Chiasma interference, reinforcement and the evolution of chromosomal inversions

Øystein Kapperud

Master of Science Thesis

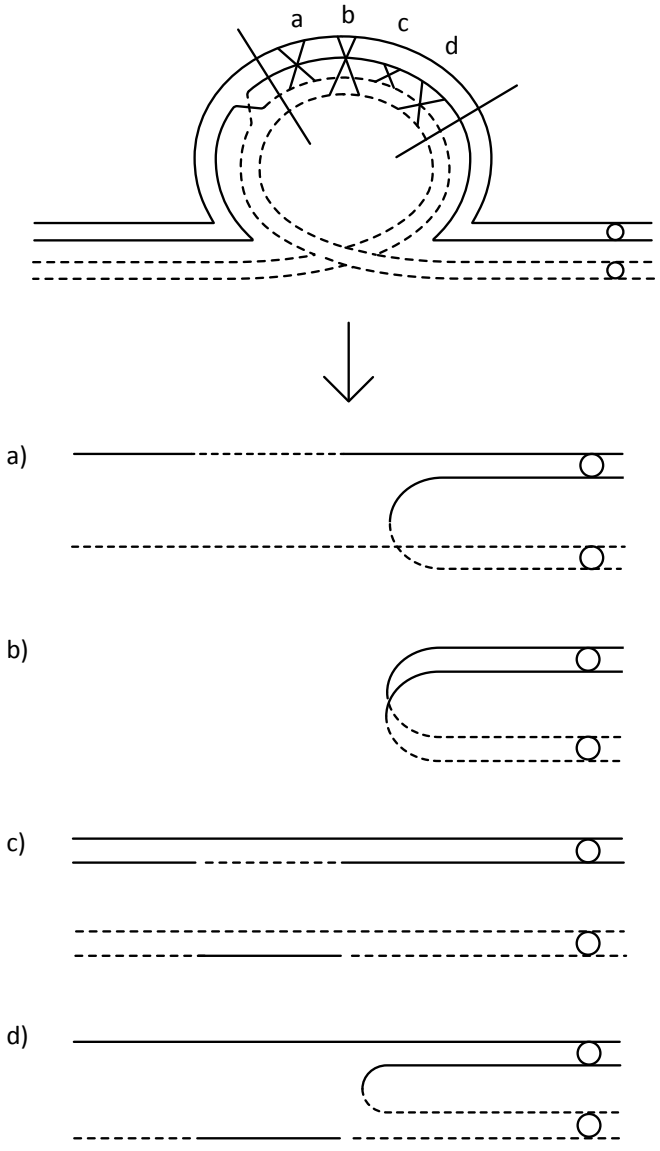
Centre for Ecological and Evolutionary Synthesis
Department of Biosciences

UNIVERSITY OF OSLO

15.12.2018

Chiasma interference, reinforcement and the evolution of chromosomal inversions

Øystein Kapperud



2018

Chiasma interference, reinforcement and the evolution of chromosomal inversions

Øystein Kapperud

Title page illustration by Sofie Ensby Rostad

http://www.duo.uio.no

Print: Reprosentralen, Universitetet i Oslo

Abstract

I present a model of chiasma interference and recombination in homo- and heterokaryotypes and a model of nonrandom mating. These two models form the basis of a general-purpose deterministic multilocus evolution simulator, which I use to show how chromosomal inversions can enhance and accelerate reinforcement in various scenarios. I conclude that an inversion can initiate a snowball effect of increasing pre- and postzygotic isolation, even when there is some recombination in heterokaryotypes.

Acknowledgements

Figures 2.2-2.3 and 2.6-2.9 were digitized by Sofie Ensby Rostad.

I thank Arcadi Navarro, Maria Servedio, Georg Kapperud, Øistein Holen, Mark Ravinet, Anna Runemark, Fabrice Eroukhmanoff and Glenn-Peter Sætre for stimulating exchanges and discussions; and my supervisors, Fabrice Eroukhmanoff and Glenn-Peter Sætre, for encouraging feedback and sound advice.

Table of contents

1 Introduction

When hybrids of two imperfectly isolated species have reduced fitness, there can be selection for preference alleles that cause their bearer to prefer mates with conspecific traits (Dobzhansky 1937). Once a contentious topic, such *reinforcement* of prezygotic isolation has in later decades been amply documented in the wild, and it is now widely accepted that it occurs (Noor 1999, Servedio and Noor 2003). A key insight of Felsenstein (1981) is that recombination between trait indicator loci and postzygotic isolation loci can make the conspecific trait a less reliable indicator of a mate's fitness, so that the advantage of preferring it is diminished. Accordingly, theoretical studies have shown that reinforcement occur less readily when recombination between pre- and postzygotic isolation loci is high (Felsentein 1981, Servedio and Kirkpatrick 1997, Servedio 2000, Servedio and Sætre 2003).

Chromosomal inversions partly suppress recombination in heterokaryotypes (e.g. Coyne et al. 1991, 1993, Navarro and Ruiz 1997, Jaarola et al. 1998) and are often found to be linked to loci involved in differentiation of alternative mating strategies (Tuttle et al. 2016, Lamichhaney et al. 2016, Wang et al. 2013), local adaptation (Etges and Levitan 2004, Sinclair-Waters et al. 2018) and pre- and postzygotic reproductive isolation between species or subpopulations (Noor et al. 2001, Feder et al. 2003, Ayala et al. 2013, Poelstra et al. 2014). For these reasons, it has been suggested that inversions can enhance reinforcement by capturing, and reducing recombination between, pre- and postzygotic isolation loci in parapatry (Tricket and Butlin 1994, Dagilis and Kirkpatrick 2016). Models of this process that incorporate realistic interactions between underdominance and recombination in heterokaryotypes have, however, not yet been explored.

Although the effects of chiasma interference on underdominance and recombination in inversion heterokaryotypes have been recognized (Navarro et al. 1997), general mathematical expressions of gamete proportions are lacking. In order to fill this gap, I will in chapter 2 of this thesis suggest a generalization of the *counting models* of chiasma interference, and use it do derive exact expressions of gamete proportions in homokaryotypes and inversion heterokaryotypes for two types of chromosomal inversions. This model and a model of nonrandom mating (chapter 3) together form the basis of a general-purpose deterministic multilocus evolution simulator that automatically generates and numerically solves appropriate recurrence equations based on the given input. The program is quite general, and can in principle be used to calculate the equilibrium state of any system that can be expressed in terms of an unbounded set of subpopulations, chromosomes and loci, with any pattern of migration between subpopulations, any environmental or genotypic fitness interactions, and any pattern of mating that corresponds to the model in chapter 3[1]. In chapter 4, I will use it to show how chromosomal inversions can enhance and accelerate reinforcement in a variety of scenarios. I conclude (chapter 5) that an inversion can initiate a snowball effect of increasing postzygotic (Navarro and Barton 2003) as well as prezygotic isolation, even when there is some recombination in heterokaryotypes.

[1] I aim to publish an extended version of this program with a full user manual at some later time, but I have made the source code for the current version available as an online appendix, along with an example input file (online appendix and Appendix B, this volume). Note that although all the operations needed to run the simulations discussed in this text have been extensively tested, some of the remaining ones should be regarded as being at the beta stage at this moment.

2 Recombination in homo- and heterokaryotypes under a general model of chiasma interference

2.1 Recombination and interference: terminology and introduction

The effect of chiasma interference on recombination in homokaryotypes is extensively discussed in the literature (section 2.13), but the results are rarely, if ever, incorporated into evolutionary models. For inversion heterokaryotypes, little work has been done apart from Navarro et al.'s (1997) approximate expressions of recombination rates for a maximum of two loci in restricted locations. In this chapter, I will suggest a model of chiasma interference that unifies the various *counting interference models* (section 2.13) into a single framework, and use it to derive exact expressions for gamete proportions in homo- and heterokaryotypes for an indefinite number of loci at any location on the chromosome.

2.1.1 Basic terminology

Since discussions of recombination and related issues risk being obfuscated by inconsistent and ambiguous terminology, I will begin this chapter by carefully laying out my own. A *tetrad* is a bundle of four *chromatids*, originating as a duplication of each of the two *parental homologues*, with one *pair* of chromatids denoted *sister chromatids* if they are derived from the same homologue and *non-sister chromatids* if they are not. Exchanges of genetic material – or *crossing over* – occur as a series of *chiasmata* (singular *chiasma*) or *chiasma events* distributed along the tetrad according to a model of *chiasma interference,* so that each such event *involves* two *non-sister chromatids,* or, equivalently, one out of the four possible *non-sister chromatid pairs*, chosen according to a model of *chromatid interference*.[2] A chromatid is said to *be recombinant* or *show recombination* in a given interval if it is *involved* in an odd number of chiasma events within that interval (note that this is not the same as saying that an odd number of chiasma events occur within the interval), and to *be non-recombinant* or show *non-recombination* in the opposite case. A *recombination pattern* is a set of Boolean variables that represent the *recombination status* – recombination (1) or non-recombination (0) – in each of the $\mathfrak{n}$ adjoining and non-overlapping intervals, so that if the set of all intervals – the *region of interest* – is $\{\mathbb{i}_0, \mathbb{i}_1, \mathbb{i}_2 \dots \mathbb{i}_{\mathfrak{n}-1}\}$, then we can denote a recombination pattern *r* as

$$\boldsymbol{r} = \{r(\mathbb{i}_0), r(\mathbb{i}_1), r(\mathbb{i}_2) \dots r(\mathbb{i}_{\mathfrak{n}-1})\}$$

where

$$r(\mathbb{i}) = \begin{cases} 1, & recombination\ in\ \mathbb{i} \\ 0, & non-recombination\ in\ \mathbb{i} \end{cases}$$

[2] This is sometimes referred to as the *four-strand model of recombination*, to distinguish it from the simplified *two-strand* (sometimes *one-strand*) *model* that is implicit in most textbook accounts (see Speed 1995 for a discussion of these two models). *Strand* is here synonymous with chromatid.

Accordingly, a chromatid is said to *show* or *have* recombination pattern $\boldsymbol{r}$ if the presence or absence of recombination in all the intervals of the region of interest correspond to $\boldsymbol{r}$. I will use the random vector $\boldsymbol{R} = \{R(𝕚_0), R(𝕚_1), R(\,𝕚_2)\,\ldots R(𝕚_{𝕟-1})\}$ to represent the recombination pattern of a randomly chosen chromatid, so that I can write e.g. "the probability of observing a chromatid with recombination pattern $\boldsymbol{r}$ is p" as

$$\Pr\{\boldsymbol{R} = \boldsymbol{r}\} = p$$

or "the probability of observing recombination status $r(𝕚)$ in interval $𝕚$ is p" as

$$\Pr\{R(𝕚) = r(𝕚)\} = p$$

or "the probability of observing recombination in interval $𝕚$ is p" as

$$\Pr\{R(𝕚) = 1\} = p$$

For notational simplicity, I will sometimes use the shorthand forms $R_k = R(𝕚_k)$ and $r_k = r(𝕚_k)$ when discussing recombination for a set of indexed intervals.

Meiosis results in a set of haploid *gametes* or *products of meiosis*. For *inversion homokaryotypes* – chromosomes that are not heterozygous for a chromosomal inversion – all chromatids have an equal chance of becoming a gamete, whereas this is not always the case for inversion heterokaryotypes, as we shall see. A gamete can be represented in two ways: as a recombination pattern (in which case I will say that the gamete show or have the recombination pattern in question), or as a haplotype – a set of alleles, one for each marker or loci that demarcate the boundaries between intervals. I will use the former representation in this chapter; an algorithm for converting to the latter for a given diploid genotype is given in the *Chromosome_diplotype* method *calculate_gamete_frequencies* in the main program (online appendix).

I will assign *directionality* to the region of interest by denoting interval $𝕚_0$ as the *the leftmost interval* and $𝕚_{𝕟-1}$ as the *rightmost interval*. More generally, I will say that any interval $𝕚_x$ is positioned to the *left* of interval $𝕚_y$ if $x < y$, and to *the right* of $𝕚_y$ if $x > y$. Furthermore, for any interval $𝕚_x$, I will for $0 < x < 𝕟 - 1$ denote the boundary that is shared with $𝕚_{x-1}$ as the *left boundary* of $𝕚_x$, and the boundary that is shared with $𝕚_{x+1}$ as the *right boundary* of $𝕚_x$. In the special case of $𝕚_0$, I will refer to the boundary that is shared with $𝕚_1$ as the *right boundary* of $𝕚_0$, and the boundary that is not shared with any other interval as the *left boundary* of $𝕚_0$, which is also the *left boundary* of the whole region of interest, or *the leftmost boundary*. Similary, in the special case of $𝕚_{𝕟-1}$ we have that the *left boundary* is the one shared with $𝕚_{𝕟-2}$, and the *right boundary,* a.k.a *rightmost boundary*, is the one not shared with any other interval. The directionality also applies to events within and between intervals, so that I will say, for example, that a chiasma event occur to the *left* of another chiasma event if the former occur closer to the *left* boundary within the same interval or in an interval closer to the *leftmost* boundary, and vice versa. If the region of interest is defined as in the preceding, then the *reversed region of interest* is the same region redefined so that the meaning of *left* and *right* is interchanged. A particular calculation, algorithm or analysis is then said to be *direction reversible* if it gives the same result regardless of whether it is performed on the region of interest or on the reversed region of interest.

2.1.2 Chromatid interference

To get from a model of *chiasma* interference – determining how the chiasma events are distributed in space – to recombination pattern probabilities, one needs to assume a model of *chromatid* interference – determining the chromatid involvement probabilities for each event. If a given chromatid pair is *less* likely to be involved in a chiasma event if it was involved in a neighboring event, we say that there is *positive chromatid interference,* and in the opposite case (*more* rather than *less* likely) we say that there is *negative chromatid interference*. If, on the other hand, the involvement probabilities are independent, we say that there is *no chromatid interference* or *independent chromatid involvement.* The evidence for chromatid interference is inconsistent and ambiguous (Zhao et al. 1995b), and almost all models, partly for this reason and partly for reasons of simplicity, assume independent involvement. I will here do so as well, though I will in the final section suggest a simple way to extend one of my algorithms so as to account for chromatid interference.

2.1.3 Chiasma interference

A *renewal process* (see Ross 2014, chapter 7) is in the following defined as a stochastic process that represents the number of events (of a given type) that occur in an interval of a given length, when the distances between neighboring events are all drawn from the same *interarrival distance distribution*. I will say that the events *occur independently* or *are independent* if the probability of observing a given number of events in any given interval is independent of the number of events in all other disjoint intervals. The renewal process that possesses this attribute is known as the *Poisson process* (Ross 2014, definition 5.2, theorem 5.1, proposition 5.1 and section 5.2.2), in which the interarrival distance distribution is exponential, and the number of events in an interval of a given (genetic) length follows a Poisson distribution, so that

$$\Pr\{\text{y } events\ in\ interval\ \mathbb{i}\} = \frac{e^{-\mu}\mu^{y}}{y!}$$

where μ is the expected (i.e. average) number of events in interval $\mathbb{i}$. If the chiasma events occur independently, they are distributed according to a Poisson process, and we say that there is *no chiasma interference.* This seems to be the case in some organisms, e.g. *Schizosaccharomyces pombe* (Munz 1994), but these are exceptional (see Berchowitz and Copenhaver 2010 for a review). I will refer to the model that assume no chiasma interference as *the Poisson interference model*, as first described in Haldane (1919). If chiasma events do not occur independently – i.e. if the probability of observing a given number of chiasma events in one interval depends on the number of chiasma events in a disjoint interval – we say that there is *chiasma interference.* A common measure of the degree of interference is the *coefficient of coincidence* (Muller 1916, Foss et al. 1993), one version of which is defined as

$$C(L_0, L_1) = \frac{\Pr\{R(\mathbb{i}_0) = 1, R(\mathbb{i}_1) = 1\}}{\Pr\{R(\mathbb{i}_0) = 1\}\Pr\{R(\mathbb{i}_1) = 1\}}$$

where $\mathbb{i}_0$ and $\mathbb{i}_1$ are two adjoining intervals of genetic length L_0 and L_1 (This is equivalent to S_3 in Foss et al. 1993, except that they impose the additional constraint that $L_0 = L_1$). Note that if the chiasma events occur independently, then $\Pr\{R(\mathbb{i}_0) = 1, R(\mathbb{i}_1) = 1\} = \Pr\{R(\mathbb{i}_0) = 1\}\Pr\{R(\mathbb{i}_1) = 1\}$ and $C(L_0, L_1) = 1$ for all values of L_0 and L_1. It is common to distinguish between *positive chiasma interference,* in which a chiasma in one location impede (i.e. makes less likely) the generation of a

chiasma in a nearby location ($C(L_0, L_1)$*<1* for small L_0, L_1), and *negative chiasma interference,* in which a chiasma in one location facilitates (i.e. makes more likely) the generation of a chiasma in a nearby location ($C(L_0, L_1)$*>1* for small L_0, L_1). In practice, chiasma interference is almost always positive and gradually decreasing with distance, i.e. $C(L_0, L_1)$ is typically close to zero for small L_0, L_1 and approaches 1 as L_0, L_1 are increased (Berchowitz and Copenhaver 2010, Foss et al. 1993).

The existence of chiasma interference implies that the information that a chiasma has occurred at one location on a chromosome must somehow propagate to nearby locations where it interferes with (i.e. influences the probability of) the generation of other chiasmata. I will refer to this information as *the interference signal.* Exactly how the interference signal manifests itself physically is poorly understood (Hillers 2004, Berchowitz and Copenhaver 2010); suggestions include a hypothetical polymer that grows out from each chiasma (King and Mortimer 1990), and the build-up and release (at chiasma locations) of physical stress along the chromosome (Kleckner et al. 2004, Wang et al. 2015). Foss et al. (1993) suggests a model, henceforth referred to as the *pure counting model*, where *intermediate events* occur independently (i.e. according to a Poisson process) along the chromosome, but each such event is subsequently resolved as either a chiasma event – with both crossing over and gene conversion – or what I will refer to as a *dummy event* – with gene conversion but not crossing over – in a strict sequence so that there are always m dummy events between each pair of chiasma events). Letting *C* denote an intermediate event, C_x a chiasma event and C_0 a dummy event, the counting model hence postulate that for *m = 2* the sequence …$C_xC_0C_0C_xC_0C_0C_xC_0C_0$… will be repeated along the span of the region of interest. This implies a chiasma interarrival distance distribution equal to the sum of *m+1* exponential distributions (m dummy events plus one chiasma event); in the literature on chiasma interference, this distribution is variously referred to as the Gamma distribution (e.g. Cobbs 1978, McPeek and Speed 1995), the chi-squared distribution (e.g. Zhao et al. 1995a), and the Erlang distribution (Nolan 2017). Interference in the pure counting model depends on genetic, as opposed to physical, distance and is always positive (Lange et al. 1997) and stronger for higher *m* (Foss et al. 1993, Navarro et al. 1997; see figure 4.8, chapter 4, in this text). Note that the chiasma events occur independently if *m = 0*, because then all (independent) intermediate events are resolved as chiasma events. The Poisson interference model, in which there is no chiasma interference, can therefore be thought of as a special case of the counting model.

There are two ways of interpreting the pure counting model; either as a literal description of the physical manifestation of the interference signal – a "machine that can count", in the words of Foss and Stahl (1995) – or as a mathematical abstraction that, disregarding gene conversions, models the distribution of chiasma events without making any claims as to how interference actually works. The latter interpretation is foreshadowed in the mathematically equivalent models of Cobbs (1978) and Stam (1979), among others (see McPeek and Speed 1995 for a brief historical overview). Foss et al. (1993), however, clearly favor the former interpretation, and they support this view by presenting a remarkably good fit between the predicted and observed coefficients of coincidence for *Drosophila* and *Neurospora* when *m,* crucially, is independently estimated from the ratio of the number of gene conversions to the number of chiasmata in a different dataset. Additional evidence for the accuracy of the pure counting model's prediction of the distribution of chiasma events in at least some species is given in Lande and Stahl (1993), Zhao et al. (1995), and McPeek and Speed (1995). One prediction of the literal "machine that can count" interpretation – that a region enclosed by two chiasma events will be enriched for gene conversions without crossing over – is, however, not fulfilled in *Saccharomyces cerevisiae* (Foss and Stahl 1995). Stahl et al. (2004) and Malkova et al. (2004) suggest that this discrepancy is due to the presence of two distinct and independent types of chiasmata, with and without interference, in *S. cerevisiae* and some other organisms, and they provide evidence to that effect (see also Berchowitz and Copenhaver 2010). I will refer to the chiasma interference model that incorporates the none-interference type in addition to the

interference (counting) type of chiasma events as the *two-pathway counting model.* This model has proved a significantly better fit to data from e.g. *Arabidopsis thaliana* (Copenhaver et al. 2002) and humans (Housworth and Stahl 2003), compared to the pure counting model. Another suggested generalization, anticipated in Foss et al. (1993), is the *Poisson-skip model* of Lange et al. (1997), in which the number of dummy events between each chiasma event is drawn from a probability distribution. Among the advantages of this model is that allows for more fine-tuned modelling of interference strengths, and that it allows for negative as well as positive chiasma interference. For clarity, I will henceforth refer to the Poisson-skip model as the *stochastic*[3] *counting model*.

In keeping with the deficiency of evidence for any of the proposed theories (Hillers 2004, Berchowitz and Copenhaver 2010), I will in this text remain agnostic about the physical manifestation of the interference signal and, disregarding gene conversions altogether, treat the dummy events as useful mathematical abstraction that may or may not actually correspond to sites of gene conversion without crossing over. On that basis, I will now suggest a further generalization of the four models introduced in this section – the *Poisson model, the pure counting model, the two-pathway counting model* and the *stochastic counting model* – which I will refer to as the *general counting model* or just *the general model* for short.

2.1.4 Outline of a general counting model of chiasma interference

As in the two-pathway counting model, I postulate two mutually independent chiasma-generating pathways. The *type I pathway* is without chiasma interference, and generate only *type I chiasma events* according to a Poisson process. The *type II pathway* is (potentially) with interference, and generate *type II intermediate events* according to a Poisson process; these are subsequently resolved as either *type II chiasma events* or *type II dummy events,* where the number of the latter following each instance of the former is drawn from a user-defined probability distribution. That is, the type II chiasma events occur according to the stochastic counting model. It will be convenient to give the different events symbols for easier reference, in particular when these have to be incorporated into mathematical expression. In the following I will denote the type I chiasma events, type II chiasma events, type II dummy events, and type II intermediate events as X', X'', O'', and C'', respectively (note that the C'' events comprises the union of all X'' and O'' events.). The union of all chiasma events of either type will be denoted X, and the union of all events – the *intermediate events* (without the *type II* qualifier) – will be denoted C. The same symbols with a subscript indicating an interval will serve as random variables representing the number of the event in question within that interval. Hence, I can, for example, write "the probability of observing one or more type I chiasma events in interval $\mathbb{i}$ is p" as $\Pr\{X'_{\mathbb{i}} > 0\} = p$, or "the expected (i.e. average) number of type II intermediate events in interval $\mathbb{i}$ is λ" as $E[C''_{\mathbb{i}}] = \lambda$.

In the homokaryotype case, the general model can be described by three user-defined sets of parameters. $\boldsymbol{\mu} = \{\mu_0, \mu_1, \mu_2, \dots, \mu_{\mathfrak{n}-1}\}$ and $\boldsymbol{\lambda} = \{\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{\mathfrak{n}-1}\}$ are the expected number of type I chiasma events and type II intermediate events, respectively, in each of the $\mathfrak{n}$ intervals, so that $\mu_k = E\left[X'_{\mathbb{i}_k}\right]$, $\lambda_k = E\left[C''_{\mathbb{i}_k}\right] for\ k = 0,1,2 \dots \mathfrak{n}-1$. Keep in mind that the $X'_{\mathbb{i}}$ and $C''_{\mathbb{i}}$ events are both Poisson distributed. The third set of parameters is the distribution of probabilities for observing a given number of consecutive O'' events following a X'' event. I will call this distribution the *intervening O''events distribution,* and denote it $\boldsymbol{\gamma} = \{\gamma_0, \gamma_1, \gamma_2 \dots \gamma_m\}$, so that $\gamma_q, for\ q = 0,1,2 \dots m,\ \sum_{q=0}^{m} \gamma_q = 1$, gives the probability of observing q consecutive O'' events following a X'' events, or, equivalently, q intervening O'' events between a pair of X'' events. One *cycle* hence consists of one X'' event and q

[3] The terms *deterministic* and *stochastic* are in this chapter used to differentiate between models based on probability distributions that respectively do have or do not have the full probability mass ($= 1$) distributed to a single value. All these models are deterministic in the sense that no random number generators are involved.

O'' events, where q is drawn from $\boldsymbol{\gamma}$, so that the expected total number of C'' events in a cycle is $\sum_{q=0}^{m}(q+1)\gamma_q$. I will assume that there is a user-defined finite upper limit, m, on the possible number of intervening O'' events, so that $\gamma_q = 0\ for\ q > m$. These three sets of parameters unambiguously determine the genetic lengths of each interval in units of *Morgan (*or *centiMorgan*), as I will show in section 2.1.6. Note that the general counting model is equivalent to the stochastic counting model in the special case where $\mu_k = 0\ for\ k = 0,1,2 \ldots \mathfrak{n} - 1$, i.e. when there are no type I chiasma events; and to the two-pathway counting model in the special case where $\gamma_m = 1, \gamma_q = 0\ for\ q \neq m$, i.e. when there are strictly m intervening O'' events between all consecutive pairs of X'' events. The parameters for the general model in homokaryotypes is summarized in table 2.1; figure 2.1 presents the general model, the two-pathway counting model, the stochastic counting model, the pure counting model, and the Poisson model as a model hierarchy where each downwards arrow pointing from model *a* to model *b* indicate which parameters you have to restrict (and how) in model *a* to get model *b* as a special case. I will in this text sometimes refer to all these collectively as *the counting models*, and use the term *pure counting model* when referring specifically to the model suggested by Foss et al. (1993).

Symbol	Short description
m	The maximum number of O'' events between any two X'' events (implicit in $\boldsymbol{\gamma}$)
$\boldsymbol{\gamma}$	The intervening O''events distribution
$\boldsymbol{\mu}$	The expected number of X'events for each interval
$\boldsymbol{\lambda}$	The expected number of C''events for each interval

Table 2.1: Input parameters in the general interference model.

