoobscura*”. *Genetics* 94.2, pp. 445–466.

Dagilis, A. J. and M. Kirkpatrick (2016). “Prezygotic isolation, mating preferences, and the evolution of chromosomal inversions”. *Evolution* 70.7, pp. 1465–1472.

Dobzhansky, T. (1947). “Genetics of natural populations. XIV. A response of certain gene arrangements in the third chormosome of Drosophila pseudoobscura to natural selection”. *Genetics* 32.2, pp. 142–160.

Dobzhansky, T. (1955). “An extreme case of heterosis in a central american population of Drosophila Tropicalis”. *Proceedings of the National Academy of Sciences*.

Dobzhansky, T. and T. G. Dobzhansky (1970). *Genetics of the Evolutionary Process*. Columbia University Press.

Dyer, K. A., B. Charlesworth, and J. Jaenike (2007). “Chromosome-wide linkage disequilibrium as a consequence of meiotic drive”. *Proceedings of the National Academy of Sciences* 104.5, pp. 1587–1592.

Dyer, K. A. and D. W. Hall (2019). “Fitness consequences of a non-recombining *sex-ratio* drive chromosome can explain its prevalence in the wild”. *Proceedings of the Royal Society B: Biological Sciences* 286.1917, p. 20192529.

Dyson, E. A. and G. D. D. Hurst (2004). “Persistence of an extreme sex-ratio bias in a natural population”. *Proceedings of the National Academy of Sciences* 101.17, pp. 6520–6523.

Edelstein-Keshet, L. (1987). *Mathematical Models in Biology*. SIAM.

Edwards, A. W. F. (1961). “The population genetics of “sex-ratio” in *Drosophila pseudoobscura*”. *Heredity* 16.3, pp. 291–304.

Entani, T. et al. (1999). “Centromeric localization of an S-RNase gene in Petunia hybrida Vilm.” *Theoretical and Applied Genetics* 99.3, pp. 391–397.

Faria, R., P. Chaube, et al. (2019). “Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes”. *Molecular Ecology* 28.6, pp. 1375–1393.

Faria, R., K. Johannesson, et al. (2019). “Evolving inversions”. *Trends in Ecology & Evolution* 34.3, pp. 239–248.

Feder, J. L., R. Gejji, et al. (2011). “Adaptive chromosomal divergence driven by mixed geographic mode of evolution”. *Evolution* 65.8, pp. 2157–2170.

Feder, J. L. and P. Nosil (2009). “Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions?” *Evolution* 63.12, pp. 3061–3075.

Felsenstein, J. (1974). “The evolutionary advantage of recombination”. *Genetics* 78.2, pp. 737–756.

Finnegan, S. R. (2020). “The effect of sex-ratio meiotic drive on sex, survival, and size in the Malaysian stalk-eyed fly, Teleopsis dalmanni”. PhD thesis. University College London.

Finnegan, S. R., L. Nitsche, et al. (2019). “Does meiotic drive alter male mate preference?” *Behavioral Ecology* 31.1. Ed. by P. Smiseth, pp. 194–201.

Finnegan, S. R., N. J. White, et al. (2019). “Meiotic drive reduces egg-to-adult viability in stalk-eyed flies”. *Genetics* 89, pp. 171–192.

Fisher, R. A. (1930). *The genetical theory of natural selection.*

Fishman, L. and J. H. Willis (2005). “A novel meiotic drive locus almost completely distorts segregation in *Mimulus* (monkeyflower) hybrids”. *Genetics* 169.1, pp. 347–353.

Frank, S. A. (1991). “Haldane’s Rule: A Defense of the Meiotic Drive Theory”. *Evolution* 45.7, pp. 1714–1717.

Galizi, R. et al. (2014). “A synthetic sex ratio distortion system for the control of the human malaria mosquito”. *Nature Communications* 5.1, pp. 1–8.

Gileva, E. A. (1987). “Meiotic drive in the sex chromosome system of the varying lemming, *Dicrostonyx torquatus Pall.* (*Rodentia, Microtinae*)”. *Heredity* 59.3, pp. 383–389.

Haag-Liautard, C. et al. (2007). “Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila”. *Nature* 445.7123, pp. 82–85.

Hager, E. R. et al. (2022). “A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes”. *Science* 377.6604, pp. 399–405.

Haigh, J. (1978). “The accumulation of deleterious genes in a population—Muller’s Ratchet”. *Theoretical Population Biology* 14.2, pp. 251–267.

Haldane, J. B. S. (1957). “The conditions for coadaptation in polymorphism for inversions”. *Journal of Genetics* 55.1, pp. 218–225.

Hale, D. W. (1986). “Heterosynapsis and suppression of chiasmata within heterozygous pericentric inversions of the Sitka deer mouse”. *Chromosoma* 94.6, pp. 425–432.

Haller, B. C. and P. W. Messer (2022). “SLiM 4: Multispecies Eco-Evolutionary Modeling”. *The American Naturalist*, E000–E000.

Hamilton, W. D. (1967). “Extraordinary sex ratios”. *Science* 156.3774, pp. 477–488.

Hardy, I. C. and R. A. Boulton (2019). “Sex Allocation, Sex Ratios and Reproduction”. *Encyclopedia of Animal Behavior*. 2nd. Elsevier, pp. 464–471.

Harringmeyer, O. S. and H. E. Hoekstra (2022). “Chromosomal inversion polymorphisms shape the genomic landscape of deer mice”. *Nature Ecology & Evolution*, pp. 1–15.

Hartwell, L. (2000). *Genetics: From Genes to Genomes*. McGraw-Hill.

Hatcher, M. J. (2000). “Persistence of selfish genetic elements: population structure and conflict”. *Trends in Ecology & Evolution* 15.7, pp. 271–277.

Hatcher, M. J. et al. (1999). “Population dynamics under parasitic sex ratio distortion”. *Theoretical Population Biology* 56.1, pp. 11–28.

Hays, G. C. et al. (2017). “Population viability at extreme sex-ratio skews produced by temperature-dependent sex determination”. *Proceedings of the Royal Society B: Biological Sciences* 284.1848, p. 20162576.

Helleu, Q., P. R. Gérard, and C. Montchamp-Moreau (2015). “Sex Chromosome Drive”. *Cold Spring Harbor Perspectives in Biology* 7.2, a017616.

Hickey, W. A. and G. B. Craig (1966). “Genetic distortion of sex ratio in a mosquito, *Aedes aegypti*”. *Genetics* 53, pp. 1177–1196.

Hill, W. G. and A. Robertson (1966). “The effect of linkage on limits to artificial selection”. *Genetical Research* 8.3, pp. 269–294.

Hitchcock, T. J. and A. Gardner (2020). “A gene’s-eye view of sexual antagonism”. *Proceedings of the Royal Society B: Biological Sciences* 287.1932, p. 20201633.

Holman, L. et al. (2015). “Coevolutionary dynamics of polyandry and sex-linked meiotic drive”. *Evolution* 69.3, pp. 709–720.

Huang, K., R. L. Andrew, et al. (2020). “Multiple chromosomal inversions contribute to adaptive divergence of a dune sunflower ecotype”. *Molecular Ecology* 29.14, pp. 2535–2549.

Huang, K. and L. H. Rieseberg (2020). “Frequency, origins, and evolutionary role of chromosomal inversions in plants”. *Frontiers in Plant Science* 11.

Jaenike, J. (1996). “Sex-ratio meiotic drive in the *Drosophila quinaria Group*”. *The American Naturalist* 148.2, pp. 237–254.

James, A. C. and J. Jaenike (1990). ““Sex ratio” meiotic drive in *Drosophila testacea*”. *Genetics* 126, pp. 651–656.

Jay, P. et al. (2022). “Sheltering of deleterious mutations explains the stepwise extension of recombination suppression on sex chromosomes and other supergenes”. *PLOS Biology* 20.7, e3001698.

Jones, F. C. et al. (2012). “The genomic basis of adaptive evolution in threespine sticklebacks”. *Nature* 484.7392, pp. 55–61.

Joron, M. et al. (2011). “Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry”. *Nature* 477.7363, pp. 203–206.

Karageorgiou, C. et al. (2019). “Long-read based assembly and synteny analysis of a reference Drosophila subobscura genome reveals signatures of structural evolution driven by inversions recombination-suppression effects”. *BMC Genomics* 20.1, p. 223.

Keais, G. L., S. Lu, and S. J. Perlman (2020). “Autosomal suppression and fitness costs of an old driving X chromosome in *Drosophila testacea*”. *Journal of Evolutionary Biology* 00, pp. 1–10.

Kern, P. et al. (2015). “Double trouble: combined action of meiotic drive and *Wolbachia* feminization in *Eurema* butterflies”. *Biology Letters* 11.5, p. 20150095.

Kim, K.-W. et al. (2021). *Stepwise evolution of a butterfly supergene via duplication and inversion.* preprint. Evolutionary Biology.

Kimura, M. and T. Ohta (1970). “Probability of fixation of a mutant gene in a finite population when selective advantage decreases with time”. *Genetics* 65.3, pp. 525–534.

Kimura, M. and T. Ohta (1973). “The age of a neutral mutant persisting in a finite population”. *Genetics* 75.1, pp. 199–212.

Kirkpatrick, M. (2010). “How and why chromosome inversions evolve”. *PLOS Biology* 8.9, e1000501.

Kirkpatrick, M. and N. Barton (2006). “Chromosome inversions, local adaptation and speciation”. *Genetics* 173.1, pp. 419–434.

Kirkpatrick, M., T. Johnson, and N. Barton (2002). “General models of multilocus evolution”. *Genetics* 161.4, pp. 1727–1750.

Koch, E. L. et al. (2021). “Genetic variation for adaptive traits is associated with polymorphic inversions in Littorina saxatilis”. *Evolution Letters* 5.3, pp. 196–213.

Kruger, A. N. et al. (2019). “A neofunctionalized X-linked ampliconic gene family Is essential for male fertility and equal sex ratio in mice”. *Current Biology* 29.21, 3699–3706.e5.

1936 Lamichhaney, S. et al. (2016). “Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax)”. *Nature Genetics* 48.1, pp. 84–
1938 88.

Larner, W. et al. (2019). “An X-linked meiotic drive allele has strong, recessive fitness
1940 costs in female *Drosophila pseudoobscura*”. *Proceedings of the Royal Society B: Biological Sciences* 286.1916, p. 20192038.

1942 Larracuente, A. M. and D. C. Presgraves (2012). “The selfish *Segregation Distorter* gene complex of *Drosophila melanogaster*”. *Genetics* 192.1, pp. 33–53.

1944 Lee, C.-R. et al. (2017). “Young inversion with multiple linked QTLs under selection in a hybrid zone”. *Nature Ecology & Evolution* 1.5, pp. 1–13.

1946 Lenormand, T. (2002). “Gene flow and the limits to natural selection”. *Trends in Ecology & Evolution* 17.4, pp. 183–189.

1948 Lenormand, T. and S. P. Otto (2000). “The Evolution of Recombination in a Heterogeneous Environment”. *Genetics* 156.1, pp. 423–438.

1950 Lenormand, T. and D. Roze (2022). “Y recombination arrest and degeneration in the absence of sexual dimorphism”. *Science* 375.6581, pp. 663–666.

1952 Li, J. et al. (2016). “Genetic architecture and evolution of the S locus supergene in Primula vulgaris”. *Nature Plants* 2.12, pp. 1–7.

1954 Lindholm, A. K. et al. (2016). “The ecology and evolutionary dynamics of meiotic drive”. *Trends in Ecology & Evolution* 31.4, pp. 315–326.

1956 Ling, A. and R. Cordaux (2010). “Insertion sequence inversions mediated by ectopic recombination between terminal inverted repeats”. *PLOS ONE* 5.12, e15654.

Lowry, D. B. and J. H. Willis (2010). “A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation”. *PLOS Biology* 8.9, e1000500.

Mackintosh, C., A. Pomiankowski, and M. F. Scott (2021). “X-linked meiotic drive can boost population size and persistence”. *Genetics* 217.1, iyaa018.

Manna, F., G. Martin, and T. Lenormand (2011). “Fitness Landscapes: An Alternative Theory for the Dominance of Mutation”. *Genetics* 189.3, pp. 923–937.

Manser, A., B. König, and A. K. Lindholm (2020). “Polyandry blocks gene drive in a wild house mouse population”. *Nature Communications* 11.1, p. 5590.

Manser, A., A. K. Lindholm, et al. (2017). “Sperm competition suppresses gene drive among experimentally evolving populations of house mice”. *Molecular Ecology* 26.20, pp. 5784–5792.

Maynard Smith, J. (1998). *Evolutionary Genetics*. Oxford University Press.

Meade, L. et al. (2020). “Maintenance of fertility in the face of meiotic drive”, p. 49.

Mérot, C. et al. (2020). “A roadmap for understanding the evolutionary significance of structural genomic variation”. *Trends in Ecology & Evolution* 35.7, pp. 561–572.

Muller, H. J. (1932). “Some genetic aspects of sex”. *The American Naturalist* 66.703, pp. 118–138.

Muller, H. J. (1964). “The relation of recombination to mutational advance”. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 1.1, pp. 2–9.

Navarro, A., A. Barbadilla, and A. Ruiz (2000). “Effect of inversion polymorphism on the neutral nucleotide variability of linked chromosomal regions in Drosophila”. *Genetics* 155.2, pp. 685–698.

Navarro, A., E. Betrán, et al. (1997). "Recombination and Gene Flux Caused by Gene Conversion and Crossing Over in Inversion Heterokaryotypes". *Genetics* 146.2, pp. 695–709.

Nei, M., K.-I. Kojima, and H. E. Schaffer (1967). "Frequency changes of new inversions in populations under mutation-selection equilibria". *Genetics* 57.4, pp. 741–750.

Ohta, T. (1971). "Associative overdominance caused by linked detrimental mutations*". *Genetics Research* 18.3, pp. 277–286.

Olds-Clarke, P. and L. R. Johnson (1993). "*t* haplotypes in the mouse compromise sperm flagellar function". *Developmental Biology* 155.1, pp. 14–25.

Otto, S. P. and T. Day (2011). *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press.

Ovaskainen, O. and B. Meerson (2010). "Stochastic models of population extinction". *Trends in Ecology & Evolution* 25.11, pp. 643–652.

Parker, G. A. (1990). "Sperm competition games: raffles and roles". *Proceedings of the Royal Society of London. Series B: Biological Sciences* 242.1304, pp. 120–126.

Patten, M. M. (2019). "The X chromosome favors males under sexually antagonistic selection". *Evolution* 73.1, pp. 84–91.

Peichel, C. L. and D. A. Marques (2017). "The genetic and molecular architecture of phenotypic diversity in sticklebacks". *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1713, p. 20150486.

Pénisson, S. et al. (2013). "Lineage dynamics and mutation–selection balance in non-adapting asexual populations". *Journal of Statistical Mechanics: Theory and Experiment* 2013.01, P01013.

Pinzone, C. A. and K. A. Dyer (2013). "Association of polyandry and *sex-ratio* drive prevalence in natural populations of *Drosophila neotestacea*". *Proceedings of the Royal Society B: Biological Sciences* 280.1769, p. 20131397.

Poelstra, J. W. et al. (2014). "The genomic landscape underlying phenotypic integrity in the face of gene flow in crows". *Science* 344.6190, pp. 1410–1414.

Pomiankowski, A. and L. D. Hurst (1999). "Driving sexual preference". *Trends in Ecology & Evolution* 14.11, pp. 425–426.

Price, T. A. R., D. J. Hodgson, et al. (2008). "Selfish genetic elements promote polyandry in a fly". *Science* 322.5905, pp. 1241–1243.

Price, T. A. R., A. Bretman, et al. (2014). "Does polyandry control population sex ratio via regulation of a selfish gene?" *Proceedings of the Royal Society B: Biological Sciences* 281.1783, p. 20133259.

Price, T. A. R., G. D. D. Hurst, and N. Wedell (2010). "Polyandry prevents extinction". *Current Biology* 20.5, pp. 471–475.

Proulx, S. R. and H. Teotónio (2022). "Selection on modifiers of genetic architecture under migration load". *PLOS Genetics* 18.9, e1010350.

Prowse, T. A. et al. (2019). "A Y-chromosome shredding gene drive for controlling pest vertebrate populations". *eLife* 8. Ed. by P. W. Messer and D. Tautz, e41873.

Rathje, C. C. et al. (2019). "Differential sperm motility mediates the sex ratio drive shaping mouse sex chromosome evolution". *Current Biology* 29.21, 3692–3698.e4.

Rice, W. R. (1984). "Sex chromosomes and the evolution of sexual dimorphism". *Evolution* 38.4, pp. 735–742.

Roberts, P. A. (1976). “genetics of chromosome aberration”. *Genetics and biology of Drosophila*.

Runge, J.-N. and A. K. Lindholm (2018). “Carrying a selfish genetic element predicts increased migration propensity in free-living wild house mice”. *Proceedings of the Royal Society B: Biological Sciences* 285.1888, p. 20181333.

Samuk, K. et al. (2017). “Gene flow and selection interact to promote adaptive divergence in regions of low recombination”. *Molecular Ecology* 26.17, pp. 4378–4390.

Sandler, L., Y. Hiraizumi, and I. Sandler (1959). “Meiotic drive in natural populations of *Drosophila melanogaster*. I. the cytogenetic basis of segregation-distortion”. *Genetics* 44.2, pp. 233–250.

Santos, M. (2009). “Recombination Load in a Chromosomal Inversion Polymorphism of Drosophila subobscura”. *Genetics* 181.2, pp. 803–809.

Schaal, S. M., B. C. Haller, and K. E. Lotterhos (2022). “Inversion invasions: when the genetic basis of local adaptation is concentrated within inversions in the face of gene flow”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377.1856, p. 20210200.

Schimenti, J. (2000). “Segregation distortion of mouse t haplotypes: the molecular basis emerges”. *Trends in Genetics* 16.6, pp. 240–243.

Sharp, N. P. and S. P. Otto (2016). “Evolution of sex: Using experimental genomics to select among competing theories”. *BioEssays* 38.8, pp. 751–757.

Stenløkk, K. et al. (2022). “The emergence of supergenes from inversions in Atlantic salmon”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 377.1856, p. 20210195.

Sturtevant, A. H. and G. W. Beadle (1936). “The relations of inversions in the X chromosome of Drosophila melanogaster to crossing over and disjunction”. *Genetics* 21.5, pp. 554–604.

Sturtevant, A. H. (1917). “Genetic Factors Affecting the Strength of Linkage in Drosophila”. *Proceedings of the National Academy of Sciences* 3.9, pp. 555–558.

Sutter, A. et al. (2019). “No selection for change in polyandry under experimental evolution”. *Journal of Evolutionary Biology* 32.7, pp. 717–730.

Tao, Y. et al. (2007). “A sex-ratio meiotic drive system in *Drosophila simulans*. I: an autosomal suppressor”. *PLoS Biology* 5.11. Ed. by D. Barbash, e292.

Taylor, D. R., M. J. Saur, and E. Adams (1999). “Pollen performance and sex-ratio evolution in a dioecious plant”. *Evolution* 53.4, pp. 1028–1036.

Taylor, J. E. and J. Jaenike (2002). “Sperm competition and the dynamics of X chromosome drive: stability and extinction”. *Genetics* 160, pp. 1721–1731.

Taylor, J. E. and J. Jaenike (2003). “Sperm competition and the dynamics of X chromosome drive in finite and structured populations”. *Ann. Zool. Fennici.* 40, pp. 195–206.