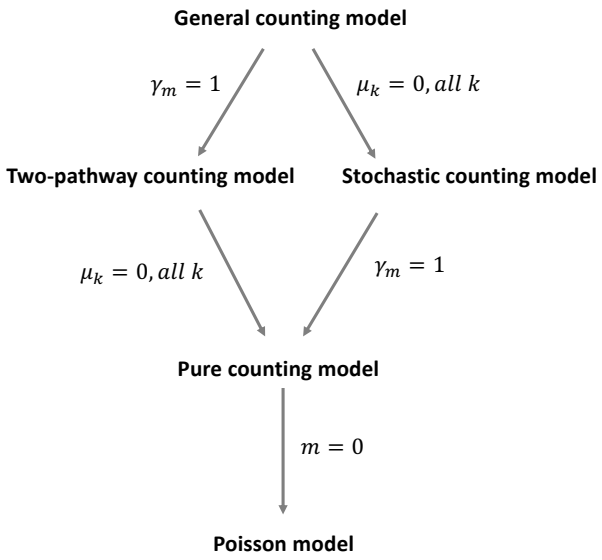


Figure 2.1: The hierarchy of counting models.

2.1.5 Stationarity in the general model

In the following, I will define the *phase* at any location on a tetrad as the number of O'' events between that location and the nearest X'' event to the *right*. To make this less abstract, you can imagine moving along the tetrad from left to right and keeping track of the phase by counting

downwards for each O'' event you encounter. When you reach 0, you know that the next C'' event will be a X'' event, and when you pass it, you immediately draw a number between 0 and m according to the intervening O'' event distribution ($\boldsymbol{\gamma}$), which gives you the number of O'' event between your current position and the next X'' event to the right. You then proceed by counting downwards from that number, until you reach the next X'' event and draw a new number of intervening O'' events, and so on. I will furthermore define $Q_{\mathbb{i}^l}$ and $Q_{\mathbb{i}^r}$ as random variables representing the phase at the left and right boundary, respectively, of interval $\mathbb{i}$, and more generally Q_a as a random variable representing the phase at a location a. A *phase distribution* is now a probability distribution that gives the probabilities of observing the individual possible phases at a given location. The type I events do not affect the phase, so they will be ignored in this section.

Now observe that if we define the vectors[4]

$$\boldsymbol{\pi}_{\mathbb{i}^l} = (\Pr\{Q_{\mathbb{i}^l} = 0\} \quad \Pr\{Q_{\mathbb{i}^l} = 1\} \quad \Pr\{Q_{\mathbb{i}^l} = 2\} \quad \dots \quad \Pr\{Q_{\mathbb{i}^l} = m\})$$
$$\boldsymbol{\pi}_{\mathbb{i}^r} = (\Pr\{Q_{\mathbb{i}^r} = 0\} \quad \Pr\{Q_{\mathbb{i}^r} = 1\} \quad \Pr\{Q_{\mathbb{i}^r} = 2\} \quad \dots \quad \Pr\{Q_{\mathbb{i}^r} = m\})$$

as the phase distributions at the left and right boundary, respectively, of interval $\mathbb{i}$, then

$$\boldsymbol{\pi}_{\mathbb{i}^r} = \sum_{c=0}^{\infty} \frac{e^{\lambda}\lambda^c}{c!} \boldsymbol{\pi}_{\mathbb{i}^l} \boldsymbol{P}^c$$

where $\boldsymbol{P}$ is an *m+1, m+1* matrix with (zero-indexed) element *[i,j]* given by

$$\boldsymbol{P}[i,j] = \begin{cases} 1, & i \neq 0; j = i-1 \\ \gamma_j, & i = 0; \\ 0, & otherwise; \end{cases} \qquad for\ i,j = 0,1,2 \dots m$$

and

$$\lambda = E[C_{\mathbb{i}}''],$$

For example, if *m=4*, then

$$\boldsymbol{P} = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$\boldsymbol{P}$ is a transition matrix that transforms the phase distribution at a given location a into the phase distribution at a location b, when there is one C'' event between a and b, and $\boldsymbol{P^c}$ (the matrix multiplication of c instances of $\boldsymbol{P}$) is the transition matrix when there are c C'' events between a and b. Hence,

$$(\Pr\{Q_{\mathbb{i}^r} = 0 | C_{\mathbb{i}}'' = c\} \quad \Pr\{Q_{\mathbb{i}^r} = 1 | C_{\mathbb{i}}'' = c\} \quad \Pr\{Q_{\mathbb{i}^r} = 2 | C_{\mathbb{i}}'' = c\} \quad \dots \quad \Pr\{Q_{\mathbb{i}^r} = m | C_{\mathbb{i}}'' = c\}) = \boldsymbol{\pi}_{\mathbb{i}^l} \boldsymbol{P}^c$$

and by conditioning on the number of C'' events in $\mathbb{i}$,

[4] See Appendix A for a note on vector and matrix notation

$$\boldsymbol{\pi}_{\mathbb{I}^r} = \sum_{c=0}^{\infty} \frac{e^{\lambda}\lambda^c}{c!}\boldsymbol{\pi}_{\mathbb{I}^l}\boldsymbol{P}^c$$

which we can interpret as a uniform continuous Markov chain with imbedded transition matrix $\boldsymbol{P}$. I will now define *the stationary phase distribution*, $\boldsymbol{\pi} = (\pi_0 \quad \pi_1 \quad \pi_2 \quad \dots \quad \pi_m),\ \sum_{q=0}^{m}\pi_q = 1,$ as the phase distribution that satisfy the equation

$$\boldsymbol{\pi} = \sum_{c=0}^{\infty} \frac{e^{\lambda}\lambda^c}{c!}\boldsymbol{\pi}\boldsymbol{P}^c$$

for all $\lambda \in \mathbb{R}_{\geq 0}$. That is, the $\boldsymbol{\pi}$ for which $\boldsymbol{\pi}_{\mathbb{I}^r} = \boldsymbol{\pi}_{\mathbb{I}^l} = \boldsymbol{\pi}$ regardless of the size of the interval. But if $\boldsymbol{\pi}_{\mathbb{I}^l} = \boldsymbol{\pi}_{\mathbb{I}^r}$ is true after one transition, then, by induction, it is true after any number of transitions, so

$$\boldsymbol{\pi} = \sum_{c=0}^{\infty} \frac{e^{\lambda}\lambda^c}{c!}\boldsymbol{\pi}\boldsymbol{P}^c \rightarrow \boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P}$$

which means we can find $\boldsymbol{\pi}$ by solving the equation set

$$\begin{gathered}\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{P} \\ \sum_{q=0}^{m}\pi_q = 1\end{gathered}$$

(the latter is true because the phase probabilities must sum to 1.) This gives

$$\begin{aligned}\pi_0 &= \pi_0\gamma_0 + \pi_1 \\ \pi_1 &= \pi_0\gamma_1 + \pi_2 \\ \pi_2 &= \pi_0\gamma_2 + \pi_3 \\ &\dots \\ \pi_{m-2} &= \pi_0\gamma_{m-2} + \pi_{m-1} \\ \pi_{m-1} &= \pi_0\gamma_{m-1} + \pi_m \\ \pi_m &= \pi_0\gamma_m\end{aligned}$$

Or, with some algebra,

$$\begin{aligned}\pi_m &= \pi_0\gamma_m \\ \pi_{m-1} &= \pi_0(\gamma_m + \gamma_{m-1}) \\ \pi_{m-2} &= \pi_0(\gamma_m + \gamma_{m-1} + \gamma_{m-2}) \\ &\dots \\ \pi_2 &= \pi_0(\gamma_m + \gamma_{m-1} + \gamma_{m-2} + \cdots + \gamma_2) \\ \pi_1 &= \pi_0(\gamma_m + \gamma_{m-1} + \gamma_{m-2} + \cdots + \gamma_2 + \gamma_1) \\ \pi_0 &= \pi_0(\gamma_m + \gamma_{m-1} + \gamma_{m-2} + \cdots + \gamma_2 + \gamma_1 + \gamma_0)\end{aligned}$$

(Note that $\sum_{q=0}^{m}\gamma_q = 1$.) By summing the above equations,

$$\sum_{q=0}^{m}\pi_q = 1 = \pi_0\sum_{k=0}^{m}\sum_{q=k}^{m}\gamma_q = \pi_0\sum_{q=0}^{m}(q+1)\gamma_q$$

so

$$\pi_0 = \frac{1}{\sum_{q=0}^{m}(q+1)\gamma_q}$$

and in general

$$\pi_k = \frac{\sum_{q=k}^{m}\gamma_q}{\sum_{q=0}^{m}(q+1)\gamma_q} \qquad for\ k = 0,1,2,3 \ldots m$$

This distribution is also derived (by a different argument) in Lange et al. (1997); the stationary phase distribution is the same in my model as in their stochastic counting model, because the type I events (that are lacking in the latter model) do not affect the phase.

The stationary phase distribution is important because if it is given that the phase distribution at any given location *a* corresponds to this distribution, then the phase distribution at all other locations *b* also corresponds to this distribution (if this is not clear, think of the left boundary of the interval 𝕚 in the derivation above as location *a* and the right boundary as location *b*). This means that if we *assume stationarity*, then all regions that are described by the same set of **λ, σ** and **γ** values are fungible, regardless of their position on the tetrad. It also means the *left* and *right* boundaries of any interval or region have the same phase distribution, which means that it does not matter which we call *left* and which we call *right.* This makes the calculation of gamete probabilities in homokaryotypes direction reversible, in the sense defined in the introduction.

Cobbs' (1978) and Stam's (1979) early analyses of (the mathematical equivalent of) the pure counting model were complicated by the assumption of a non-stationary phase distribution (though they both considered the stationary phase distribution as a special case), which were motivated by Mather's (1938) hypothesis that the process of chiasma generation starts at the centromere and proceed from there in both directions, implying that the phase at the centromere is non-stationary. This hypothesis was in turn motivated by early results indicating that the interference signal is blocked by the centromere, meaning that chiasma on one side of the centromere does not interfere with chiasma on the other. However, a more recent analysis by Colombo and Jones (1997) indicate that the results cited by Mather are merely statistical artefacts, and that, in contradiction to Mather's predictions, interference does in fact work across the centromere in the same way as in other regions. This leaves it unclear exactly how and where the process of chiasma generation starts; or as Colombo and Jones (1997, p. 226) put it, "If chiasma formations does not start from the centromere, or end against the centromere, where does it start from or end? Or, to put it bluntly, does it start from anywhere?" Given that we do not know the answers to these questions (Hillers 2004, Berchowitz and Copenhaver 2010), it seems to me that assuming a stationary phase distribution, as almost all more recent treatments of chiasma interference do, is as plausible and parsimonious as any other alternative. I will therefore do so in the rest of this text. The predictions of a stationary and non-stationary model are in any case not very different (Lande and Stahl 1993).

2.1.6 Genetic distances

The terms *Morgan* and *centiMorgan* are sometimes subject to confusion in the non-technical literature. The standard definition, as used in e.g. Lande and Stahl (1993), Foss et al. (1993), and Navarro and Ruiz (1997), is that an interval is of length *y* Morgans (100*y* centiMorgans) – or, equivalently, that the genetic distance between the two boundaries enclosing the interval is *y* Morgans – if a single randomly chosen chromatid is on average involved in *y* chiasma events within that interval (note that this is not the same as saying that *y* chiasma events occur within the

interval). Since a single chiasma event always strictly involves two out of four chromatids, we have that

$$Length\ of\ \mathbb{i}\ in\ Morgans = \frac{1}{2}E[X_{\mathbb{i}}] = \frac{1}{2}(E[X'_{\mathbb{i}}] + E[X''_{\mathbb{i}}])$$

The expected number of C'' events in a single cycle is $\sum_{q=0}^{m}(q+1)\gamma_q$. Since there per definition is always exactly one X'' event in a cycle,

$$E[X''_{\mathbb{i}}] = \frac{\lambda}{\sum_{q=0}^{m}(q+1)\gamma_q}$$

and accordingly

$$Length\ of\ \mathbb{i}\ in\ Morgans = \frac{1}{2}\left(\mu + \frac{\lambda}{\sum_{q=0}^{m}(q+1)\gamma_q}\right)$$

where

$$\mu = E[X'_{\mathbb{i}}], \lambda = E[C''_{\mathbb{i}}]$$

It will be convenient in this chapter to represent the length of the intervals as $E[X'_{\mathbb{i}}]$, $E[C''_{\mathbb{i}}]$ and $\boldsymbol{\gamma}$ without continuously converting to and from units of *Morgan*. You can use the equation above to make your own conversions whenever you see fit.

2.1.7 Mather's equation

The following useful expression was first proved by Mather (1938), and so is commonly referred to as *Mather's equation*:

$$\Pr\{R(\mathbb{i}) = 1|X_{\mathbb{i}} = x\} = \begin{cases} 1/2, & x = 1,2,3\ldots \\ 0, & x = 0 \end{cases}$$

and, since $\Pr\{R(\mathbb{i}) = 0|X_{\mathbb{i}} = x\} = 1 - \Pr\{R(\mathbb{i}) = 1|X_{\mathbb{i}} = x\}$,

$$\Pr\{R(\mathbb{i}) = 0|X_{\mathbb{i}} = x\} = \begin{cases} 1/2, & x = 1,2,3\ldots \\ 1, & x = 0 \end{cases}$$

That is, the probability of observing recombination in interval $\mathbb{i}$ given the number of chiasma events, *x,* in that interval is 0 if *x = 0* and $1/2$ if *x > 0*. It follows, as Mather also noted, that

$$\begin{aligned} \Pr\{R(\mathbb{i}) = 1\} &= \sum_{x=0}^{\infty} \Pr\{R(\mathbb{i}) = 1|X_{\mathbb{i}} = x\} \Pr\{X_{\mathbb{i}} = x\} \\ &= \sum_{x=1}^{\infty} \Pr\{R(\mathbb{i}) = 1|X_{\mathbb{i}} = x\} \Pr\{X_{\mathbb{i}} = x\} \\ &= \frac{1}{2}\sum_{x=1}^{\infty} \Pr\{X_{\mathbb{i}} = x\} \end{aligned}$$

$$= \frac{1}{2}\Pr\{X_{𝕚} > 0\}$$
$$= \frac{1}{2}(1 - \Pr\{X_{𝕚} = 0\})$$

and

$$\Pr\{R(𝕚) = 0\} = 1 - \frac{1}{2}(1 - \Pr\{X_{𝕚} = 0\}) = \frac{1}{2}(1 + \Pr\{X_{𝕚} = 0\})$$

For the benefit of your intuition, I will include here a simple proof of Mather's equation, different from the one given in Mather (1938), but similar to the one given in Cobbs (1978). I first postulate that

$$\Pr\{R(𝕚) = 1 | X_{𝕚} = x\} = \boldsymbol{v_0 P^x w} \quad for\ x = 0,1,2,3\ldots$$

where

$$\boldsymbol{v_0} = (1 \quad 0)$$
$$\boldsymbol{P} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$
$$\boldsymbol{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

To see why this equation is true, let $\{M_x\}_{x \in \mathbb{Z}_{\geq 0}}$ be a stochastic process that represent the recombination status of interval 𝕚 when there are *x* chiasma events in 𝕚. If there are no chiasma events in 𝕚, then all chromatid must have recombination status 0 (i.e. non-recombination), so

$$\boldsymbol{v_0} = (1 \quad 0) = (\Pr\{M_0 = 0\} \quad \Pr\{M_0 = 1\})$$

We can now interpret $\boldsymbol{P}$ as

$$\boldsymbol{P} = \begin{pmatrix} \Pr\{M_{x+1} = 0 | M_x = 0\} & \Pr\{M_{x+1} = 1 | M_x = 0\} \\ \Pr\{M_{x+1} = 0 | M_x = 1\} & \Pr\{M_{x+1} = 1 | M_x = 1\} \end{pmatrix} \quad for\ x = 0,1,2\ldots$$

i.e. $\boldsymbol{P}$ is the Markovian transition matrix for $\{M_x\}_{x \in \mathbb{Z}_{\geq 0}}$ where a transition correspond to a chiasma event. We can see that $\boldsymbol{P} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ because an additional chiasma event will involve a randomly chosen chromatid with probability $1/2$, and given that the chromatid is involved, its recombination status will *flip* (go from 0 to 1 or 1 to 0) with probability 1; and given that it is *not* involved, its recombination status will stay the same with probability 1. Hence,

$$(\Pr\{M_x = 0\} \quad \Pr\{M_x = 1\}) = \boldsymbol{v_0 P^x} \quad for\ x = 0,1,2\ldots$$

and

$$\Pr\{R(𝕚) = 1 | X_{𝕚} = x\} = \Pr\{M_x = 1\} = \boldsymbol{v_0 P^x w} \quad for\ x = 0,1,2\ldots$$

Since, as we can confirm by simple matrix multiplication, $\boldsymbol{P} = \boldsymbol{P}^2$, it follows that

$$\boldsymbol{P}^x = \boldsymbol{P}^{x-2}\boldsymbol{P}^2 = \boldsymbol{P}^{x-2}\boldsymbol{P} = \boldsymbol{P}^{x-3}\boldsymbol{P}^2 = \boldsymbol{P}^{x-3}\boldsymbol{P} = \cdots = \boldsymbol{P}^2 = \boldsymbol{P} \quad for\ x = 1,2,3 \ldots$$

so

$$\Pr\{R(\mathbb{i}) = 1 | X_{\mathbb{i}} = x\} = \begin{cases} \boldsymbol{v_0 P w}, & x = 1,2,3 \ldots \\ \boldsymbol{v_0 I w}, & x = 0 \end{cases}$$

or

$$\Pr\{R(\mathbb{i}) = 1 | X_{\mathbb{i}} = x\} = \begin{cases} 1/2, & x = 1,2,3 \ldots \\ 0, & x = 0 \end{cases}$$

QED

I am now in position to state the first theorem, which gives the recombination pattern probabilities for homokaryotypes under the general model.

2.2 Homokaryotypes

2.2.1 Theorem 1: Recombination in homokaryotypes

Assuming stationarity and no chromatid interference, the probability of observing recombination pattern $\boldsymbol{r}$ on a homokaryotypic chromosome under the general chiasma interference model is given by

$$\Pr\{\boldsymbol{R}=\boldsymbol{r}\}=\boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{n}-1}\boldsymbol{M}_{\mathbb{I}_k}(\boldsymbol{r})\right)\mathbf{1}^T$$

where

$$\boldsymbol{\pi}=(\pi_0 \quad \pi_1 \quad \pi_2 \quad \dots \quad \pi_m)$$

$$\pi_l=\frac{\sum_{q=l}^{m}\gamma_q}{\sum_{q=0}^{m}(q+1)\gamma_q} \quad for\ l=0,1,2\dots m$$

$$\boldsymbol{M}_{\mathbb{I}_k}(\boldsymbol{r})=\begin{cases}\frac{1}{2}\boldsymbol{G}_{\mathbb{I}_k}, & r(\mathbb{I}_k)=1\\ \frac{1}{2}\boldsymbol{G}_{\mathbb{I}_k}+\boldsymbol{H}_{\mathbb{I}_k}, & r(\mathbb{I}_k)=0\end{cases} \qquad for\ k=0,1,2\dots\mathfrak{n}-1$$

$$\boldsymbol{G}_{\mathbb{I}_k}[i,j]=\psi_k(i,j)e^{-\mu_k}+\left(\psi_k(i,j)+\delta_{\{i\geq j\}}\frac{\lambda_k^{i-j}e^{-\lambda_k}}{(i-j)!}\right)(1-e^{-\mu_k}) \quad for\ i,j=0,1,2\dots m$$

$$\psi_k(i,j)=\sum_{n=0}^{\infty}b_n\sum_{q=j}^{m}\gamma_q\frac{\lambda_k^{i+1+n+q-j}e^{-\lambda_k}}{(i+1+n+q-j)!}$$

$$b_n=\begin{cases}\sum_{q=0}^{n-1}b_q\gamma_{n-1-q}, & n=1,2,3\dots\\ 1, & n=0\end{cases}$$

$$\boldsymbol{H}_{\mathbb{I}_k}[i,j]=\begin{cases}\frac{\lambda_k^{i-j}e^{-\lambda_k}}{(i-j)!}e^{-\mu_k}, & i\geq j\\ 0, & i<j\end{cases} \quad for\ i,j=0,1,2\dots m$$

$$\delta_{\{condition\}}=\begin{cases}1, & condition\ is\ true\\ 0, & condition\ is\ false\end{cases}$$

$$\lambda_k=E[C_{\mathbb{I}_k}'']$$

$$\mu_k=E[X_{\mathbb{I}_k}']$$

Proof:
We can interpret element *i,j* of $\boldsymbol{H}_{\mathbb{I}_k}$ as

$$\boldsymbol{H}_{\mathbb{I}_k}[i,j]=\Pr\{X_{\mathbb{I}_k}=0, Q_{\mathbb{I}_k^r}=j|Q_{\mathbb{I}_k^l}=i\}$$

i.e., element *i,j* is equal to the probability of observing no chiasma events (of either type) in $\mathbb{i}_k$ *and* phase *j* at the right boundary, *given* phase *i* at the left boundary, which in the case $i \geq j$ is equivalent to the probability of observing *i* - *j* $C''_{\mathbb{i}_k}$ events multiplied by the probability of observing 0 $X'_{\mathbb{i}_k}$ events. If $i < j$, it is impossible to arrive at phase *j* without at least one X'' event in the interval, so the probability is in this case 0. We can similarly interpret element *i,j* of matrix $\boldsymbol{G}_{\mathbb{i}_k}$ as

$$\boldsymbol{G}_{\mathbb{i}_k}[i,j] = \Pr\{X_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j | Q^l_{\mathbb{i}_k} = i\}$$

and the function $\psi_k(i,j)$ as

$$\psi_k(i,j) = \Pr\{X''_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j | Q^l_{\mathbb{i}_k} = i\}$$

To see why this is so, first note that b_n, equivalent to u_n in Lange et al. (1997), gives the probability that the *n-th* C''event to the right of any X''event is also a X''event; as in Lange et al., it is derived by recursively conditioning on the index of the last X'' event before the *n*-th C'' event, with the base case $b_0 = 1$. Now note that to get at least one $X''_{\mathbb{i}_k}$ event and phase *j* at the right boundary, given phase *i* at the left boundary, one must observe in interval $\mathbb{i}_k$ (*i* O'' events to get to the first X''event) + (the first X'' event) + (*n* additional C''events of which the last is a X'' event) + (*q* - *j* additional O''events, where *q* is the number of O''events in the rightmost cycle of the interval, to end up in phase *j*), in total *i + 1 + n + q – j* C'' events. Hence, by summing over all possible values of *n* and *q*,

$$\psi_k(i,j) = \sum_{n=0}^{\infty} b_n \sum_{q=j}^{m} \gamma_q \frac{\lambda_k^{i+1+n+q-j} e^{-\lambda_k}}{(i+1+n+q-j)!} = \Pr\{X''_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j | Q^l_{\mathbb{i}_k} = i\}$$

which is similar to the $1_{\{j>0\}}$ term in equation 6 in Lange et al. (1997) (though note that my 'phase' is defined differently from their 'state').

By conditioning on the presence ($X'_{\mathbb{i}_k} > 0$) or absence ($X'_{\mathbb{i}_k} = 0$) of X' events in the interval,

$$\Pr\left\{X_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} = \Pr\left\{X''_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} \Pr\{X'_{\mathbb{i}_k} = 0\} + \Pr\left\{Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} \Pr\{X'_{\mathbb{i}_k} > 0\}$$

Since

$$\begin{aligned}\Pr\left\{Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} &= \Pr\left\{X''_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} + \Pr\left\{X''_{\mathbb{i}_k} = 0, Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} \\ &= \psi_k(i,j) + \delta_{\{i \geq j\}} \frac{\lambda_k^{i-j} e^{-\lambda_k}}{(i-j)!}\end{aligned}$$

and

$$\begin{aligned}\Pr\{X'_{\mathbb{i}_k} > 0\} &= (1 - e^{-\mu_k}) \\ \Pr\{X'_{\mathbb{i}_k} = 0\} &= e^{-\mu_k}\end{aligned}$$

we now have that

$$\Pr\left\{X_{\mathbb{i}_k} > 0, Q^r_{\mathbb{i}_k} = j \middle| Q^l_{\mathbb{i}_k} = i\right\} = \psi_k(i,j) e^{-\mu_k} + \left(\psi_k(i,j) + \delta_{\{i \geq j\}} \frac{\lambda_k^{i-j} e^{-\lambda_k}}{(i-j)!}\right)(1 - e^{-\mu_k}) = \boldsymbol{G}_{\mathbb{i}_k}[i,j]$$

For comparison, my $\boldsymbol{G}_{\mathbb{i}_k}$ and $\boldsymbol{H}_{\mathbb{i}_k}$ are for the general model what $\sum_{x=1}^{\infty} \boldsymbol{D}_x(y)$ and $\boldsymbol{D_0}(y)$ are for the pure counting model in Zhao et al. (1995a), what $\sum_{x=1}^{\infty} \boldsymbol{D}(x,m,p,y)$ and $\boldsymbol{D}(0,m,p,y)$ are for the two-

pathway counting model in Copenhaver et al. (2002) and what $P(t) - \tilde{Q}(t)$ and $\tilde{Q}(t)$ are for the stochastic counting model in Lange et al. (1997), with various notational differences.