Thompson, M. J. and C. D. Jiggins (2014). “Supergenes and their role in evolution”. *Heredity* 113.1, pp. 1–8.

Tigano, A. and V. L. Friesen (2016). “Genomics of local adaptation with gene flow”. *Molecular Ecology* 25.10, pp. 2144–2164.

Trickett, A. J. and R. K. Butlin (1994). “Recombination suppressors and the evolution of new species”. *Heredity* 73.4, pp. 339–345.

Turner, B. C. and D. D. Perkins (1979). “Spore killer, a chromosomal factor in *Neurospora* that kills meiotic products not containing it”. *Genetics* 93.3, pp. 587–606.

Úbeda, F. and D. Haig (2005). “On the evolutionary stability of mendelian segregation”. *Genetics* 170.3, pp. 1345–1357.

Uecker, H. and J. Hermisson (2011). “On the Fixation Process of a Beneficial Mutation in a Variable Environment”. *Genetics* 188.4, pp. 915–930.

Unckless, R. L. and A. G. Clark (2014). “Sex-ratio meiotic drive and interspecific competition”. *Journal of Evolutionary Biology* 27.8, pp. 1513–1521.

Villoutreix, R. et al. (2021). “Inversion breakpoints and the evolution of supergenes”. *Molecular Ecology* 30.12, pp. 2738–2755.

Wasserman, M. (1968). “Recombination-induced chromosomal heterosis”. *Genetics* 58.1, pp. 125–139.

Wedell, N. (2013). “The dynamic relationship between polyandry and selfish genetic elements”. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368.1613, p. 20120049.

Wellenreuther, M. and L. Bernatchez (2018). “Eco-evolutionary genomics of chromosomal inversions”. *Trends in Ecology & Evolution* 33.6, pp. 427–440.

Werren, J. H. (2011). “Selfish genetic elements, genetic conflict, and evolutionary innovation”. *Proceedings of the National Academy of Sciences* 108.Supplement 2, pp. 10863–10870.

West, S. (2009). *Sex Allocation*. Princeton University Press.

White, M. J. D. (1978). *Modes of Speciation*. W. H. Freeman.

Windbichler, N., P. A. Papathanos, and A. Crisanti (2008). “Targeting the X chromosome during spermatogenesis Induces Y chromosome transmission ratio distortion and early dominant embryo lethality in *Anopheles gambiae*”. *PLOS Genetics* 4.12, e1000291.

Yeaman, S. (2015). “Local adaptation by alleles of small effect”. *The American Naturalist* 186.S1, S74–S89.

Yeaman, S. and M. C. Whitlock (2011). “The genetic architecture of adaptation under migration–selection balance”. *Evolution* 65.7, pp. 1897–1911.

Appendix 1

Alternative form of density dependence

In the main text, we assumed that competition for resources among adults is a source of density dependent selection by reducing the survival or fecundity of adult females. The assumption is that the density dependence is generated by the population size (αN), but not by the birth rate (b). Here, we explore an alternative form of density dependence in which competition for resources can cause the population size to be depressed as population birth rate increases. For instance, if the density dependence is defined by

$$(1 - b\alpha N), \tag{A1}$$

then increasing the birth rate does not always increase population size (Figure A1). Without meiotic drive, the equilibrium population size is

$$\hat{N}|_{p=0} = \frac{b-2}{b^2\alpha}, \tag{A2}$$

which now includes a quadratic term in b not present in (Eq 8). Thus, when birth rates are very high, the equilibrium population size decreases because competition becomes more intense. For example, if competition is a function of the number of

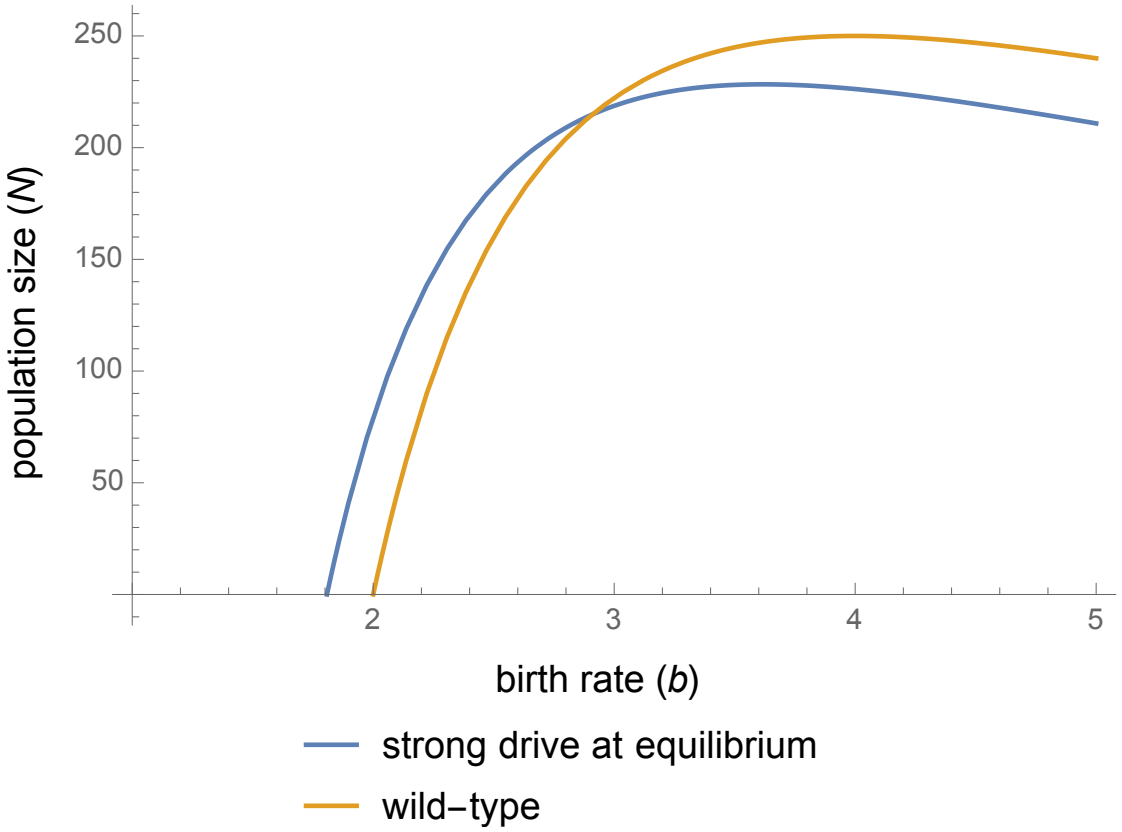


Figure A1: Equilibrium population size given density dependence is based on the the intrinsic birth rate (b). As before, meiotic drive allows the population to persist with lower birth rates ($b < 2$). But with higher values of the birth rate ($b > 3$), meiotic drive reduces population size. Parameter values: $s_m = 0$, $c = 1$, $h = 0.1$, $s_f = 0.8$, $\lambda_f = 1$, $\alpha = 10^{-3}$.

juveniles $J = bN$, then high birth rates both increase the number of juveniles, J, and increase the strength of competition among them.

As in our main results, we find that the intrinsic birth rate must be at least two for wild-type populations to persist whereas populations with drive can persist with a lower intrinsic birth rate (Figure A1). However, meiotic drive does not always increase population size in this scenario because increasing the effective birth rate by biasing the sex ratio towards females does not always lead to larger populations. Thus, some forms of density dependence could mean that increased birth rates do not increase population size, in which case the effect of meiotic drive on boosting the effective birth rate may change. However, we expect that increased birth rates will increase population size in most models of intraspecific competition.

Mathematica notebook

This section contains a pdf version of a Mathematica notebook containing derivations of results in Chapter 2.

X-linked meiotic drive can boost population size and persistence

Carl Mackintosh, Andrew Pomiankowski, and Michael F Scott

Overview

This notebook analyses the eco-evolutionary dynamics of X-linked meiotic drive.

In this file, we assume that the birth rate declines linearly with population size. This assumption means that increased birth rates will increase the population size, which we expect to be generally true. Different assumptions about density-dependence may mean that increases in birth rate can actually decrease population size. We explore one such possibility in another Supplementary notebook. Further supplementary material contains python code used to conduct stochastic simulations of population size and persistence when there is X-linked meiotic drive.

Since the spread of meiotic drive alleles may be affected the degree of polyandry via sperm competition, we investigate cases where females mate once ($\lambda_f = 1$), twice ($\lambda_f = 2$), or many times ($\lambda_f \to \infty$), each of which is presented in a separate section. The first section (Common Notation and Substitutions), should be entered before evaluating any of these sections.

Common Notation and Substitutions [ENTER FIRST]

We track adult (diploid) genotypes at an X linked locus that experiences meiotic drive in males. There are therefore 5 genotypes: 3 female genotypes and 2 male genotypes. In this notebook we use upper case D for the wildtype allele (which could also be called S_t or X) and lower case d to give the meiotic drive allele (which could also be called S_r or X_d). The 5 adult genotypes are therefore
females:
$S_tS_t = XX$= DDf
$S_rS_t = X_dX$ = Ddf
$S_rS_r = X_dX_d$ = ddf
males:
S_tY = XY = Dm
S_rY = X_dY = dm
such that
ijf is the density of females with genotype ij
im is the density of males with genotype i

We assume that the X-linked meiotic drive allele biases transmission in males, such that a $\left(\frac{(1+\delta)}{2}\right)$ fraction of gametes produced by a meiotic drive male (dm) carry the drive allele (d).

2130

Furthermore, we assume that the ejaculate size of meiotic drive males, relative to wildtype males, is c.

We use the following substitutions to convert between these genotype densities and the total density of males (maleNum) and females (femaleNum), and the frequency of the drive allele among males (pm) and females (pf), and the inbreeding coefficient (F). These quantities are defined by substitutionEquations, subeq can be used to convert from the genotype densities into male/female densities and allele frequencies.

```
In[•]:= substitutionEquations = {
          femaleNum == DDf + Ddf + ddf,
          maleNum == (dm + Dm),
          pm == dm/(dm + Dm),
          ((1 - pf)^2 (1 - F) + F (1 - pf)) == DDf/(DDf + Ddf + ddf),
          2 pf (1 - pf) (1 - F) == Ddf/(DDf + Ddf + ddf),
          (pf)^2 (1 - F) + F pf == ddf/(DDf + Ddf + ddf)
        };
      subeq = Simplify[Solve[substitutionEquations, {dm, Dm, DDf, Ddf, ddf}]]
Out[•]= {{dm → maleNum pm, Dm → maleNum - maleNum pm,
    DDf → -femaleNum (-1 + pf) (1 + (-1 + F) pf),
    Ddf → 2 (-1 + F) femaleNum (-1 + pf) pf, ddf → femaleNum pf (F + pf - F pf)}}
```

In the absence of any competition for resources, each female will survive to produce b surviving juveniles. Density dependence acts to reduce the survival of females to reproduction, survival of juveniles, and/or female fecundity (which are all modelled equivalently) by a factor of (1-α N) where N is the total density of males and females (maleNum+femaleNum). That is α is the per adult competitive effect on reproductive rate. totAdults gives the total population density of males and females.

```
In[•]:= totAdults = DDf + Ddf + ddf + Dm + dm;
```

Before juveniles reach adulthood, they experience selection according to their genotype. The relative fitness of different genotypes is given by
$W_{XX} = 1$
$W_{X_d X} = 1 - h\,sf$
$W_{X_d X_d} = 1 - sf$
$W_{XY} = 1$
$W_{X_d Y} = 1 - sm$
We use ijfprime and imprime to specify the densities of each genotype after one generation (recursion/difference equations for change in genotype densities). For example, DDfprime is the density of females with genotype DD after one generation.

Allele Frequency and Population Size Dynamics - females mate with one male ($\lambda_f = 1$)

Recursion equations

We consider each pairwise combination of male and female genotypes and the probability that each mating type produces each offspring genotype. We assume random mating so that each mating combination happens in proportion to the frequency of males with each genotype (e.g., $\frac{dm}{(Dm+dm)}$ is the frequency of drive males among males). These equations track all mating combinations that produce each juvenile genotype. The rate of birth and density-dependence competition (b*(1-α totAdults)) is not directly dependent on genotype.

```
In[•]:= DDfjuveniles = b * (1 - α totAdults) (DDf * Dm / (Dm + dm) * 1/2 +
          Ddf/2 * Dm / (Dm + dm) * 1/2);

        Ddfjuveniles = b * (1 - α totAdults) (DDf * dm / (Dm + dm) * ((1 + δ)/2) +
          Ddf/2 (Dm / (Dm + dm) * 1/2 + dm / (Dm + dm) * ((1 + δ)/2)) +
          ddf * Dm / (Dm + dm) * 1/2);

        ddfjuveniles = b * (1 - α totAdults) (Ddf/2 * dm / (Dm + dm) * ((1 + δ)/2) +
          ddf * dm / (Dm + dm) * ((1 + δ)/2));

        Dmjuveniles = b * (1 - α totAdults) (DDf * Dm / (Dm + dm) * 1/2 +
          Ddf/2 * (Dm / (Dm + dm) * 1/2 + dm / (Dm + dm) * ((1 - δ)/2)) +
          DDf * dm / (Dm + dm) * ((1 - δ)/2));

        dmjuveniles = b * (1 - α totAdults) (Ddf/2 * (dm / (Dm + dm) * ((1 - δ)/2) + Dm / (Dm + dm) * 1/2) +
          ddf * (Dm / (Dm + dm) 1/2 + dm / (Dm + dm) * ((1 - δ)/2)));
```

Before the next generation of adults, viability selection acts according to genotype.

```
In[•]:= DDfprime = DDfjuveniles * 1;
        Ddfprime = Ddfjuveniles * (1 - h * sf);
        ddfprime = ddfjuveniles * (1 - sf);
        Dmprime = Dmjuveniles * 1;
        dmprime = dmjuveniles * (1 - sm);
```

2132

Equilibrium allele frequency and population size

We first look at only the difference in allele frequency and inbreeding coefficient to calculate the polymorphic allele frequency equilibrium (when changes in allele frequency and inbreeding coefficient are 0)

In[•]:=
```
solveme =
  Simplify[{(dmprime/(Dmprime + dmprime)) - dm/(dm + Dm),
    (2 ddfprime + Ddfprime)/(2 (ddfprime + Ddfprime + DDfprime)) -
     (2 ddf + Ddf)/(2 (ddf + Ddf + DDf)),
    (4 ddfprime DDfprime - Ddfprime^2)/((2 ddfprime + Ddfprime) (Ddfprime + 2 DDfprime)) -
     (4 ddf DDf - Ddf^2)/((2 ddf + Ddf) (Ddf + 2 DDf))} /. subeq];
equilibrium = Solve[solveme[[1]] == {0, 0, 0}, {pm, pf, F}] // Simplify
```

Out[•]=
$$\left\{\left\{\text{pm} \to \frac{(-1+\text{sm})\,(\delta-\text{sm}\,(1+\delta)+\text{h sf}\,(-2+\text{sm}-\delta+\text{sm}\,\delta))}{2\,\text{sf}\,(-1+\text{sm})\,(1+\delta)+\text{h sf}\,(-2+\text{sm})\,(-2+\text{sm}-\delta+\text{sm}\,\delta)-\text{sm}\,(\text{sm}-\delta+\text{sm}\,\delta)},\right.\right.$$
$$\text{pf} \to \frac{\delta-\text{sm}\,(1+\delta)+\text{h sf}\,(-2+\text{sm}-\delta+\text{sm}\,\delta)}{2\,\text{sf}\,(-(-1+\text{sm})\,(1+\delta)+\text{h}\,(-2+\text{sm}-\delta+\text{sm}\,\delta))},$$
$$\text{F} \to -\left(\left(-2\,\text{h sf}^2\,(-1+\text{sm})\,(1+\delta)\,(-2+\text{sm}-\delta+\text{sm}\,\delta)+\text{h}^2\,\text{sf}^2\,(-2+\text{sm}-\delta+\text{sm}\,\delta)^2+\right.\right.$$
$$(\text{sm}-\delta+\text{sm}\,\delta)\,(\delta+2\,\text{sf}\,(-1+\text{sm})\,(1+\delta)-\text{sm}\,(1+\delta))\Big)\Big/$$
$$\left(\text{h}^2\,\text{sf}^2\,(-2+\text{sm}-\delta+\text{sm}\,\delta)^2+(\text{sm}-\delta+\text{sm}\,\delta)^2-\right.$$
$$\left.\left.\left.\left.2\,\text{sf}\,(2\,(-1+\text{sm})\,(1+\delta)+\text{h}\,(-2+\text{sm}-\delta+\text{sm}\,\delta)^2)\right)\right)\right\}\right\}$$

The full system is described by

In[•]:=
```
solvemeFull =
  {DDfprime - DDf, Ddfprime - Ddf, ddfprime - ddf, Dmprime - Dm, dmprime - dm} /.
    subeq[[1]] // Simplify;
```

We can then solve for the population sizes. First we find population sizes when drive alleles are absent (pm=pf=0) or fixed (pm=pf=1):

```
In[•]:= noDrivePopSizeSol = Solve[
          (solvemeFull /. pm → 0 /. pf → 0) == {0, 0, 0, 0, 0}, {maleNum, femaleNum}][[2]]
        fixedDrivePopSizeSol = Solve[(solvemeFull /. pm → 1 /. pf → 1) == {0, 0, 0, 0, 0},
            {maleNum, femaleNum}][[2]] // Simplify
        NnoDrive = (maleNum + femaleNum) /. noDrivePopSizeSol // FullSimplify
        Nfixed = (maleNum + femaleNum) /. fixedDrivePopSizeSol // FullSimplify
```

$$\text{Out[•]=}\ \left\{\text{maleNum} \to \frac{-2+b}{2\,b\,\alpha},\ \text{femaleNum} \to \frac{-2+b}{2\,b\,\alpha}\right\}$$

$$\text{Out[•]=}\ \left\{\text{maleNum} \to -\frac{(-1+sm)\ (-1+\delta)\ \left(2+b\ \left(-1+sf\right)\ (1+\delta)\right)}{b\ \left(-1+sf\right)\ \alpha\ (1+\delta)\ \left(-2+sf+sm+sf\,\delta-sm\,\delta\right)},\right.$$
$$\left.\text{femaleNum} \to \frac{2+b\ \left(-1+sf\right)\ (1+\delta)}{b\,\alpha\ \left(-2+sf+sm+sf\,\delta-sm\,\delta\right)}\right\}$$

$$\text{Out[•]=}\ \frac{-2+b}{b\,\alpha}$$

$$\text{Out[•]=}\ \frac{1+\frac{2}{b\ (-1+sf)\ (1+\delta)}}{\alpha}$$

Now for the case where a polymorphic equilibrium has been reached

```
In[•]:= simpleSolvemeFull = Simplify[(solvemeFull)];
        equilPopSizeSol =
         Simplify[Solve[Flatten[Simplify[(simpleSolvemeFull /. equilibrium)]] ==
            {0, 0, 0, 0, 0}, {femaleNum, maleNum}]]
        (maleNum + femaleNum) /. equilPopSizeSol // FullSimplify
```

2134

Out[•]=

$$\Big\{\{\text{femaleNum}\to 0,\ \text{maleNum}\to 0\},$$
$$\Big\{\text{femaleNum}\to \Big(\text{sf}\ \Big(-(-1+\text{sm})\ (1+\delta)+\text{h}\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)\Big)$$
$$\Big(\text{b}\,\text{h}^2\,\text{sf}^2\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2+(\text{sm}-\delta+\text{sm}\,\delta)\ \Big(-\text{b}\,\delta+\text{sm}\ \Big(-4+\text{b}+\text{b}\,\delta\Big)\Big)-$$
$$2\,\text{sf}\ \Big(-4\ (-1+\text{sm})\ (1+\delta)+2\,\text{b}\ (-1+\text{sm})\ (1+\delta)-$$
$$2\,\text{h}\ \Big(-2+\text{sm}\Big)\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)+\text{b}\,\text{h}\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2\Big)\Big)\Big/$$
$$\Big(\text{b}\,\alpha\ \Big(-\text{sm}\ (\text{sm}\ (-1+\delta)-\delta)\ (\text{sm}-\delta+\text{sm}\,\delta)^2+\text{h}^2\,\text{sf}^3\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2$$
$$\Big(-(-1+\text{sm})\ (1+\delta)+\text{h}\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)\Big)+$$
$$\text{sf}^2\ \Big(8\ (-1+\text{sm})^2\ (1+\delta)^2+2\,\text{h}\ \Big(8-11\,\text{sm}+3\,\text{sm}^2\Big)\ (1+\delta)\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)-$$
$$\text{h}^2\ \Big(-8-\text{sm}\ \Big(-6+\delta\Big)+\text{sm}^2\ (-1+\delta)\Big)\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2\Big)+$$
$$\text{sf}\ (\text{sm}-\delta+\text{sm}\,\delta)\ \Big(2\,\text{h}\,\text{sm}^3\ \Big(-1+\delta^2\Big)+\delta\ \Big(1+\delta-\text{h}\ \Big(2+\delta\Big)\Big)+$$
$$\text{sm}^2\ \Big(-5-4\,\delta+\delta^2+\text{h}\ \Big(9-5\,\delta^2\Big)\Big)+\text{sm}\ \Big(5+3\,\delta-2\,\delta^2+2\,\text{h}\ \Big(-5+\delta+2\,\delta^2\Big)\Big)\Big)\Big)\Big),$$
$$\text{maleNum}\to -\Big(\Big(\Big(\text{b}\,\text{h}^2\,\text{sf}^2\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2+(\text{sm}-\delta+\text{sm}\,\delta)\ \Big(-\text{b}\,\delta+\text{sm}\ \Big(-4+\text{b}+\text{b}\,\delta\Big)\Big)-$$
$$2\,\text{sf}\ \Big(-4\ (-1+\text{sm})\ (1+\delta)+2\,\text{b}\ (-1+\text{sm})\ (1+\delta)-$$
$$2\,\text{h}\ \Big(-2+\text{sm}\Big)\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)+\text{b}\,\text{h}\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2\Big)\Big)$$
$$\Big(\text{sm}\ (\text{sm}\ (-1+\delta)-\delta)\ (\text{sm}-\delta+\text{sm}\,\delta)^2+\text{sf}^2\ \Big(-4\ (-1+\text{sm})^2\ (1+\delta)^2+$$
$$\text{h}^2\ \Big(-2+\text{sm}\Big)\ \Big(2+\text{sm}\ (-1+\delta)-\delta\Big)\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2+$$
$$2\,\text{h}\ (-1+\text{sm})\ (1+\delta)\ \Big(-8-2\,\delta+\delta^2+\text{sm}\ \Big(8+3\,\delta-2\,\delta^2\Big)+\text{sm}^2\ \Big(-2-\delta+\delta^2\Big)\Big)\Big)-$$
$$2\,\text{sf}\ (\text{sm}-\delta+\text{sm}\,\delta)\ \Big(\text{h}\,\text{sm}^3\ \Big(-1+\delta^2\Big)+\delta\ \Big(1+\delta-\text{h}\ \Big(2+\delta\Big)\Big)-$$
$$\text{sm}^2\ \Big(2+\delta-\delta^2+\text{h}\ \Big(-4+\delta+3\,\delta^2\Big)\Big)+\text{sm}\ \Big(2-2\,\delta^2+\text{h}\ \Big(-4+3\,\delta+3\,\delta^2\Big)\Big)\Big)\Big)\Big)\Big/$$
$$\Big(\text{b}\,\alpha\ \Big(-\text{sm}\ (\text{sm}\ (-1+\delta)-\delta)\ (\text{sm}-\delta+\text{sm}\,\delta)^4+\text{h}^4\,\text{sf}^5\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^4$$
$$\Big(-(-1+\text{sm})\ (1+\delta)+\text{h}\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)\Big)-$$
$$\text{h}^2\,\text{sf}^4\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2\ \Big(-12\ (-1+\text{sm})^2\ (1+\delta)^2+$$
$$\text{h}^2\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2\ \Big(\text{sm}^2\ (-1+\delta)-2\ \Big(6+\delta\Big)+\text{sm}\ \Big(8+\delta\Big)\Big)-2\,\text{h}\ (-1+\text{sm})$$
$$(1+\delta)\ \Big(24+14\,\delta+\delta^2+\text{sm}^2\ \Big(4+5\,\delta+\delta^2\Big)-\text{sm}\ \Big(20+19\,\delta+2\,\delta^2\Big)\Big)\Big)+$$
$$2\,\text{sf}^3\ \Big(-16\ (-1+\text{sm})^3\ (1+\delta)^3-\text{h}^2\ (-1+\text{sm})\ (1+\delta)\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^2$$
$$\Big(16\ \Big(3+\delta\Big)+\text{sm}^2\ (11+7\,\delta)-\text{sm}\ \Big(40+23\,\delta\Big)\Big)+$$
$$\text{h}^3\ \Big(-2+\text{sm}-\delta+\text{sm}\,\delta\Big)^3\ \Big(8\ \Big(2+\delta\Big)+\text{sm}^2\ \Big(11+7\,\delta-4\,\delta^2\Big)+$$
$$2\,\text{sm}^3\ \Big(-1+\delta^2\Big)+\text{sm}\ \Big(-20-15\,\delta+2\,\delta^2\Big)\Big)-4\,\text{h}\ (-1+\text{sm})^2\ (1+\delta)^2$$
$$\Big(2\ \Big(12+8\,\delta+\delta^2\Big)+\text{sm}^2\ \Big(5+7\,\delta+2\,\delta^2\Big)-\text{sm}\ \Big(22+23\,\delta+4\,\delta^2\Big)\Big)\Big)+$$
$$\text{sf}\ (\text{sm}-\delta+\text{sm}\,\delta)^2\ \Big(4\,\text{h}\,\text{sm}^4\ (-1+\delta)\ (1+\delta)^2+\text{sm}\,\delta^2\ \Big(3-13\,\text{h}+3\,\delta-7\,\text{h}\,\delta\Big)+$$
$$\delta^2\ \Big(-1-\delta+\text{h}\ \Big(2+\delta\Big)\Big)+3\,\text{sm}^2\ (1+\delta)\ \Big(3-\delta^2+\text{h}\ \Big(-6+3\,\delta+5\,\delta^2\Big)\Big)-$$
$$\text{sm}^3\ (1+\delta)\ \Big(9-\delta^2+\text{h}\ \Big(-17+4\,\delta+13\,\delta^2\Big)\Big)\Big)-2\,\text{sf}^2\ (\text{sm}-\delta+\text{sm}\,\delta)$$
$$\Big(3\,\text{h}^2\,\text{sm}^5\ (-1+\delta)\ (1+\delta)^3-\text{h}\,\text{sm}^4\ (1+\delta)^2\ \Big(12+3\,\delta-\delta^2+\text{h}\ \Big(-24+3\,\delta+13\,\delta^2\Big)\Big)+$$
$$\delta\ \Big(2\ (1+\delta)^2-\text{h}^2\ \Big(-2+\delta\Big)\ \Big(2+\delta\Big)^2+\text{h}\ \Big(-8-10\,\delta-\delta^2+\delta^3\Big)\Big)+$$
$$\text{sm}^3\ (1+\delta)\ \Big(-2\ \Big(7+8\,\delta+\delta^2\Big)+\text{h}\ \Big(64+68\,\delta+8\,\delta^2-4\,\delta^3\Big)+$$
$$\text{h}^2\ \Big(-74-55\,\delta+31\,\delta^2+22\,\delta^3\Big)\Big)+\text{sm}^2\ \Big(2\ (1+\delta)^2\ \Big(14+3\,\delta\Big)+\text{h}^2\ \Big(104+$$
$$166\,\delta+37\,\delta^2-45\,\delta^3-18\,\delta^4\Big)+3\,\text{h}\ \Big(-36-69\,\delta-37\,\delta^2-2\,\delta^3+2\,\delta^4\Big)\Big)+$$
$$\text{sm}\ \Big(-2\ (1+\delta)^2\ \Big(7+3\,\delta\Big)+\text{h}\ \Big(56+110\,\delta+62\,\delta^2+4\,\delta^3-4\,\delta^4\Big)+$$
$$\text{h}^2\ \Big(-56-84\,\delta-22\,\delta^2+17\,\delta^3+7\,\delta^4\Big)\Big)\Big)\Big)\Big)\Big\}\Big\}$$

```
Out[•]= {0, (b h^2 sf^2 (-2 + sm + (-1 + sm) δ)^2 + (sm + (-1 + sm) δ) ((-4 + b) sm + b (-1 + sm) δ) -
        2 sf (-4 (-1 + sm) (1 + δ) + 2 b (-1 + sm) (1 + δ) -
           2 h (-2 + sm) (-2 + sm + (-1 + sm) δ) + b h (-2 + sm + (-1 + sm) δ)^2)) /
      (b α (-4 sf (-1 + sm) (1 + δ) - 2 h sf (-2 + sm + (-1 + sm) δ)^2 +
        h^2 sf^2 (-2 + sm + (-1 + sm) δ)^2 + (sm + (-1 + sm) δ)^2))}
```

Verify this is an equilibrium by seeing if the difference equations are 0 at this point

```
In[•]:= Simplify[solvemeFull /. subeq /. equilibrium /. equilPopSizeSol[[2]]]
Out[•]= {{{0, 0, 0, 0, 0}}}
```

Define the φ and ψ terms and show these can be used to express the within-sex allele frequencies.

```
In[•]:= ϕ = δ - (sm + sm δ + h sf (2 + δ - sm (1 + δ)));
        φ = sm (1 + δ) - (δ + 2 sf (-1 + sm) (1 + δ) + h sf (2 + δ - sm (1 + δ)));
        pfeq = ϕ / (ϕ + φ) // FullSimplify;
        pfeq - pf /. equilibrium[[1]] // Factor
        pmeq = (1 - sm) ϕ / ((1 - sm) ϕ + φ) // Simplify;
        pmeq - pm /. equilibrium[[1]] // Factor
Out[•]= 0
Out[•]= 0
```

Then find the equilibrium population size and polymorphism and show it can be written with b = b* as in the manuscript

```
In[•]:= Neq = Simplify[((femaleNum + maleNum) /. equilPopSizeSol[[2]])] // Simplify
        Factor[(b (1 + ϕ pfeq/2) ((1 - pmeq) / (1 - pfeq)) - 2) / (α b (1 + ϕ pfeq/2) ((1 - pmeq) / (1 - pfeq))) - Neq] // Simplify
        bstar = Factor[(b (1 + ϕ pfeq/2) ((1 - pmeq)/(1 - pfeq)))];
        Factor[(bstar - 2)/(bstar α) - Neq]
Out[•]= (b h^2 sf^2 (-2 + sm - δ + sm δ)^2 + (sm - δ + sm δ) (-b δ + sm (-4 + b + b δ)) -
        2 sf (-4 (-1 + sm) (1 + δ) + 2 b (-1 + sm) (1 + δ) -
           2 h (-2 + sm) (-2 + sm - δ + sm δ) + b h (-2 + sm - δ + sm δ)^2)) /
      (b α (h^2 sf^2 (-2 + sm - δ + sm δ)^2 + (sm - δ + sm δ)^2 -
        2 sf (2 (-1 + sm) (1 + δ) + h (-2 + sm - δ + sm δ)^2)))
Out[•]= 0
Out[•]= 0
```

Now, we show that b* is increased by a factor that is the mean fitness of females at the polymorphic equilibrium (pm, pf) relative to the female population size without drive ($\frac{b-2}{2\,b\,\alpha}$).

2136

In[•]:=
```
meanF =
  Simplify[(DDfjuveniles * 1 + Ddfjuveniles * (1 - h * sf) + ddfjuveniles * (1 - sf)) /
      femaleNum /. subeq /. equilibrium /. noDrivePopSizeSol];
b meanF - bstar // Simplify
```

Out[•]=
```
{{0}}
```

Finally, we also show that, when drive is fixed, the population size can be given by b~, as presented in the text.

In[•]:=
```
bTilde = b (1 - sf) (1 + δ);
Factor[((bTilde - 2)/(bTilde α)) - Nfixed]
```

Out[•]=
```
0
```

Stability

There are at least two immediately apparent equilibria: when drive is absent from the population and when drive is fixed in the population.

We first find the Jacobian for the system:

In[•]:=
```
recursions = {Dmprime, dmprime, DDfprime, Ddfprime, ddfprime};
jacob = Simplify[Transpose[{D[recursions, Dm], D[recursions, dm],
     D[recursions, DDf], D[recursions, Ddf], D[recursions, ddf]}]];
```

And then we assess the conditions under which the drive-absent and drive-fixed equilibria are stable/instable.

drive absent (pm=pf=0) equilibrium

To evaluate the stability of the equilibrium where drive is absent we consider the eigenvalues of the Jacobian at this point (substituting pf->0, pm->0, and the equilibrium population size without drive). The equilibrium is unstable whenever it has an eigenvalue with absolute value > 1.

In[•]:=
```
Simplify[jacob /. subeq /. pf → 0 /. pm → 0 /. noDrivePopSizeSol][[1]];
charpolynomial0 = Factor[Det[% - λ * IdentityMatrix[5]]]
```

Out[•]=
$$-\frac{1}{4}\lambda^2\left(-4+b+2\lambda\right)\left(-1+h\,sf+sm-h\,sf\,sm-\delta+h\,sf\,\delta+sm\,\delta-h\,sf\,sm\,\delta-\lambda+h\,sf\,\lambda+2\lambda^2\right)$$

The factor $-\frac{1}{4}\lambda^2\left(-4+b+2\lambda\right)$ gives two 0 eigenvalues and and eigenvalue of $\frac{4-b}{2}$, corresponding to the change in overall population size. Thus, we can see that the demographic dynamics are orthogonal to the evolutionary dynamics.

Here, we divide through by this factor to simplify the equation and evaluate the evolutionary dynamics. We then write the characteristic polynomial in the form
$a\lambda^2 + b\lambda + c$
finding the a, b, and c coefficients.

```
In[•]:= Collect[charpolynomial0 / (-1/4 λ^2 (-4 + b + 2 λ)), λ, Simplify];
        Coefficient[%, λ^2];
        charpolySimpleACoef0 = %% / % (*simplify so that the a coefficient is 1*)
        acoef0 = Coefficient[charpolySimpleACoef0, λ^2];
        bcoef0 = Coefficient[charpolySimpleACoef0, λ];
        ccoef0 = charpolySimpleACoef0 - bcoef0 * λ - acoef0 * λ^2 // Simplify;
Out[•]= 1/2 (-(-1 + h sf) (-1 + sm) (1 + δ) + (-1 + h sf) λ + 2 λ^2)
```

The discriminant must be non-negative for the eigenvalues to be real

```
In[•]:= bcoef0^2 - 4 acoef0 ccoef0 // FullSimplify
Out[•]= 1/4 (-1 + h sf) (-1 + h sf + 8 (-1 + sm) (1 + δ))
```

We can see that this is always true as the two terms in the product are always ≤ 0.

The conditions for the above characteristic polynomial p to have roots between -1 and 1 are:
1. $p(1) > 0$
2. $p(-1) > 0$
3. $-2 < b < 2$

When any of these are false, the largest eigenvalue is > 1, and the equilibrium is unstable.
Conditions 2 and 3 are always true:

```
In[•]:= FullSimplify[(charpolySimpleACoef0 /. λ → -1) > 0,
          {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= True

In[•]:= FullSimplify[-2 < bcoef0 < 2, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= True
```

But condition 1 fails when the following inequality is true, giving us the instability condition.

```
In[•]:= unstabConditions0 = FullSimplify[
          (charpolySimpleACoef0 /. λ → 1) < 0,  {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= sm + sm δ + h sf (2 + δ - sm (1 + δ)) < δ
```

which can be re-written using the w fitness notation

```
In[•]:= Simplify[unstabConditions0, {wXdY == 1 - sm, wXdX == 1 - h sf, wXdXd == 1 - sf}]
Out[•]= wXdX + wXdX wXdY (1 + δ) > 2
```

drive fixed (pm=pf=1) equilibrium

We proceed similarly to before, but now considering the other equilibrium:

2138

```
In[•]:= Simplify[jacob /. subeq /. pf → 1 /. pm → 1 /. fixedDrivePopSizeSol][[1]];
        charpolynomial1 = Factor[Det[% - λ * IdentityMatrix[5]]]
```

$$\text{Out}[\bullet]= \frac{1}{4\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)}\,(4-\text{b}+\text{b sf}-\text{b}\,\delta+\text{b sf}\,\delta-2\,\lambda)\,\lambda^2$$
$$(-1+\text{h sf}-\lambda+\text{h sf}\,\lambda+\text{sm}\,\lambda-\text{h sf sm}\,\lambda-\delta\,\lambda+\text{h sf}\,\delta\,\lambda+\text{sm}\,\delta\,\lambda-\text{h sf sm}\,\delta\,\lambda+$$
$$2\,\lambda^2-2\,\text{sf}\,\lambda^2-2\,\text{sm}\,\lambda^2+2\,\text{sf sm}\,\lambda^2+2\,\delta\,\lambda^2-2\,\text{sf}\,\delta\,\lambda^2-2\,\text{sm}\,\delta\,\lambda^2+2\,\text{sf sm}\,\delta\,\lambda^2)$$

The factor $\frac{(4-\text{b}+\text{b sf}-\text{b}\,\delta+\text{b sf}\,\delta-2\,\lambda)\,\lambda^2}{4\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)}$ gives two 0 eigenvalues and and eigenvalue of $2+\frac{1}{2}\,\text{b}\,(-1+\text{sf})\,(1+\delta)$, corresponding to the stability of the ecological equilibrium population size. Thus, we can see that the demographic dynamics are orthogonal to the evolutionary dynamics.

Here, we divide through by this factor to simplify the equation and evaluate the evolutionary dynamics. We then write the characteristic polynomial in the form
$a\lambda^2+b\lambda+c$
finding the a, b, and c coefficients.

```
In[•]:= Simplify[charpolynomial1 / ((4 - b + b sf - b δ + b sf δ - 2 λ) λ^2 / (4 (-1 + sf) (-1 + sm) (1 + δ)))]
        Collect[%, λ, Simplify]
```

$$\text{Out}[\bullet]= -1+(-1+\text{sm})\,(1+\delta)\,\lambda+2\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)\,\lambda^2+\text{h}\,(\text{sf}-\text{sf}\,(-1+\text{sm})\,(1+\delta)\,\lambda)$$

$$\text{Out}[\bullet]= -1+\text{h sf}+(1-\text{h sf})\,(-1+\text{sm})\,(1+\delta)\,\lambda+2\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)\,\lambda^2$$

```
In[•]:= Simplify[charpolynomial1 / ((4 - b + b sf - b δ + b sf δ - 2 λ) λ^2 / (4 (-1 + sf) (-1 + sm) (1 + δ)))];
        Collect[%, λ, Simplify];
        Coefficient[%, λ^2];
        charpolySimpleACoef1 = Collect[%% / %, λ, Simplify]
        acoef1 = Coefficient[charpolySimpleACoef1, λ^2];
        bcoef1 = Coefficient[charpolySimpleACoef1, λ];
        ccoef1 = charpolySimpleACoef1 - bcoef1 * λ - acoef1 * λ^2;
```

$$\text{Out}[\bullet]= \frac{-1+\text{h sf}}{2\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)}+\frac{(1-\text{h sf})\,\lambda}{2\,(-1+\text{sf})}+\lambda^2$$

The discriminant is nonnegative and so the eigenvalues are real:

```
In[•]:= bcoef1^2 - 4 acoef1 ccoef1
```

$$\text{Out}[\bullet]= \frac{(1-\text{h sf})^2}{4\,(-1+\text{sf})^2}-\frac{2\,(-1+\text{h sf})}{(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)}$$

The conditions for the above characteristic polynomial p to have roots between -1 and 1 are:
1. $p(1) > 0$
2. $p(-1) > 0$
3. $-2 < b$
4. $b < 2$

When any of these are false, the largest eigenvalue is > 1, and the equilibrium is unstable.