From Mather's equation, we can deduce that

$$\Pr\left\{R_k = 1, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\} = \frac{1}{2}\Pr\left\{X_{\mathbb{I}_k} > 0, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\}$$

$$\Pr\left\{R_k = 0, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\} = \frac{1}{2}\Pr\left\{X_{\mathbb{I}_k} > 0, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\} + \Pr\left\{X_{\mathbb{I}_k} = 0, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\}$$

so

$$\Pr\left\{R_k = r_k, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\} = \boldsymbol{M}_{\mathbb{I}_k}(\boldsymbol{r})[i, j]$$

where

$$\boldsymbol{M}_{\mathbb{I}_k}(\boldsymbol{r}) = \begin{cases} \frac{1}{2}\boldsymbol{G}_{\mathbb{I}_k}, & r_k = 1 \\ \frac{1}{2}\boldsymbol{G}_{\mathbb{I}_k} + \boldsymbol{H}_{\mathbb{I}_k}, & r_k = 0 \end{cases} \qquad for\ k = 0,1,2 \ldots \mathfrak{n} - 1$$

If we assume stationarity, then

$$\Pr\left\{Q_{\mathbb{I}_0^l} = i\right\} = \pi_i \qquad for\ i = 0,1,2 \ldots m$$

so

$$\sum_{i=0}^{m} \Pr\left\{Q_{\mathbb{I}_0^l} = i\right\} \Pr\left\{R_0 = r_0, Q_{\mathbb{I}_0^r} = j | Q_{\mathbb{I}_0^l} = i\right\} = \sum_{i=0}^{m} \pi_i \Pr\left\{R_0 = r_0, Q_{\mathbb{I}_0^r} = j | Q_{\mathbb{I}_0^l} = i\right\}$$
$$= \Pr\{R_0 = r_0, Q_{\mathbb{I}_0^r} = j\} = \left(\boldsymbol{\pi}\boldsymbol{M}_{\mathbb{I}_0}(\boldsymbol{r})\right)[j]$$

Since $Q_{\mathbb{I}_k^r} = Q_{\mathbb{I}_{k+1}^l}$ we have

$$\begin{aligned}
\Pr\{R_0 = r_0, R_1 = r_1, Q_{\mathbb{I}_1^r} = j\} &= \sum_{i=0}^{m} \Pr\left\{R_0 = r_0, R_1 = r_1, Q_{\mathbb{I}_1^l} = i, Q_{\mathbb{I}_1^r} = j\right\} \\
&= \sum_{i=0}^{m} \Pr\left\{R_0 = r_0, Q_{\mathbb{I}_1^l} = i\right\} \Pr\left\{R_1 = r_1, Q_{\mathbb{I}_1^r} = j | R_0 = r_0, Q_{\mathbb{I}_1^l} = i\right\} \\
&= \sum_{i=0}^{m} \Pr\left\{R_0 = r_0, Q_{\mathbb{I}_1^l} = i\right\} \Pr\left\{R_1 = r_1, Q_{\mathbb{I}_1^r} = j | Q_{\mathbb{I}_1^l} = i\right\} \\
&= \sum_{i=0}^{m} \Pr\{R_0 = r_0, Q_{\mathbb{I}_0^r} = i\} \Pr\left\{R_1 = r_1, Q_{\mathbb{I}_1^r} = j | Q_{\mathbb{I}_1^l} = i\right\} \\
&= \left(\boldsymbol{\pi}\boldsymbol{M}_{\mathbb{I}_0}(\boldsymbol{r})\boldsymbol{M}_{\mathbb{I}_1}(\boldsymbol{r})\right)[j]
\end{aligned}$$

(Note that $\Pr\left\{R_1 = r_1, Q_{\mathbb{I}_1^r} = j | R_0 = r_0, Q_{\mathbb{I}_1^l} = i\right\} = \Pr\left\{R_1 = r_1, Q_{\mathbb{I}_1^r} = j | Q_{\mathbb{I}_1^l} = i\right\}$, because R_0 is irrelevant if the phase at the left boundary of $\mathbb{I}_1$ is known.)

Let S_1 be the statement $\Pr\{R_0 = r_0, R_1 = r_1, Q_{\mathbb{I}_1^r} = j\} = \left(\boldsymbol{\pi}\boldsymbol{M}_{\mathbb{I}_0}(\boldsymbol{r})\boldsymbol{M}_{\mathbb{I}_1}(\boldsymbol{r})\right)[j]$, and assume that the induction hypothesis S_{k-1} defined as

$$\Pr\{R_0 = r_0, R_1 = r_1, \dots, R_{k-1} = r_{k-1}, Q_{\mathbb{I}_{k-1}^r} = j\} = \left(\boldsymbol{\pi}\boldsymbol{M}_{\mathbb{I}_0}(r) \dots \boldsymbol{M}_{\mathbb{I}_{k-1}}(r)\right)[j]$$

is true. Now,

$$\Pr\{R_0 = r_0, R_1 = r_1, \dots, R_k = r_k, Q_{\mathbb{I}_k^r} = j\} = \sum_{i=0}^{m} \Pr\{R_0 = r_0, R_1 = r_1, \dots, R_k = r_k, Q_{\mathbb{I}_k^l} = i, Q_{\mathbb{I}_k^r} = j\}$$

$$= \sum_{i=0}^{m} \Pr\left\{R_0 = r_0, \dots, R_{k-1} = r_{k-1}, Q_{\mathbb{I}_k^l} = i\right\} \Pr\left\{R_k = r_k, Q_{\mathbb{I}_k^r} = j | R_0 = r_0, \dots, R_{k-1} = r_{k-1}, Q_{\mathbb{I}_k^l} = i\right\}$$

$$= \sum_{i=0}^{m} \Pr\left\{R_0 = r_0, \dots, R_{k-1} = r_{k-1}, Q_{\mathbb{I}_k^l} = i\right\} \Pr\left\{R_k = r_k, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\}$$

$$= \sum_{i=0}^{m} \Pr\{R_0 = r_0, \dots, R_{k-1} = r_{k-1}, Q_{\mathbb{I}_{k-1}^r} = i\} \Pr\left\{R_k = r_k, Q_{\mathbb{I}_k^r} = j | Q_{\mathbb{I}_k^l} = i\right\}$$

$$= \left(\boldsymbol{\pi}\boldsymbol{M}_{\mathbb{I}_0}(r) \dots \boldsymbol{M}_{\mathbb{I}_{k-1}}(r)\boldsymbol{M}_{\mathbb{I}_k}(r)\right)[j]$$

where the last equality follows from the induction hypothesis. Since S_1 is true and S_{k-1} implies S_k, is follows that S_{n-1} must be true. Hence,

$$\Pr\{\boldsymbol{R} = \boldsymbol{r}\} = \sum_{j=0}^{m} \Pr\{\boldsymbol{R} = \boldsymbol{r}, Q_{\mathbb{I}_{n-1}^r} = j\} = \boldsymbol{\pi}\boldsymbol{M}_{\mathbb{I}_0}(\boldsymbol{r})\boldsymbol{M}_{\mathbb{I}_1}(\boldsymbol{r})\boldsymbol{M}_{\mathbb{I}_2}(\boldsymbol{r}) \dots \boldsymbol{M}_{\mathbb{I}_{n-1}}(\boldsymbol{r})\mathbf{1}^T$$

QED

Nolan (2017) derives a closed-form expression for the recombination pattern probabilities under the counting model. Building on the results from that paper, I will in the next theorem do the same for the more general two-pathway counting model.

2.2.2 Theorem 2: A closed-form version of the *G* matrices

In the special case where $\gamma_m = 1, \gamma_q = 0 \; for \; q \neq m$ (i.e. the two-pathway counting model), the matrix $\boldsymbol{G}_{\mathbb{I}_k}$ can be written in closed form as

$$\boldsymbol{G}_{\mathbb{I}_k}[i,j] = \psi_k(i,j)e^{-\mu_k} + \left(\psi_k(i,j) + \delta_{\{i \geq j\}} \frac{\lambda_k^{i-j} e^{-\lambda_k}}{(i-j)!}\right)(1 - e^{-\mu_k})$$

where

$$\psi_k(i,j) = e^{-\lambda_k}\left(\delta_{\{i \geq j\}}\left[f_{i-j,m+1}(\lambda_k) - \frac{\lambda_k^{i-j}}{(i-j)!}\right] + \delta_{\{i<j\}} f_{m+1+i-j,m+1}(\lambda_k)\right)$$

$$f_{r,q}(\lambda) = \frac{1}{q}\sum_{j=0}^{q-1} e^{\lambda cos\left(\frac{2\pi j}{q}\right)} \cos\left[\lambda \sin\left(\frac{2\pi j}{q}\right) - \frac{2\pi r j}{q}\right] \quad for \; r = 0,1,2 \ldots q-1$$

Proof:
The key to this theorem is the relation

$$\sum_{k=0}^{\infty} \frac{\lambda^{qk+r}}{(qk+r)!} = \frac{1}{q}\sum_{j=0}^{q-1} e^{\lambda \cos\left(\frac{2\pi j}{q}\right)} \cos\left[\lambda \sin\left(\frac{2\pi j}{q}\right) - \frac{2\pi r j}{q}\right] \quad for \; r = 0,1,2,3 \ldots q-1$$

which is derived in Erdelyi (1955) and Nolan (2017). If $\gamma_m = 1, \gamma_q = 0 \; for \; q \neq m$, then

$$\psi_k(i,j) = \sum_{n=0}^{\infty} b_n \sum_{q=j}^{m} \gamma_q \frac{\lambda_k^{i+1+n+q-j} e^{-\lambda_k}}{(i+1+n+q-j)!}$$

reduces to

$$\psi_k(i,j) = \sum_{c=0}^{\infty} \frac{\lambda_k^{c(m+1)+m+1+i-j} e^{-\lambda_k}}{(c(m+1)+m+1+i-j)!}$$

which we can now write as

$$\begin{aligned}
\psi_k(i,j) &= \sum_{c=0}^{\infty} \frac{\lambda_k^{c(m+1)+m+1+i-j} e^{-\lambda_k}}{(c(m+1)+m+1+i-j)!} \\
&= \delta_{\{i \geq j\}}\left(\sum_{c=0}^{\infty} \frac{\lambda_k^{c(m+1)+i-j} e^{-\lambda_k}}{(c(m+1)+i-j)!} - \frac{\lambda_k^{i-j} e^{-\lambda_k}}{(i-j)!}\right) + \delta_{\{i<j\}}\left(\sum_{c=0}^{\infty} \frac{\lambda_k^{c(m+1)+m+1+i-j} e^{-\lambda_k}}{(c(m+1)+m+1+i-j)!}\right) \\
&= e^{-\lambda_k}\left(\delta_{\{i \geq j\}}\left[f_{i-j,m+1}(\lambda_k) - \frac{\lambda_k^{i-j}}{(i-j)!}\right] + \delta_{\{i<j\}} f_{m+1+i-j,m+1}(\lambda_k)\right)
\end{aligned}$$

where

$$f_{r,q}(\lambda) = \sum_{k=0}^{\infty} \frac{\lambda^{qk+r}}{(qk+r)!} = \frac{1}{q}\sum_{j=0}^{q-1} e^{\lambda\cos\left(\frac{2\pi j}{q}\right)}\cos\left[\lambda\sin\left(\frac{2\pi j}{q}\right) - \frac{2\pi r j}{q}\right] \qquad for\ r = 0,1,2 \ldots q-1$$

(note that f is not defined for $r \geq q$)
QED

Theorem 2 provides a closed-form version of the two-pathway counting model. In the special case where $\gamma_m = 1, \gamma_q = 0\ for\ q \neq m$ and $\mu_k = 0\ for\ all\ k$, we get the simple counting model, and element *i,j* of my matrix $\boldsymbol{G}_{\mathbb{i}_k}$ reduces to

$$\boldsymbol{G}_{\mathbb{i}_k}[i,j] = e^{-\lambda_k}\left(\delta_{\{i\geq j\}}\left[f_{i-j,m+1}(\lambda_k) - \frac{\lambda_k^{i-j}}{(i-j)!}\right] + \delta_{\{i<j\}} f_{m+1+i-j,m+1}(\lambda_k)\right)$$

which is equivalent to $(\boldsymbol{D}_{\infty}(u) - \boldsymbol{D}_{\boldsymbol{0}}(u))[i.j]$ in Nolan (2017) (though note that his 'phases' are reversed with respect to mine, and that his *m* is equal to my *m+1*. Also note that Nolan's equations, like mine, implicitly assume crossing over only between non-sister chromatids, his claim to the contrary notwithstanding). If we in addition set *m = 0*, then the chiasma events occur independently, and we get the Poisson model. My expression now reduces to

$$Pr\{\boldsymbol{R} = \boldsymbol{r}\} = \prod_{k=0}^{\mathfrak{n}-1} M_{\mathbb{i}_k}(\boldsymbol{r})$$

where

$$M_{\mathbb{i}_k}(\boldsymbol{r}) = \begin{cases} \frac{1}{2}\left(1 - e^{\lambda_k}\right), & r(\mathbb{i}_k) = 1 \\ \frac{1}{2}\left(1 + e^{\lambda_k}\right), & r(\mathbb{i}_k) = 0 \end{cases}$$

Note that $M_{\mathbb{i}_k}(r)$ is reduced to a scalar. Haldane (1919) previously proved the relation $\Pr\{R(\mathbb{i}) = 1\} = \frac{1}{2}\left(1 - e^{\lambda_k}\right)$ (and $\Pr\{R(\mathbb{i}) = 0\} = 1 - \Pr\{R(\mathbb{i}) = 1\} = \frac{1}{2}\left(1 + e^{\lambda_k}\right)$) for the Poisson model by summing the probabilities that a single randomly chosen chromatid is involved is an odd number of chiasma events, and identifying the Taylor series expansion for sinus hyperbolicus.[5] It also follows directly from the relation $Pr\{R(\mathbb{i}) = 1\} = \frac{1}{2}(1 - \Pr\{X_{\mathbb{i}} = 0\})$, which, as we saw previously, follows from Mather's equation. Hence, the reduced expression above tells us that for the Poisson model, we can calculate the probabilities of observing $r(\mathbb{i}_k)$ for each interval individually and then simply multiply the results to get the probability of observing the full pattern $\boldsymbol{r}$. This is because in the Poisson model the number of chiasma events in each interval is independent of the number of chiasma events in all other intervals, and so the recombination rates are also independent. This is a

[5] His derivation can be expressed in a single line as

$$Pr\{R(\mathbb{i}) = 1\} = e^{-\frac{\lambda}{2}}\sum_{n=1}^{\infty}\frac{\left(\frac{\lambda}{2}\right)^{2n-1}}{(2n-1)!} = e^{-\frac{\lambda}{2}}\sinh\left(\frac{\lambda}{2}\right) = e^{-\frac{\lambda}{2}}\left(\frac{e^{\frac{\lambda}{2}} - e^{-\frac{\lambda}{2}}}{2}\right) = \frac{1}{2}\left(1 - e^{-\lambda}\right)$$

well-known property of the Poisson interference model, and the main reason why it is so commonly used.

2.2.3 Theorem 3: The coefficient of coincidence for the general model

The coefficient of coincidence for the general model is given by

$$C(L_0, L_1) = \frac{\frac{1}{2}(\Pr\{R(\mathbb{i}_0) = 1\} + \Pr\{R(\mathbb{i}_1) = 1\} - \Pr\{R(\mathbb{i}_{0\cap 1}) = 1\})}{\Pr\{R(\mathbb{i}_0) = 1\}\Pr\{R(\mathbb{i}_1) = 1\}} \qquad (1)$$

where $\mathbb{i}_{0\cap 1}$ is the interval between the left boundary of $\mathbb{i}_0$ and the right boundary of $\mathbb{i}_1$, and

$$\Pr\{R(\mathbb{i}) = 1\} = \frac{1}{2}\left(1 - e^{-\mu_{\mathbb{i}}} \sum_{q=0}^{m} \pi_q \sum_{c=0}^{q} \frac{e^{-\lambda_{\mathbb{i}}}\lambda_{\mathbb{i}}^{c}}{c!}\right), \qquad for\ \mathbb{i} = \mathbb{i}_0, \mathbb{i}_1, \mathbb{i}_{0\cap 1} \qquad (2)$$

$$\lambda_{\mathbb{i}} = E[C_{\mathbb{i}}'']$$

$$\mu_{\mathbb{i}} = E[X_{\mathbb{i}}']$$

$$E\left[C_{\mathbb{i}_{0\cap 1}}''\right] = E\left[C_{\mathbb{i}_0}''\right] + E\left[C_{\mathbb{i}_1}''\right]$$

$$E\left[X_{\mathbb{i}_{0\cap 1}}'\right] = E\left[X_{\mathbb{i}_0}'\right] + E\left[X_{\mathbb{i}_1}'\right]$$

Proof:
Equation (1) follows from the definition of $C(L_0, L_1)$ (section 2.1.3) and the relation

$$\begin{aligned}\Pr\{R(\mathbb{i}_{0\cap 1}) = 1\} &= \Pr\{R(\mathbb{i}_0) = 1\ XOR\ R(\mathbb{i}_1) = 1\} \\ &= \Pr\{R(\mathbb{i}_0) = 1\} + \Pr\{R(\mathbb{i}_1) = 1\} - 2\Pr\{R(\mathbb{i}_0) = 1, R(\mathbb{i}_1) = 1\}\end{aligned}$$

(Foss et al. 1993). To get equation (2), first note that

$$\Pr\{X_{\mathbb{i}} = 0\} = \Pr\{X_{\mathbb{i}}' = 0\}\Pr\{X_{\mathbb{i}}'' = 0\}$$

where

$$\Pr\{X_{\mathbb{i}}' = 0\} = e^{-\mu_{\mathbb{i}}}$$

$$\Pr\{X_{\mathbb{i}}'' = 0\} = \sum_{q=0}^{m} \pi_q \sum_{c=0}^{q} \frac{e^{-\lambda_{\mathbb{i}}}\lambda_{\mathbb{i}}^{c}}{c!}$$

The latter expression sums over all possible phases q at the left boundary of $\mathbb{i}$ weighted by their stationary probabilities (π_q), and multiplies each term with the probability of observing between 0 and q C'' events in the interval, which by the definition of phase means all consecutive C'' events up to, but not including, the first X'' event. Equation (2) now follows from Mather's equation,

$$\Pr\{R(\mathbb{i}) = 1\} = \frac{1}{2}(1 - \Pr\{X_{\mathbb{i}} = 0\})$$

When $E\left[X'_{\mathbb{i}_0}\right] = E\left[X'_{\mathbb{i}_1}\right] = 0$ and $\gamma_m = 1$ (pure counting model), then $\pi_q = \frac{1}{m+1}\ for\ q = 0,1,2 \ldots m$, and the expression reduces to

$$\begin{aligned}\Pr\{R(\mathbb{i}) = 1\} &= \frac{1}{2}\left(1 - \sum_{q=0}^{m}\frac{1}{m+1}\sum_{c=0}^{q}\frac{e^{-\lambda_{\mathbb{i}}}\lambda_{\mathbb{i}}^{c}}{c!}\right)\\ &= \frac{1}{2}\left(1 - \sum_{c=0}^{m}\frac{e^{-\lambda_{\mathbb{i}}}\lambda_{\mathbb{i}}^{c}}{c!}\left(\frac{m+1-c}{m+1}\right)\right)\\ &= \frac{1}{2}\left(1 - \sum_{c=0}^{m}\frac{e^{-\lambda_{\mathbb{i}}}\lambda_{\mathbb{i}}^{c}}{c!}\left(1 - \frac{c}{m+1}\right)\right)\end{aligned}$$

which is Foss et al.'s (1993) expression for the pure counting model. When in addition m=0,

$$\Pr\{R(\mathbb{i}) = 1\} = \frac{1}{2}\left(1 - e^{-\lambda}\right)$$

and

$$\begin{aligned}C(L_0, L_1) &= \frac{\frac{1}{2}\left(\Pr\{R(\mathbb{i}_0) = 1\} + \Pr\{R(\mathbb{i}_1) = 1\} - \Pr\{R(\mathbb{i}_{0\cap 1}) = 1\}\right)}{\Pr\{R(\mathbb{i}_0) = 1\}\Pr\{R(\mathbb{i}_1) = 1\}}\\ &= \frac{\left(1 - e^{-\lambda_{\mathbb{i}_0}}\right) + \left(1 - e^{-\lambda_{\mathbb{i}_1}}\right) - \left(1 - e^{-\left(\lambda_{\mathbb{i}_0} + \lambda_{\mathbb{i}_1}\right)}\right)}{\left(1 - e^{-\lambda_{\mathbb{i}_0}}\right)\left(1 - e^{-\lambda_{\mathbb{i}_1}}\right)}\\ &= \frac{1 - e^{-\lambda_{\mathbb{i}_0}} - e^{-\lambda_{\mathbb{i}_1}} + e^{-\left(\lambda_{\mathbb{i}_0} + \lambda_{\mathbb{i}_1}\right)}}{1 - e^{-\lambda_{\mathbb{i}_0}} - e^{-\lambda_{\mathbb{i}_1}} + e^{-\left(\lambda_{\mathbb{i}_0} + \lambda_{\mathbb{i}_1}\right)}} = 1\end{aligned}$$

showing, as expected, that for the Poisson interference model the coefficient of coincidence is always 1 regardless of L_0 and L_1.

2.3 Inversion heterokaryotypes

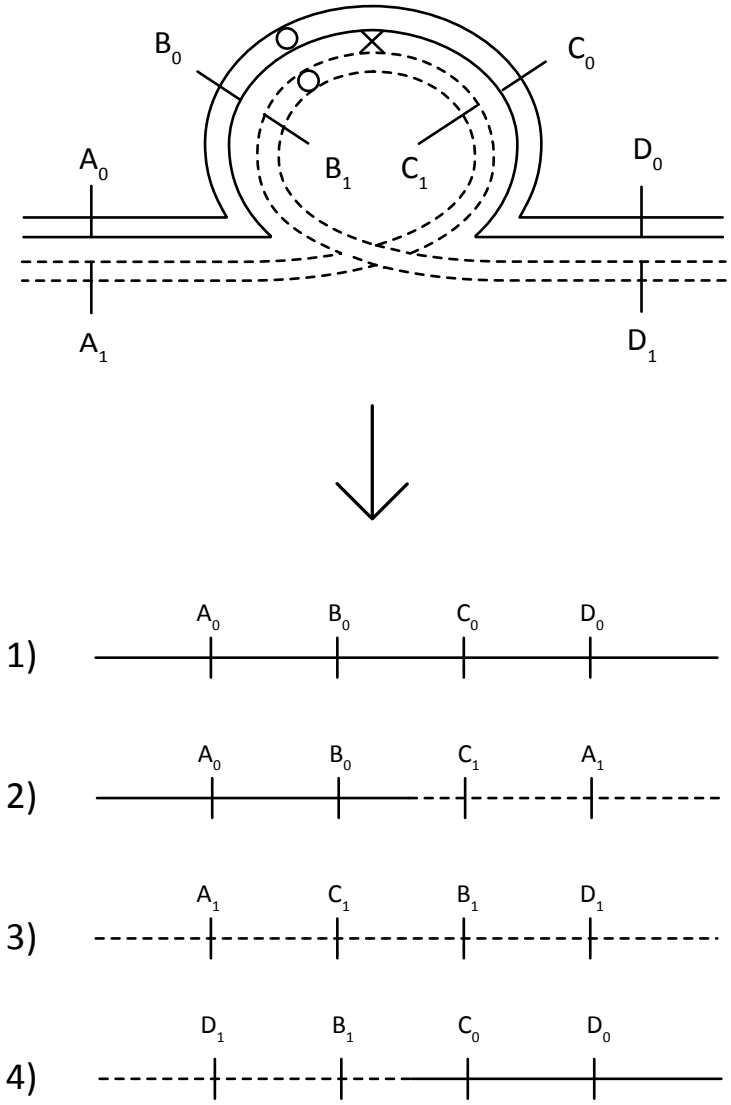
2.3.1 Chiasma inhibition

During meiosis for chromosomes heterozygous for a chromosomal inversion, the inverted region can form a homosynaptic *inversion loop* (figure 2.2) inside and outside of which the formation of chiasmata is partly or fully inhibited (Coyne et al. 1991, 1993, Navarro and Ruiz 1997, Jaarola et al. 1998, Anton et al. 2005, Pegueroles et al. 2010, del Priore and Pigozzi 2015). I will henceforth define the *inhibition factor* for interval $\mathbb{i}_k$, denoted $d_k \in \mathbb{R}_{\geq 0}$, so that if the expected number of X' and C'' events in interval $\mathbb{i}_k$ is μ_k and λ_k, respectively, in a homokaryotype, then the corresponding values in a heterokaryotype are $d_k\mu_k$ and $d_k\lambda_k$. The two boundaries of the inverted region, the *breakpoint boundaries,* can for our purpose be thought of as loci that serve as interval boundaries in the same way that other loci do, so that e.g. the interval between the left breakpoint boundary and the leftmost loci within the inverted region (or the right breakpoint boundary if there are no loci within the inverted region) is an interval in the same respect that the intervals enclosed by two non-

breakpoint loci are. Since inversions affect the rate of chiasma generation also in interval outside of the inverted region (e.g. Pegueroles et al. 2010), d_k also applies for such intervals. I will adopt the convention that $E[]$ always refer to expectation in a homokaryotype, so that e.g. $E[C_k'']$ and $d_k E[C_k'']$ gives the expected number of type II intermediate events in interval k in a homokaryoype and heterokaryotype, respectively. This notation allows different intervals within and outside of the inverted region to be experience inhibition to different degrees, depending, for example, on their relative distance from the breakpoints. If $\mathbb{h}$ denotes the inverted region, $\{\mathbb{i}_k\}_{k=0,1,2\ldots\mathfrak{h}-1}$ the set of all intervals within the inverted region, and $\mu_{\mathbb{h}}$ and $\lambda_{\mathbb{h}}$ the expected number of X' and C'' events, respectively, in the region corresponding to the inverted region in a homokaryotype, then

$$d_{\mathbb{h}} = \sum_{k=0}^{\mathfrak{h}-1} d_k \left(\frac{\mu_k + \lambda_k}{\mu_{\mathbb{h}} + \lambda_{\mathbb{h}}}\right)$$

is the inversion factor for the inverted region as a whole, which is equivalent to d in Navarro and Ruiz (1997). From the degree of underdominance (see discussion below) of mostly laboratory-induced pericentric inversions of different genetic lengths in a large *Drosophila* dataset, the authors of that paper estimated that a $d_{\mathbb{h}}$ value of about 0.25 gave the overall best fit. In other words, the expected number of chiasma events within the inverted region in heterokarytypes is overall about a quarter of that in the corresponding region in homokaryotypes, which, as they point out, coincide with the independent previous estimate of Novitski and Braver (1954), also for *Drosophila.* The authors also note, however, that the individual $d_{\mathbb{h}}$ values of each inversion seem to depend on the genetic length of the inversion, and in particular that chiasma formation is perfectly suppressed ($d_{\mathbb{h}} \approx 0$) when the inversion is short, a point to which I will return.