Now checking these conditions for instability by reversing the inequalities required for stability

```
In[•]:= r1 = FullSimplify[(charpolySimpleACoef1 /. λ → 1) < 0,
           {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
         r2 = FullSimplify[(charpolySimpleACoef1 /. λ → -1) < 0,
           {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
         r3 = FullSimplify[bcoef1 < -2, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
         r4 = FullSimplify[bcoef1 > 2, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= δ + 2 sf (-1 + sm) (1 + δ) + h sf (2 + δ - sm (1 + δ)) < sm (1 + δ)

Out[•]= 2 + 3 δ + sf (-2 - (2 + h) δ + (2 + h) sm (1 + δ)) < 3 sm (1 + δ)

Out[•]= 3 + h sf < 4 sf

Out[•]= False
```

Region r3 is independent of drive and relates only to female heterozygote and drive homozygote fitness (it can be rewritten as 1 – s_f < (1 – hs_f)/4).

The region defined by r1 is included in all the other regions, so we need only satisfy this inequality

```
In[•]:= Reduce[{! r1, r2, r3, 1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= False
```

Thus, the instability conditions are:

```
In[•]:= unstabConditions1 = r1
Out[•]= δ + 2 sf (-1 + sm) (1 + δ) + h sf (2 + δ - sm (1 + δ)) < sm (1 + δ)
```

which can be re-written using the w fitness notation

```
In[•]:= FullSimplify[unstabConditions1, {wXdY == 1 - sm, wXdX == 1 - h sf, wXdXd == 1 - sf}]
Out[•]= 2 wXdXd wXdY (1 + δ) < wXdX (1 + wXdY + wXdY δ)
```

Critical birth rate

The critical birth rate is the lowest birth rate for which a population still exists (equilibrium population size, N, bigger than 0).

We first find the critical birth rate when drive is absent, polymorphic, or fixed.

2140

```
In[•]:= bcritNoDrive =
         b /. Solve[Simplify[(maleNum + femaleNum) /. noDrivePopSizeSol] == 0, b][[1]]
        bcritEquil = b /. Solve[Neq == 0, b][[1]]
        bcritFixed = b /. Solve[Nfixed == 0, b][[1]]
```

Out[•]= 2

$$\text{Out[•]= } -\Big(\big(4\,\big(-2\,sf+4\,h\,sf+2\,sf\,sm-4\,h\,sf\,sm-sm^2+h\,sf\,sm^2-2\,sf\,\delta+2\,h\,sf\,\delta+sm\,\delta+2\,sf\,sm\,\delta-3\,h\,sf\,sm\,\delta-sm^2\,\delta+h\,sf\,sm^2\,\delta\big)\big)\Big/\big(4\,sf-8\,h\,sf+4\,h^2\,sf^2-4\,sf\,sm+8\,h\,sf\,sm-4\,h^2\,sf^2\,sm+sm^2-2\,h\,sf\,sm^2+h^2\,sf^2\,sm^2+4\,sf\,\delta-8\,h\,sf\,\delta+4\,h^2\,sf^2\,\delta-2\,sm\,\delta-4\,sf\,sm\,\delta+12\,h\,sf\,sm\,\delta-6\,h^2\,sf^2\,sm\,\delta+2\,sm^2\,\delta-4\,h\,sf\,sm^2\,\delta+2\,h^2\,sf^2\,sm^2\,\delta+\delta^2-2\,h\,sf\,\delta^2+h^2\,sf^2\,\delta^2-2\,sm\,\delta^2+4\,h\,sf\,sm\,\delta^2-2\,h^2\,sf^2\,sm\,\delta^2+sm^2\,\delta^2-2\,h\,sf\,sm^2\,\delta^2+h^2\,sf^2\,sm^2\,\delta^2\big)\Big)$$

$$\text{Out[•]= } -\frac{2}{(-1+sf)\,(1+\delta)}$$

Given that drive is polymorphic (ϕ>0,φ>0), the critical birth rate is always smaller than it would have been without drive

Given that drive reaches fixation (ϕ>0,φ<0), the critical birth rate is always smaller than it would have been without drive

```
In[•]:= Reduce[{bcritNoDrive < bcritEquil,
          0 < sf < 1, 0 < h < 1, 0 < sm < 1, 0 < b, 0 < δ < 1, ϕ > 0, φ > 0}]
        Reduce[{bcritNoDrive < bcritFixed, 0 < sf < 1, 0 < h < 1,
          0 < sm < 1, 0 < b, 0 < δ < 1, ϕ > 0, φ < 0}]
```

Out[•]= False

Out[•]= False

The following functions define population sizes in these different cases for plotting (the same as the expressions derived above).

```
In[•]:= func[{sf_, sm_, h_, δ_, α_, c_, b_}] :=
         If[
          (sf (-(-1 + sm) (1 + δ) + h (-2 + sm - δ + sm δ))
                (b h^2 sf^2 (-2 + sm - δ + sm δ)^2 + (sm - δ + sm δ) (-b δ + sm (-4 + b + b δ)) -
                  2 sf (-4 (-1 + sm) (1 + δ) + 2 b (-1 + sm) (1 + δ) -
                      2 h (-2 + sm) (-2 + sm - δ + sm δ) + b h (-2 + sm - δ + sm δ)^2)))/
            (b α (-sm (sm (-1 + δ) - δ) (sm - δ + sm δ)^2 + h^2 sf^3 (-2 + sm - δ + sm δ)^2
                  (-(-1 + sm) (1 + δ) + h (-2 + sm - δ + sm δ)) +
                sf^2 (8 (-1 + sm)^2 (1 + δ)^2 + 2 h (8 - 11 sm + 3 sm^2) (1 + δ) (-2 + sm - δ + sm δ) -
                    h^2 (-8 - sm (-6 + δ) + sm^2 (-1 + δ)) (-2 + sm - δ + sm δ)^2) +
                sf (sm - δ + sm δ) (2 h sm^3 (-1 + δ^2) + δ (1 + δ - h (2 + δ)) +
                    sm^2 (-5 - 4 δ + δ^2 + h (9 - 5 δ^2)) + sm (5 + 3 δ - 2 δ^2 + 2 h (-5 + δ + 2 δ^2)))))) -
           ((b h^2 sf^2 (-2 + sm - δ + sm δ)^2 + (sm - δ + sm δ) (-b δ + sm (-4 + b + b δ)) -
                2 sf (-4 (-1 + sm) (1 + δ) + 2 b (-1 + sm) (1 + δ) -
                    2 h (-2 + sm) (-2 + sm - δ + sm δ) + b h (-2 + sm - δ + sm δ)^2)))
```

$$\Big(\mathrm{sm}\,(\mathrm{sm}\,(-1+\delta)-\delta)\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+\mathrm{sf}^2\Big(-4\,(-1+\mathrm{sm})^2\,(1+\delta)^2+h^2\,(-2+\mathrm{sm})\,(2+\mathrm{sm}\,(-1+\delta)-\delta)\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+2\,h\,(-1+\mathrm{sm})\,(1+\delta)\,(-8-2\,\delta+\delta^2+\mathrm{sm}\,(8+3\,\delta-2\,\delta^2)+\mathrm{sm}^2\,(-2-\delta+\delta^2))\Big)-2\,\mathrm{sf}\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)\,(h\,\mathrm{sm}^3\,(-1+\delta^2)+\delta\,(1+\delta-h\,(2+\delta))-\mathrm{sm}^2\,(2+\delta-\delta^2+h\,(-4+\delta+3\,\delta^2))+\mathrm{sm}\,(2-2\,\delta^2+h\,(-4+3\,\delta+3\,\delta^2)))\Big)\Big/\Big(b\,\alpha\,\Big(-\mathrm{sm}\,(\mathrm{sm}\,(-1+\delta)-\delta)\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^4+h^4\,\mathrm{sf}^5\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^4\,(-(-1+\mathrm{sm})\,(1+\delta)+h\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta))-h^2\,\mathrm{sf}^4\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2\,(-12\,(-1+\mathrm{sm})^2\,(1+\delta)^2+h^2\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2\,(\mathrm{sm}^2\,(-1+\delta)-2\,(6+\delta)+\mathrm{sm}\,(8+\delta))-2\,h\,(-1+\mathrm{sm})\,(1+\delta)\,(24+14\,\delta+\delta^2+\mathrm{sm}^2\,(4+5\,\delta+\delta^2)-\mathrm{sm}\,(20+19\,\delta+2\,\delta^2)))+2\,\mathrm{sf}^3\,(-16\,(-1+\mathrm{sm})^3\,(1+\delta)^3-h^2\,(-1+\mathrm{sm})\,(1+\delta)\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2\,(16\,(3+\delta)+\mathrm{sm}^2\,(11+7\,\delta)-\mathrm{sm}\,(40+23\,\delta))+h^3\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^3\,(8\,(2+\delta)+\mathrm{sm}^2\,(11+7\,\delta-4\,\delta^2)+2\,\mathrm{sm}^3\,(-1+\delta^2)+\mathrm{sm}\,(-20-15\,\delta+2\,\delta^2))-4\,h\,(-1+\mathrm{sm})^2\,(1+\delta)^2\,(2\,(12+8\,\delta+\delta^2)+\mathrm{sm}^2\,(5+7\,\delta+2\,\delta^2)-\mathrm{sm}\,(22+23\,\delta+4\,\delta^2)))+\mathrm{sf}\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2\,(4\,h\,\mathrm{sm}^4\,(-1+\delta)\,(1+\delta)^2+\mathrm{sm}\,\delta^2\,(3-13\,h+3\,\delta-7\,h\,\delta)+\delta^2\,(-1-\delta+h\,(2+\delta))+3\,\mathrm{sm}^2\,(1+\delta)\,(3-\delta^2+h\,(-6+3\,\delta+5\,\delta^2))-\mathrm{sm}^3\,(1+\delta)\,(9-\delta^2+h\,(-17+4\,\delta+13\,\delta^2)))-2\,\mathrm{sf}^2\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)\,(3\,h^2\,\mathrm{sm}^5\,(-1+\delta)\,(1+\delta)^3-h\,\mathrm{sm}^4\,(1+\delta)^2\,(12+3\,\delta-\delta^2+h\,(-24+3\,\delta+13\,\delta^2))+\delta\,(2\,(1+\delta)^2-h^2\,(-2+\delta)\,(2+\delta)^2+h\,(-8-10\,\delta-\delta^2+\delta^3))+\mathrm{sm}^3\,(1+\delta)\,(-2\,(7+8\,\delta+\delta^2)+h\,(64+68\,\delta+8\,\delta^2-4\,\delta^3)+h^2\,(-74-55\,\delta+31\,\delta^2+22\,\delta^3))+\mathrm{sm}^2\,(2\,(1+\delta)^2\,(14+3\,\delta)+h^2\,(104+166\,\delta+37\,\delta^2-45\,\delta^3-18\,\delta^4)+3\,h\,(-36-69\,\delta-37\,\delta^2-2\,\delta^3+2\,\delta^4))+\mathrm{sm}\,(-2\,(1+\delta)^2\,(7+3\,\delta)+h\,(56+110\,\delta+62\,\delta^2+4\,\delta^3-4\,\delta^4)+h^2\,(-56-84\,\delta-22\,\delta^2+17\,\delta^3+7\,\delta^4)))))<0,$$

$$\mathrm{Null},\Big(\mathrm{sf}\,(-(-1+\mathrm{sm})\,(1+\delta)+h\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta))\,(b\,h^2\,\mathrm{sf}^2\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)\,(-b\,\delta+\mathrm{sm}\,(-4+b+b\,\delta))-2\,\mathrm{sf}\,(-4\,(-1+\mathrm{sm})\,(1+\delta)+2\,b\,(-1+\mathrm{sm})\,(1+\delta)-2\,h\,(-2+\mathrm{sm})\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)+b\,h\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2))\Big)\Big/\Big(b\,\alpha\,\Big(-\mathrm{sm}\,(\mathrm{sm}\,(-1+\delta)-\delta)\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+h^2\,\mathrm{sf}^3\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2\,(-(-1+\mathrm{sm})\,(1+\delta)+h\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta))+\mathrm{sf}^2\,(8\,(-1+\mathrm{sm})^2\,(1+\delta)^2+2\,h\,(8-11\,\mathrm{sm}+3\,\mathrm{sm}^2)\,(1+\delta)\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)-h^2\,(-8-\mathrm{sm}\,(-6+\delta)+\mathrm{sm}^2\,(-1+\delta))\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2)+\mathrm{sf}\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)\,(2\,h\,\mathrm{sm}^3\,(-1+\delta^2)+\delta\,(1+\delta-h\,(2+\delta))+\mathrm{sm}^2\,(-5-4\,\delta+\delta^2+h\,(9-5\,\delta^2))+\mathrm{sm}\,(5+3\,\delta-2\,\delta^2+2\,h\,(-5+\delta+2\,\delta^2)))\Big)\Big)-\Big((b\,h^2\,\mathrm{sf}^2\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)\,(-b\,\delta+\mathrm{sm}\,(-4+b+b\,\delta))-2\,\mathrm{sf}\,(-4\,(-1+\mathrm{sm})\,(1+\delta)+2\,b\,(-1+\mathrm{sm})\,(1+\delta)-2\,h\,(-2+\mathrm{sm})\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)+b\,h\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2))\Big(\mathrm{sm}\,(\mathrm{sm}\,(-1+\delta)-\delta)\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+\mathrm{sf}^2\Big(-4\,(-1+\mathrm{sm})^2\,(1+\delta)^2+h^2\,(-2+\mathrm{sm})\,(2+\mathrm{sm}\,(-1+\delta)-\delta)\,(-2+\mathrm{sm}-\delta+\mathrm{sm}\,\delta)^2+2\,h\,(-1+\mathrm{sm})\,(1+\delta)\,(-8-2\,\delta+\delta^2+\mathrm{sm}\,(8+3\,\delta-2\,\delta^2)+\mathrm{sm}^2\,(-2-\delta+\delta^2))\Big)-2\,\mathrm{sf}\,(\mathrm{sm}-\delta+\mathrm{sm}\,\delta)\,(h\,\mathrm{sm}^3\,(-1+\delta^2)+\delta\,(1+\delta-h\,(2+\delta))-$$

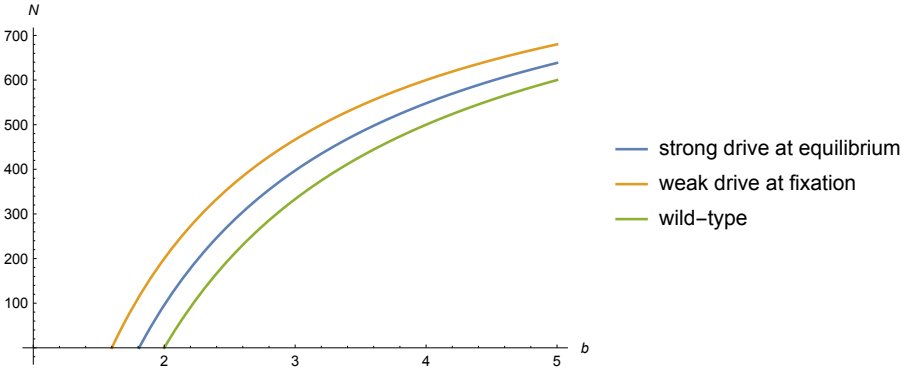
2142

```
        sm^2 (2 + δ - δ^2 + h (-4 + δ + 3 δ^2)) + sm (2 - 2 δ^2 + h (-4 + 3 δ + 3 δ^2)))))) /
   (b α (-sm (sm (-1 + δ) - δ) (sm - δ + sm δ)^4 + h^4 sf^5 (-2 + sm - δ + sm δ)^4
      (-(-1 + sm) (1 + δ) + h (-2 + sm - δ + sm δ)) -
     h^2 sf^4 (-2 + sm - δ + sm δ)^2 (-12 (-1 + sm)^2 (1 + δ)^2 +
        h^2 (-2 + sm - δ + sm δ)^2 (sm^2 (-1 + δ) - 2 (6 + δ) + sm (8 + δ)) -
        2 h (-1 + sm) (1 + δ) (24 + 14 δ + δ^2 + sm^2 (4 + 5 δ + δ^2) - sm (20 + 19 δ + 2 δ^2))) +
     2 sf^3 (-16 (-1 + sm)^3 (1 + δ)^3 - h^2 (-1 + sm) (1 + δ) (-2 + sm - δ + sm δ)^2
        (16 (3 + δ) + sm^2 (11 + 7 δ) - sm (40 + 23 δ)) +
        h^3 (-2 + sm - δ + sm δ)^3 (8 (2 + δ) + sm^2 (11 + 7 δ - 4 δ^2) +
           2 sm^3 (-1 + δ^2) + sm (-20 - 15 δ + 2 δ^2)) - 4 h (-1 + sm)^2 (1 + δ)^2
         (2 (12 + 8 δ + δ^2) + sm^2 (5 + 7 δ + 2 δ^2) - sm (22 + 23 δ + 4 δ^2))) +
     sf (sm - δ + sm δ)^2 (4 h sm^4 (-1 + δ) (1 + δ)^2 + sm δ^2 (3 - 13 h + 3 δ - 7 h δ) +
        δ^2 (-1 - δ + h (2 + δ)) + 3 sm^2 (1 + δ) (3 - δ^2 + h (-6 + 3 δ + 5 δ^2)) -
        sm^3 (1 + δ) (9 - δ^2 + h (-17 + 4 δ + 13 δ^2))) - 2 sf^2 (sm - δ + sm δ)
      (3 h^2 sm^5 (-1 + δ) (1 + δ)^3 - h sm^4 (1 + δ)^2 (12 + 3 δ - δ^2 + h (-24 + 3 δ + 13 δ^2)) +
        δ (2 (1 + δ)^2 - h^2 (-2 + δ) (2 + δ)^2 + h (-8 - 10 δ - δ^2 + δ^3)) + sm^3 (1 + δ)
         (-2 (7 + 8 δ + δ^2) + h (64 + 68 δ + 8 δ^2 - 4 δ^3) + h^2 (-74 - 55 δ + 31 δ^2 + 22 δ^3)) +
        sm^2 (2 (1 + δ)^2 (14 + 3 δ) + h^2 (104 + 166 δ + 37 δ^2 - 45 δ^3 - 18 δ^4) +
           3 h (-36 - 69 δ - 37 δ^2 - 2 δ^3 + 2 δ^4)) + sm (-2 (1 + δ)^2 (7 + 3 δ) + h
            (56 + 110 δ + 62 δ^2 + 4 δ^3 - 4 δ^4) + h^2 (-56 - 84 δ - 22 δ^2 + 17 δ^3 + 7 δ^4)))))))]
func2[{b_, α_}] := If[(-2 + b)/(b α) > 0, (-2 + b)/(b α), Null]
func3[{sf_, sm_, h_, δ_, α_, c_, b_}] :=
 If[(2 + b (-1 + sf) (1 + δ))/(b α (-2 + sf + sm + sf δ - sm δ)) - ((-1 + sm) (-1 + δ) (2 + b (-1 + sf) (1 + δ)))/(b (-1 + sf) α (1 + δ) (-2 + sf + sm + sf δ - sm δ)) > 0,
  (2 + b (-1 + sf) (1 + δ))/(b α (-2 + sf + sm + sf δ - sm δ)) - ((-1 + sm) (-1 + δ) (2 + b (-1 + sf) (1 + δ)))/(b (-1 + sf) α (1 + δ) (-2 + sf + sm + sf δ - sm δ)), Null]
```

Produce plot

In[•]:=
```
testParameters = {h → 0.1, α → 1/1000, δ → 1, c → 1, sm → 0, sf → 0.8};
weakDriveParams = {h → 0, α → 10^(-3), δ → 0.25, c → 1, sm → 0, sf → 0};
Plot[{func[({sf, sm, h, δ, α, c, b} /. testParameters)],
  func3[({sf, sm, h, δ, α, c, b} /. weakDriveParams)],
  func2[({b, α} /. testParameters)]}, {b, 1, 5}, AxesLabel → {b, N}, PlotLegends →
  {"strong drive at equilibrium", "weak drive at fixation", "wild-type"}]
```

Out[•]=



Allele Frequency and Population Size Dynamics - females mate with two males ($\lambda_f = 2$)

Recursion equations

In this section, each female mates with two males. The probability of choosing each male genotype is assumed to be random such that it occurs in proportion to the proportion of males with each genotype. These recursion/difference equations track the mating combinations of female and two male genotypes that can produce offspring of each genotype.