Figur 2.2: An inversion loop for a pericentric inversion. The lines with the same line type (dotted or dashed) are sister chromatids, and the circles represent the centromere. The X symbolizes a chiasmata, which involves *the chromatids touched by the upper and lower arms of the X. Gametes that show recombination in the interval comprising the full inverted region are unbalanced, as illustrated here by gamete 2) and 4), which lack locus D and A, respectively.*

2.3.2 Interference across breakpoint boundaries

When calculating recombination pattern probabilities for inversion heterokaryotypes with more than one distinct interval within the inverted region, we are confronted with a dilemma that to my knowledge is not explicitly addressed in the literature. If we imagine four intervals, $\mathbb{i}_0, \mathbb{i}_1, \mathbb{i}_2, \mathbb{i}_3$, of which one, $\mathbb{i}_0$, is to the left of the inverted region, two, $\mathbb{i}_1$ and $\mathbb{i}_2$, is inside the inverted region, and one, $\mathbb{i}_3$, is to the right of the inverted region, then their order from left to right will be $\mathbb{i}_0, \mathbb{i}_1, \mathbb{i}_2, \mathbb{i}_3$ in one parental homologue but $\mathbb{i}_0, \mathbb{i}_2, \mathbb{i}_1, \mathbb{i}_3$ in the other. It is therefore not obvious how the interference signal generated by a chiasma in, say, interval $\mathbb{i}_0$, will travel through an inversion loop; will it travel through $\mathbb{i}_1$ before $\mathbb{i}_2$ or $\mathbb{i}_2$ before $\mathbb{i}_1$ (figure 2.3)? Will it somehow split up and travel simultaneously through $\mathbb{i}_1$ and $\mathbb{i}_2$? Will it continue out through the other inversion breakpoint or loop around and

somehow cause additional interference in $\mathbb{i}_0$? At our current state of knowledge, these questions have no obvious answers. One possibility is that the interference signal is simply blocked by the inversion breakpoints, so that, in our example, the number of chiasmata in interval $\mathbb{i}_1$ is dependent on the number of chiasmata in interval $\mathbb{i}_2$ (and vice versa), but independent of the number of chiasmata in intervals $\mathbb{i}_0$ and $\mathbb{i}_3$. In support of this idea, Gorlov and Borodin (1995) found no evidence of chiasma interference from one loop to the other in a double heterozygote for two partly overlapping inversions in mice, possibly because the synaptomenal complex is initiated independently inside and outside of the inverted region. Mary et al. (2016), however, did find interference from one side of an inversion to another, although in that study the region was prevented from forming an inversion loop, which might be relevant. As more research is needed to settle this question (see chapter 5), I have included in the program a user-defined parameter α for which the value 1 indicate normal interference across breakpoint boundaries, and 0 indicate no interference across breakpoint boundaries, i.e. that the phase distribution is reset to the stationary probabilities to the right of both breakpoints regardless of the what has happened to the left of those points. When $\alpha = 1$, the program will assume that the interference signal travels in either of the two directions illustrated in figure 2.3 with equal probability, so that the final recombination pattern probabilities are given by the average of the values calculated in each of these scenarios. Note that when all intervals of interest are located inside the inverted region, the results for $\alpha = 1$ and $\alpha = 0$ will be the same.

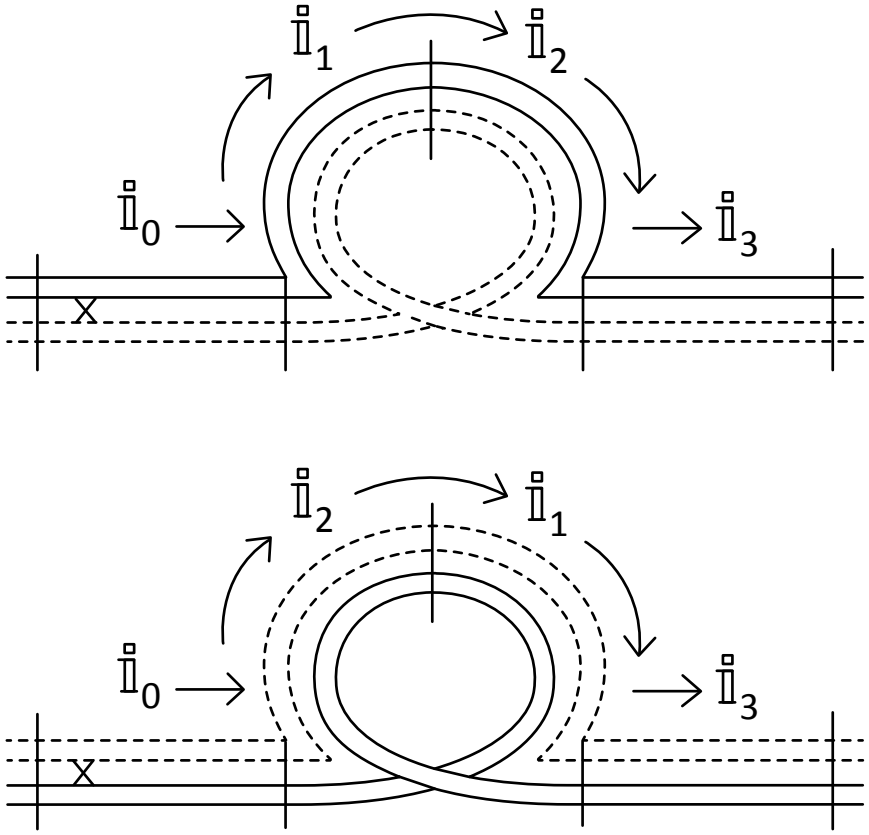


Figure 2.3: The problem of interference across breakpoint boundaries. Does the interference signal from a chiasma event (the x) in $\mathbb{i}_0$ go through $\mathbb{i}_1$ (top) or $\mathbb{i}_2$ (bottom)? Or is it blocked by the breakpoint boundaries?

2.3.3 Sterility in pericentric and paracentric inversion heterokaryotypes

It is common to distinguish between *pericentric inversions*, which include the centromere, and *paracentric inversions,* which do not. In individuals heterozygous for both of these types of inversion, crossing over within the inverted region results in a proportion of *unbalanced gametes* – gametes that do not have the full set of genetic material (Figures 2.2, 2.7, 2.8, 2.9). In the following, all such gametes are assumed to produce inviable zygotes – i.e. the program automatically assign all offspring inheriting one or two such gametes a zygote fitness of 0 in the next generation – and the probability of observing such gametes for any given individual or genotype will be referred to as that individual or genotype's *sterility*. In the rest of this chapter, I will derive expressions for the sterility and recombination pattern probabilities for individuals heterozygous for both types of inversions. Since the mathematics differ between the two, they will be considered in turn.

2.3.4 Terminology for pericentric inversions

For pericentric inversion heterokaryotypes, I will divide the region of interest into three subregions for easier reference. *The left region* is the region comprising all intervals to the left of the inverted region, *the inverted region* is, as before, the region comprising all intervals captured by the inversion, and *the right region* is the region comprising all intervals to the right of the inverted region. These three regions will be denoted $\mathbb{d}$, $\mathbb{h}$ and $\mathbb{r}$, respectively, so that e.g. $E[C_{\mathbb{h}}'']$ and $d_{\mathbb{h}}E[C_{\mathbb{h}}'']$ denotes the expected number of type II intermediate events in the inverted region as a whole in a homokaryotype and heterokaryotype, respectively. The number of intervals in $\mathbb{d}$, $\mathbb{h}$, and $\mathbb{r}$ will be denoted $\mathfrak{d}$, $\mathfrak{h}$, and $\mathfrak{r}$, so that the total number of intervals is $\mathfrak{n} = \mathfrak{d} + \mathfrak{h} + \mathfrak{r}$. As before, $\mathbb{i}_k$ refer to zero-indexed interval number *k* from the left (one-indexed interval *k + 1*), so from the preceding we can unambiguously deduce that $\mathbb{i}_k, k = 0,1,2 \ldots \mathfrak{d} - 1$ is an interval in the left region, $\mathbb{i}_k, k = \mathfrak{d}, \mathfrak{d} + 1, \mathfrak{d} + 2 \ldots \mathfrak{d} + \mathfrak{h} - 1$ is an interval in the inverted region, and $\mathbb{i}_k, k = \mathfrak{d} + \mathfrak{h}, \mathfrak{d} + \mathfrak{h} + 1, \mathfrak{d} + \mathfrak{h} + 2 \ldots \mathfrak{n} - 1$ is an interval in the right region. For easier reading, I will when convenient adopt the convention that an $\mathbb{d}$, $\mathbb{h}$, and $\mathbb{r}$ with subscript index *k* refer to the same interval as $\mathbb{i}_k$, the only difference being the additional explicit (and redundant) information about which of the three subregions the interval belongs to. For example, $\mathbb{i}_k$ and $\mathbb{h}_k$ refer to the same interval, zero-indexed number *k* from the left in the region of interest, but in the latter case you can immediately see that the interval is in the inverted region without having to scrutinize the index.

For pericentric inversion heterokaryotypes, a chromatid is, as illustrated in figure 2.2, unbalanced if and only if it shows recombination in the inverted region. Since by assumption all chromatids, unbalanced or not, have an equal chance of becoming gametes, the sterility, ζ, of such a chromosome is simply given by $\zeta = \Pr\{R(\mathbb{h}) = 1\}$. Hence we have the following theorem:

2.3.5 Theorem 4: The sterility of pericentric inversion heterokaryotypes

The sterility of an individual heterozygous for a pericentric inversion is given by

$$\zeta = \Pr\{R(\mathbb{h}) = 1\} = \frac{1}{2}\left(1 - e^{-d_{\mathbb{h}}\mu_{\mathbb{h}}} \sum_{q=0}^{m} \pi_q \sum_{c=0}^{q} \frac{e^{-d_{\mathbb{h}}\lambda_{\mathbb{h}}}(d_{\mathbb{h}}\lambda_{\mathbb{h}})^c}{c!}\right)$$

Proof:
This follows from the discussion in theorem 3 above.

2.3.6 *d* values in the Coyne/Navarro and Ruiz dataset

Navarro and Ruiz's (1997) figure 1 uses the dataset discussed in section 2.3.1 to plot each individual pericentric inversion's sterility against its genetic length, together with a line for their expression of the sterility with the value of $d_{\mathbb{h}}$ that gives the best fit overall. I will here instead use their dataset to plot the individual $d_{\mathbb{h}}$ values against the genetic length, which for my purpose will prove more useful. Assuming, for simplicity, a Poisson model of interference (Navarro and Ruiz also make this assumption, and note that using a more realistic model of interference makes little difference) and combining with the expression for genetic lengths (section 2.1.6), theorem 4 reduces to

$$\zeta = \frac{1}{2}(1 - e^{-2d_{\mathbb{h}}L})$$

where L is the genetic length of the inverted region in Morgans (this is equivalent to Navarro and Ruiz's equation 2, except the latter disregard the probability of more than two chiasma events in the inverted region). Solving for $d_{\mathbb{h}}$ gives

$$d_{\mathbb{h}} = -\frac{\ln(1 - 2\zeta)}{2L}$$

Using the values for ζ and L in the Coyne et al. (1993)/Navarro and Ruiz (1997) dataset now give the scatterplot in figure 2.4. Quite a few of the inversions in the set actually have slightly or significantly negative sterility values, meaning that the fertility is higher in heterokaryotypes than in homokaryotypes. The $d_{\mathbb{h}}$ values cannot be negative, so this must be due to sampling errors or pleiotropic effects (see Coyne et al. 1993 for details on the calculation of the sterility values). I therefore plot the $d_{\mathbb{h}}$ values in question at zero. Note that these all cluster in the lower range of inversion lengths, possibly because short inversion heterokaryotypes fail to form a homosynaptic loop (Coyne and Orr 2004, Anton et al. 2005), and that there generally seems to be a tendency for higher $d_{\mathbb{h}}$ values with longer inversions, meaning that shorter inversions suppress chiasma formation more than longer ones.

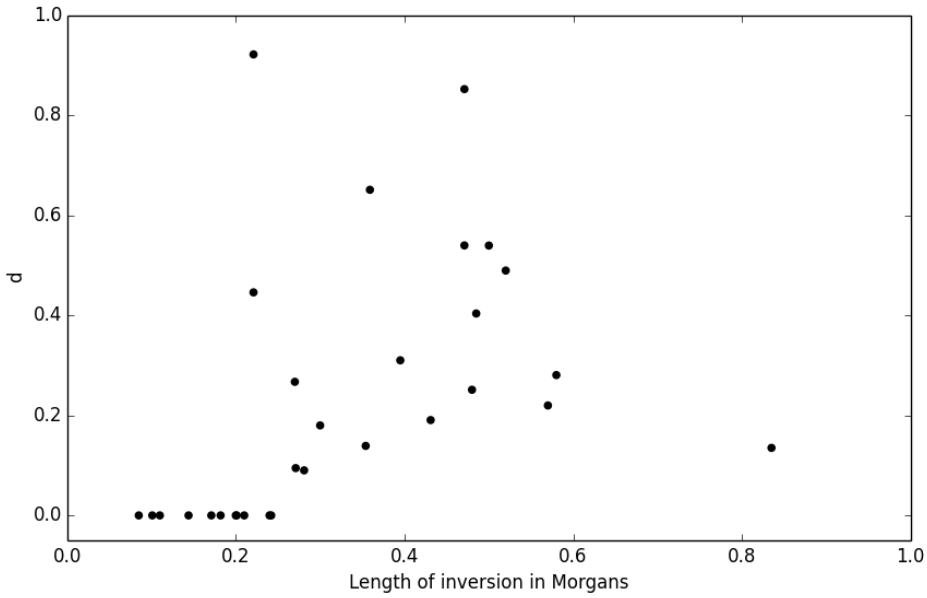


Figure 2.4: The individual $d_{\mathbb{h}}$ values for the inversions in the dataset from Coyne et al. (1993) and Navarro and Ruiz (1997). $d_{\mathbb{h}}$ values for inversions with negative sterility are set to zero for reasons explained in the main text.

2.3.7 Theorem 5: Recombination in pericentric inversion heterokaryotypes

The probability of observing recombination pattern *r* on a chromosome with a pericentric inversion is given by

$$\Pr\{\boldsymbol{R}_{pericentric} = \boldsymbol{r}\} = \boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{b}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)\boldsymbol{Q}^{1-\alpha}\left(\prod_{k=\mathfrak{b}}^{\mathfrak{b}+\mathfrak{h}-1}\boldsymbol{M}_{\mathbb{h}_k}(r)\right)\boldsymbol{Q}^{1-\alpha}\left(\prod_{k=\mathfrak{b}+\mathfrak{h}}^{\mathfrak{n}-1}\boldsymbol{M}_{\mathbb{r}_k}(r)\right)\boldsymbol{1}^T$$

where

$$\boldsymbol{M}_{\mathbb{i}_k}(r) = \begin{cases} \frac{1}{2}\boldsymbol{G}_{\mathbb{i}_k}, & r(\mathbb{i}_k) = 1 \\ \frac{1}{2}\boldsymbol{G}_{\mathbb{i}_k} + \boldsymbol{H}_{\mathbb{i}_k}, & r(\mathbb{i}_k) = 0 \end{cases} \qquad \mathbb{i} = \mathbb{d}, \mathbb{h}, \mathbb{r}$$

$$\boldsymbol{G}_{\mathbb{i}_k}[i,j] = \psi_k(i,j)e^{-d_k\mu_k} + \left(\psi_{\boldsymbol{k}}(i,j) + \delta_{\{j\geq i\}}\frac{(d_k\lambda_k)^{i-j}e^{-d_k\lambda_k}}{(i-j)!}\right)(1-e^{-d_k\mu_k}), \qquad for\ \ i,j = 0,1,2\ldots m$$

$$\psi_k(i,j) = \sum_{n=0}^{\infty} b_n \sum_{q=j}^{m}\gamma_q\frac{(d_k\lambda_k)^{i+1+n+q-j}e^{-d_k\lambda_k}}{(i+1+n+q-j)!}$$

$$\boldsymbol{H}_{\mathbb{i}_k}[i,j] = \begin{cases} \frac{(d_k\lambda_k)^{i-j}e^{-d_k\lambda_k}}{(i-j)!}e^{-(d_k\mu_k)}, & i \geq j \\ 0, & i < j \end{cases} \quad for\ i,j = 0,1,2\ldots m$$

$$\boldsymbol{Q}[i,j] = \boldsymbol{\pi}[j] = \pi_j \quad i,j = 0,1,2\ldots m$$

Proof:
Apart from the inhibition factors (*d*), the only difference between theorems 1 and 5 is the matrix $\boldsymbol{Q}$ that is inserted into the latter equation at the positions corresponding to the breakpoint boundaries. The parameter α indicate the presence ($\alpha = 1$) or absence ($\alpha = 0$) of chiasma interference across the breakpoint boundaries, so in the former case $\boldsymbol{Q}^{1-\alpha} = \boldsymbol{Q}^0 \equiv \boldsymbol{I}_{m+1}$ (i.e. the identity matrix). In the absence of interference across the breakpoint boundaries ($\alpha = 0$), $\boldsymbol{Q}$ serves to redistribute the phase probabilities according to the stationary distribution ($\boldsymbol{\pi}$). We can see why this is so by decomposing $\boldsymbol{Q}$ into the two matrices $\boldsymbol{Q}'$ and $\boldsymbol{Q}''$ so that

$$\boldsymbol{Q} = \boldsymbol{Q}'\boldsymbol{Q}''$$

where

$$\boldsymbol{Q}'[i,j] = \begin{cases} 1, & j = 0 \\ 0, & j > 0 \end{cases} \quad i,j = 0,1,2\ldots m$$

$$\boldsymbol{Q}''[i,j] = \begin{cases} \pi_j, & i = 0 \\ 0, & i > 0 \end{cases} \quad i,j = 0,1,2\ldots m$$

That is, $\boldsymbol{Q}'$ maps all phase probabilities (arbitrarily) to phase 0, whereas $\boldsymbol{Q}''$maps them from 0 to *j* in proportion to the stationary probabilities. This is equivalent to making the chiasma events in the three subregions mutually independent, as we can see from the following argument (*j = 0,1,2…m* for all relevant equations):

$$\left(\boldsymbol{\pi}\prod_{k=0}^{\mathfrak{d}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)[j] = \Pr\{R(\mathbb{d}_0) = r_0, \dots, R(\mathbb{d}_{\mathfrak{d}-1}) = r_{\mathfrak{d}-1}, Q_{\mathbb{d}^r} = j\}$$

$$\left(\boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{d}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)\boldsymbol{Q}'\right)[j] = \begin{cases}\Pr\{R(\mathbb{d}_0) = r_0, \dots, R(\mathbb{d}_{\mathfrak{d}-1}) = r_{\mathfrak{d}-1}\}, & j = 0\\ 0, & j = 1,2,3 \dots m\end{cases}$$

$$\left(\boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{d}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)\boldsymbol{Q}'\boldsymbol{Q}''\right)[j] = \left(\boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{d}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)\boldsymbol{Q}\right)[j]$$
$$= \Pr\{R(\mathbb{d}_0) = r_0, \dots, R(\mathbb{d}_{\mathfrak{d}-1}) = r_{\mathfrak{d}-1}\}\pi_j$$

$$\left(\boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{d}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)\boldsymbol{Q}\left(\prod_{k=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1}\boldsymbol{M}_{\mathbb{h}_k}(r)\right)\boldsymbol{Q}\right)[j]$$
$$= \Pr\{R(\mathbb{d}_0) = r_0, \dots, R(\mathbb{d}_{\mathfrak{d}-1}) = r_{\mathfrak{d}-1}\}\Pr\{R(\mathbb{h}_{\mathfrak{d}}) = r_{\mathfrak{d}}, \dots, R\left(\mathbb{h}_{\mathfrak{d}+\mathfrak{h}-1}\right) = r_{\mathfrak{d}+\mathfrak{h}-1}\}\,\pi_j$$

$$\boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{d}-1}\boldsymbol{M}_{\mathbb{d}_k}(r)\right)\boldsymbol{Q}\left(\prod_{k=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1}\boldsymbol{M}_{\mathbb{h}_k}(r)\right)\boldsymbol{Q}\left(\prod_{k=\mathfrak{d}+\mathfrak{h}}^{\mathfrak{n}-1}\boldsymbol{M}_{\mathbb{r}_k}(r)\right)\boldsymbol{1}^T$$
$$= \Pr\{R(\mathbb{d}_0) = r_0, \dots, R(\mathbb{d}_{\mathfrak{d}-1}) = r_{\mathfrak{d}-1}\}\Pr\{R(\mathbb{h}_{\mathfrak{d}}) = r_{\mathfrak{d}}, \dots, R\left(\mathbb{h}_{\mathfrak{d}+\mathfrak{h}-1}\right) = r_{\mathfrak{d}+\mathfrak{h}-1}\}\Pr\{R\left(\mathbb{r}_{\mathfrak{d}+\mathfrak{h}}\right) = r_{\mathfrak{d}+\mathfrak{h}}, \dots, R(\mathbb{r}_{\mathfrak{n}-1}) = r_{\mathfrak{n}-1}\}$$

Note that when $\alpha = 1$, the program executes theorem 5 once with the intervals in the original order and once with the intervals ordered so that order inside the inverted region is reversed (see figure 2.3), and the final gamete proportions (haplotype representation) resulting from the two calculations are averaged. This is also the case for theorem 7.

QED

In the main program, the probabilities for all recombination patterns, unbalanced or not, in a pericentric heterokaryotype, are calculated using theorem 5. Unbalanced chromatids are subsequently recognized and pooled into in a single category *unbalanced*. The probability of observing a chromatid in category *unbalanced* is accordingly equal to the sum of the probabilities of observing each individual unbalanced chromatid, which is, of course, equal to the sterility.

A chromatid is unbalanced if and only if it shows recombination in an odd number of intervals within the inverted region (see figure 2.5 for an induction proof). Hence, the function

$$\varphi(\boldsymbol{r}) = \begin{cases} 1, & \sum_{i=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} r(\mathbb{h}_i)\ is\ even \\ 0, & \sum_{i=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} r(\mathbb{h}_i)\ is\ odd \end{cases}$$

is 1 if $\boldsymbol{r}$ is balanced and 0 if $\boldsymbol{r}$ is unbalanced, and can therefore be used to place patterns in the appropriate category. Note that the probabilities of observing a chromatid with a certain recombination pattern is given as a proportion of all chromatids, balanced or unbalanced, and not, as in e.g. Navarro et al. (1997), as a proportion of balanced chromatids only.

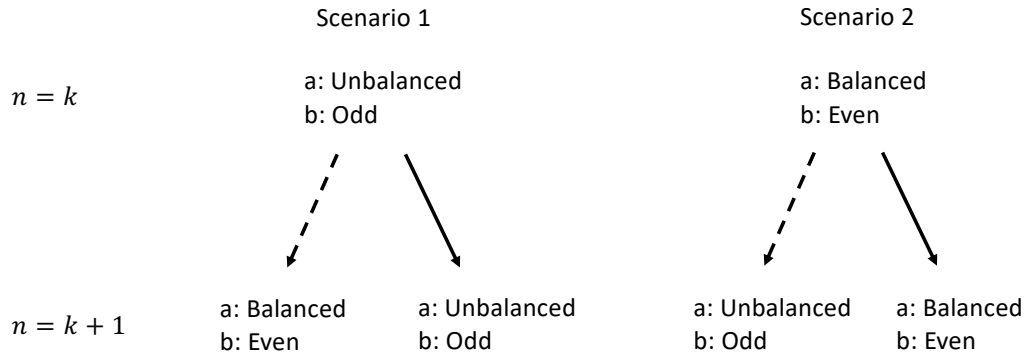


Figure 2.5: An induction proof of the statement (S) that a chromatid is unbalanced if and only if it shows recombination in an odd number of intervals within the inverted region. Keep in mind that an interval shows recombination if and only if it is involved in an odd number of chiasma events within the interval, and that a chromatid (in a heterokaryotype) is unbalanced if and only if it show recombination in the inverted region. Note first that the statement is trivially true when the inverted region consists of only a single interval (S_1). Now assume for the sake of argument that the boundaries of the inverted region can be moved to include additional intervals, and that the statement is true when the inverted region consist of n = k intervals (S_k). Hence, we must consider two scenarios: either the chromatid is unbalanced (involved in an odd number of chiasma events within the inverted region), which, the induction hypothesis states, must mean that it also shows recombination in an odd number of intervals within the inverted region (scenario 1), or the chromatid is balanced (involved in an even number of chiasma events within the inverted region), meaning that it show recombination in an even number of intervals within the inverted region (scenario 2). The figure shows the effect of including an additional recombinant (dashed arrow) or non-recombinant (solid arrow) interval in both of these scenarios (a: unbalanced/balanced chromatid, b: odd/even number of recombinant interval in the inverted region). Note that the statement is true for n = k + 1. Hence, since S_1 is true and S_k implies S_{k+1}, the statement must be true for all n > 0.

2.3.8 Terminology for paracentric inversions

For paracentric inversion karyotypes, I will divide the region of interest into four subregions, from left to right: *the distal region* comprises all intervals to the left of the inverted region, *the inverted region* is as before, the *proximal region* comprises all intervals between the inverted region and the centromere (which, for our purpose, serve as a loci), and the *right region* comprises all intervals to the right of the proximal region. These will in the following be denoted $\mathbb{d}$, $\mathbb{h}$, $\mathbb{p}$, and $\mathbb{r}$, respectively, with the number of intervals in each denoted $\mathfrak{d}$, $\mathfrak{h}$, $\mathfrak{p}$ and $\mathfrak{r}$. Otherwise the notation is the same as for pericentric inversions.