2144

```
In[•]:= DDfjuveniles2 =
  b * (1 - α totAdults) (DDf (Dm^2 / (2 (Dm + dm)^2) + 2 Dm dm / (Dm + dm)^2 (1/2 / (1 + c))) +
    Ddf (Dm^2 / (Dm + dm)^2 (1/2) + 2 Dm dm / (Dm + dm)^2 (1/2 / (1 + c))));
Ddfjuveniles2 =
  b * (1 - α totAdults) (DDf (dm^2 / (Dm + dm)^2 ((1 + δ)/2) + 2 dm Dm / (Dm + dm)^2 (c ((1+δ)/2) / (1 + c))) +
    Ddf (Dm^2 / (Dm + dm)^2 (1/2) + dm^2 / (Dm + dm)^2 ((1 + δ)/2) + 2 Dm dm / (Dm + dm)^2 ((1/2 + c ((1+δ)/2)) / (1 + c))) +
    ddf (Dm^2 / (Dm + dm)^2 1/2 + 2 Dm dm / (Dm + dm)^2 1/2 / (1 + c)));
ddfjuveniles2 =
  b * (1 - α totAdults) (Ddf (dm^2 / (Dm + dm)^2 ((1 + δ)/2) + 2 Dm dm / (Dm + dm)^2 (c ((1+δ)/2) / (1 + c))) +
    ddf (dm^2 / (Dm + dm)^2 ((1 + δ)/2) + 2 Dm dm / (Dm + dm)^2 (c ((1+δ)/2) / (1 + c))));
Dmjuveniles2 = b * (1 - α totAdults)
  (DDf (Dm^2 / (Dm + dm)^2 1/2 + dm^2 / (Dm + dm)^2 ((1 - δ)/2) + 2 Dm dm / (Dm + dm)^2 ((1/2 + c ((1-δ)/2)) / (1 + c))) +
    Ddf (Dm^2 / (Dm + dm)^2 (1/2) + dm^2 / (Dm + dm)^2 (((1 - δ)/2)) + 2 Dm dm / (Dm + dm)^2 ((1/2 + c ((1-δ)/2)) / (1 + c))));
dmjuveniles2 = b * (1 - α totAdults)
  (Ddf (Dm^2 / (Dm + dm)^2 (1/2) + dm^2 / (Dm + dm)^2 ((1 - δ)/2) + 2 Dm dm / (Dm + dm)^2 ((1/2 + c ((1-δ)/2)) / (1 + c))) +
    ddf (Dm^2 / (Dm + dm)^2 (1/2) + dm^2 / (Dm + dm)^2 ((1 - δ)/2) + 2 Dm dm / (Dm + dm)^2 ((1/2 + c ((1-δ)/2)) / (1 + c))));
```

Before the next generation of adults, viability selection acts according to genotype.

```
In[•]:= DDfprime2 = DDfjuveniles2 * 1;
Ddfprime2 = Ddfjuveniles2 * (1 - h * sf);
ddfprime2 = ddfjuveniles2 * (1 - sf);
Dmprime2 = Dmjuveniles2 * 1;
dmprime2 = dmjuveniles2 * (1 - sm);
```

Equilibrium allele frequency and population size

The eco-evolutionary dynamics are at equilibrium when the difference equations (solvemeFull) are equal to 0 (no change). We use subeq to convert the difference equations so that they track the genotype frequency in males (pm), females (pf), inbreeding coefficient (F), and the number of males (maleNum) and females (femaleNum).

```
In[•]:= solvemeFull = {DDfprime2 - DDf, Ddfprime2 - Ddf, ddfprime2 - ddf,
          Dmprime2 - Dm, dmprime2 - dm} /. subeq[[1]] // Simplify;
```

When drive is absent (pm=0, pf=0) or fixed (pm=1, pf=1), the allele frequencies cannot change. We can get the equilibrium population size (maleNum and femaleNum) for these evolutionary equilibria.

```
In[•]:= noDrivePopSizeSol2 =
         Solve[(solvemeFull /. pm → 0 /. pf → 0) == {0, 0, 0, 0, 0}, {maleNum, femaleNum}]
        fixedDrivePopSizeSol2 =
         Solve[(solvemeFull /. pm → 1 /. pf → 1) == {0, 0, 0, 0, 0}, {maleNum, femaleNum}]
```

$$\text{Out[•]= } \left\{\{\text{maleNum} \to 0,\ \text{femaleNum} \to 0\},\ \left\{\text{maleNum} \to \frac{-2+b}{2\,b\,\alpha},\ \text{femaleNum} \to \frac{-2+b}{2\,b\,\alpha}\right\}\right\}$$

$$\text{Out[•]= } \left\{\{\text{maleNum} \to 0,\ \text{femaleNum} \to 0\},\ \left\{\text{maleNum} \to \frac{(-1+sm)\,(-1+\delta)\,\left(2-b+b\,sf-b\,\delta+b\,sf\,\delta\right)}{b\,\left(-1+sf\right)\,\alpha\,(1+\delta)\,\left(2-sf-sm-sf\,\delta+sm\,\delta\right)},\right.\right.$$
$$\left.\left.\text{femaleNum} \to \frac{2-b+b\,sf-b\,\delta+b\,sf\,\delta}{b\,\alpha\,\left(-2+sf+sm+sf\,\delta-sm\,\delta\right)}\right\}\right\}$$

We also expect drive can reach an intermediate equilibrium allele frequency. We make it easier to solve for this equilibrium allele frequency by focussing on the difference in allele frequency and allele frequency between generations.

```
In[•]:= solveme =
          Simplify[{{(dmprime2/(Dmprime2 + dmprime2)) - pm,
              (2 ddfprime2 + Ddfprime2)/(2 (ddfprime2 + Ddfprime2 + DDfprime2)) - pf,
              (4 ddfprime2 DDfprime2 - Ddfprime2^2)/((2 ddfprime2 + Ddfprime2) (Ddfprime2 + 2 DDfprime2)) - F} /.
            {pm → dm/(dm + Dm), pf → (2 ddf + Ddf)/(2 (ddf + Ddf + DDf)),
             F → (-Ddf^2 + 4 ddf DDf)/((2 ddf + Ddf) (Ddf + 2 DDf))} /. subeq];
```

However, we are only able to get a result for the polymorphic equilibrium when we further assume that there is no selection in adult males (sm=0).

```
In[•]:= equilibrium = Solve[(solveme[[1]] /. sm → 0) == {0, 0, 0}, {pm, pf, F}] // Simplify
```

$$\text{Out[•]= } \left\{\left\{\text{pm} \to \right.\right.$$
$$-\frac{1}{4\,(-1+c)\,sf\,\left(-1-\delta+h\,\left(2+\delta\right)\right)}\left(4\,c\,sf-8\,c\,h\,sf-\delta+c\,\delta+4\,c\,sf\,\delta+h\,sf\,\delta-5\,c\,h\,sf\,\delta+\right.$$
$$\sqrt{\left(-8\,(-1+c)\,sf\,\left(-1-\delta+h\,\left(2+\delta\right)\right)\,\left(1+h\,sf+c\,\left(-1-2\,\delta+h\,sf\,\left(3+2\,\delta\right)\right)\right)+\left.\left.\left(\left(-1+h\,sf\right)\,\delta+c\,\left(\delta+sf\,\left(4-8\,h+4\,\delta-5\,h\,\delta\right)\right)\right)^2\right)},\ \text{pf} \to$$
$$-\frac{1}{4\,(-1+c)\,sf\,\left(-1-\delta+h\,\left(2+\delta\right)\right)}\left(4\,c\,sf-8\,c\,h\,sf-\delta+c\,\delta+4\,c\,sf\,\delta+h\,sf\,\delta-5\,c\,h\,sf\,\delta+\right.$$
$$\left.\sqrt{\left(-8\,(-1+c)\,sf\,\left(-1-\delta+h\,\left(2+\delta\right)\right)\,\left(1+h\,sf+c\,\left(-1-2\,\delta+h\,sf\,\left(3+2\,\delta\right)\right)\right)+\left(\left(-1+h\,sf\right)\,\delta+c\,\left(\delta+sf\,\left(4-8\,h+4\,\delta-5\,h\,\delta\right)\right)\right)^2\right)},$$

2146

$$
\begin{aligned}
F \to \Big(&-2\,sf+4\,c\,sf-2\,c^2\,sf+4\,h\,sf-8\,c\,h\,sf+4\,c^2\,h\,sf+8\,sf^2+4\,c\,sf^2+4\,c^2\,sf^2-\\
&18\,h\,sf^2-36\,c\,h\,sf^2-10\,c^2\,h\,sf^2+8\,h^2\,sf^2+48\,c\,h^2\,sf^2+8\,c^2\,h^2\,sf^2+8\,h\,sf^3+\\
&24\,c\,h\,sf^3-14\,h^2\,sf^3-36\,c\,h^2\,sf^3-14\,c^2\,h^2\,sf^3+4\,h^3\,sf^3-8\,c\,h^3\,sf^3+4\,c^2\,h^3\,sf^3+\\
&4\,c\,h^2\,sf^4+12\,c^2\,h^2\,sf^4+2\,h^3\,sf^4+4\,c\,h^3\,sf^4-6\,c^2\,h^3\,sf^4-9\,sf\,\delta+10\,c\,sf\,\delta-\\
&c^2\,sf\,\delta+12\,h\,sf\,\delta-8\,c\,h\,sf\,\delta-4\,c^2\,h\,sf\,\delta+8\,sf^2\,\delta-4\,c\,sf^2\,\delta+4\,c^2\,sf^2\,\delta-13\,h\,sf^2\,\delta-\\
&38\,c\,h\,sf^2\,\delta+3\,c^2\,h\,sf^2\,\delta+48\,c\,h^2\,sf^2\,\delta+16\,c^2\,h^2\,sf^2\,\delta+8\,h\,sf^3\,\delta+32\,c\,h\,sf^3\,\delta-\\
&8\,c^2\,h\,sf^3\,\delta-11\,h^2\,sf^3\,\delta-42\,c\,h^2\,sf^3\,\delta-27\,c^2\,h^2\,sf^3\,\delta+4\,h^3\,sf^3\,\delta-8\,c\,h^3\,sf^3\,\delta+\\
&4\,c^2\,h^3\,sf^3\,\delta+4\,c\,h^2\,sf^4\,\delta+20\,c^2\,h^2\,sf^4\,\delta+h^3\,sf^4\,\delta+6\,c\,h^3\,sf^4\,\delta-7\,c^2\,h^3\,sf^4\,\delta+\delta^2-\\
&2\,c\,\delta^2+c^2\,\delta^2-2\,sf\,\delta^2+10\,c\,sf\,\delta^2-h\,sf\,\delta^2-6\,c\,h\,sf\,\delta^2-9\,c^2\,h\,sf\,\delta^2-8\,c\,sf^2\,\delta^2+\\
&4\,h\,sf^2\,\delta^2-2\,c\,h\,sf^2\,\delta^2+14\,c^2\,h\,sf^2\,\delta^2-h^2\,sf^2\,\delta^2+10\,c\,h^2\,sf^2\,\delta^2+7\,c^2\,h^2\,sf^2\,\delta^2+\\
&8\,c\,h\,sf^3\,\delta^2-8\,c^2\,h\,sf^3\,\delta^2-2\,h^2\,sf^3\,\delta^2-10\,c\,h^2\,sf^3\,\delta^2-12\,c^2\,h^2\,sf^3\,\delta^2+h^3\,sf^3\,\delta^2-\\
&2\,c\,h^3\,sf^3\,\delta^2+c^2\,h^3\,sf^3\,\delta^2+8\,c^2\,h^2\,sf^4\,\delta^2+2\,c\,h^3\,sf^4\,\delta^2-2\,c^2\,h^3\,sf^4\,\delta^2+\\
&3\,sf\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}+\\
&c\,sf\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-\\
&4\,h\,sf\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-\\
&4\,c\,h\,sf\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-\\
&2\,h\,sf^2\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}+\\
&2\,c\,h\,sf^2\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}+\\
&4\,h^2\,sf^2\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}+4\,c\,h^2\,sf^2\\
&\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-\\
&h^2\,sf^3\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-3\,c\,h^2\,sf^3\\
&\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-\\
&\delta\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}+\\
&c\,\delta\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}+\\
&2\,sf\,\delta\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-4\,c\,h\,sf\,\delta\\
&\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+\\
&\quad\big((-1+h\,sf)\,\delta+c\,(\delta+sf\,(4-8\,h+4\,\delta-5\,h\,\delta))\big)^2\Big)}-\\
&2\,h\,sf^2\,\delta\,\sqrt{\Big(-8\,(-1+c)\,sf\,\big(-1-\delta+h\,(2+\delta)\big)\,\big(1+h\,sf+c\,\big(-1-2\,\delta+h\,sf\,(3+2\,\delta)\big)\big)+
\end{aligned}
$$

$$\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\Big)+2\,\mathrm{c\,h\,sf}^2\,\delta$$
$$\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)+$$
$$\mathrm{h}^2\,\mathrm{sf}^2\,\delta\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)+3\,\mathrm{c\,h}^2\,\mathrm{sf}^2\,\delta$$
$$\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)-2\,\mathrm{c\,h}^2\,\mathrm{sf}^3\,\delta$$
$$\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)\Big)\Big/$$
$$\Big(2\left(\left(-1+\mathrm{c}\right)^2\delta^2+\mathrm{h}^3\,\mathrm{sf}^4\left(1+\mathrm{c}\left(3+2\,\delta\right)\right)^2+\mathrm{sf}\left(1-2\,\delta+\mathrm{h}\left(-2+2\,\delta-3\,\delta^2\right)+\right.\right.$$
$$\left.2\,\mathrm{c}\left(-1+2\,\delta^2+\mathrm{h}\left(2+2\,\delta+\delta^2\right)\right)+\mathrm{c}^2\left(1+2\,\delta-\mathrm{h}\left(2+6\,\delta+7\,\delta^2\right)\right)\right)-\mathrm{h\,sf}^3$$
$$\left(2\,\mathrm{c}\left(-4\left(1+\delta\right)+\mathrm{h}\left(3-2\,\delta^2\right)+\mathrm{h}^2\left(6+6\,\delta+\delta^2\right)\right)+\mathrm{h}\left(-7-6\,\delta+\mathrm{h}\left(10+6\,\delta+\delta^2\right)\right)+\right.$$
$$\left.\mathrm{c}^2\left(-8\left(1+\delta\right)+\mathrm{h}^2\left(10+14\,\delta+5\,\delta^2\right)+\mathrm{h}\left(17+22\,\delta+8\,\delta^2\right)\right)\right)+\mathrm{sf}^2$$
$$\left(8\left(1+\delta\right)-5\,\mathrm{h}\left(5+4\,\delta\right)+\mathrm{h}^2\left(20+12\,\delta+3\,\delta^2\right)+2\,\mathrm{c}\left(4\left(1+\delta\right)+\mathrm{h}^2\left(12+12\,\delta+\delta^2\right)-\right.\right.$$
$$\left.\left.\left.\mathrm{h}\left(15+18\,\delta+4\,\delta^2\right)\right)+\mathrm{c}^2\,\mathrm{h}\left(-9-8\,\delta+4\,\delta^2+\mathrm{h}\left(20+28\,\delta+11\,\delta^2\right)\right)\right)\right)\Big)\Big\},$$

$$\Big\{\mathrm{pm}\to\frac{1}{4\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)}\Big(-4\,\mathrm{c\,sf}+8\,\mathrm{c\,h\,sf}+\delta-\mathrm{c}\,\delta-$$
$$4\,\mathrm{c\,sf}\,\delta-\mathrm{h\,sf}\,\delta+5\,\mathrm{c\,h\,sf}\,\delta+$$
$$\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)\Big),$$

$$\mathrm{pf}\to\frac{1}{4\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)}\Big(-4\,\mathrm{c\,sf}+8\,\mathrm{c\,h\,sf}+\delta-\mathrm{c}\,\delta-$$
$$4\,\mathrm{c\,sf}\,\delta-\mathrm{h\,sf}\,\delta+5\,\mathrm{c\,h\,sf}\,\delta+$$
$$\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)\Big),$$