As illustrated in figures 2.7, 2.8, and 2.9, different number of chiasmata in the inverted and proximal regions produce different proportions of tetrads with five different possible *configurations*

(figure 2.5). In the following, a *no bridge* tetrad is a tetrad that do not form chromosome bridges at either anaphase I or II, a *single anaphase I bridge* tetrad is a tetrad in which one chromatid pair form a chromosome bridge at anaphase I, a *double anaphase I bridge* tetrad is a tetrad in which two chromatid pairs (i.e. all four chromatids) form anaphase I bridges, and a *single* and *double anaphase II bridge* tetrad is the same for anaphase II. A to my knowledge unprecedented expression for the proportions of these five configurations given the number of chiasma events in the inverted and proximal regions is included in theorem 6. The chromatids involved in a chromosome bridge break randomly at the respective anaphase, and so become unbalanced in the sense defined above. The unbalanced chromatids generated in paracentric inversion heterokaryotypes differ in details from the ones generated in pericentric inversion heterokaryotypes (compare figures 2.2 and 2.7), but for our purpose they can be treated as equivalent once they become gametes; i.e. all offspring resulting from one or two unbalanced gametes have zygote fitness 0 in either case.

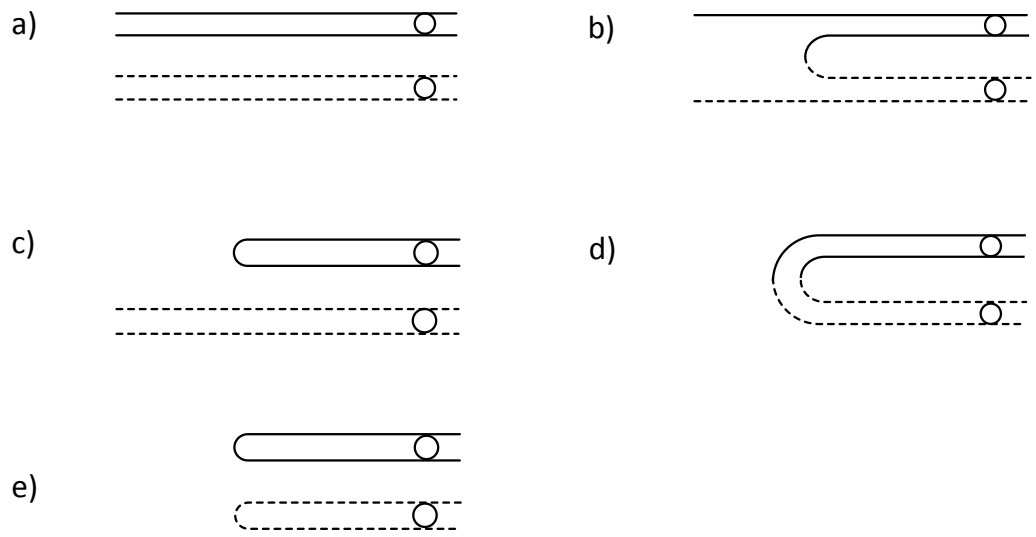


Figure 2.6: The five types of tetrad configurations. a) no bridge, b) single anaphase I bridge, c) single anaphase II bridge, d) double anaphase I bridge, e) double anaphase II bridge. Acentric fragments (see figure 2.7) are not shown. Chromatids involved in a bridge break randomly at the given anaphase, and are necessarily unbalanced.

Up until now I have only considered cases where all chromatids have an equal chance of becoming gametes, so that I have treated "the probability of observing recombination pattern $\boldsymbol{r}$ is p", "the probability of observing a chromatid with recombination pattern $\boldsymbol{r}$ is p", and "the probability of observing a gamete with recombination pattern $\boldsymbol{r}$ is p" as equivalent. I will now consider a case where this is no longer true. In the *linear meiosis* of females of *Drosophila* (Sturtevant and Beadle 1936, Roberts 1976) and *Sciara* (Carson 1946), the two unbalanced chromatids in an anaphase I tetrad are retained in the polar bodies, so the remaining two balanced chromatids are the only ones that can become gametes. Since a single chiasma event in the inverted region always generates an anaphase I tetrad (figure 2.7), this means that additional chiasma events in either the inverted or proximal region is required to produce unbalanced gametes (figures 2.8 and 2.9). Hence, the sterility is significantly reduced. Furthermore, the set of recombination patterns that end up in an anaphase I tetrad is not a representative sample of the full set, so the gamete proportions are also affected. Navarro et al. (1997) provides approximate expressions for the sterility and gamete proportions in paracentric inversion heterokaryotypes with linear meiosis under the Poisson and pure counting interference models for a maximum of two loci in either the distal, inverted or proximal region. I will here present exact infinite series expressions for the sterility and gamete proportions under the

general interference model for an indefinite number of loci in each of the regions on the chromosome. I begin with the former, which is theorem 6.

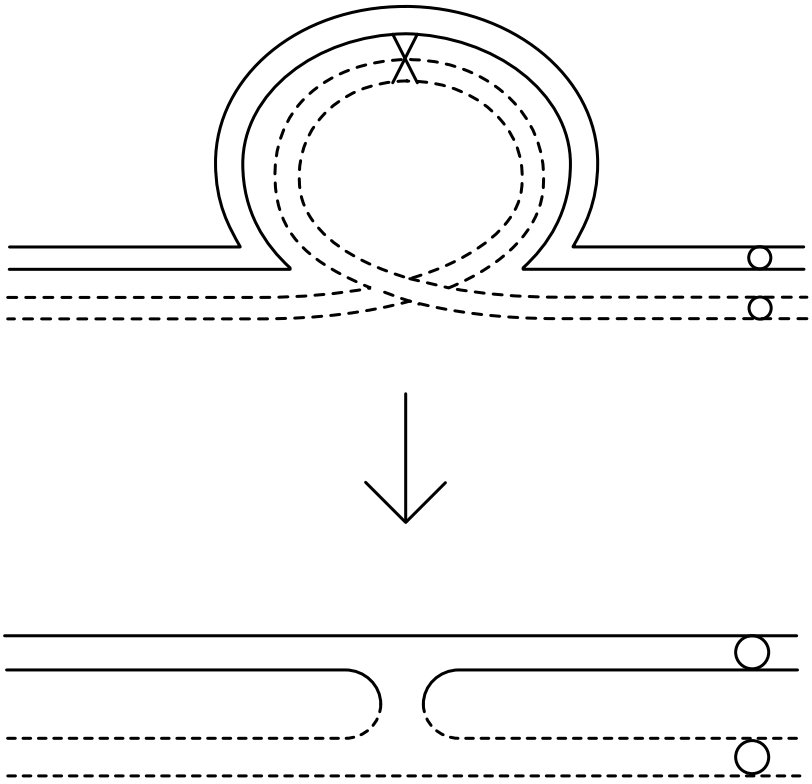


Figure 2.7: An inversion loop for a paracentric inversion. A single chiasma in the inverted region creates a tetrad with an anaphase I bridge and an acentric fragment (meaning that it is not connected to the centromere) which fails to segregate.

2.3.9 Theorem 6: The sterility of paracentric heterokaryotypes with linear meiosis

The sterility of a paracentric inversion heterokaryotype with linear meiosis is given by

$$\zeta_{paracentric\ linear} = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty}[\boldsymbol{\pi}\boldsymbol{D}_{\mathbb{h}}(x_1)\boldsymbol{Q}^{1-\alpha}\boldsymbol{D}_{\mathbb{p}}(x_2)\mathbf{1}^T]\boldsymbol{v}_{\mathbf{0,0}}\boldsymbol{T}_{\mathbb{h}}^{x_1}\boldsymbol{T}_{\mathbb{p}}^{x_2}\boldsymbol{w}_{\boldsymbol{u}}$$

where

$$\boldsymbol{D}_{\mathbb{i}}(0)[i,j] = \begin{cases} \frac{(d_{\mathbb{i}}\lambda_{\mathbb{i}})^{i-j}e^{-d_{\mathbb{i}}\lambda_{\mathbb{i}}}}{(i-j)!}e^{-d_{\mathbb{i}}\mu_{\mathbb{i}}}, & i \geq j \\ 0, & i < j \end{cases} \qquad for\ i,j = 0,1,2\ldots m;\ \mathbb{i} = \mathbb{h},\mathbb{p}$$

$$\boldsymbol{D}_{\mathbb{i}}(x)[i,j] = \sum_{l=0}^{x-1}\sum_{n=x-l-1}^{z}\sum_{q=0}^{j} g_n(x-l-1)\gamma_q \frac{e^{-d_{\mathbb{i}}(\lambda_{\mathbb{i}}+\mu_{\mathbb{i}})}(d_{\mathbb{i}}[\lambda_{\mathbb{i}}+\mu_{\mathbb{i}}])^h}{h!}\binom{h}{l}\left(\frac{\mu_{\mathbb{i}}}{\lambda_{\mathbb{i}}+\mu_{\mathbb{i}}}\right)^l\left(\frac{\lambda_{\mathbb{i}}}{\lambda_{\mathbb{i}}+\mu_{\mathbb{i}}}\right)^{h-l}$$

$$+\delta_{\{i\geq j\}}\frac{e^{-d_{\mathbb{i}}\mu_{\mathbb{i}}}(d_{\mathbb{i}}\mu_{\mathbb{i}})^x}{x!}\frac{e^{-d_{\mathbb{i}}\lambda_{\mathbb{i}}}(d_{\mathbb{i}}\lambda_{\mathbb{i}})^{i-j}}{(i-j)!}, \qquad for\ x = 1,2,3\ldots;\ i,j = 0,1,2\ldots m;\ \mathbb{i} = \mathbb{h},\mathbb{p}$$

$z = (x-l-1)(m+1)$

$h = i+1+l+n+q-j$

$$g_n(s) = \begin{cases} \sum_{k=s-1}^{n-1} g_k(s-1)\gamma_{n-1-k}, & n \geq s \neq 0 \\ 1, & n = s = 0 \\ 0, & n \neq 0; s = 0 \\ 0, & n < s \end{cases}$$

$\boldsymbol{v}_{\mathbf{0,0}} = (1 \quad 0 \quad 0 \quad 0 \quad 0)$

$$\boldsymbol{T}_{\mathbb{h}} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{T}_{\mathbb{p}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\boldsymbol{w}_{\boldsymbol{u}} = \begin{pmatrix} 0 \\ 0 \\ 1/2 \\ 1 \\ 1 \end{pmatrix}$$

$d_{\mathbb{i}}\lambda_{\mathbb{i}} = d_{\mathbb{i}}E[C_{\mathbb{i}}''] \; for \; \mathbb{i} = \mathbb{h}, \mathbb{p}$

$d_{\mathbb{i}}\mu_{\mathbb{i}} = d_{\mathbb{i}}E[X_{\mathbb{i}}'] \; for \; \mathbb{i} = \mathbb{h}, \mathbb{p}$

Proof:
The basic structure of this equation is

$$\zeta = \Pr\{unbalanced\ gamete\} = \sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} \Pr\{X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2\} \Pr\{unbalanced\ gamete | X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2\}$$

i.e. it finds the probability of an unbalanced gamete by conditioning on the number of chiasma events in each region. $\Pr\{X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2\} = \boldsymbol{\pi D}_{\mathbb{h}}(x_1)\boldsymbol{Q}^{1-\alpha}\boldsymbol{D}_{\mathbb{p}}(x_2)\mathbf{1}^T$ is a special case of a more general expression that is shown in the proof for theorem 7. I therefore focus here on the conditional probability of observing an unbalanced gamete, $\Pr\{unbalanced\ gamete | X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2\}$, which I will denote ζ_{x_1,x_2}. For small values of x_1 and x_2, we can find ζ_{x_1,x_2} through the (rather tedious) process of drawing all possible tetrads and counting the resulting balanced and unbalanced gametes, but for higher values this approach becomes unmanageable. Navarro et al. (1997) provide a table of ζ_{x_1,x_2} values (denoted $1-\rho_{ij}$ in that text) for x_1 and x_2 ranging from 0 to 2, but do not derive a general expression.

To find one, consider first a stochastic process $\{Z(x_1, x_2)\}_{(x_1,x_{i_2})\in\mathbb{Z}_{\geq 0}}$ that represents the configuration of a tetrad when $X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2$ (i.e. when there are x_1 chiasma events in the inverted region and x_2 chiasma events in the proximal region), so that

$$Z(x_1, x_2) = \begin{cases} 0, & no\ bridge \\ 1, & single\ anaphase\ I\ bridge \\ 2, & single\ anaphase\ II\ bridge \\ 3, & double\ anaphase\ I\ bridge \\ 4, & double\ anaphase\ II\ bridge \end{cases}$$

We now define the matrices $\boldsymbol{T}_{\mathbb{h}}$ and $\boldsymbol{T}_{\mathbb{p}}$ so that

$$\boldsymbol{T}_{\mathbb{h}}[i,j] = \Pr\{Z(x_1+1,0) = j | Z(x_1,0) = i\}, \qquad x_1 = 0,1,2,\dots; i,j = 0,1,2,3,4$$
$$\boldsymbol{T}_{\mathbb{p}}[i,j] = \Pr\{Z(x_1,x_2+1) = j | Z(x_1,x_2) = i\}, \qquad x_1,x_2 = 0,1,2,\dots; i,j = 0,1,2,3,4$$

and the vector $\boldsymbol{v}_{\mathbf{0,0}}$ so that

$$\boldsymbol{v}_{0,0} = (\Pr\{Z(0,0)=0\} \quad \Pr\{Z(0,0)=1\} \quad \Pr\{Z(0,0)=2\} \quad \Pr\{Z(0,0)=3\} \quad \Pr\{Z(0,0)=4\})$$

That is, the $\boldsymbol{T}_{\mathbb{i}}$ matrices are the Markovian transition matrices whose element *i,j* give the probability of observing a tetrad in configuration *j* after considering an additional chiasma event at the right end of region $\mathbb{i}$ ($= \mathbb{h}, \mathbb{p}$), given that the tetrad was in configuration *i* before considering that chiasma event and that there are no other chiasma event further to the right; and $\boldsymbol{v}_{\mathbf{0,0}}$ is the tetrad configuration distribution when there are no chiasma events in either interval. Accordingly,

$$\begin{aligned}\left(\boldsymbol{v}_{x_1-1,0}\boldsymbol{T}_{\mathbb{h}}\right)[j] &= \sum_{i=0}^{4} \Pr\{Z(x_1,0)=j|Z(x_1-1,0)=i\}\Pr\{Z(x_1-1,0)=i\} \\ &= \Pr\{Z(x_1,0)=j\} \\ &= \boldsymbol{v}_{x_1,0}[j]\end{aligned}$$

so

$$\begin{aligned}\boldsymbol{v}_{x_1,0} &= \boldsymbol{v}_{x_1-1,0}\boldsymbol{T}_{\mathbb{h}} = \left(\boldsymbol{v}_{x_1-2,0}\boldsymbol{T}_{\mathbb{h}}\right)\boldsymbol{T}_{\mathbb{h}} \\ &= \boldsymbol{v}_{x_1-2,0}\boldsymbol{T}_{\mathbb{h}}^2 = \left(\boldsymbol{v}_{x_1-3,0}\boldsymbol{T}_{\mathbb{h}}\right)\boldsymbol{T}_{\mathbb{h}}^2 \\ &= \boldsymbol{v}_{x_1-3,0}\boldsymbol{T}_{\mathbb{h}}^3 \\ &\ldots \\ &= \boldsymbol{v}_{0,0}\boldsymbol{T}_{\mathbb{h}}^{x_1} \quad for\ x_1 = 0,1,2\ldots\end{aligned}$$

Furthermore,

$$\left(\boldsymbol{v}_{x_1,x_2-1}\boldsymbol{T}_{\mathbb{p}}\right)[j] = \sum_{i=0}^{4} \Pr\{Z(x_1,x_2)=j|Z(x_1,x_2-1)=i\}\Pr\{Z(x_1,x_2-1)=i\} = \boldsymbol{v}_{\boldsymbol{x_1,x_2}}[j]$$
$$for\ x_1 = 0,1,2,3\ldots; x_2 = 1,2,3\ldots; j = 0,1,2,3,4$$

so it follows that

$$\boldsymbol{v}_{x_1,x_2} = \boldsymbol{v}_{0,0}\boldsymbol{T}_{\mathbb{h}}^{x_1}\boldsymbol{T}_{\mathbb{p}}^{x_2}, \qquad for\ x_1, x_2 = 0,1,2\ldots$$

where

$$\boldsymbol{v}_{\boldsymbol{x_1,x_2}} = \begin{pmatrix}\Pr\{Z(x_1,x_2)=0\} & \Pr\{Z(x_1,x_2)=1\} & \Pr\{Z(x_1,x_2)=2\} & \Pr\{Z(x_1,x_2)=3\} & \Pr\{Z(x_1,x_2)=4\}\end{pmatrix}$$

If there are 0 chiasma events in both the inverted and proximal regions, then all tetrads must necessarily be in state *no bridge,* so

$$\boldsymbol{v}_{0,0} = \begin{pmatrix}1 & 0 & 0 & 0 & 0\end{pmatrix}$$

We can find $\boldsymbol{T}_{\mathbb{h}}$ and $\boldsymbol{T}_{\mathbb{p}}$ simply by considering the resulting tetrads for the four possible combinations of non-sister strand involvements when an extra chiasma event is added at the right end of the appropriate interval of a tetrad that is originally in the given configuration (see figures 2.8 and 2.9). This results in

$$\boldsymbol{T}_{\mathbb{h}} = \begin{pmatrix}0 & 1 & 0 & 0 & 0 \\ 1/4 & 1/2 & 0 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

$$\boldsymbol{T}_{\mathbb{p}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

(Cobbs (1978) uses a similar line of reasoning to find the probability distribution of parental ditype, tetratype, and non-parental ditype tetrads given the number of chiasma events in a single colinear interval; note the similarity between his matrix $\boldsymbol{T}$ and my $\boldsymbol{T}_{\mathbb{h}}$.)

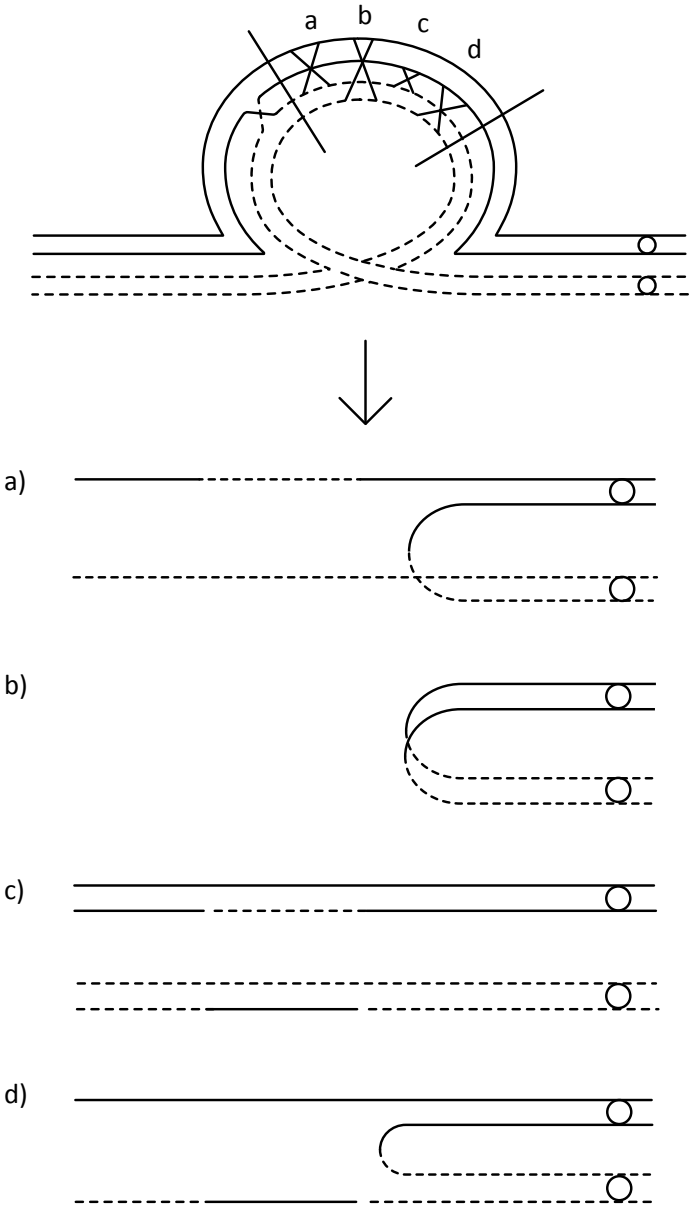


Figure 2.8: The leftmost X represent an anaphase I bridge, and the X's marked a, b, c, d represent the four possible non-sister chromatid pair involvements in a single chiasma event. The resulting configurations for each pair is shown below. The figure shows that $Pr\{Z(x_1 + 1{,}0) = 0 | Z(x_1, 0) = 1\} = 1/4$, $Pr\{Z(x_1 + 1{,}0) = 1 | Z(x_1, 0) = 1\} = 1/2$, *and* $Pr\{Z(x_1 + 1{,}0) = 3 | Z(x_1, 0) = 1\} = 1/4$, *which gives the second row (zero-indexed row 1) in matrix* $\boldsymbol{T}_{\mathbb{h}}$. *The remaining rows are found in the same manner.*

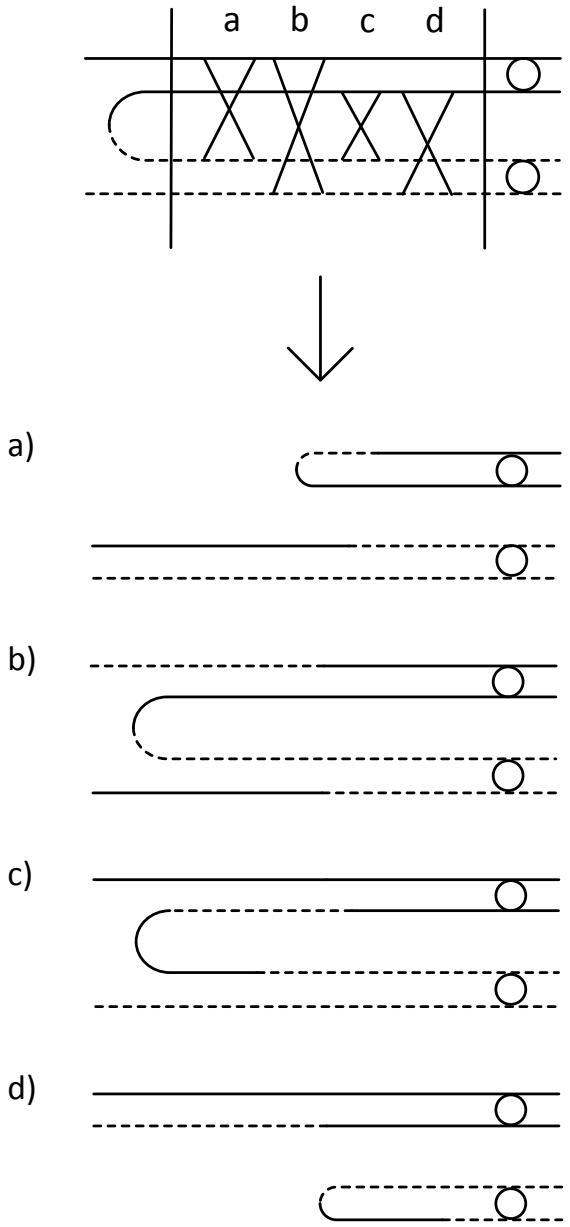


Figure 2.9: The effect of a single additional chiasma in the proximal region when the tetrad is originally in the configuration single anaphase I bridge. The figure shows that $Pr\{Z(x_1, x_2 + 1) = 1 | Z(x_1, x_2) = 1\} = 1/2$ *and* $Pr\{Z(x_1, x_2 + 1) = 2 | Z(x_1, x_2) = 1\} = 1/2$*, which give the second row (zero-indexed row 1) of matrix* $\boldsymbol{T}_{\mathbb{p}}$*. The remaining rows are found in the same manner.*

Now that we have an expression for the distribution of tetrad configurations given the number of chiasma events in each region, we only need to weight each configuration according to its proportion of unbalanced gametes. Since unbalanced chromatids in an anaphase I tetrad are retained in the polar bodies (meaning that anaphase I tetrads produce only balanced gametes), the weight vector becomes

$$\boldsymbol{w}_u = \begin{pmatrix} 0 \\ 0 \\ 1/2 \\ 1 \\ 1 \end{pmatrix}$$

so

$$\zeta_{x_1x_2} = \Pr\{unbalanced\ gamete | X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2\} = \boldsymbol{v}_{0,0}\boldsymbol{T}_{\mathbb{h}}^{x_1}\boldsymbol{T}_{\mathbb{p}}^{x_2}\boldsymbol{w}_u$$

For illustration, inserting $x_1, x_2 = 0,1,2$ generates the values in table 2 in Navarro et al. (1997):

$$\begin{aligned}
\zeta_{0,0} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{0}\boldsymbol{T}_{\mathbb{p}}^{0}\boldsymbol{w_u} = 0\\
\zeta_{0,1} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{0}\boldsymbol{T}_{\mathbb{p}}^{1}\boldsymbol{w_u} = 0\\
\zeta_{0,2} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{0}\boldsymbol{T}_{\mathbb{p}}^{2}\boldsymbol{w_u} = 0\\
\zeta_{1,0} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{1}\boldsymbol{T}_{\mathbb{p}}^{0}\boldsymbol{w_u} = 0\\
\zeta_{1,1} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{1}\boldsymbol{T}_{\mathbb{p}}^{1}\boldsymbol{w_u} = 1/4\\
\zeta_{1,2} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{1}\boldsymbol{T}_{\mathbb{p}}^{2}\boldsymbol{w_u} = 1/8\\
\zeta_{2,0} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{2}\boldsymbol{T}_{\mathbb{p}}^{0}\boldsymbol{w_u} = 1/4\\
\zeta_{2,1} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{2}\boldsymbol{T}_{\mathbb{p}}^{1}\boldsymbol{w_u} = 3/8\\
\zeta_{2,2} &= \boldsymbol{v_{0,0}}\boldsymbol{T}_{\mathbb{h}}^{2}\boldsymbol{T}_{\mathbb{p}}^{2}\boldsymbol{w_u} = 5/16
\end{aligned}$$

$\Pr\{balanced\ gamete | X_{\mathbb{h}} = x_1, X_{\mathbb{p}} = x_2\}$, denoted ρ_{ij} in Navarro et al., is of course equal to $1 - \zeta_{x_1x_2}$.

QED

2.3.10 Prologue to theorem 7

For the next theorem we will need some additional terminology. I will say that a chromatid is in *state* s_j: $\boldsymbol{r}_{s_j} = \boldsymbol{r}, t_{s_j} = t$ if it shows recombination pattern $\boldsymbol{r}$ *and* is in a tetrad with configuration t, and I will refer to $\boldsymbol{r}_{s_j}$ as the recombination pattern associated with state s_j, and t_{s_j} as the tetrad configuration associated with state s_j. The *statespace* is hence the set of all possible combinations of recombination patterns and tetrad configuration, with two exceptions. Firstly, I will collapse tetrad configurations *double anaphase I bridge* and *double anaphase II bridge* into the single configuration *double bridge.* This is because *double anaphase I bridge* and *double anaphase II bridge* tetrads both produce only unbalanced gametes, which means that for our purposes they are indistinguishable, and from matrix $\boldsymbol{T}_{\mathbb{p}}$ we can deduce that the two form an *absorbing class* – meaning that no number of additional chiasma events at the right end of the proximal region (or in the right region) can transform a double anaphase I or II bridge tetrad into any other configuration. That is to say, if we were to redefine the stochastic process in theorem 6 so that

$$Z(x_1, x_2) = \begin{cases} 0, & no\ bridge \\ 1, & single\ anaphase\ I\ bridge \\ 2, & single\ anaphase\ II\ bridge \\ 3, & double\ bridge \end{cases}$$

$$\boldsymbol{v_{0,0}} = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{T}_{\mathbb{h}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/4 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{T}_{\mathbb{p}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{w}_{\boldsymbol{u}} = \begin{pmatrix} 0 \\ 0 \\ 1/2 \\ 1 \end{pmatrix}$$

then the resulting value for ζ would be unaffected. Secondly, for reasons that will become clear in the proof for theorem 6, all states with unbalanced recombination patterns that also show recombination in any interval with index higher than $\mathfrak{d} + \mathfrak{h} - 1$ (i.e. the intervals in the proximal and rightmost regions) are excluded.

As an example, with two intervals in the inverted region and one in the proximal region and zero intervals in the distal and rightmost regions (i.e. $\mathfrak{d} = 0, \mathfrak{h} = 2, \mathfrak{p} = 1, \mathfrak{r} = 0$), which means that there is one loci of interest in the inverted region and none in either of the other regions (note that this implies that $\mathbb{i}_2 = \mathbb{p}_2 = \mathbb{p}$), we get the (arbitrarily numbered) statespace:

s_0: $\boldsymbol{r}_{s_0} = \{r_{s_0}(\mathbb{h}_0) = 0, r_{s_0}(\mathbb{h}_1) = 0, r_{s_0}(\mathbb{p}_2) = 0\}, t_{s_0} = 0$
s_1: $\boldsymbol{r}_{s_1} = \{r_{s_1}(\mathbb{h}_0) = 0, r_{s_1}(\mathbb{h}_1) = 0, r_{s_1}(\mathbb{p}_2) = 0\}, t_{s_1} = 1$
s_2: $\boldsymbol{r}_{s_2} = \{r_{s_2}(\mathbb{h}_0) = 0, r_{s_2}(\mathbb{h}_1) = 0, r_{s_2}(\mathbb{p}_2) = 0\}, t_{s_2} = 2$
s_3: $\boldsymbol{r}_{s_3} = \{r_{s_3}(\mathbb{h}_0) = 0, r_{s_3}(\mathbb{h}_1) = 0, r_{s_3}(\mathbb{p}_2) = 1\}, t_{s_3} = 0$
s_4: $\boldsymbol{r}_{s_4} = \{r_{s_4}(\mathbb{h}_0) = 0, r_{s_4}(\mathbb{h}_1) = 0, r_{s_4}(\mathbb{p}_2) = 1\}, t_{s_4} = 1$
s_5: $\boldsymbol{r}_{s_5} = \{r_{s_5}(\mathbb{h}_0) = 0, r_{s_5}(\mathbb{h}_1) = 0, r_{s_5}(\mathbb{p}_2) = 1\}, t_{s_5} = 2$
s_6: $\boldsymbol{r}_{s_6} = \{r_{s_6}(\mathbb{h}_0) = 0, r_{s_6}(\mathbb{h}_1) = 1, r_{s_6}(\mathbb{p}_2) = 0\}, t_{s_6} = 1$
s_7: $\boldsymbol{r}_{s_7} = \{r_{s_7}(\mathbb{h}_0) = 0, r_{s_7}(\mathbb{h}_1) = 1, r_{s_7}(\mathbb{p}_2) = 0\}, t_{s_7} = 2$
s_8: $\boldsymbol{r}_{s_8} = \{r_{s_8}(\mathbb{h}_0) = 0, r_{s_8}(\mathbb{h}_1) = 1, r_{s_8}(\mathbb{p}_2) = 0\}, t_{s_8} = 3$
s_9: $\boldsymbol{r}_{s_9} = \{r_{s_9}(\mathbb{h}_0) = 1, r_{s_9}(\mathbb{h}_1) = 0, r_{s_9}(\mathbb{p}_2) = 0\}, t_{s_9} = 1$
s_{10}: $\boldsymbol{r}_{s_{10}} = \{r_{s_{10}}(\mathbb{h}_0) = 1, r_{s_{10}}(\mathbb{h}_1) = 0, r_{s_{10}}(\mathbb{p}_2) = 0\}, t_{s_{10}} = 2$
s_{11}: $\boldsymbol{r}_{s_{11}} = \{r_{s_{11}}(\mathbb{h}_0) = 1, r_{s_{11}}(\mathbb{h}_1) = 0, r_{s_{11}}(\mathbb{p}_2) = 0\}, t_{s_{11}} = 3$
s_{12}: $\boldsymbol{r}_{s_{12}} = \{r_{s_{12}}(\mathbb{h}_0) = 1, r_{s_{12}}(\mathbb{h}_1) = 1, r_{s_{12}}(\mathbb{p}_2) = 0\}, t_{s_{12}} = 0$
s_{13}: $\boldsymbol{r}_{s_{13}} = \{r_{s_{13}}(\mathbb{h}_0) = 1, r_{s_{13}}(\mathbb{h}_1) = 1, r_{s_{13}}(\mathbb{p}_2) = 0\}, t_{s_{13}} = 1$
s_{14}: $\boldsymbol{r}_{s_{14}} = \{r_{s_{14}}(\mathbb{h}_0) = 1, r_{s_{14}}(\mathbb{h}_1) = 1, r_{s_{14}}(\mathbb{p}_2) = 0\}, t_{s_{14}} = 2$
s_{15}: $\boldsymbol{r}_{s_{15}} = \{r_{s_{15}}(\mathbb{h}_0) = 1, r_{s_{15}}(\mathbb{h}_1) = 1, r_{s_{15}}(\mathbb{p}_2) = 1\}, t_{s_{15}} = 0$
s_{16}: $\boldsymbol{r}_{s_{16}} = \{r_{s_{16}}(\mathbb{h}_0) = 1, r_{s_{16}}(\mathbb{h}_1) = 1, r_{s_{16}}(\mathbb{p}_2) = 1\}, t_{s_{16}} = 1$
s_{17}: $\boldsymbol{r}_{s_{17}} = \{r_{s_{17}}(\mathbb{h}_0) = 1, r_{s_{17}}(\mathbb{h}_1) = 1, r_{s_{17}}(\mathbb{p}_2) = 1\}, t_{s_{17}} = 2$

where

$$t = \begin{cases} 0, & no\ bridge \\ 1, & single\ anaphase\ I\ bridge \\ 2, & single\ anaphase\ II\ bridge \\ 3, & double\ bridge \end{cases}$$

Note that there are no states for unbalanced patterns in *no bridge* tetrads or balanced patterns in *double bridge* tetrads; this is because if the tetrad is in the *no bridge* configuration, then none of the patterns associated with the tetrad can be unbalanced, and, similarly, if the tetrad is in the *double*

bridge configuration, then none of the patterns associated with the tetrad can be balanced, so these combinations are impossible and hence not included in the statespace. The number of states – or, equivalently, the *size* of the statespace – will henceforth be denoted $\mathfrak{s}$; in the example above $\mathfrak{s} = 18$. In the main program, the statespace is generated automatically in the class *Karyotype* method *generate_statespace.*

Similarly to theorem 6, we can now describe the basic structure of theorem 7 as consisting of three distinct parts: a stochastic process with one transition matrix for each interval gives the conditional state probability distribution; conditioning on the number of chiasma events in each intervals then gives the unconditional state probability distribution; and weighting the unconditional state probability distribution appropriately finally extracts the information of interest. As before, I will first give the theorem in full for easier reference, before providing a proof.

2.3.11 Theorem 7: Recombination in paracentric heterokaryotypes with linear meiosis

Let the statespace be as defined above. The probability of observing a gamete with recombination pattern $\boldsymbol{r}$ for a paracentric inversion heterokaryotype with linear meiosis is now given by

$$\Pr\{\boldsymbol{R}_{paracentric\ linear} = \boldsymbol{r}\} = \boldsymbol{v}\boldsymbol{w}_{\boldsymbol{r}}^{T}$$

where

$$\boldsymbol{w}_{\boldsymbol{r}}[i] = \begin{cases} 2, & \boldsymbol{r}_{s_i} = \boldsymbol{r}; \varphi(\boldsymbol{r}_{s_i}) = 1; t_{s_i} = 1 \\ 0, & \boldsymbol{r}_{s_i} = \boldsymbol{r}; \varphi(\boldsymbol{r}_{s_i}) = 0; t_{s_i} = 1 \\ 1, & \boldsymbol{r}_{s_i} = \boldsymbol{r}; t_{s_i} \neq 1 \\ 0, & \boldsymbol{r}_{s_i} \neq \boldsymbol{r} \end{cases} \qquad for\ i = 0,1,2 \dots \mathfrak{s} - 1$$

$$\varphi(\boldsymbol{r}) = \begin{cases} 1, & \sum_{i=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} r(\mathbb{h}_i)\ is\ even \\ 0, & \sum_{i=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} r(\mathbb{h}_i)\ is\ odd \end{cases}$$

$$\boldsymbol{v} = \sum_{x_0=0}^{\infty} \sum_{x_1=0}^{\infty} \dots \sum_{x_{\mathfrak{n}-1}=0}^{\infty} \mathbb{D}(x_0, x_1 \dots x_{\mathfrak{n}-1})\, \boldsymbol{v}_{x_0,x_1\dots x_{\mathfrak{n}-1}}$$

$$\mathbb{D}(x_0, x_1 \dots x_{\mathfrak{n}-1}) = \boldsymbol{\pi}\left(\prod_{k=0}^{\mathfrak{d}-1} \boldsymbol{D}_{\mathbb{d}_k}(x_k)\right)\boldsymbol{Q}^{1-\alpha}\left(\prod_{k=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} \boldsymbol{D}_{\mathbb{h}_k}(x_k)\right)\boldsymbol{Q}^{1-\alpha}\left(\prod_{k=\mathfrak{d}+\mathfrak{h}}^{\mathfrak{d}+\mathfrak{h}+\mathfrak{p}-1} \boldsymbol{D}_{\mathbb{p}_k}(x_k)\right)\left(\prod_{k=\mathfrak{d}+\mathfrak{h}+\mathfrak{p}}^{\mathfrak{n}-1} \boldsymbol{D}_{\mathbb{r}_k}(x_k)\right)\boldsymbol{1}^{T}$$

$$\boldsymbol{\pi} = (\pi_0 \quad \pi_1 \quad \pi_2 \quad \dots \quad \pi_m)$$

$$\pi_l = \frac{\sum_{q=l}^{m} \gamma_q}{\sum_{q=0}^{m}(q+1)\gamma_q} \quad for\ l = 0,1,2 \dots m$$

$$\boldsymbol{Q}[i,j] = \pi_j \quad for\ i,j = 0,1,2 \dots m$$

$$\boldsymbol{v}_{\boldsymbol{x_0,x_1\dots x_{n-1}}} = \boldsymbol{v}_{\boldsymbol{0,0,\dots,0}}\left(\prod_{k=0}^{\mathfrak{d}-1} \boldsymbol{P}_{\mathbb{d}_k}^{x_k}\right)\left(\prod_{k=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} \boldsymbol{P}_{\mathbb{h}_k}^{x_k}\right)\left(\prod_{k=\mathfrak{d}+\mathfrak{h}}^{\mathfrak{d}+\mathfrak{h}+\mathfrak{p}-1} \boldsymbol{P}_{\mathbb{p}_k}^{x_k}\right)\left(\prod_{k=\mathfrak{d}+\mathfrak{h}+\mathfrak{p}}^{\mathfrak{n}-1} \boldsymbol{P}_{\mathbb{r}_k}^{x_k}\right)$$

$$\boldsymbol{D}_{\mathbb{i}_k}(0)[i,j] = \begin{cases} \frac{(d_k\lambda_k)^{i-j}e^{-d_k\lambda_k}}{(i-j)!}e^{-d_k\mu_k}, & i \geq j \\ 0, & otherwise \end{cases} \qquad for\ i,j = 0,1,2 \dots m$$

$$\boldsymbol{D}_{\mathbb{i}_k}(x)[i,j] = \sum_{l=0}^{x-1} \sum_{n=x-l-1}^{z} \sum_{q=j}^{m} g_n(x-l-1)\gamma_q \frac{e^{-d_k(\lambda_k+\mu_k)}(d_k[\lambda_k+\mu_k])^h}{h!} \binom{h}{l} \left(\frac{\mu_k}{\lambda_k+\mu_k}\right)^l \left(\frac{\lambda_k}{\lambda_k+\mu_k}\right)^{h-l}$$

$$+\delta_{\{i\geq j\}} \frac{e^{-d_k\mu_k}(d_k\mu_k)^x}{x!} \frac{e^{-d_k\lambda_k}(d_k\lambda_k)^{i-j}}{(i-j)!}, \qquad for\ x = 1,2,3\ldots;\ i,j = 0,1,2\ldots m$$

$z = (x-l-1)(m+1)$

$h = i+1+l+n+q-j$

$$g_n(s) = \begin{cases} \sum_{k=s-1}^{n-1} g_k(s-1)\gamma_{n-1-k}, & n \geq s \neq 0 \\ 1, & n = s = 0 \\ 0, & n \neq 0; s = 0 \\ 0, & n < s \end{cases}$$

$\lambda_k = E\left[C''_{\mathbb{i}_k}\right];\ \mu_k = E[X'_{\mathbb{i}_k}]$

$$\boldsymbol{v}_{0,0,\ldots,0}[j] = \begin{cases} 1, & r_{s_j}(\mathbb{i}_l) = 0\ for\ all\ l; t_{s_j} = 0 \\ 0, & otherwise \end{cases} \qquad for\ j = 0,1,2\ldots \mathfrak{s}-1$$

$$\boldsymbol{P}_{\mathbb{d}_k}[i,j] = \begin{cases} \frac{1}{2}, & \theta(i,j,k) = 1; t_{s_i} = t_{s_j} = 0 \\ 0, & otherwise \end{cases}$$

$for\ k = 0,2,3\ldots \mathfrak{d}-1; i,j = 0,1,2\ldots \mathfrak{s}-1$

$$\boldsymbol{P}_{\mathbb{h}_k}[i,j] = \begin{cases} 1/2, & \theta(i,j,k) = 1;\ t_{s_i} = 0; t_{s_j} = 1 \\ 1/4, & \theta(i,j,k) = 1; \varphi\left(\boldsymbol{r}_{s_j}\right) = 1; t_{s_i} = 1; t_{s_j} = 0 \\ 1/4, & \theta(i,j,k) = 1;\ t_{s_i} = 1; t_{s_j} = 1 \\ 1/4, & \theta(i,j,k) = 1;\ \varphi\left(\boldsymbol{r}_{s_j}\right) = 0; t_{s_i} = 1; t_{s_j} = 3 \\ 1/2, & \theta(i,j,k) = 1;\ t_{s_i} = 3; t_{s_j} = 1 \\ 0, & otherwise \end{cases}$$

$for\ k = \mathfrak{d}+1, \mathfrak{d}+2\ldots \mathfrak{d}+\mathfrak{h}-1; i,j = 0,1,2\ldots \mathfrak{s}-1$

$$\boldsymbol{P}_{\mathbb{p}_k}[i,j]=\begin{cases}1/2, & \theta(i,j,k)=1;\ t_{s_i}=t_{s_j}=0\\ 1/2, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=0;\ r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)\ for\ all\ l\ ;t_{s_i}=t_{s_j}=1\\ 1/2, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=0;\ r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)\ for\ all\ l;t_{s_i}=1;\ t_{s_j}=2\\ 1/4, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=1;\theta(i,j,k)=1;t_{s_i}=t_{s_j}=1\\ 1/4, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=1;\theta(i,j,k)=1;t_{s_i}=1;t_{s_j}=2\\ 1, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=0;\ r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)\ for\ all\ l\ ;t_{s_i}=2;t_{s_j}=1\\ 1/2, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=1;\ \theta(i,j,k)=1\ ;t_{s_i}=2;t_{s_j}=1\\ 1, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=0;\ r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)\ for\ all\ l\ ;t_{s_i}=t_{s_j}=3\\ 0, & otherwise\end{cases}$$

$$for\ k=\mathfrak{d}+\mathfrak{h},\mathfrak{d}+\mathfrak{h}+1\ldots\mathfrak{d}+\mathfrak{h}+\mathfrak{p}-1;i,j=0,1,2\ldots\mathfrak{s}-1$$

$$\boldsymbol{P}_{\mathbb{r}_k}[i,j]=\begin{cases}1, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=0;\ r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)\ for\ all\ l;t_{s_i}=t_{s_j}\\ 1/2, & \varphi(\boldsymbol{r}_{s_i})=\varphi\left(\boldsymbol{r}_{s_j}\right)=1;\theta(i,j,k)=1;t_{s_i}=t_{s_j}\\ 0, & otherwise\end{cases}$$

$$for\ k=\mathfrak{d}+\mathfrak{h}+\mathfrak{p},\mathfrak{d}+\mathfrak{h}+\mathfrak{p}+1\ldots\mathfrak{n}-1;i,j=0,1,2\ldots\mathfrak{s}-1$$

$$\theta(i,j,k)=\begin{cases}1, & r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)\ for\ l\neq k;r_{s_i}(\mathbb{i}_l)=r_{s_j}(\mathbb{i}_l)=0\ for\ l>k\\ 0, & otherwise\end{cases}$$

Proof:
Consider a stochastic process $\{S(x_0,x_1,x_2\ldots x_{\mathfrak{n}-1})\}_{(x_0,x_1,x_2\ldots x_{\mathfrak{n}-1})\in\mathbb{Z}_{\geq 0}}$ where $S(x_0,x_1,x_2\ldots x_{\mathfrak{n}-1})$ represents the state of a chromatid (recombination pattern and tetrad configuration) when there are x_0 chiasma events in $\mathbb{i}_0$, x_1 chiasma events in $\mathbb{i}_1$, etc, so that

$$S(x_0,x_1,x_2\ldots x_{\mathfrak{n}-1})=j,\qquad the\ process\ is\ in\ state\ s_j;\ for\ j=0,1,2\ldots\mathfrak{s}-1$$

If we now define the vector $\boldsymbol{v}_{0,0,\ldots,0}$ so that

$$\boldsymbol{v}_{0,0,\ldots,0}[j]=\Pr\{S(0,0,\ldots,0)=j\}\qquad j=0,1,2\ldots\mathfrak{s}-1$$

and the matrices $\boldsymbol{P}_{\mathbb{i}_k}, k=0,1,2\ldots\mathfrak{n}-1;\ \mathbb{i}=\mathbb{d},\mathbb{h},\mathbb{p},\mathbb{r}$, so that

$$\boldsymbol{P}_{\mathbb{i}_k}[i,j]=\Pr\left\{S(x_0,x_1,\ldots,x_k+1,0,0\ldots,0)=j\middle|S\left((x_0,x_1,\ldots,x_k,0,0\ldots,0)\right)=i\right\}$$
$$for\ \ x_0,x_1\ldots x_k=0,1,2\ldots;\ i,j=0,1,2\ldots\mathfrak{s}-1$$

then using the same argument as in the proof for theorem 6, we get

$$\boldsymbol{v}_{x_0,x_1,\ldots,x_{\mathfrak{n}-1}}=\boldsymbol{v}_{0,0\ldots0}\boldsymbol{P}_{\mathbb{i}_0}^{x_0}\boldsymbol{P}_{\mathbb{i}_1}^{x_1}\ldots\boldsymbol{P}_{\mathbb{i}_{\mathfrak{n}-1}}^{x_{\mathfrak{n}-1}}$$

where

$$\boldsymbol{v}_{x_0,x_1,\ldots,x_{\mathfrak{n}-1}}[j] = \Pr\{S(x_0, x_1, \ldots, x_{\mathfrak{n}-1}) = j\} \quad for\ j = 0,1,2 \ldots \mathfrak{s} - 1;\ x_0, x_1, \ldots, x_{\mathfrak{n}-1} = 0,1,2 \ldots$$

i.e. the $\boldsymbol{P}_{\mathbb{i}_k}$ matrices are transition matrices in the same sense that the $\boldsymbol{T}_{\mathbb{i}}$ matrices are in theorem 6. To get a more intuitive understanding of these matrices, note that the element *i,j* in each $\boldsymbol{P}_{\mathbb{i}_k}$ gives the probability of observing a chromatid with recombination pattern $\boldsymbol{r}_{s_j}$ in a tetrad with configuration t_{s_j} when considering an additional chiasma event in interval $\mathbb{i}_k$, given that the chromatid had recombination pattern $\boldsymbol{r}_{s_i}$ and was in a tetrad with configuration t_{s_i} before considering that chiasma event, assuming that there are no chiasma events in any intervals with index higher than *k*. You might (erroneously) conclude from this that the validity of the theorem rests on the rather restrictive assumption that the chiasma events occur in a strict temporal sequence from left to right, so that, say, no chiasma events can occur in $\mathbb{i}_0$ at a time later than the occurrence of any chiasma event in $\mathbb{i}_1$. To see why this is not so, imagine that all the chiasma events in our region of interest have already occurred, in whatever temporal sequence, and that we after-the-fact assign to a miniscule daemon the task of calculating the state probability distribution given the number of chiasma events in each interval. Being miniscule, our daemon decides to approach this task as a journey rightwards from the leftmost to the rightmost boundary of the region of interest, during which he abides by the following procedure:

1. At the beginning of the journey, assume that no chiasma events have occurred in any of the intervals, and estimate the state probability distribution accordingly
2. Every time a chiasma event is encountered, update the estimate by performing a linear transformation on the former estimate using the transition matrix for the current interval.
3. At the end of the journey, report the final estimate.

The 'initial estimate' of the conditional state probability distribution in point 1 is by definition $\boldsymbol{v}_{0,0,\ldots,0}$, and performing in sequence x_0 linear transformation using $\boldsymbol{P}_{\mathbb{i}_0}$, then x_1 linear transformations using $\boldsymbol{P}_{\mathbb{i}_1}$, and so on, results in a 'final estimate' $\boldsymbol{v}_{0,0\ldots0}\boldsymbol{P}_{\mathbb{i}_0}^{x_0}\boldsymbol{P}_{\mathbb{i}_1}^{x_1} \ldots \boldsymbol{P}_{\mathbb{i}_{\mathfrak{n}-1}}^{x_{\mathfrak{n}-1}}$ which is equal to $\boldsymbol{v}_{x_0,x_1,\ldots,x_{\mathfrak{n}-1}}$ regardless of the temporal sequence in which the chiasma events originally occurred.

This thought experiment also provides an explanation as to why we without losing information can simplify the statespace by removing the states with unbalanced recombination patterns that also show recombination in any of the intervals rightwards of the inverted region. Once the daemon has left the inverted region, it has already determined which chromatids will be unbalanced (because no number of chiasma events outside of the inverted region can 'rebalance' an unbalanced chromatid), and the only relevant additional information about those chromatids is for our purpose the configurations of their associated tetrads. The transition matrices for intervals rightwards of the inverted region therefore perform the appropriate transformation on the tetrad of states of with unbalanced patterns, but leaves the pattern itself unchanged (e.g. the condition $\varphi(\boldsymbol{r}_{s_i}) = \varphi\left(\boldsymbol{r}_{s_j}\right) = 0;\ r_{s_i}(\mathbb{i}_l) = r_{s_j}(\mathbb{i}_l)\ for\ all\ l\ ; t_{s_i} = 2; t_{s_j} = 1$ in $\boldsymbol{P}_{\mathbb{p}_k}$).

The *Karyotype* method *calculate_transition_matrices* in the main program generates the $\boldsymbol{P}_{\mathbb{i}_k}$ matrices by looping over all states in the statespace (`state_from` in the excerpt below), and checking its tetrad configuration (`if state_from[-1] == 0:`, etc), whether it is balanced (`if is_balanced(state_from):`) and whether it shows recombination in the current interval (`if state_from[i] == False:`). The states are stored as lists of length $\mathfrak{n} + 1$ where the first $\mathfrak{n}$ elements indicate recombination (*True*) or nonrecombination (*False*) for each of the $\mathfrak{n}$ intervals, and the last element indicate the tetrad configuration as a number from 0 to 3 (see the definition of the statespace above). By considering the effect of each the four possible chromatid involvement pairs in the additional chiasma event, the four possible states (*state_to1, state_to2,* etc) to which the original

state can be transformed are found and assigned an equal probability of 0.25 (because of the assumption of no *chromatid* interference). Each state's unique index is then found (`from_index = statespace.index(state_from),` etc), and used to set the element of the matrix to the appropriate value. The excerpt below show the part of the loop that calculates a matrix for interval with index `i` in the inverted region (the matrices for the other regions are handled separately; `i` in the code is equivalent to *k* in the theorem). The first line (`if sum(state_from[i+1:intervals_n]) == 0:`) tell the program to ignore states that show recombination in intervals to the right of *i*. In the case of a *state_from* with configuration *no bridge,* the procedure is relatively straightforward: the *state_from* chromatid must necessarily be balanced, the new configuration will always be *single anaphase 1 bridge* (figure 2.7), and the new chromatid will show recombination or non-recombination with equal probability, regardless of whether the original chromatid (associated with *state_from*) showed recombination (see matrix $\boldsymbol{P}$ in the proof of Mather's equation, section 2.1.7). The case of a *state_from* with *anaphase 1 bridge* configuration (`elif state_from[-1] == 1:`) is more complicated, but it can in general be solved through the following line of thinking. First, consider whether or not the chromatid is balanced (`if is_balanced(state_from):`). This determines whether or not the chromatid is involved in the bridge. Then consider what happens with the recombination status and the tetrad configuration for each of the four possible chromatid pair involvements (use figure 2.8 as aid). The transition probabilities for the other tetrad configurations and regions are found in the same way. The comment after each assignment of recombination status to the new state (e.g. *#1* in `state_to1[i] = True #1`) indicate which of the conditions listed in the proof that corresponds to that transition, numbered so that the top entry is number 1, and lower entries have increasingly higher numbers. For example, the line `state_to1[i] = True #1` below, indicate that this transition corresponds to the first condition for a matrix in the inverted region, which is

$$\theta(i,j,k) = 1;\ t_{s_i} = 0; t_{s_j} = 1$$

The definitions of $\boldsymbol{P}_{\mathbb{I}_k}$ given in the proof can therefore be thought as condensed summaries of the calculations in the *calculate_transition_matrices* method, which are more intuitive, but less concise.