$$\mathrm{F}\to\Big(-2\,\mathrm{sf}+4\,\mathrm{c\,sf}-2\,\mathrm{c}^2\,\mathrm{sf}+4\,\mathrm{h\,sf}-8\,\mathrm{c\,h\,sf}+4\,\mathrm{c}^2\,\mathrm{h\,sf}+8\,\mathrm{sf}^2+4\,\mathrm{c\,sf}^2+$$
$$4\,\mathrm{c}^2\,\mathrm{sf}^2-18\,\mathrm{h\,sf}^2-36\,\mathrm{c\,h\,sf}^2-10\,\mathrm{c}^2\,\mathrm{h\,sf}^2+8\,\mathrm{h}^2\,\mathrm{sf}^2+48\,\mathrm{c\,h}^2\,\mathrm{sf}^2+$$
$$8\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^2+8\,\mathrm{h\,sf}^3+24\,\mathrm{c\,h\,sf}^3-14\,\mathrm{h}^2\,\mathrm{sf}^3-36\,\mathrm{c\,h}^2\,\mathrm{sf}^3-$$
$$14\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^3+4\,\mathrm{h}^3\,\mathrm{sf}^3-8\,\mathrm{c\,h}^3\,\mathrm{sf}^3+4\,\mathrm{c}^2\,\mathrm{h}^3\,\mathrm{sf}^3+4\,\mathrm{c\,h}^2\,\mathrm{sf}^4+$$
$$12\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^4+2\,\mathrm{h}^3\,\mathrm{sf}^4+4\,\mathrm{c\,h}^3\,\mathrm{sf}^4-6\,\mathrm{c}^2\,\mathrm{h}^3\,\mathrm{sf}^4-9\,\mathrm{sf}\,\delta+10\,\mathrm{c\,sf}\,\delta-$$
$$\mathrm{c}^2\,\mathrm{sf}\,\delta+12\,\mathrm{h\,sf}\,\delta-8\,\mathrm{c\,h\,sf}\,\delta-4\,\mathrm{c}^2\,\mathrm{h\,sf}\,\delta+8\,\mathrm{sf}^2\,\delta-4\,\mathrm{c\,sf}^2\,\delta+$$
$$4\,\mathrm{c}^2\,\mathrm{sf}^2\,\delta-13\,\mathrm{h\,sf}^2\,\delta-38\,\mathrm{c\,h\,sf}^2\,\delta+3\,\mathrm{c}^2\,\mathrm{h\,sf}^2\,\delta+48\,\mathrm{c\,h}^2\,\mathrm{sf}^2\,\delta+$$
$$16\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^2\,\delta+8\,\mathrm{h\,sf}^3\,\delta+32\,\mathrm{c\,h\,sf}^3\,\delta-8\,\mathrm{c}^2\,\mathrm{h\,sf}^3\,\delta-11\,\mathrm{h}^2\,\mathrm{sf}^3\,\delta-$$
$$42\,\mathrm{c\,h}^2\,\mathrm{sf}^3\,\delta-27\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^3\,\delta+4\,\mathrm{h}^3\,\mathrm{sf}^3\,\delta-8\,\mathrm{c\,h}^3\,\mathrm{sf}^3\,\delta+4\,\mathrm{c}^2\,\mathrm{h}^3\,\mathrm{sf}^3\,\delta+$$
$$4\,\mathrm{c\,h}^2\,\mathrm{sf}^4\,\delta+20\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^4\,\delta+\mathrm{h}^3\,\mathrm{sf}^4\,\delta+6\,\mathrm{c\,h}^3\,\mathrm{sf}^4\,\delta-7\,\mathrm{c}^2\,\mathrm{h}^3\,\mathrm{sf}^4\,\delta+$$
$$\delta^2-2\,\mathrm{c}\,\delta^2+\mathrm{c}^2\,\delta^2-2\,\mathrm{sf}\,\delta^2+10\,\mathrm{c\,sf}\,\delta^2-\mathrm{h\,sf}\,\delta^2-6\,\mathrm{c\,h\,sf}\,\delta^2-$$
$$9\,\mathrm{c}^2\,\mathrm{h\,sf}\,\delta^2-8\,\mathrm{c\,sf}^2\,\delta^2+4\,\mathrm{h\,sf}^2\,\delta^2-2\,\mathrm{c\,h\,sf}^2\,\delta^2+14\,\mathrm{c}^2\,\mathrm{h\,sf}^2\,\delta^2-$$
$$\mathrm{h}^2\,\mathrm{sf}^2\,\delta^2+10\,\mathrm{c\,h}^2\,\mathrm{sf}^2\,\delta^2+7\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^2\,\delta^2+8\,\mathrm{c\,h\,sf}^3\,\delta^2-8\,\mathrm{c}^2\,\mathrm{h\,sf}^3\,\delta^2-$$
$$2\,\mathrm{h}^2\,\mathrm{sf}^3\,\delta^2-10\,\mathrm{c\,h}^2\,\mathrm{sf}^3\,\delta^2-12\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^3\,\delta^2+\mathrm{h}^3\,\mathrm{sf}^3\,\delta^2-2\,\mathrm{c\,h}^3\,\mathrm{sf}^3\,\delta^2+$$
$$\mathrm{c}^2\,\mathrm{h}^3\,\mathrm{sf}^3\,\delta^2+8\,\mathrm{c}^2\,\mathrm{h}^2\,\mathrm{sf}^4\,\delta^2+2\,\mathrm{c\,h}^3\,\mathrm{sf}^4\,\delta^2-2\,\mathrm{c}^2\,\mathrm{h}^3\,\mathrm{sf}^4\,\delta^2-$$
$$3\,\mathrm{sf}\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$
$$\left.\left(\left(-1+\mathrm{h\,sf}\right)\delta+\mathrm{c}\left(\delta+\mathrm{sf}\left(4-8\,\mathrm{h}+4\,\delta-5\,\mathrm{h}\,\delta\right)\right)\right)^2\right)-$$
$$\mathrm{c\,sf}\sqrt{\left(-8\left(-1+\mathrm{c}\right)\mathrm{sf}\left(-1-\delta+\mathrm{h}\left(2+\delta\right)\right)\left(1+\mathrm{h\,sf}+\mathrm{c}\left(-1-2\,\delta+\mathrm{h\,sf}\left(3+2\,\delta\right)\right)\right)+\right.}$$

2148

$$
\begin{aligned}
&\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\Big)+\\
&4h\,sf\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+\\
&4c\,h\,sf\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+\\
&2h\,sf^2\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-\\
&2c\,h\,sf^2\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-\\
&4h^2\,sf^2\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-4c\,h^2\,sf^2\\
&\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+\\
&h^2\,sf^3\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+3c\,h^2\,sf^3\\
&\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+\\
&\delta\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-\\
&c\,\delta\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-\\
&2\,sf\,\delta\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+4c\,h\,sf\,\delta\\
&\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+\\
&2h\,sf^2\,\delta\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-2c\,h\,sf^2\,\delta\\
&\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-\\
&h^2\,sf^2\,\delta\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)-3c\,h^2\,sf^2\,\delta\\
&\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)+2c\,h^2\,sf^3\,\delta\\
&\sqrt{\left(-8\left(-1+c\right)sf\left(-1-\delta+h\left(2+\delta\right)\right)\left(1+h\,sf+c\left(-1-2\delta+h\,sf\left(3+2\delta\right)\right)\right)+\right.}\\
&\left.\left(\left(-1+h\,sf\right)\delta+c\left(\delta+sf\left(4-8h+4\delta-5h\delta\right)\right)\right)^2\right)\Big)\Big/\\
&\Big(2\left(\left(-1+c\right)^2\delta^2+h^3\,sf^4\left(1+c\left(3+2\delta\right)\right)^2+sf\left(1-2\delta+h\left(-2+2\delta-3\delta^2\right)+\right.\right.\\
&\quad\left.2c\left(-1+2\delta^2+h\left(2+2\delta+\delta^2\right)\right)+c^2\left(1+2\delta-h\left(2+6\delta+7\delta^2\right)\right)\right)-h\,sf^3\\
&\quad\left(2c\left(-4\left(1+\delta\right)+h\left(3-2\delta^2\right)+h^2\left(6+6\delta+\delta^2\right)\right)+h\left(-7-6\delta+h\left(10+6\delta+\delta^2\right)\right)+\right.\\
&\quad\left.\left.c^2\left(-8\left(1+\delta\right)+h^2\left(10+14\delta+5\delta^2\right)+h\left(17+22\delta+8\delta^2\right)\right)\right)+sf^2
\end{aligned}
$$

$$\left(8\ (1+\delta)-5\,h\ (5+4\,\delta)+h^2\left(20+12\,\delta+3\,\delta^2\right)+2\,c\left(4\ (1+\delta)+h^2\left(12+12\,\delta+\delta^2\right)-h\left(15+18\,\delta+4\,\delta^2\right)\right)+c^2\,h\left(-9-8\,\delta+4\,\delta^2+h\left(20+28\,\delta+11\,\delta^2\right)\right)\right)\Big)\Big)\Big\}\Big\}$$

Stability

There are at least two immediately apparent equilibria: when drive is absent from the population and when drive is fixed in the population.

We first find the Jacobian for the system:

```
In[•]:= recursions2 = {Dmprime2, dmprime2, DDfprime2, Ddfprime2, ddfprime2};
jacob2 = Simplify[Transpose[{D[recursions2, Dm], D[recursions2, dm],
      D[recursions2, DDf], D[recursions2, Ddf], D[recursions2, ddf]}]];
```

And then we assess the conditions under which the drive-absent and drive-fixed equilibria are stable/instable.

drive absent (pm=pf=0) equilibrium

To evaluate the stability of the equilibrium where drive is absent we consider the eigenvalues of the Jacobian at this point (substituting pf->0, pm->0, and the equilibrium population size without drive). The equilibrium is unstable whenever it has an eigenvalue with absolute value > 1.

```
In[•]:= Simplify[jacob2 /. subeq[[1]] /. pf → 0 /. pm → 0 /. noDrivePopSizeSol2[[2]]];
charpolynomial02 = Factor[Det[% - λ * IdentityMatrix[5]]]
```

$$\text{Out}[\bullet]= -\frac{1}{4\ (1+c)}\lambda^2\left(-4+b+2\,\lambda\right)\left(-2\,c+2\,c\,h\,sf+2\,c\,sm-2\,c\,h\,sf\,sm-2\,c\,\delta+2\,c\,h\,sf\,\delta+2\,c\,sm\,\delta-2\,c\,h\,sf\,sm\,\delta-\lambda-c\,\lambda+h\,sf\,\lambda+c\,h\,sf\,\lambda+2\,\lambda^2+2\,c\,\lambda^2\right)$$

The factor $-\frac{\lambda^2\left(-4+b+2\,\lambda\right)}{4\ (1+c)}$ gives two 0 eigenvalues and and eigenvalue of $\frac{4-b}{2}$, corresponding to the change in overall population size. Thus, we can see that the demographic dynamics are orthogonal to the evolutionary dynamics.

Here, we divide through by this factor to simplify the equation and evaluate the evolutionary dynamics. We then write the characteristic polynomial in the form
$a\lambda^2 + b\lambda + c$
finding the a, b, and c coefficients.

```
In[•]:= Collect[charpolynomial02 / (-λ^2 (-4 + b + 2 λ) / (4 (1 + c))), λ, Simplify];
Coefficient[%, λ^2];
charpolySimpleACoef02 = Collect[%% / %, λ, Simplify]
acoef02 = Coefficient[charpolySimpleACoef02, λ^2];
bcoef02 = Coefficient[charpolySimpleACoef02, λ];
ccoef02 = charpolySimpleACoef02 - bcoef02 * λ - acoef02 * λ^2 // Simplify;
```

$$\text{Out}[\bullet]= -\frac{c\left(-1+h\,sf\right)\ (-1+sm)\ (1+\delta)}{1+c}+\frac{1}{2}\left(-1+h\,sf\right)\lambda+\lambda^2$$

2150

Now check the discriminant

```
In[•]:= bcoef02^2 - 4 acoef02 ccoef02 // FullSimplify
```

$$\text{Out}[\bullet]= \frac{1}{4}\left(-1+\text{h sf}\right)\left(-1+\text{h sf}+\frac{16\,\text{c}\,(-1+\text{sm})\,(1+\delta)}{1+\text{c}}\right)$$

We can see that this is always positive as the two terms in the product are always ≤ 0.

The conditions for the above characteristic polynomial p to have roots between -1 and 1 are:
1. $p(1) > 0$
2. $p(-1) > 0$
3. $-2 < b < 2$

When any of these are false, the largest eigenvalue is > 1, and the equilibrium is unstable.
Condition 3 is always true:

```
In[•]:= FullSimplify[(charpolySimpleACoef02 /. λ → 1) < 0,
         {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        FullSimplify[(charpolySimpleACoef02 /. λ → -1) < 0,
         {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        FullSimplify[-2 < bcoef02 < 2, {1 > δ >= 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
```

$$\text{Out}[\bullet]= \frac{1}{2}\left(1+\text{h sf}-\frac{2\,\text{c}\left(-1+\text{h sf}\right)\,(-1+\text{sm})\,(1+\delta)}{1+\text{c}}\right) < 0$$

$$\text{Out}[\bullet]= \frac{1}{2}\left(3-\text{h sf}-\frac{2\,\text{c}\left(-1+\text{h sf}\right)\,(-1+\text{sm})\,(1+\delta)}{1+\text{c}}\right) < 0$$

```
Out[•]= True
```

The second of these inequalities implies the first and so we need only the first (only one condition is needed for instability).

```
In[•]:= unstabConditions02 = FullSimplify[
          (charpolySimpleACoef02 /. λ → 1) < 0,  {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
```

$$\text{Out}[\bullet]= \frac{1}{2}\left(1+\text{h sf}-\frac{2\,\text{c}\left(-1+\text{h sf}\right)\,(-1+\text{sm})\,(1+\delta)}{1+\text{c}}\right) < 0$$

which can be re-written using the w fitness notation

```
In[•]:= Simplify[unstabConditions02, {wXdX == 1 - h sf, wXdXd == 1 - sf, wXdY == 1 - sm}]
```

$$\text{Out}[\bullet]= (1+\text{c})\left(-2+\text{wXdX}+\text{c}\left(-2+\text{wXdX}+2\,\text{wXdX}\,\text{wXdY}\,(1+\delta)\right)\right) > 0$$

drive fixed (pm=pf=1) equilibrium

We proceed similarly to before, but now considering the other equilibrium:

```
In[•]:= Simplify[jacob2 /. subeq[[1]] /. pf → 1 /. pm → 1 /. fixedDrivePopSizeSol2[[2]]];
        charpolynomial12 = Factor[Det[% - λ * IdentityMatrix[5]]]
```

$$\text{Out}[\bullet]= \frac{1}{4\,(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}\,(4-b+b\,sf-b\,\delta+b\,sf\,\delta-2\,\lambda)\,\lambda^2$$
$$(-2+2\,h\,sf-\lambda-c\,\lambda+h\,sf\,\lambda+c\,h\,sf\,\lambda+sm\,\lambda+c\,sm\,\lambda-h\,sf\,sm\,\lambda-c\,h\,sf\,sm\,\lambda-\delta\,\lambda-$$
$$c\,\delta\,\lambda+h\,sf\,\delta\,\lambda+c\,h\,sf\,\delta\,\lambda+sm\,\delta\,\lambda+c\,sm\,\delta\,\lambda-h\,sf\,sm\,\delta\,\lambda-c\,h\,sf\,sm\,\delta\,\lambda+2\,\lambda^2+$$
$$2\,c\,\lambda^2-2\,sf\,\lambda^2-2\,c\,sf\,\lambda^2-2\,sm\,\lambda^2-2\,c\,sm\,\lambda^2+2\,sf\,sm\,\lambda^2+2\,c\,sf\,sm\,\lambda^2+2\,\delta\,\lambda^2+$$
$$2\,c\,\delta\,\lambda^2-2\,sf\,\delta\,\lambda^2-2\,c\,sf\,\delta\,\lambda^2-2\,sm\,\delta\,\lambda^2-2\,c\,sm\,\delta\,\lambda^2+2\,sf\,sm\,\delta\,\lambda^2+2\,c\,sf\,sm\,\delta\,\lambda^2)$$

The factor $\frac{(4-b+b\,sf-b\,\delta+b\,sf\,\delta-2\,\lambda)\,\lambda^2}{4\,(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}$ gives two 0 eigenvalues and and eigenvalue of $2+\frac{1}{2}\,b\,(-1+sf)\,(1+\delta)$, corresponding to the stability of the ecological equilibrium population size. Thus, we can see that the demographic dynamics are orthogonal to the evolutionary dynamics.

Here, we divide through by this factor to simplify the equation and evaluate the evolutionary dynamics. We then write the characteristic polynomial in the form
$a\lambda^2 + b\lambda + c$
finding the a, b, and c coefficients.

```
In[•]:= Collect[charpolynomial12 / ((4 - b + b sf - b δ + b sf δ - 2 λ) λ^2 / (4 (1 + c) (-1 + sf) (-1 + sm) (1 + δ))), λ, Simplify];
        Coefficient[%, λ^2];
        charpolySimpleACoef12 = Collect[%% / %, λ, Simplify]
        acoef12 = Coefficient[charpolySimpleACoef12, λ^2];
        bcoef12 = Coefficient[charpolySimpleACoef12, λ];
        ccoef12 = charpolySimpleACoef12 - bcoef12 * λ - acoef12 * λ^2;
```

$$\text{Out}[\bullet]= \frac{-1+h\,sf}{(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}+\frac{(1-h\,sf)\,\lambda}{2\,(-1+sf)}+\lambda^2$$

The discriminant is nonnegative and so the eigenvalues are real:

```
In[•]:= bcoef12^2 - 4 acoef12 ccoef12
```

$$\text{Out}[\bullet]= \frac{(1-h\,sf)^2}{4\,(-1+sf)^2}-\frac{4\,(-1+h\,sf)}{(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}$$

The conditions for the above characteristic polynomial p to have roots between -1 and 1 are:
1. $p(1) > 0$
2. $p(-1) > 0$
3. $-2 < b$
4. $b < 2$

When any of these are false, the largest eigenvalue is > 1, and the equilibrium is unstable.