The algorithm can easily be adapted to account for chromatid interference as well as chiasma interference; simply expand the statespace to include a binary indicator of whether or not the chromatid associated with each state was involved in the previous chiasma event, and multiply each transition probability with an appropriate factor based on this information.

```
…
if sum(state_from[i+1:intervals_n]) == 0:
       if state_from[-1] == 0: # no bridge
              state_to1 = copy.copy(state_from)
              state_to2 = copy.copy(state_from)
              state_to1[i] = True #1
              state_to2[i] = False #1
              state_to1[-1] = 1
              state_to2[-1] = 1
              from_index = statespace.index(state_from)
              to_index1 = statespace.index(state_to1)
              to_index2 = statespace.index(state_to2)
              matrix[from_index][to_index1] = 0.5
              matrix[from_index][to_index2] = 0.5
       elif state_from[-1] == 1: # single a1 bridge
              state_to1 = copy.copy(state_from)
              state_to2 = copy.copy(state_from)
              state_to3 = copy.copy(state_from)
              state_to4 = copy.copy(state_from)

              if is_balanced(state_from): #balanced
```

```
                if state_from[i] == False:
                        state_to1[i] = False #2
                        state_to1[-1] = 0

                        state_to2[i] = True #3
                        state_to2[-1] = 1

                        state_to3[i] = False #3
                        state_to3[-1] = 1

                        state_to4[i] = True #4
                        state_to4[-1] = 3
                else:
                        state_to1[i] = True #2
                        state_to1[-1] = 0

                        state_to2[i] = True #3
                        state_to2[-1] = 1

                        state_to3[i] = False #3
                        state_to3[-1] = 1

                        state_to4[i] = False #4
                        state_to4[-1] = 3

        else: #unbalanced
                if state_from[i] == False:
                        state_to1[i] = True #2
                        state_to1[-1] = 0

                        state_to2[i] = True #3
                        state_to2[-1] = 1

                        state_to3[i] = False #3
                        state_to3[-1] = 1

                        state_to4[i] = False #4
                        state_to4[-1] = 3

                else:
                        state_to1[i] = False #2
                        state_to1[-1] = 0

                        state_to2[i] = False #3
                        state_to2[-1] = 1

                        state_to3[i] = True #3
                        state_to3[-1] = 1

                        state_to4[i] = True #4
                        state_to4[-1] = 3

        from_index = statespace.index(state_from)
        to_index1 = statespace.index(state_to1)
        to_index2 = statespace.index(state_to2)
        to_index3 = statespace.index(state_to3)
        to_index4 = statespace.index(state_to4)

        matrix[from_index][to_index1] = 0.25
        matrix[from_index][to_index2] = 0.25
        matrix[from_index][to_index3] = 0.25
        matrix[from_index][to_index4] = 0.25


elif state_from[-1] == 3: # double bridge
        state_to1 = copy.copy(state_from)
        state_to2 = copy.copy(state_from)
        state_to1[i] = True #5
        state_to1[-1] = 1
        state_to2[i] = False #5
```

```
state_to2[-1] = 1

from_index = statespace.index(state_from)

to_index1 = statespace.index(state_to1)
to_index2 = statespace.index(state_to2)

matrix[from_index][to_index1] = 0.5
matrix[from_index][to_index2] = 0.5
```

If there are no chiasma events in any of the intervals, then there can be no chromatids showing recombination in any of the intervals, and all tetrads must be in configuration *no bridge.* Hence,

$$\boldsymbol{v}_{\mathbf{0},\mathbf{0},\ldots,\mathbf{0}}[j] = \begin{cases} 1, & r_{s_j}(\mathbb{i}_l) = 0 \; for \; all \; l; t_{s_j} = 0 \\ 0, & otherwise \end{cases} \qquad for \; j = 0,1,2 \ldots \mathfrak{s} - 1$$

and

$$\boldsymbol{v}_{\boldsymbol{x_0},\boldsymbol{x_1},\ldots,\boldsymbol{x_{\mathfrak{n}-1}}} = \boldsymbol{v}_{\mathbf{0},\mathbf{0}\ldots\mathbf{0}} \boldsymbol{P}_{\mathbb{i}_\mathbf{0}}^{x_0} \boldsymbol{P}_{\mathbb{i}_\mathbf{1}}^{x_1} \ldots \boldsymbol{P}_{\mathbb{i}_{\mathbf{n}-\mathbf{1}}}^{x_{\mathfrak{n}-1}}$$

Now that we know how to calculate $\boldsymbol{v}_{\boldsymbol{x_0},\boldsymbol{x_1},\ldots,\boldsymbol{x_{\mathfrak{n}-1}}}$, we can get the unconditional state probability distribution, $\boldsymbol{v}$, by conditioning on the number of chiasma events in each interval. That is,

$$\boldsymbol{v} = \sum_{x_0=0}^{\infty} \sum_{x_1=0}^{\infty} \ldots \sum_{x_{\mathfrak{n}-1}=0}^{\infty} \Pr\{X_{\mathbb{i}_0} = x_0, X_{\mathbb{i}_1} = x_1, \ldots, X_{\mathbb{i}_{\mathfrak{n}-1}} = x_{\mathfrak{n}-1},\} \boldsymbol{v}_{\boldsymbol{x_0},\boldsymbol{x_1}\ldots\boldsymbol{x_{\mathfrak{n}-1}}}$$

My next goal is therefore to show that

$$\begin{aligned} &\Pr\{X_{\mathbb{i}_0} = x_0, X_{\mathbb{i}_1} = x_1, \ldots, X_{\mathbb{i}_{\mathfrak{n}-1}} = x_{\mathfrak{n}-1},\} = \mathbb{D}(x_0, x_1 \ldots x_{\mathfrak{n}-1}) \\ &= \boldsymbol{\pi} \left(\prod_{k=0}^{\mathfrak{d}-1} \boldsymbol{D}_{\mathbb{d}_k}(x_k) \right) \boldsymbol{Q}^{1-\alpha} \left(\prod_{k=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} \boldsymbol{D}_{\mathbb{h}_k}(x_k) \right) \boldsymbol{Q}^{1-\alpha} \left(\prod_{k=\mathfrak{d}+\mathfrak{h}}^{\mathfrak{d}+\mathfrak{h}+\mathfrak{p}-1} \boldsymbol{D}_{\mathbb{p}_k}(x_k) \right) \left(\prod_{k=\mathfrak{d}+\mathfrak{h}+\mathfrak{p}}^{\mathfrak{n}-1} \boldsymbol{D}_{\mathbb{r}_k}(x_k) \right) \mathbf{1}^T \end{aligned}$$

where

$$\boldsymbol{D}_{\mathbb{i}_k}(0)[i,j] = \begin{cases} \dfrac{(d_k\lambda_k)^{i-j} e^{-d_k\lambda_k}}{(i-j)!} e^{-d_k\mu_k}, & i \geq j \\ 0, & otherwise \end{cases} \qquad i,j = 0,1,2 \ldots m$$

$$\boldsymbol{D}_{\mathbb{i}_k}(x)[i,j] = \sum_{l=0}^{x-1} \sum_{n=x-l-1}^{z} \sum_{q=j}^{m} g_n(x-l-1)\gamma_q \frac{e^{-d_k(\lambda_k+\mu_k)}(d_k[\lambda_k+\mu_k])^h}{h!} \binom{h}{l} \left(\frac{\mu_k}{\lambda_k+\mu_k}\right)^l \left(\frac{\lambda_k}{\lambda_k+\mu_k}\right)^{h-l}$$

$$+\delta_{\{i\geq j\}} \frac{e^{-d_k\mu_k}(d_k\mu_k)^x}{x!} \frac{e^{-d_k\lambda_k}(d_k\lambda_k)^{i-j}}{(i-j)!}, \qquad for \; x = 1,2,3 \ldots; \; i,j = 0,1,2 \ldots m$$

$$z = (x-l-1)(m+1)$$

$$h = i + 1 + l + n + q - j$$

$$g_n(s) = \begin{cases} \sum_{k=s-1}^{n-1} g_k(s-1)\gamma_{n-1-k}, & n \geq s \neq 0 \\ 1, & n = s = 0 \\ 0, & n \neq 0; s = 0 \\ 0, & n < s \end{cases}$$

Element *i,j* of the $\boldsymbol{D}_{\mathbb{i}_k}(x), k = 0,1,2 \ldots \mathfrak{n} - 1; x = 0,1,2 \ldots$ matrices gives the probability of observing *x* chiasma events in $\mathbb{i}_k$ *and* phase *j* at the right boundary of $\mathbb{i}_k$ *given* phase *i* at the left boundary of $\mathbb{i}_k$ under the general chiasma interference model, i.e.

$$\boldsymbol{D}_{\mathbb{i}_k}(x)[i,j] = \Pr\{X_{\mathbb{i}_k} = x, Q_{\mathbb{i}_k^r} = j | Q_{\mathbb{i}_k^l} = i\}$$

$\boldsymbol{D}_{\mathbb{i}_k}(0)$ is therefore equivalent to $\boldsymbol{H}_{\mathbb{i}_k}$ in theorem 1. My expression for $\boldsymbol{D}_{\mathbb{i}_k}(x)[i,j], \ x = 1,2,3 \ldots$ under the general interference model builds on the corresponding expression for homokaryotypes under the combined counting model ($\gamma_m = 1; \gamma_q = 0 \ for \ q \neq m$) in Copenhaver et al. (2002). Using my notation and including the inversion factors (*d;* not present in Copenhaver et al.), the latter is given by

$$\boldsymbol{D}_{\mathbb{i}_k}(x)[i,j] = \sum_{l=0}^{x-1} \frac{e^{-d_k(\lambda_k+\mu_k)}(d_k\lambda_k + d_k\mu_k)^h}{h!} \binom{h}{l} \left(\frac{\mu_k}{\mu_k + \lambda_k}\right)^l \left(\frac{\lambda_k}{\mu_k + \lambda_k}\right)^{h-l} + \delta_{\{i \geq j\}} \frac{e^{-d_k\mu_k}(d_k\mu_k)^x}{x!} \frac{e^{-d_k\lambda_k}(d_k\lambda_k)^{i-j}}{(i-j)!}$$

where

$$h = i + 1 + l + (m+1)(x - l - 1) + m - j$$

and is derived as follows. If there are in total *x* chiasma events in the interval, then between 0 and *x* of them must be X'events. If there are $l \ (0 \leq l < x) \ X'$ events in the interval, then the total number of C events (called *Poisson events* in Copenhaver et al.) must be *h* = ($i \ O''$events before the first X'' event) + (the first X'' event) + ($l \ X'$ events [in no particular spatial order]) + ($x - l - 1$ times $m + 1$ additional C''events) + (the final $m - j \ O''$ events so as to end up in phase *j* at the right boundary), in total $h = i + 1 + l + (m+1)(x - l - 1) + m - j$ events. We can therefore get an expression for $\Pr\left\{X_{\mathbb{i}_k} = x, Q_{\mathbb{i}_k^r} = j \middle| Q_{\mathbb{i}_k^l} = i\right\}$ by summing over all possible values of *l*, as such:

$$\Pr\left\{X_{\mathbb{i}_k} = x, Q_{\mathbb{i}_k^r} = j \middle| Q_{\mathbb{i}_k^l} = i\right\} = \sum_{l=0}^{x-1} \Pr\{C_{\mathbb{i}_k} = h, X'_{\mathbb{i}_k} = l\} + \delta_{\{i \geq j\}} \frac{e^{-d_k\mu_k}(d_k\mu_k)^x}{x!} \frac{e^{-d_k\lambda_k}(d_k\lambda_k)^{i-j}}{(i-j)!}$$

The final term takes into account the special case where all the *x* chiasma events are X' events (which means that there are no X'' events). This term is non-zero only if $i \geq j$ (otherwise there must be at least one X'' events, which is a contradiction), and is in that case equal to the probability of observing *x* X' events and $i - j \ O''$ events, which is just the Poisson probabilities multiplied (Copenhaver et al. handle this special case differently). Since

$\Pr\{C_{\mathbb{I}_k} = h, X'_{\mathbb{I}_k} = l\} = \Pr\{C_{\mathbb{I}_k} = h\} \Pr\{X'_{\mathbb{I}_k} = l | C_{\mathbb{I}_k} = h\}$, and $\Pr\{C_{\mathbb{I}_k} = h\}$ and $\Pr\{X'_{\mathbb{I}_k} = l | C_{\mathbb{I}_k} = h\}$ are given by the Poisson and binomial distributions, respectively, we get

$$\begin{aligned}\Pr\{C_{\mathbb{I}_k} = h, X'_{\mathbb{I}_k} = l\} &= \Pr\{C_{\mathbb{I}_k} = h\} \Pr\{X'_{\mathbb{I}_k} = l \big| C_{\mathbb{I}_k} = h\} \\ &= \frac{e^{-d_k(\lambda_k+\mu_k)}(d_k\lambda_k + d_k\mu_k)^h}{h!} \binom{h}{l} \left(\frac{\mu_k}{\mu_k + \lambda_k}\right)^l \left(\frac{\lambda_k}{\mu_k + \lambda_k}\right)^{h-l}\end{aligned}$$

where

$$h = i + 1 + l + (m+1)(x - l - 1) + m - j$$

For the general model, the number of O'' events between each X'' event is drawn from a probability distribution, so we can no longer assume that the $(m+1)(x-l-1) - th\ C''$ event after the first X'' event will be the $(x-l-1) - th\ X''$ event in the interval. To get around this problem, I first define a function $g_n(s)$ that gives the probability that the *n-th* C''event after an X'' event is the *s-th* X'' event after that X'' event. This is achieved by defining the base case $g_0(0) = 1$, which means that the zeroth C'' event is the zeroth X'' event with probability 1, and then recursively conditioning on which X'' event is the last before the *s-th*, i.e.

$$g_n(s) = \begin{cases} \sum_{k=s-1}^{n-1} g_k(s-1)\gamma_{n-1-k}, & n \geq s \neq 0 \\ 1, & n = s = 0 \\ 0, & n \neq 0; s = 0 \\ 0, & n < s \end{cases}$$

Note the difference between $g_n(s)$ and b_n from theorem 1: the former gives the probability that the *n-th* C''event after an X'' event is strictly the *s-th* X'' event after that X'' event, whereas the latter gives the probability that the *n-th* C''event after an X'' event is any X'' event. In the main program, the $g_n(s)$ values are computed using a dynamic programming approach and stored in a two-dimensional array for easy reference.

Now that we have an expression for $g_n(x-l-1)$, i.e. the probability that the *n-th* C'' event is the last X'' event in the interval, we can sum over all possible values of *n* weighted by $g_n(x-l-1)$, and then (in a nested sum) over all possible phases *q* after the last X'' event in the interval weighted by γ_q (from the intervening O'' events distribution). The resulting values for *h,* the total number of C events in the interval , is now h = (*i* O''events before the first X'' event) + (the first X'' event) + (*l* X' events [in no particular spatial order]) + (*n* additional C''events up to and including the last X'' event) + (the final *q - j* O'' events so as to end up in phase *j* at the right boundary), in short

$$h = i + 1 + l + n + q - j$$

The final expression for $\Pr\left\{X_{\mathbb{I}_k} = x, Q_{\mathbb{I}_k^r} = j \middle| Q_{\mathbb{I}_k^l} = i\right\} = \boldsymbol{D}_{\mathbb{I}_k}(x)[i,j]$ for the general model is hence given by

$$\boldsymbol{D}_{\mathbb{i}_k}(x)[i,j] = \sum_{l=0}^{x-1} \sum_{n=x-l-1}^{z} \sum_{q=j}^{m} g_n(x-l-1)\gamma_q \frac{e^{-d_k(\lambda_k+\mu_k)}(d_k[\lambda_k+\mu_k])^h}{h!} \binom{h}{l} \left(\frac{\mu_k}{\lambda_k+\mu_k}\right)^l \left(\frac{\lambda_k}{\lambda_k+\mu_k}\right)^{h-l}$$

$$+\delta_{\{i\geq j\}} \frac{e^{-d_k\mu_k}(d_k\mu_k)^x}{x!} \frac{e^{-d_k\lambda_k}(d_k\lambda_k)^{i-j}}{(i-j)!}, \qquad for\ x = 1,2,3\ldots;\ i,j = 0,1,2\ldots m$$

$z = (x-l-1)(m+1)$

$h = i+1+l+n+q-j$

Using an induction argument similar to the one in theorem 1 gives:

$$\Pr\{X_{\mathbb{i}_0} = x_0, Q_{\mathbb{i}_0^{\mathrm{r}}} = j\} = \left(\boldsymbol{\pi}\boldsymbol{D}_{\mathbb{i}_0}(x_0)\right)[j]$$
$$\Pr\{X_{\mathbb{i}_0} = x_0, X_{\mathbb{i}_1} = x_1, Q_{\mathbb{i}_1^{\mathrm{r}}} = j\} = \left(\boldsymbol{\pi}\boldsymbol{D}_{\mathbb{i}_0}(x_0)\boldsymbol{D}_{\mathbb{i}_1}(x_1)\right)[j]$$
$$\Pr\{X_{\mathbb{i}_0} = x_0, X_{\mathbb{i}_1} = x_1, \ldots, X_{\mathbb{i}_{\mathrm{k}}} = x_{\mathbb{i}_{\mathrm{k}}}, Q_{\mathbb{i}_{\mathrm{k}}^{\mathrm{r}}} = j\} = \left(\boldsymbol{\pi}\boldsymbol{D}_{\mathbb{i}_0}(x_0)\boldsymbol{D}_{\mathbb{i}_1}(x_1)\ldots\boldsymbol{D}_{\mathbb{i}_{\mathrm{k}}}(x_{\mathrm{k}})\right)[j]$$

The $\boldsymbol{Q}$ matrix serves to break the dependence between the number of chiasma events on either side of a breakpoint barrier (if $\alpha = 1$), as in theorems 5 and 6, and matrix multiplication with the $\mathbf{1}^T$ column-vector is equivalent to summing over all values of $Q_{\mathbb{i}_{n-1}^r}$, so

$$\Pr\{X_{\mathbb{i}_0} = x_0, X_{\mathbb{i}_1} = x_1, \ldots, X_{\mathbb{i}_{n-1}} = x_{n-1},\}$$
$$= \left(\prod_{k=0}^{\mathfrak{d}-1} \boldsymbol{D}_{\mathbb{d}_k}(x_k)\right) \boldsymbol{Q}^{1-\alpha} \left(\prod_{k=\mathfrak{d}}^{\mathfrak{d}+\mathfrak{h}-1} \boldsymbol{D}_{\mathbb{h}_k}(x_k)\right) \boldsymbol{Q}^{1-\alpha} \left(\prod_{k=\mathfrak{d}+\mathfrak{h}}^{\mathfrak{d}+\mathfrak{h}+\mathfrak{p}-1} \boldsymbol{D}_{\mathbb{p}_k}(x_k)\right) \left(\prod_{k=\mathfrak{d}+\mathfrak{h}+\mathfrak{p}}^{\mathfrak{n}-1} \boldsymbol{D}_{\mathbb{r}_k}(x_k)\right) \mathbf{1}^T$$

The final step is the weighting of the vector $\boldsymbol{v}$. Because unbalanced patterns in anaphase I tetrads are retained in the polar bodies, balanced patterns in anaphase I tetrads are weighted by 2 and unbalanced patterns in anaphase I tetrads are weighted by 0 (if this is not clear, notice that eliminating unbalanced anaphase I patterns is probabilistically equivalent to turning the two unbalanced pattern chromatids of an anaphase I tetrad into copies (one of each) of the two balanced pattern chromatids in the same tetrad, and then drawing randomly (uniformly) from the resulting four balanced pattern chromatids). So

$$\Pr\{\boldsymbol{R}_{paracentric\ linear} = \boldsymbol{r}\} = \boldsymbol{v}\boldsymbol{w}_{\boldsymbol{r}}^T$$

where

$$\boldsymbol{w}_{\boldsymbol{r}}[i] = \begin{cases} 2, & \boldsymbol{r}_{s_i} = \boldsymbol{r}; \varphi(\boldsymbol{r}_{s_i}) = 1; t_{s_i} = 1 \\ 0, & \boldsymbol{r}_{s_i} = \boldsymbol{r}; \varphi(\boldsymbol{r}_{s_i}) = 0; t_{s_i} = 1 \\ 1, & \boldsymbol{r}_{s_i} = \boldsymbol{r}; t_{s_i} \neq 1 \\ 0, & \boldsymbol{r}_{s_i} \neq \boldsymbol{r} \end{cases} \qquad for\ i = 0,1,2\ldots \mathfrak{s}-1$$

Since all unbalanced gametes are equivalent for our purposes, these are lumped together into the category *'unbalanced'* in the main program (note, again, that the sum of all unbalanced patterns is equal to the sterility, ζ.)

QED

Theorem 7 gives the recombination pattern probabilities for a homokaryotype in the special case where $\mathfrak{d} = \mathfrak{n}$, i.e. where there are no intervals in the inverted, proximal or opposite regions. It is therefore more general than theorem 1, but as the latter computes much more efficiently (it does not require a nested loop over the number of chiasma events in all intervals), it should be used whenever possible. For paracentric inversion heterokaryotypes in species with *nonlinear* meiosis, meaning that all chromatids have an equal chance of becoming gametes, the sterility and the probabilities of observing each recombination pattern are the same as for a pericentric inversion. In the following, I will therefore refer to inversions in these two cases collectively as *standard inversions*, for which I will calculate recombination patterns using theorem 5. Paracentric inversions in species with linear meiosis will be referred to as *paracentric linear inversions,* for which theorem 7 is needed.

The main program automatically chooses the appropriate algorithm based on the user's input. All infinite series expressions considered in this chapter converges quickly for realistic parameter values, and so can be approximated with arbitrary accuracy in short time; all such expressions involved in the simulations discussed in this thesis have been estimated with an error no larger than between 10^{-12} and 10^{-16}. For paracentric linear inversions, the current version of the program automatically assumes that recombination occur only in females (i.e. not in males) for all chromosomes (including for inversion homokaryotypes and chromosomes without inversion polymorphism), as this is the case in *Drosophila* (Gethmann 1988), which is one of the few groups that is known to possess linear meiosis (Sturtevant and Beadle 1936, Roberts 1976). In the remaining chapters, I will also make this assumption.

3 A model of non-random mating

A commonly used model of non-random mating – often referred to as *the fixed-relative preferences model* – assumes the following:

1) The population is infinitely large
2) Generations are non-overlapping
3) Males mate indiscriminately, whereas females mate selectively
4) The relative preference of a female with a given genotype for a male with a given genotype is the same regardless of the composition of the population
5) Choosiness is cost-free, so that the proportion of pairings involving a given female genotype is equal to that genotype's frequency before mating, regardless of preference.

(e.g. Kirkpatrick 1982, Gomulkiewicz and Hastings 1990, Servedio and Kirkpatrick 1997, Servedio 2000, Servedio and Sætre 2003). Houle and Kondrashov (2002) suggest that indirect selection on the preference loci can be modelled through an infinite series of mate evaluations for which each rejection entails a fixed cost. In this chapter, I will use a similar approach to derive a costly mate choice model that reduces to the fixed-relative preference model in a special case. It will be most instructive to start from the biological interpretation and build the mathematics bottom-up from there, as follows.

Assumptions 1,2, and 3 above apply. Let the frequency of a female with genotype *a* among females before mating be f_a and the frequency of a male genotype *b* among males before mating be m_b, and assume that the females choose their mate by performing a series of *searches,* defined so that a single search ends when a female encounter a single male. Further assume that each search entails a cost c ($0 \leq c < 1$), that the probability that a female locates a male of given genotype after a single search is proportional to that genotype's frequency among mating males (m_b) and independent of the female's preference, and that a female *a accepts* a male *b* with probability p_{ab} ($0 \leq p_{ab} \leq 1, \sum_{b'} p_{ab'} > 0$) and *rejects* a male *b* with probability $1 - p_{ab}$, where p_{ab} is some function of the genotypes of *a* and *b.* Assume that males can mate an unlimited number of times and that females can mate only once, so that if the female accepts the male, the former but not the latter is removed from the mating population. This will be reasonable assumption when eggs are sufficiently costly to produce compared to sperm, or when gestation is partly or fully internal. If a female rejects a male, she will perform another search, which is identical to the first search in that the male frequencies do not change (the female frequencies might change, but I assume that this does not affect the availability of males), and so on until she either accepts a male, dies during a search, or the mating season ends (in which case she also dies without mating, as I assume non-overlapping generations). The probability of dying during a single search is proportional to the cost *c* and is the same for all searches, and the mating season is assumed to be limited to a maximum total number *k (≥1)* of searches, though the special case where *k* approaches infinity (which implies that all females search until they either die during a search or find a mate) is applicable, as discussed below.