Now checking these conditions for instability by reversing the inequalities required for stability

2152

```
In[•]:= r1 = FullSimplify[(charpolySimpleACoef12 /. λ → 1) < 0,
           {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
         r2 = FullSimplify[(charpolySimpleACoef12 /. λ → -1) < 0,
           {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
         r3 = FullSimplify[bcoef12 < -2, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
         r4 = FullSimplify[2 < bcoef12, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
```

$$\text{Out[•]= } 1+\frac{1-h\,sf}{2\,(-1+sf)}+\frac{-1+h\,sf}{(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}<0$$

$$\text{Out[•]= } 1+\frac{1-h\,sf}{2-2\,sf}+\frac{-1+h\,sf}{(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}<0$$

```
Out[•]= 3 + h sf < 4 sf

Out[•]= False
```

The second inequality implies the first again. r3 only applies if δ or c is less than 1. Otherwise, it is covered by r1.

```
In[•]:= Reduce[{! r1, r3, 1 == δ || 1 == c, 0 < δ ≤ 1, 0 < c ≤ 1, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= False
```

If δ<1 or c<1, r3 shows that the drive-fixed equilibrium is always unstable when heterozygous females have at least 4 times higher survival probability than homozygous drive females: re-writing r3 as `4 (1 - sf) < 1 - h sf`. Even if r3 is not met, the drive-fixed equilibrium can still be unstable when the r1 condition is met.

```
In[•]:= unstabConditions12 = r1 || r3
```

$$\text{Out[•]= } 1+\frac{1-h\,sf}{2\,(-1+sf)}+\frac{-1+h\,sf}{(1+c)\,(-1+sf)\,(-1+sm)\,(1+\delta)}<0\ ||\ 3+h\,sf<4\,sf$$

Allele Frequency and Population Size Dynamics - females mate with many males ($\lambda_f \to \infty$)

Recursion equations

Here we assume that each female mates many times. Therefore, females effectively sample alleles from the total sperm produced by the male population (S). In the following recursion/difference equations,
Xm are sperm with the wildtype-X allele
Xdm are sperm with the drive-X allele
Ym are sperm with the Y allele
Xf are eggs with the wildtype-X allele
Xdf are eggs with the drive-X allele
These gametes then be combined to produce the juveniles of the next generation.

```
In[•]:= S = Dm + c dm;
        Xm = Dm / (2 S);
        Ym = (Dm/2 + dm c (1 - δ)/2) / S;
        Xdm = c (1 + δ)/2 dm / S;
        Xf = DDf + Ddf/2;
        Xdf = Ddf/2 + ddf;

        DDfjuvenilesi = b (1 - α totAdults) Xm Xf;
        Ddfjuvenilesi = b (1 - α totAdults) (Xdm Xf + Xm Xdf);
        ddfjuvenilesi = b (1 - α totAdults) Xdm Xdf;
        Dmjuvenilesi = b (1 - α totAdults) Ym Xf;
        dmjuvenilesi = b (1 - α totAdults) Ym Xdf;
```

Before the next generation of adults, viability selection acts according to genotype.

```
In[•]:= DDfprimei = DDfjuvenilesi * 1;
        Ddfprimei = Ddfjuvenilesi * (1 - h * sf);
        ddfprimei = ddfjuvenilesi * (1 - sf);
        Dmprimei = Dmjuvenilesi * 1;
        dmprimei = dmjuvenilesi * (1 - sm);
```

Equilibrium allele frequency and population size

We first look at only the difference in allele frequency and inbreeding coefficient to calculate the polymorphic allele frequency equilibrium (when changes in allele frequency and inbreeding coefficient are 0)

```
In[•]:= solvemei = Simplify[
          {(dmprimei/(Dmprimei + dmprimei)) - dm/(dm + Dm),
           (2 ddfprimei + Ddfprimei)/(2 (ddfprimei + Ddfprimei + DDfprimei)) -
            (2 ddf + Ddf)/(2 (ddf + Ddf + DDf)),
           (4 ddfprimei DDfprimei - Ddfprimei^2)/((2 ddfprimei + Ddfprimei) (Ddfprimei + 2 DDfprimei)) -
            (4 ddf DDf - Ddf^2)/((2 ddf + Ddf) (Ddf + 2 DDf))} /. subeq];
        equilibriumi = Solve[solvemei[[1]] == {0, 0, 0}, {pm, pf, F}] // Simplify
```

$$\left\{\left\{pm \to -\frac{(-1+sm)\left(-1-h\,sf+c\left(-1+h\,sf\right)(-1+sm)(1+\delta)\right)}{h\,sf\left(-2+sm\right)+sm-c\left(sf\left(2+h\left(-2+sm\right)\right)-sm\right)(-1+sm)(1+\delta)},\right.\right.$$
$$pf \to \frac{-1-h\,sf+c\left(-1+h\,sf\right)(-1+sm)(1+\delta)}{2\,sf\left(-h+c\left(-1+h\right)(-1+sm)(1+\delta)\right)},$$
$$F \to \left(1-h^2\,sf^2+2\,c\left(1-sf+\left(-1+h\right)h\,sf^2\right)(-1+sm)(1+\delta)-\right.$$
$$\left.\left.c^2\left(-1+2\,sf+\left(-2+h\right)h\,sf^2\right)(-1+sm)^2(1+\delta)^2\right)\Big/\left(\left(-1+h\,sf\right)^2-2\,c\right.\right.$$
$$\left.\left.\left(-1-2\left(-1+h\right)sf+h^2\,sf^2\right)(-1+sm)(1+\delta)+c^2\left(-1+h\,sf\right)^2(-1+sm)^2(1+\delta)^2\right)\right\}\right\}$$

The full system is described by

2154

```
In[•]:= solvemeFulli = {DDfprimei - DDf, Ddfprimei - Ddf, ddfprimei - ddf,
          Dmprimei - Dm, dmprimei - dm} /. subeq[[1]] // Simplify;
```

We can then solve for the population sizes. First we find population sizes when drive alleles are absent (pm=pf=0) or fixed (pm=pf=1):

```
In[•]:= noDrivePopSizeSoli =
          Solve[(solvemeFulli /. pm → 0 /. pf → 0) == {0, 0, 0, 0, 0}, {maleNum, femaleNum}];
        fixedDrivePopSizeSoli = Solve[(solvemeFulli /. pm → 1 /. pf → 1) == {0, 0, 0, 0, 0},
          {maleNum, femaleNum}];
        NnoDrivei = (maleNum + femaleNum) /. noDrivePopSizeSoli[[2]] // FullSimplify
        Nfixedi = (maleNum + femaleNum) /. fixedDrivePopSizeSoli[[2]] // FullSimplify
```

$$\text{Out[•]=}\ \frac{-2+b}{b\,\alpha}$$

$$\text{Out[•]=}\ \frac{1+\frac{2}{b\,(-1+sf)\,(1+\delta)}}{\alpha}$$

Now for the case where a polymorphic equilibrium has been reached

```
In[•]:= equilPopSizeSoli =
          Simplify[Solve[Flatten[Simplify[(solvemeFulli /. equilibriumi /. sm → 0)]] ==
            {0, 0, 0, 0, 0}, {femaleNum, maleNum}]];
```

```
In[•]:= Neqi = (femaleNum + maleNum) /. equilPopSizeSol[[2]] // Simplify
```

$$\text{Out[•]=}\ \Big(b\,h^2\,sf^2\,(-2+sm-\delta+sm\,\delta)^2+(sm-\delta+sm\,\delta)\,\big(-b\,\delta+sm\,(-4+b+b\,\delta)\big)-2\,sf\,\big(-4\,(-1+sm)\,(1+\delta)+2\,b\,(-1+sm)\,(1+\delta)-2\,h\,(-2+sm)\,(-2+sm-\delta+sm\,\delta)+b\,h\,(-2+sm-\delta+sm\,\delta)^2\big)\Big)\Big/\Big(b\,\alpha\,\big(h^2\,sf^2\,(-2+sm-\delta+sm\,\delta)^2+(sm-\delta+sm\,\delta)^2-2\,sf\,\big(2\,(-1+sm)\,(1+\delta)+h\,(-2+sm-\delta+sm\,\delta)^2\big)\big)\Big)$$

Stability

There are at least two immediately apparent equilibria: when drive is absent from the population and when drive is fixed in the population.

We first find the Jacobian for the system:

```
In[•]:= recursionsi = {Dmprimei, dmprimei, DDfprimei, Ddfprimei, ddfprimei};
        jacobi = Simplify[Transpose[{D[recursionsi, Dm], D[recursionsi, dm],
            D[recursionsi, DDf], D[recursionsi, Ddf], D[recursionsi, ddf]}]];
```

And then we assess the conditions under which the drive-absent and drive-fixed equilibria are stable/instable.

drive absent (pm=pf=0) equilibrium

Proceed as before:

```
In[•]:= Simplify[jacobi /. subeq[[1]] /. pf → 0 /. pm → 0 /. noDrivePopSizeSoli[[2]]];
        charpolynomial0i = Factor[Det[% - λ * IdentityMatrix[5]]]
```

$$\text{Out[•]}= -\frac{1}{4}\lambda^2\left(-4+b+2\lambda\right)\left(-c+c\,h\,sf+c\,sm-c\,h\,sf\,sm-c\,\delta+c\,h\,sf\,\delta+c\,sm\,\delta-c\,h\,sf\,sm\,\delta-\lambda+h\,sf\,\lambda+2\lambda^2\right)$$

The factor $-\frac{\lambda^2\left(-4+b+2\lambda\right)}{4}$ gives two 0 eigenvalues and and eigenvalue of $\frac{4-b}{2}$, corresponding to the change in overall population size. Thus, we can see that the demographic dynamics are orthogonal to the evolutionary dynamics.

Here, we divide through by this factor to simplify the equation and evaluate the evolutionary dynamics. We then write the characteristic polynomial in the form
$a\lambda^2 + b\lambda + c$
finding the a, b, and c coefficients.

```
In[•]:= Collect[charpolynomial0i / (-1/4 λ^2 (-4 + b + 2 λ)), λ, Simplify];
        Coefficient[%, λ^2];
        charpolySimpleACoef0i = Collect[%% / %, λ, Simplify]
        acoef0i = Coefficient[charpolySimpleACoef0i, λ^2];
        bcoef0i = Coefficient[charpolySimpleACoef0i, λ];
        ccoef0i = charpolySimpleACoef0i - bcoef0i * λ - acoef0i * λ^2 // Simplify;
```

$$\text{Out[•]}= -\frac{1}{2}c\left(-1+h\,sf\right)\left(-1+sm\right)\left(1+\delta\right)+\frac{1}{2}\left(-1+h\,sf\right)\lambda+\lambda^2$$

Now check the discriminant

```
In[•]:= ccoef0i
```

$$\text{Out[•]}= -\frac{1}{2}c\left(-1+h\,sf\right)\left(-1+sm\right)\left(1+\delta\right)$$

```
In[•]:= bcoef0i^2 - 4 acoef0i ccoef0i // FullSimplify
```

$$\text{Out[•]}= \frac{1}{4}\left(-1+h\,sf\right)\left(-1+h\,sf+8\,c\left(-1+sm\right)\left(1+\delta\right)\right)$$

We can see that this is always positive as the two terms in the product are always ≤ 0.

The conditions for the above characteristic polynomial p to have roots between -1 and 1 are:
1. $p(1) > 0$
2. $p(-1) > 0$
3. $-2 < b < 2$

When any of these are false, the largest eigenvalue is > 1, and the equilibrium is unstable.
Condition 3 is always true:

2156

```
In[•]:= FullSimplify[(charpolySimpleACoef0i /. λ → 1) < 0,
         {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        FullSimplify[(charpolySimpleACoef0i /. λ → -1) < 0,
         {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        FullSimplify[-2 < bcoef0i < 2, {1 > δ >= 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= 1 + h sf < c (-1 + h sf) (-1 + sm) (1 + δ)

Out[•]= h sf + c (-1 + h sf) (-1 + sm) (1 + δ) > 3

Out[•]= True
```

The second of these inequalities implies the first and so we need only the first (only one condition is needed for instability).

```
In[•]:= unstabConditions0i = FullSimplify[
          (charpolySimpleACoef0i /. λ → 1) < 0, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= 1 + h sf < c (-1 + h sf) (-1 + sm) (1 + δ)
```

which can be re-written using the w fitness notation

```
In[•]:= Simplify[unstabConditions0i, {wXdX == 1 - h sf, wXdXd == 1 - sf, wXdY == 1 - sm}]
Out[•]= wXdX + c wXdX wXdY (1 + δ) > 2
```

drive fixed (pm=pf=1) equilibrium

We proceed similarly to before, but now considering the other equilibrium:

```
In[•]:= Simplify[jacobi /. subeq[[1]] /. pf → 1 /. pm → 1 /. fixedDrivePopSizeSoli[[2]]];
        charpolynomial1i = Factor[Det[% - λ * IdentityMatrix[5]]]
Out[•]= 1 / (4 c (-1 + sf) (-1 + sm) (1 + δ))
          (4 - b + b sf - b δ + b sf δ - 2 λ) λ^2 (-1 + h sf - c λ + c h sf λ + c sm λ -
            c h sf sm λ - c δ λ + c h sf δ λ + c sm δ λ - c h sf sm δ λ + 2 c λ^2 - 2 c sf λ^2 -
            2 c sm λ^2 + 2 c sf sm λ^2 + 2 c δ λ^2 - 2 c sf δ λ^2 - 2 c sm δ λ^2 + 2 c sf sm δ λ^2)
```

The factor $\frac{(4-b+b\,sf-b\,\delta+b\,sf\,\delta-2\,\lambda)\,\lambda^2}{4\,c\,(-1+sf)\,(-1+sm)\,(1+\delta)}$ gives two 0 eigenvalues and and eigenvalue of $2 + \frac{1}{2}\,b\,(-1 + sf)\,(1 + \delta)$, corresponding to the stability of the ecological equilibrium population size. Thus, we can see that the demographic dynamics are orthogonal to the evolutionary dynamics.

Here, we divide through by this factor to simplify the equation and evaluate the evolutionary dynamics. We then write the characteristic polynomial in the form
$a\lambda^2 + b\lambda + c$
finding the a, b, and c coefficients.

```
In[•]:= Collect[charpolynomial1i / ((4 - b + b sf - b δ + b sf δ - 2 λ) λ^2)/(4 c (-1 + sf) (-1 + sm) (1 + δ)), λ, Simplify];
        Coefficient[%, λ^2];
        charpolySimpleACoef1i = Collect[%% / %, λ, Simplify]
        acoef1i = Coefficient[charpolySimpleACoef1i, λ^2];
        bcoef1i = Coefficient[charpolySimpleACoef1i, λ];
        ccoef1i = charpolySimpleACoef1i - bcoef1i * λ - acoef1i * λ^2;
```

$$\text{Out[•]=}\quad \frac{-1+\text{h sf}}{2\,\text{c}\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)}+\frac{(1-\text{h sf})\,\lambda}{2\,(-1+\text{sf})}+\lambda^2$$

The discriminant is nonnegative and so the eigenvalues are real:

```
In[•]:= bcoef1i^2 - 4 acoef1i ccoef1i
```

$$\text{Out[•]=}\quad \frac{(1-\text{h sf})^2}{4\,(-1+\text{sf})^2}-\frac{2\,(-1+\text{h sf})}{\text{c}\,(-1+\text{sf})\,(-1+\text{sm})\,(1+\delta)}$$

The conditions for the above characteristic polynomial p to have roots between -1 and 1 are:

1. $p(1) > 0$
2. $p(-1) > 0$
3. $-2 < b$
4. $b < 2$

When any of these are false, the largest eigenvalue is > 1, and the equilibrium is unstable.

Now checking these conditions for instability by reversing the inequalities required for stability

```
In[•]:= r1 = FullSimplify[(charpolySimpleACoef1i /. λ → 1) < 0,
          {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        r2 = FullSimplify[(charpolySimpleACoef1i /. λ → -1) < 0,
          {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        r3 = FullSimplify[bcoef1i < -2, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
        r4 = FullSimplify[2 < bcoef1i, {1 > δ > 0, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
```

$$\text{Out[•]=}\quad \text{c}\,(1-\text{h sf}+\text{c}\,(1+(-2+\text{h})\,\text{sf})\,(-1+\text{sm})\,(1+\delta))>0$$

$$\text{Out[•]=}\quad \text{c}\,(-1+\text{h sf}+\text{c}\,(-3+(2+\text{h})\,\text{sf})\,(-1+\text{sm})\,(1+\delta))<0$$

$$\text{Out[•]=}\quad 3+\text{h sf}<4\,\text{sf}$$

```
Out[•]= False
```

The second inequality implies the first again. r3 only applies if δ or c is less than 1. Otherwise, it is covered by r1.

```
In[•]:= Reduce[{!r1, r3, 1 == δ || 1 == c, 0 < δ ≤ 1, 0 < c ≤ 1, 1 > h > 0, 1 > sm > 0, 0 < sf < 1}]
Out[•]= False
```

If δ<1 or c<1, r3 shows that the drive-fixed equilibrium is always unstable when heterozygous females have at least 4 times higher survival probability than homozygous drive females: re-writing r3 as $4\,(1-\text{sf}) < 1-\text{h sf}$. Even if r3 is not met, the drive-fixed equilibrium can still be unstable when the r1 condition is met.

2158

```
In[•]:= unstabConditions1i = r1 || r3
Out[•]= c (1 - h sf + c (1 + (-2 + h) sf) (-1 + sm) (1 + δ)) > 0 || 3 + h sf < 4 sf
```

Appendix 2

This section contains a pdf version of a Mathematica notebook containing derivations of results in Chapter 3.

2162

Enter me first

In[58]:=
```
subeqns2 =
  {p1g22 == 1 - p1g11 - p1g12 - p1g21, p1fA1 == p1g11 + p1g12, p1fB1 == p1g11 + p1g21,
   p1D == p1g11 p1g22 - p1g12 p1g21, p2g22 == 1 - p2g11 - p2g12 - p2g21,
   p2fA1 == p2g11 + p2g12, p2fB1 == p2g11 + p2g21, p2D == p2g11 p2g22 - p2g12 p2g21};
alleleSubs =
  Flatten[Solve[subeqns2, {p1g11, p1g12, p1g21, p1g22, p2g11, p2g12, p2g21, p2g22}]];
symmetrySubs = {m12 → m, m21 → m, s1 → s, s2 → s};
linearise[exp_, args_] :=
 With[{subs = Evaluate[Table[args[[i]] → args[[i]] ϵ, {i, 1, Length[args]}]]},
  Simplify[Normal[Series[exp /. subs, {ϵ, 0, 1}]] /. ϵ → 1]]
secondOrder[exp_, args_] :=
 With[{subs = Evaluate[Table[args[[i]] → args[[i]] ϵ, {i, 1, Length[args]}]]},
  Simplify[Normal[Series[exp /. subs, {ϵ, 0, 2}]] /. ϵ → 1]]
symmMiSubs = {m12 → m, m21 → m};
symmSelSubs = {s1 → s, s2 → s};
equilSubs =
  {p2g21 → p2g12, p1g21 → p1g12, p1g22 → 1 - 2 p1g12 - p1g11, p2g11 → 1 - 2 p2g12 - p2g22};
```

Continent-island model

Recursions

Prior to reproduction, adult frequencies are multiplied by their relative fitness...

In[•]:=
```
ClearAll[p1w11, p1w12, p1w21, p1w22]
p1w11 = (1 + s)^2;
p1w12 = (1 + s);
p1w21 = (1 + s);
p1w22 = 1;

Clear[p1Wm]
p1Wm = p1g11 p1w11 + p1g12 p1w12 + p1g21 p1w21 + p1g22 p1w22;

p1g11s = p1g11 p1w11 / p1Wm;
p1g12s = p1g12 p1w12 / p1Wm;
p1g21s = p1g21 p1w21 / p1Wm;
p1g22s = p1g22 p1w22 / p1Wm;
```

...so that their probability of reproduction can be incorporated into the reproduction phase.

In[•]:=
```
p1g11r = (1 - r) (p1g11s) + r (p1g11s^2 + p1g11s p1g12s + p1g11s p1g21s + p1g12s p1g21s);
p1g12r = (1 - r) p1g12s + r (p1g12s^2 + p1g12s p1g11s + p1g12s p1g22s + p1g11s p1g22s);
p1g21r = (1 - r) p1g21s + r (p1g21s^2 + p1g21s p1g11s + p1g21s p1g22s + p1g11s p1g22s);
p1g22r = (1 - r) p1g22s + r (p1g22s^2 + p1g22s p1g12s + p1g22s p1g21s + p1g12s p1g21s);
```

All haplotype frequencies are reduced by migration, and then $A_2\,B_2$ migrants are introduced.

```
In[•]:= p1g11m = p1g11r (1 - m) ;
        p1g12m = p1g12r (1 - m);
        p1g21m = p1g21r (1 - m);
        p1g22m = p1g22r (1 - m) + m;

        p1g11m + p1g12m + p1g21m + p1g22m // Simplify
Out[•]= 1
```

The change in frequency is given by the difference between frequencies at the beginning of the generation, and the end of it (beginning of the next)

```
In[•]:= CIdelp1g11 = Simplify[p1g11m - p1g11];
        CIdelp1g12 = Simplify[p1g12m - p1g12];
        CIdelp1g21 = Simplify[p1g21m - p1g21];
        CIdelp1g22 = Simplify[p1g22m - p1g22];
```

QLE

```
In[•]:= Dprime = p1g11m * p1g22m - p1g12m * p1g21m /. alleleSubs // Simplify;
        delD = Dprime - p1D // Simplify;
```

Continuous time QLE

by assuming all rate parameters (m, s, r) are of the same order, we can derive a continuous time approximation to the discrete time system:

```
In[•]:= dDdt = linearise[delD, {m, s, r}]
Out[•]= m p1fA1 p1fB1 - p1D r - p1D (m + 2 (-1 + p1fA1 + p1fB1) s)
```

find the equilibrium LD value in terms of allele frequencies:

```
In[•]:= DsubAF = Solve[dDdt == 0, p1D]
Out[•]= {{p1D → m p1fA1 p1fB1 / (m + r - 2 s + 2 p1fA1 s + 2 p1fB1 s)}}
```

now we repeat for the allele frequencies. since there is no difference between locus A and locus B, we can treat just one of them.

```
In[•]:= dfA1dt = linearise[p1g11m + p1g12m - p1fA1 /. alleleSubs /. p1fB1 → p1fA1, {m, s, r}]
Out[•]= -m p1fA1 + (p1D + p1fA1 - p1fA1^2) s
```

at QLE, m, s << r so that m/r, s/r << 1. therefore, we take 1/r as a small parameter in the approximation and solve to order 1/r, using the LD value we found earlier:

```
In[•]:= Normal[Series[dfA1dt /. DsubAF /. p1fB1 → p1fA1, {r, Infinity, 1}]];
        Solve[% == 0, p1fA1];
        Collect[Simplify[Normal[Series[p1fA1 /. %[[2]], {r, Infinity, 1}]], r > 0 && s > 0], r]
Out[•]= -(m - s)/s - m (m - s)/(r s)
```

now we can go back and find the equilibrium LD in terms of the model parameters

2164

```
In[•]:= Normal[Series[p1D /. DsubAF /. p1fB1 → p1fA1 /. p1fA1 → -(m - s)/s - m (m - s)/(r s),
          {r, Infinity, 1}]] // Simplify
Out[•]= {m (m - s)^2 / (r s^2)}
```

Two deme model

Recursions

We proceed as in the continent-island model, but now include fitness and frequencies in deme 2:

```
In[1]:= ClearAll[p1w11, p1w12, p1w21, p1w22, p2w11, p2w12, p2w21, p2w22]
wSubs = {p1w11 → (1 + s1)^2,
   p1w12 → (1 + s1),
   p1w21 → (1 + s1),
   p1w22 → 1,
   p2w11 → 1,
   p2w12 → (1 + s2),
   p2w21 → (1 + s2),
   p2w22 → (1 + s2)^2};

ClearAll["p1Wm", "p2Wm"]
p1Wm = p1g11 p1w11 + p1g12 p1w12 + p1g21 p1w21 + p1g22 p1w22;
p2Wm = p2g11 p2w11 + p2g12 p2w12 + p2g21 p2w21 + p2g22 p2w22;

p1g11s = p1g11 p1w11 / p1Wm;
p1g12s = p1g12 p1w12 / p1Wm;
p1g21s = p1g21 p1w21 / p1Wm;
p1g22s = p1g22 p1w22 / p1Wm;

p2g11s = p2g11 p2w11 / p2Wm;
p2g12s = p2g12 p2w12 / p2Wm;
p2g21s = p2g21 p2w21 / p2Wm;
p2g22s = p2g22 p2w22 / p2Wm;
```

```
In[14]:= p1g11r = (1 - r) (p1g11s) + r (p1g11s^2 + p1g11s p1g12s + p1g11s p1g21s + p1g12s p1g21s);
p1g12r = (1 - r) p1g12s + r (p1g12s^2 + p1g12s p1g11s + p1g12s p1g22s + p1g11s p1g22s);
p1g21r = (1 - r) p1g21s + r (p1g21s^2 + p1g21s p1g11s + p1g21s p1g22s + p1g11s p1g22s);
p1g22r = (1 - r) p1g22s + r (p1g22s^2 + p1g22s p1g12s + p1g22s p1g21s + p1g12s p1g21s);


p2g11r = (1 - r) (p2g11s) + r (p2g11s^2 + p2g11s p2g12s + p2g11s p2g21s + p2g12s p2g21s);
p2g12r = (1 - r) p2g12s + r (p2g12s^2 + p2g12s p2g11s + p2g12s p2g22s + p2g11s p2g22s);
p2g21r = (1 - r) p2g21s + r (p2g21s^2 + p2g21s p2g11s + p2g21s p2g22s + p2g11s p2g22s);
p2g22r = (1 - r) p2g22s + r (p2g22s^2 + p2g22s p2g12s + p2g22s p2g21s + p2g12s p2g21s);
```

```
In[22]:= p1g11m = (p1g11r * (1 - m21) + m21 p2g11r);
         p1g12m = (p1g12r * (1 - m21) + m21 p2g12r);
         p1g21m = (p1g21r * (1 - m21) + m21 p2g21r);
         p1g22m = (p1g22r * (1 - m21) + m21 p2g22r);
         p2g11m = (p2g11r * (1 - m12) + m12 p1g11r);
         p2g12m = (p2g12r * (1 - m12) + m12 p1g12r);
         p2g21m = (p2g21r * (1 - m12) + m12 p1g21r);
         p2g22m = (p2g22r * (1 - m12) + m12 p1g22r);


In[30]:= delp1g11 = Simplify[p1g11m - p1g11];
         delp1g12 = Simplify[p1g12m - p1g12];
         delp1g21 = Simplify[p1g21m - p1g21];
         delp1g22 = Simplify[p1g22m - p1g22];

         delp2g11 = Simplify[p2g11m - p2g11];
         delp2g12 = Simplify[p2g12m - p2g12];
         delp2g21 = Simplify[p2g21m - p2g21];
         delp2g22 = Simplify[p2g22m - p2g22];

In[38]:= postSelectionSubs = {D1s → (s1g11 s1g22 - s1g12 s1g21),
            D2s → (s2g11 s2g22 - s2g12 s2g21), s1g11 → p1g11 p1w11 / wbar1,
            s1g12 → p1g12 p1w12 / wbar1, s1g21 → p1g21 p1w21 / wbar1, s1g22 → p1g22 p1w22 / wbar1,
            s2g11 → p2g11 p2w11 / wbar2, s2g12 → p2g12 p2w12 / wbar2,
            s2g21 → p2g21 p2w21 / wbar2, s2g22 → p2g22 p2w22 / wbar2};
```

Verify that the final recursions can be written in the simplified form given in the main text:

```
In[39]:= p1X11 = ((1 - m21) (s1g11 - r D1s) + m21 (s2g11 - r D2s));
         p1X12 = ((1 - m21) (s1g12 + r D1s) + m21 (s2g12 + r D2s));
         p1X21 = ((1 - m21) (s1g21 + r D1s) + m21 (s2g21 + r D2s));
         p1X22 = ((1 - m21) (s1g22 - r D1s) + m21 (s2g22 - r D2s));
         p2X11 = ((1 - m12) (s2g11 - r D2s) + m12 (s1g11 - r D1s));
         p2X12 = ((1 - m12) (s2g12 + r D2s) + m12 (s1g12 + r D1s));
         p2X21 = ((1 - m12) (s2g21 + r D2s) + m12 (s1g21 + r D1s));
         p2X22 = ((1 - m12) (s2g22 - r D2s) + m12 (s1g22 - r D1s));
         recursions = {p1g11m, p1g12m, p1g21m, p1g22m, p2g11m, p2g12m, p2g21m, p2g22m};
         simplifiedRecursions = {p1X11, p1X12, p1X21, p1X22, p2X11, p2X12, p2X21, p2X22};
         (recursions - simplifiedRecursions) /. postSelectionSubs /. postSelectionSubs /.
            wbar1 → p1Wm /. wbar2 → p2Wm // Simplify
Out[49]= {0, 0, 0, 0, 0, 0, 0, 0}
```

The following matrices correspond to M_{11} and M_{22} as written in the main text :

```
In[50]:= mat = {{(1 - m21) ((1 + s1)^2) / wbar1, m12 ((1 + s1)^2) / wbar1}, {m21 / wbar2, (1 - m12) / wbar2}};
         mat2 = {{(1 - m21) / wbar1, m12 / wbar1}, {m21 (1 + s2)^2 / wbar2, (1 + s2)^2 (1 - m12) / wbar2}};
         charpoly = Simplify[Det[mat - IdentityMatrix[2] * λ]];
         charpoly2 = Simplify[Det[mat2 - IdentityMatrix[2] * λ]];
```

2166

QLE

Continuous time

in the QLE approximation, we assume that r >> m, s, so that m/r, s/r << 1. so, we take 1/r as a small parameter in the model.

so, the allele frequencies at QLE are of the form f1A1 = F0 + F1/r + O(1/r^2) etc.

first, we convert the model to continuous time

In[72]:=
```
df1A1 = linearise[p1g11m + p1g12m - p1fA1 /. alleleSubs /. wSubs, {m12, m21, s1, s2, r}];
df2A1 = linearise[p2g11m + p2g12m - p2fA1 /. alleleSubs /. wSubs, {m12, m21, s1, s2, r}];
df1B1 = linearise[p1g11m + p1g21m - p1fB1 /. alleleSubs /. wSubs, {m12, m21, s1, s2, r}];
df2B1 = linearise[p2g11m + p2g21m - p2fB1 /. alleleSubs /. wSubs, {m12, m21, s1, s2, r}];
dD1 = linearise[
   p1g11m p1g22m - p1g12m p1g21m - p1D /. alleleSubs /. wSubs, {m12, m21, s1, s2, r}];
dD2 = linearise[
   p2g11m p2g22m - p2g12m p2g21m - p2D /. alleleSubs /. wSubs, {m12, m21, s1, s2, r}];
rsubs = {p1fA1 → f10 + f11 / r,
  p2fA1 → f20 + f21 / r, p1fB1 → g10 + g11 / r , p2fB1 → g20 + g21 / r}
```

Out[78]=
$$\left\{\text{p1fA1} \to \text{f10} + \frac{\text{f11}}{\text{r}}, \text{p2fA1} \to \text{f20} + \frac{\text{f21}}{\text{r}}, \text{p1fB1} \to \text{g10} + \frac{\text{g11}}{\text{r}}, \text{p2fB1} \to \text{g20} + \frac{\text{g21}}{\text{r}}\right\}$$

at QLE, close to D = 0, the rate of generation of LD due to selection is equal to the rate of decay caused by recombination:

In[79]:=
```
p1D /. Solve[(dD1 /. p1D → 0 /. p2D → 0) == r p1D, p1D] // Simplify;
p2D /. Solve[(dD2 /. p1D → 0 /. p2D → 0) == r p2D, p2D] // Simplify;
{%%, %} /. rsubs;
Coefficient[%, 1 / r] / r // Simplify;
{D1sub, D2sub} = % // Flatten
```

Out[83]=
$$\left\{\frac{(\text{f10} - \text{f20})\,(\text{g10} - \text{g20})\,\text{m21}}{\text{r}}, \frac{(\text{f10} - \text{f20})\,(\text{g10} - \text{g20})\,\text{m12}}{\text{r}}\right\}$$

at equilibrium the following expressions are equal to 0

In[84]:=
```
dAlleleFreq = {df1A1, df2A1 , df1B1, df2B1};
```

using this value of LD, determine the linear term in the allele frequencies, which are the allele frequencies when the system is in full linkage equilibrium (ie D = 0)

In[85]:=
```
dAlleleFreq /. rsubs /. p1D → D1sub /. p2D → D2sub // Simplify;
% /. r → 1 / R // Simplify;
% /. R → 0;
freqLinearTermSubs = Simplify[Solve[% == 0, {f10, f20, g10, g20}][[16]]]
```

Out[88]=
$$\left\{f10 \to \frac{1}{2}\left(1 - \frac{2\,m21}{s1} + \frac{\sqrt{4\,m12\,m21 + s1\,s2}}{\sqrt{s1}\ \sqrt{s2}}\right),\ f20 \to \frac{2\,m12 + s2 - \frac{\sqrt{s2}\ \sqrt{4\,m12\,m21+s1\,s2}}{\sqrt{s1}}}{2\,s2},\right.$$
$$\left. g10 \to \frac{1}{2}\left(1 - \frac{2\,m21}{s1} + \frac{\sqrt{4\,m12\,m21 + s1\,s2}}{\sqrt{s1}\ \sqrt{s2}}\right),\ g20 \to \frac{2\,m12 + s2 - \frac{\sqrt{s2}\ \sqrt{4\,m12\,m21+s1\,s2}}{\sqrt{s1}}}{2\,s2}\right\}$$

now find the O(1/r) term

In[89]:=
```
dAlleleFreq /. rsubs /. p1D → D1sub /. p2D → D2sub // Simplify;
% /. r → 1 / R // Simplify;
D[%, R] /. R → 0 // Simplify
freqFirstOrderTermSubs = Simplify[Solve[% == 0, {f11, f21, g11, g21}]] // Flatten
```

Out[91]=
$$\{f11\,(-m21 + s1 - 2\,f10\,s1) + m21\,(f21 + (f10 - f20)\,(g10 - g20)\,s1),$$
$$f11\,m12 - (f10 - f20)\,(g10 - g20)\,m12\,s2 - f21\,(m12 + s2 - 2\,f20\,s2),$$
$$g11\,(-m21 + s1 - 2\,g10\,s1) + m21\,(g21 + (f10 - f20)\,(g10 - g20)\,s1),$$
$$g11\,m12 - (f10 - f20)\,(g10 - g20)\,m12\,s2 - g21\,(m12 + s2 - 2\,g20\,s2)\}$$

Out[92]=
$$\left\{f11 \to \frac{(f10 - f20)\,(g10 - g20)\,m21\,((-1 + 2\,f20)\,s1\,s2 + m12\,(-s1 + s2))}{m12\,(s1 - 2\,f10\,s1) + (-1 + 2\,f20)\,(m21 + (-1 + 2\,f10)\,s1)\,s2},\right.$$
$$f21 \to \frac{(f10 - f20)\,(g10 - g20)\,m12\,((-1 + 2\,f10)\,s1\,s2 + m21\,(-s1 + s2))}{m12\,(s1 - 2\,f10\,s1) + (-1 + 2\,f20)\,(m21 + (-1 + 2\,f10)\,s1)\,s2},$$
$$g11 \to \frac{(f10 - f20)\,(g10 - g20)\,m21\,((-1 + 2\,g20)\,s1\,s2 + m12\,(-s1 + s2))}{m12\,(s1 - 2\,g10\,s1) + (-1 + 2\,g20)\,(m21 + (-1 + 2\,g10)\,s1)\,s2},$$
$$\left. g21 \to \frac{(f10 - f20)\,(g10 - g20)\,m12\,((-1 + 2\,g10)\,s1\,s2 + m21\,(-s1 + s2))}{m12\,(s1 - 2\,g10\,s1) + (-1 + 2\,g20)\,(m21 + (-1 + 2\,g10)\,s1)\,s2}\right\}$$

In[129]:=
```
qleAlleleFreqs = Simplify[rsubs /. freqFirstOrderTermSubs /. freqLinearTermSubs,
    s1 > 0 && s2 > 0 && m12 > 0 && m21 > 0 && r > 0] // Flatten;
qleD = FullSimplify[
  {p1D → D1sub, p2D → D2sub} /. freqLinearTermSubs, s1 > 0 && s2 > 0 && m12 > 0 && m21 > 0]
```

Out[130]=
$$\left\{p1D \to \frac{m21\,\left(m12\,s1 + m21\,s2 - \sqrt{s1\,s2\,(4\,m12\,m21 + s1\,s2)}\right)^2}{r\,s1^2\,s2^2},\right.$$
$$\left. p2D \to \frac{m12\,\left(m12\,s1 + m21\,s2 - \sqrt{s1\,s2\,(4\,m12\,m21 + s1\,s2)}\right)^2}{r\,s1^2\,s2^2}\right\}$$

In[94]:=
```
FullSimplify[D1sub /. freqLinearTermSubs, s1 > 0 && s2 > 0 && m12 > 0 && m21 > 0]
```

Out[94]=
$$\frac{m21\,\left(m12\,s1 + m21\,s2 - \sqrt{s1\,s2\,(4\,m12\,m21 + s1\,s2)}\right)^2}{r\,s1^2\,s2^2}$$

2168

In[99]:=
```
simpleLinearTerms = Simplify[
   freqLinearTermSubs /. m12 → s2 l2 /. m21 → s1 l1, s1 > 0 && s2 > 0 && l1 > 0 && l2 > 0];
Simplify[freqFirstOrderTermSubs /. m12 → s2 l2 /. m21 → s1 l1,
  s1 > 0 && s2 > 0 && l1 > 0 && l2 > 0];
{f11 / (g10 - g20) /. % /. simpleLinearTerms,
  f21 / (g10 - g20) /. % /. simpleLinearTerms,
  g11 / (f10 - f20) /. % /. simpleLinearTerms,
  g21 / (f10 - f20) /. % /. simpleLinearTerms} // Simplify
```

Out[101]=
$$\left\{l1\left(s1-\frac{l2\,(s1+s2)}{\sqrt{1+4\,l1\,l2}}\right),\ -l2\,s2+\frac{l1\,l2\,(s1+s2)}{\sqrt{1+4\,l1\,l2}},\right.$$
$$\left.l1\left(s1-\frac{l2\,(s1+s2)}{\sqrt{1+4\,l1\,l2}}\right),\ -l2\,s2+\frac{l1\,l2\,(s1+s2)}{\sqrt{1+4\,l1\,l2}}\right\}$$

Invasion probabilities

The QLE results don't fully capture the dynamics, so to calculate the invasion probability we use numerically solved haplotypes frequencies:

In[243]:=
```
secondOrderDels =
  secondOrder[{delp1g11, delp1g12, delp2g22, delp2g12} /. wSubs /. equilSubs /.
    equilSubs, {m12, m21, s1, s2, r}];
secondOrderFreqs[M12_, M21_, S1_, S2_, R_] :=
 NSolve[(secondOrderDels /. m12 → M12 /. m21 → M21 /. s1 → S1 /. s2 → S2 /. r → R) == 0 &&
   0 ≤ p1g11 ≤ 1 && 0 ≤ p1g12 ≤ 1 && 0 ≤ p2g22 ≤ 1 && 0 ≤ p2g12 ≤ 1 && p1g11 * p2g22 > 0,
  {p1g11, p1g12, p2g22, p2g12}, Reals, WorkingPrecision → 7]


branchinginv1[M12_, M21_, S1_, S2_, R_] :=

 With[{freqs = Flatten[secondOrderFreqs[M12, M21, S1, S2, R]]},
  ((1 - z1) + (1 - z2)) /.
   FindRoot[({Exp[-(mat[[1, 1]] (1 - z1) + mat[[1, 2]] (1 - z2))], Exp[-(mat[[2, 1]] (1 - z1) +
                       mat[[2, 2]] (1 - z2))]} /. wbar1 → p1Wm /. wbar2 → p2Wm /. wSubs /.
              m12 → M12 /. m21 → M21 /. s1 → S1 /. s2 → S2 /. r → R /. equilSubs /.
        equilSubs /. freqs) == {z1, z2}, {{z1, 0}, {z2, 0}}]]

branchinginv1weighted[M12_, M21_, S1_, S2_, R_] :=

 With[{freqs = Flatten[secondOrderFreqs[M12, M21, S1, S2, R]]},
  ((1 - z1) * p1g11 + (1 - z2) * p2g11) /.
     FindRoot[({Exp[-(mat[[1, 1]] (1 - z1) + mat[[1, 2]] (1 - z2))],
                    Exp[-(mat[[2, 1]] (1 - z1) + mat[[2, 2]] (1 - z2))]} /. wbar1 → p1Wm /.
                 wbar2 → p2Wm /. wSubs /. m12 → M12 /. m21 → M21 /. s1 → S1 /.
              s2 → S2 /. r → R /. equilSubs /. equilSubs /. freqs) ==
       {z1, z2}, {{z1, 0}, {z2, 0}}] /. equilSubs /. freqs]
```