Now let $M_{ab}(k, c)$ be the proportion *a/b*-couples among all mated couples for the given value of *k* and *c*. For *k* = 1, this variable is proportional to the probability of randomly picking a female with genotype *a* among all mating females, times the probability that the female survives the search, times the probability that the female encounter a male with genotype *b*, times the probability that the female accepts the male, i.e.

$$M_{ab}(k = 1, c) \propto f_a(1 - c)m_b p_{ab}$$

(males may or may not experience a cost for each search; If the cost is the same for all males, then the relative proportions of males stay the same.) For *k=2*, the female *a* must either survive, encounter and accept the male *b* in the first search *or* survive and reject the male she encounter in the first search before surviving, encountering and accepting the male *b* in the second search, i.e.

$$M_{ab}(k=2,c) \propto f_a\left[(1-c)m_b p_{ab} + (1-c)\left(\sum_{b' \in \boldsymbol{B}} m_{b'}(1-p_{ab'})\right)(1-c)m_b p_{ab}\right]$$

where $\boldsymbol{B}$ is the set of all male genotypes. If we now define

$$N_a = \sum_{b' \in \boldsymbol{B}} m_{b'} p_{ab'} = 1 - \sum_{b' \in \boldsymbol{B}} m_{b'}(1-p_{ab'})$$

and

$$Z_a = (1-N_a)(1-c)$$

then N_a gives the probability that female *a* will settle on a mate after a single search (this probability is the same for all searches since the availability of males does not change), and $Z_a = (1-N_a)(1-c)$ gives is the probability that the female will survive the search $(1-c)$ and *not* settle on a mate $(1-N_a)$. We can now write

$$M_{ab}(k=2,c) \propto f_a(1-c)m_b p_{ab}(1+Z_a)$$

Using the same reasoning for *k = 3* gives

$$\begin{aligned} M_{ab}(k=3,c) &\propto f_a[(1-c)m_b p_{ab} + (1-c)Z_a m_b p_{ab} + (1-c)Z_a^2 m_b p_{ab}] \\ &= f_a(1-c)m_b p_{ab}\sum_{l=0}^{2} Z_a^l \end{aligned}$$

and in general

$$M_{ab}(k,c) \propto f_a(1-c)m_b p_{ab}\sum_{l=0}^{k-1} Z_a^l \qquad (3.1)$$

This equation is visualized in figure 3.1.

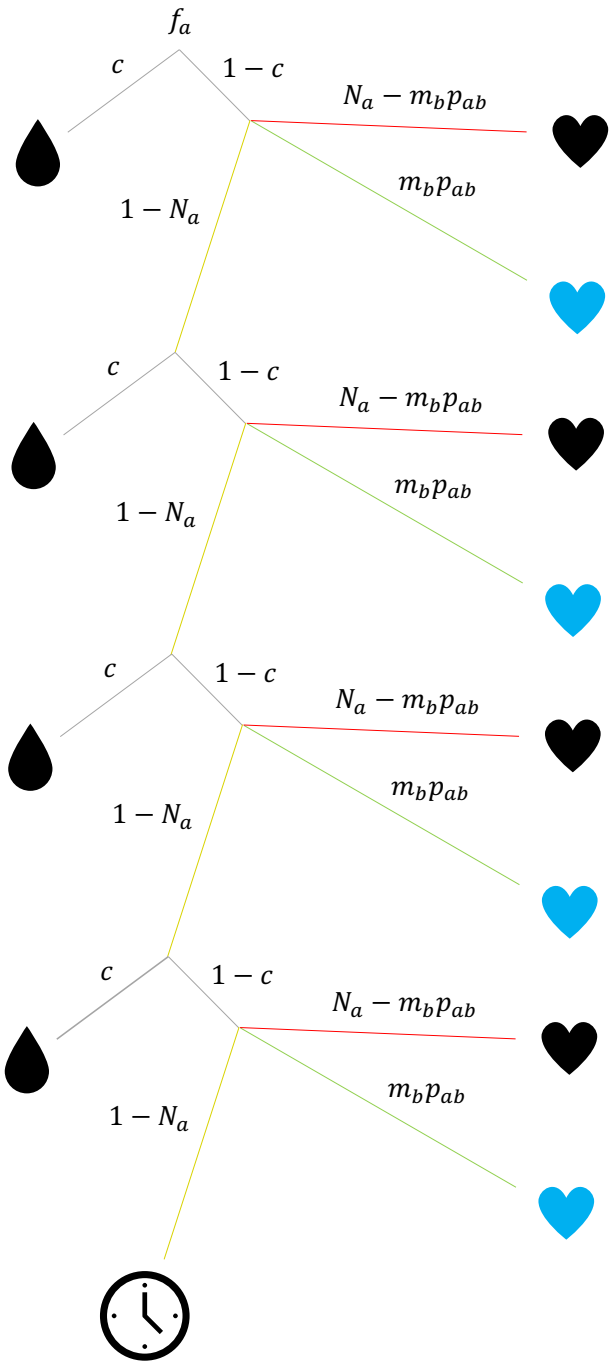


Figure 3.1: A probability tree showing the probability that a female with genotype a *will mate with a male with genotype* b *(here k = 4). Grey lines: the female searches and either dies (probability c) or survives (probability 1-c). Green line: The female meets and accepts a male with genotype* b *($m_b p_{ab}$). Red line: The female meets and accepts a male with a genotype different from* b *($N_a - m_b p_{ab}$). Yellow line: The female does not accept the male she meets ($1 - N_a$). All black symbols: the female is removed from the mating pool without mating with a male with genotype* b*. Black tear: the female dies. Black heart: The female mates with a different male. Black clock: mating season ends. Blue heart: female* a *and male* b *mates. Equation 3.1 is the sum of all paths leading to a blue heart.*

Note that since $0 < N_a \leq 1$ and $0 \leq c < 1$ (I imposed these restriction in the introduction to the model above), it follows that $0 \leq Z_a < 1$, so since $\sum_{l=0}^{k-1} Z_a^l$ is a geometric series, we get

$$M_{ab}(k,c) \propto f_a(1-c)m_b p_{ab} \frac{(1-Z_a^k)}{1-Z_a}$$

which, since $(1-c)$ is a constant, implies that

$$M_{ab}(k,c) \propto f_a m_b p_{ab} \frac{(1-Z_a^k)}{1-Z_a} = f_a m_b p_{ab} \frac{(1-Z_a^k)}{N_a + c(1-N_a)}$$

or, to be precise,

$$M_{ab}(k,c) = \frac{1}{Q}\left(\frac{f_a m_b p_{ab}(1-Z_a^k)}{N_a + c(1-N_a)}\right)$$

where Q is the normalizing constant,

$$Q = \sum_{a' \in \boldsymbol{A}} \sum_{b' \in \boldsymbol{B}} f_{a'} m_{b'} p_{a'b'} \frac{\left(1-Z_{a'}^k\right)}{N_{a'} + c(1-N_{a'})}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are the set of all female and male genotypes, respectively (the cost of the first search cancels out, as it is the same for all females.) Note that $Z_a^k = [(1-N_a)(1-c)]^k$ is the probability that a female a will perform and survive k searches without choosing a mate, and that $1 - Z_a^k$ is the probability of the negation of that.

Now define the *mating fitness* of female a (F_a^f) and male b (F_b^m) as

$$F_a^f = \frac{f_a^*}{f_a}$$

$$F_b^m = \frac{m_b^*}{m_b}$$

where the frequencies with and without the * are the frequencies after and before mating, respectively. We can find f_a^* and m_b^* by summing over all couples involving a and b, respectively. Hence,

$$f_a^* = \sum_{b' \in \boldsymbol{B}} M_{ab'}(k,c) = \frac{1}{Q} \sum_{b' \in \boldsymbol{B}} \frac{f_a m_{b'} p_{ab'}(1-Z_a^k)}{N_a + c(1-N_a)} = \frac{1}{Q}\left[\frac{f_a N_a (1-Z_a^k)}{N_a + c(1-N_a)}\right]$$

so

$$F_a^f \propto \frac{N_a(1-Z_a^k)}{N_a + c(1-N_a)}$$

Similarly,

$$m_b^* = \sum_{a' \in \boldsymbol{A}} M_{a'b}(k,c) = \frac{1}{Q} \sum_{a' \in \boldsymbol{A}} f_{a'} m_b p_{a'b} \frac{\left(1-Z_{a'}^k\right)}{N_{a'} + c(1-N_{a'})} = \sum_{a' \in \boldsymbol{A}} \frac{f_{a'}^* m_b p_{a'b}}{N_{a'}}$$

so

$$F_b^m = \sum_{a' \in \boldsymbol{A}} \frac{f_{a'}^* p_{a'b}}{N_{a'}}$$

Note that these expressions are all frequency-dependent, and that this attribute is an inherent property of the model rather than an additional assumption. For example, we can from the expression for F_a^f deduce that the mating fitness of a female is lower if she has a preference for rare males (i.e. $1 - N_a$ and Z_a are large, N_a is small), because she will endure a larger cost for additional searches and risk ending up without any mates at all at the end of the mating season. Also note that the disadvantage in preferring rare males is more severe for higher *c* and lower *k*. Similarly, we see from the expression for F_b^m that the mating fitness of a male is higher if he is strongly preferred by females who survive the searching procedure.

When *k* approaches infinity, meaning that all females perform repeated searches until they either find a mate or die during a search, Z_a^k $(0 \leq Z_a < 1)$ approaches zero, so

$$M_{ab}(k \to \infty, c) \propto \frac{f_a m_b p_{ab}}{N_a + c(1 - N_a)}$$

and

$$F_a^f \propto \frac{N_a}{N_a + c(1 - N_a)}$$

The latter expression is equivalent to equation 2.3 in Houle and Kondrashov (2002). In the fixed-relative preferences model, the proportion of pairings between a female *a* and a male *b* is given by $\frac{f_a m_b p_{ab}}{N_a}$ where f_a, m_b, and N_a are as defined above and p_{ab} is the relative preference of female *a* for male *b* (see e.g. Gomulkiewicz and Hastings 1990, section *The Model*). In my model, when $k \to \infty$ and $c = 0$, M_{ab} reduces to

$$M_{ab}(k \to \infty, c = 0) \propto \frac{f_a m_b p_{ab}}{N_a}$$

and since $\sum_{b' \in \boldsymbol{B}} \frac{f_a m_{b'} p_{ab'}}{N_a} = f_a$ and $\sum_{a' \in \boldsymbol{A}} f_{a'} = 1$, it follows that

$$Q(k \to \infty, c = 0) = \sum_{a' \in \boldsymbol{A}} \sum_{b' \in \boldsymbol{B}} \frac{f_{a'} m_{b'} p_{a'b'}}{N_{a'}} = 1$$

so

$$M_{ab}(k \to \infty, c = 0) = \frac{f_a m_b p_{ab}}{N_a}$$

which means that my model reduces to the fixed-relative preference model in the special case where $k \to \infty, c = 0$. (p_{ab} is not restricted to be between 0 and 1 in the fixed-relative preference model, but scaling the preferences does not change the result when $k \to \infty, c = 0$). Also note that when $k \to \infty, c = 0$,

$$F_a^f = \frac{N_a}{N_a} = 1$$

which means all females have the same fitness regardless of preferences, as there is no selection against preferring rare males. Furthermore, when $p_{a'b'} = 1\ for\ all\ a' \in \boldsymbol{A}, all\ b' \in \boldsymbol{B}$, meaning that all females will accept any male on first encounter, we get for all values of *k* and *c*

$$M_{ab}(k, c) \propto f_a m_b$$

which is the standard definition of random mating.

From the recombination pattern probabilities derived in the previous chapter, it is easy to derive the proportions of each genotype generated by a given couple (see the methods called *calculate_gamete_frequencies* in classes *Chromosome_diplotype* and *Genotype,* and the method *find_offspring* in class *Couple* in the program). The frequency of a given genotype *g* in the next generation can then be calculated as $\sum_{a\in\boldsymbol{A}} \sum_{b\in\boldsymbol{B}} y_{abg} M_{ab}$ where y_{abg} is the proportion of *g* genotypes produced by the couple *a/b.* In the program, this calculation is done using matrix operations on *Numpy arrays* (see method *run* in class *Simulation*), which is computationally highly efficient (Langtangen 2008)

4 Reinforcement and the evolution of chromosomal inversions

4.1 The model

With *reinforcement*, I here mean the evolution of selective mating preferences in response to selection against interspecific mating (Servedio and Noor 2003). Many models of reinforcement can be derived from a single general model, referred to as the *PTMN-model,* that assumes four diallelic loci: a preference locus *P* and a trait locus *T* – collectively the *prezygotic isolation loci* – and two directly selected loci, *M* and *N* – collectively the *postzygotic isolation loci*. The preference of a given individual *a* for another individual *b* of the opposite sex is determined by *a*'s genotype at the *P* locus and *b's* genotype at the *T* locus; hence, purely assortative mating, as in Felsenstein's (1981) model, occurs when there is initial total linkage disequilibrium and no recombination, henceforth *perfect linkage,* between *P* and *T* (Kirkpatrick and Ravigne 2002 p. 25-26 make the similar point that "assortment and mating preferences can be treated as a single form of prezygotic isolation by regarding assortment as the special case where a mating preference acts on itself."). The directly selected *M* and *N* loci can be either locally adapted or epistatic or both. These loci set up the selection pressure against interspecific mating by contributing to hybrid inviability, but do not themselves influence the bearer's mating preference. Models where the trait locus itself interact epistatically with another locus, as in Dagilis and Kirkpatrick's (2016) preference-trait model, can be derived by assuming perfect linkage between *T* and *M*.

Tricket and Butlin (1994) used deterministic simulations of the Felsenstein (1981) model to show that reinforcement can occur more readily when an inversion captures the *P/T* and *M* loci. Dagilis and Kirkpatrick's (2016) analysis quantified this effect in their similar assortative mating model, as well as the effect of an inversion capturing the *P* and *T/M* loci in their preference-trait model and some other models that cannot easily be expressed in terms of the general *PTMN*-model. Using an approach that differs in several respects (to be discussed) from the ones in Tricket and Butlin (1994) and Dagilis and Kirkpatrick (2016), I will in this chapter further explore the effect of both standard and paracentric linear inversions on reinforcement.

In addition to *P, T, M* and *N,* my model includes two bialleleic loci, *[* and *],* corresponding to the breakpoints of a chromosomal inversion. In the case of a paracentric linear inversion, it includes one additional monalleleic loci, *@,* corresponding to the centromere. I will assume that the inversion spans the four ordinary loci (*P, T, M,* and *N*), so that the ordering is *[PTMN]* and *[PTMN]@* for standard and paracentric linear inversions, respectively. All loci are autosomal and located on the same chromosome. A single parameter *L* gives the genetic length in Morgans (see section 2.1.6) of all intervals in homokaryotypes except the interval between *N* and *],* which for reasons that will become clear is always set to have genetic length *0*. Hence, the total length of the inverted region is *4L*, and the length of the proximal region is *L* (see figure 4.1). The alleles $[_0$ and $]_0$, $[_1$ and $]_1$ represent the ancestral and diverged arrangement, respectively, so that e.g. the *genotype key* $[_{01}P_{01}T_{01}M_{01}N_{01}]_{01}$ indicate an inversion heterokaryotype heterozygous for all loci. The *allele indices* in the genotype keys are position sensitive, with the first and second position at each loci referring to the allele at the two parental homologues. From this we can, for example, infer that an individual with genotype key $[_{01}P_{01}T_{01}M_{01}N_{01}]_{01}$ has the two *parental haplotypes* $[_0P_0T_0M_0N_0]_0$ and $[_1P_1T_1M_1N_1]_1$. I will often refer to the diverged arrangement simply as *the inversion*. Hence, the *frequency of the inversion*, is simply the frequency of the $[_1$ allele (or, equivalently, the frequency of the $]_1$ allele or the $[_1]_1$ haplotype. These are all equivalent because $[_1$ and $]_1$ always occur together.)

For simplicity, I will assume that the inversion suppresses chiasma formation to an equal degree, *d,* in all intervals, so that $d_k = d_{\mathbb{h}} = d$ for all intervals $\mathbb{h}_k$ in the inverted region. Since studies have shown that chiasmata formation is also suppressed in a region stretching a fair distance outside of the inversion itself (e.g. Pegueroles et al. 2010), I will set $d_{\mathbb{p}} = d$ where $\mathbb{p}$ is the proximal region. Hence, in heterokaryotypes the inverted and proximal regions measure *4dL* and *dL* Morgans, respectively. Since the pure counting model provides a convenient and unambiguous scale of interference, with stronger interference for higher *m* (Foss et al. 1993, Navarro et al. 1997), I will use it to define three interference conditions: no interference *(m=0),* moderate interference *(m=3),* and strong interference *(m=7)*. For comparison, chiasmata in *Schizosaccharomyces pombe*, *Neurospora, Drosophila,* and *Mus* have been found to be approximately distributed according to a pure counting model with *m = 0, m = 2, m = 4,* and *m = 6,* respectively (Munz 1994, Foss et al. 1993, Lange et al. 1997). All the models mentioned in this chapter, except my own, implicitly or explicitly assume no chiasma interference. Like Servedio and Sætre (2003), but unlike Servedio and Kirkpatrick (1997) and Servedio (2000), I assume diploidy.

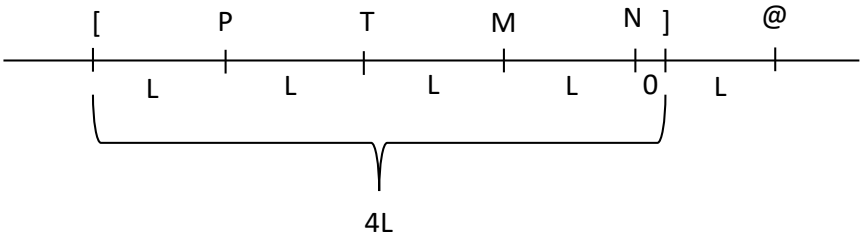


Figure 4.1: The relative positions of the different loci and the genetic length of each interval.

The total population is divided into two distinct habitats, which I will refer to as *habitat 0* and *habitat 1*. The demes in these two habitats are assumed to have diverged in allopatry, so that they are initially fixed for the T_0, M_0, and N_0 (habitat 0); and T_1, M_1, and N_1 (habitat 1) alleles, respectively. The P_1, $[_1$ and $]_1$ alleles will be introduced during the runs and are initially not present, so the genotypes $[_{00}P_{00}T_{00}M_{00}N_{00}]_{00}$ and $[_{00}P_{00}T_{11}M_{11}N_{11}]_{00}$ start with frequency 1.0 in their respective habitats. As in Servedio and Kirkpatrick (1997), Servedio (2000), and Servedio and Sætre (2003), I assume that the preference locus is expressed only in females, and that the trait alleles are adapted to the habitats with the corresponding indices, and expressed and selected for only in males. In particular, a male with diplotype T_{ij} , *i,j = 0,1* in habitat *k* gets a *fitness contribution*,

$$F\left(T_{ij}\ in\ habitat\ k\right) = \begin{cases} 1 + s_T, & i = j = k \\ 1, & i \neq j \\ \dfrac{1}{1 + s_T}, & i = j \neq k \end{cases}$$

whereas females gets a fitness contribution of *1* regardless of habitat or genotype at the *T* locus. In all the runs discussed in this chapter, $S_T = 0.2$.

Mating occur only within habitats. Couple proportions are determined by the model in chapter 3, with the acceptance probabilities p_{ab} (i.e. the probability that a female of genotype *a* will accept a male of genotype *b* at any given encounter) given as in the following table,

a↓/b→	T_{00}	T_{01}	T_{11}
P_{00}	$(1+\beta_0)y$	y	$\frac{y}{1+\beta_0}$
P_{01}	$\frac{y}{\frac{1}{2}\left(\frac{1}{1+\beta_0}+1+\beta_1\right)}$	y	$\frac{y}{\frac{1}{2}\left(1+\beta_0+\frac{1}{1+\beta_1}\right)}$
P_{11}	$\frac{y}{1+\beta_1}$	y	$(1+\beta_1)y$

Table 4.1: the preference p_{ab} of a female with genotype a *for a male with genotype* b

where

$$y = \frac{\tau}{\max(1+\beta_0, 1+\beta_1)}, \qquad 0 < \tau \leq 1$$

The factor *y* ensures that all probabilities are between 0 and 1. The proportion, among all couples after pairing, of the *couple* made up of female *a* and male *b* is now given by

$$M_{ab}(k,c) = \frac{1}{Q}\left(\frac{f_a m_b p_{ab}(1-Z_a^k)}{N_a + c(1-N_a)}\right)$$

where all symbols are as defined in chapter 3. Random mating occurs when $\beta_0 = \beta_1 = 0$, and *symmetrical mating* occurs whenever $\beta_0 = \beta_1$. When $k \to \infty, c = 0$, the model reduces to the fixed-relative preference model with Gomulkiewicz and Hastings' (1990) parametrization. I will throughout assume symmetrical mating with $\beta_0 = \beta_1 = 0.1$ and $k \to \infty$. Unless otherwise noted, *c = 0*.

The fitness contribution from the *M* and *N* loci is for both sexes given by

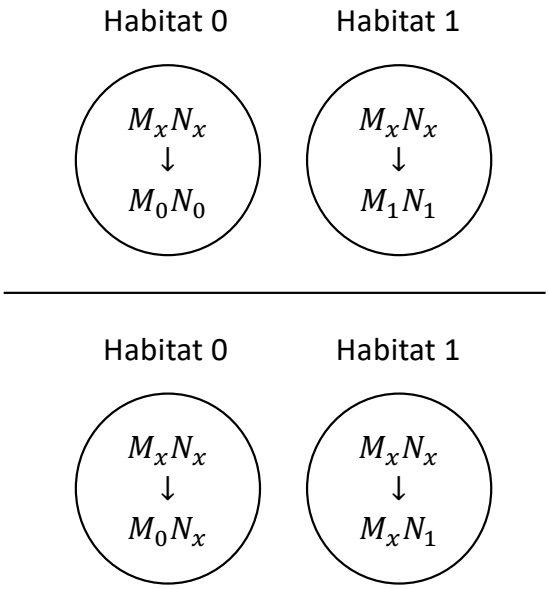
$$F(M_{ij}N_{kl}) = \begin{cases} 1, & i=j=k=l \\ 1-hS_E, & i+j+k+l = 1 \text{ or } i+j+k+l = 3 \\ 1-S_E, & i+j+k+l = 2 \end{cases}$$

$$for\ i,j,k,l = 0,1$$

In other words, the two *pure diplotypes* ($M_{00}N_{00}$ and $M_{11}N_{11}$) at these two loci gets a fitness contribution of 1, diplotypes with all allele indices the same except one gets a fitness contribution of $1-hS_E$, and diplotypes with two of each allele index gets a fitness contribution of $1-S_E$. This is equivalent to the parametrization in Servedio and Sætre (2003). Here, and in the following, a *pure* haplotype or diplotype is one that has the same allele index for all loci of a given set, meaning that all alleles originated from the same habitat. In all my simulations, $h = 0.5$ and $S_E = 0.5$. The interaction between the fitness contribution from *T* and the fitness contribution from *M/N* is multiplicative.

The *M* and *N* alleles interact epistatically, but are not differently selected in the two habitats. This does not necessarily mean that they must have diverged by drift in allopatry; if both loci originally were fixed for a now lost third ancestral allele that were replaced in the two habitats by different positively and universally selected alleles (M_0 and N_0 versus M_1 and N_1) with the same non-epistatic fitness, then they would be neutral in this regard after secondary contact. Note that this account differs from the typical textbook model of *Dobzhansky-Muller incompatibilities* (e.g. Futuyma 2013), in which only one of the alleles is replaced in each habitat. The difference between these two models will prove important later (see chapter 5); I will use the terms *four-allele model* and *two-allele model* to distinguish them (see figure 4.2). In general, when an allele, haplotype or diplotype is

adapted to the genetic background or local environment in one habitat but not the other, I will say that it is *differentially adaptive.*



Figur 4.2: Two models of the evolution of incompatible alleles in allopatry. Top: four-allele model. Bottom: two-allele model. M_x and N_x are ancestral and universally compatible. Derived 0-index alleles are compatible with the ancestral alleles and other 0-index alleles, but are incompatible with 1-index alleles. Vice versa for 1-index alleles.

Migration begins at *secondary contact*. In the main program, migration rates are given for *n* habitats as the *migration matrix* $\boldsymbol{t}$,

$$\boldsymbol{t}[i,j] = t_{ij}, \qquad for\ i,j = 0,1,2 \ldots n-1$$

where

$$\sum_k t_{kj} = 1, \qquad for\ j = 0,1,2 \ldots n-1$$

so that a representative sample of habitat *i* before each *migration event* make up a proportion t_{ij} of habitat *j* after the event. With two habitats, we get the migration matrix $\begin{pmatrix} 1-t_{10} & t_{01} \\ t_{10} & 1-t_{01} \end{pmatrix}$. In this text, I will mostly consider the case of *symmetrical migration,* which for two habitats can be expressed by the single parameter *t,* so that

$$\boldsymbol{t} = \begin{pmatrix} 1-t & t \\ t & 1-t \end{pmatrix}$$

On one occasion I will also consider the case of *one-way migration* (also known as a continent-island model), so that

$$\boldsymbol{t} = \begin{pmatrix} 1 & t \\ 0 & 1-t \end{pmatrix}$$

An *equilibrium* is reached when all the genotype frequencies changes less than a predetermined value, Δ, from one generation to the next. At this point, an *equilibrium action* is performed, e.g. introduce a new mutation, change migration rates, or end the simulation. A *scenario* is a description of the order in which the different equilibria actions occur. I will consider four different scenarios (the delta values in parenthesis indicate the condition for performing each action; these are the same as those used by Servedio and Sætre 2003 in what I call the control scenario):

Control scenario (no inversion)

1) Secondary contact (start migration)
2) P_1 is introduced in habitat 1 ($\Delta = 10^{-12}$)
3) End ($\Delta = 10^{-10}$)

Scenario 1:

1) Secondary contact
2) P_1 is introduced in habitat 1 ($\Delta = 10^{-12}$)
3) New inversion captures the haplotype $P_1T_1M_1N_1$ in habitat 1 ($\Delta = 10^{-10}$)
4) End ($\Delta = 10^{-10}$)

Scenario 2:

1) Secondary contact
2) P_1 is introduced in habitat 1 ($\Delta = 10^{-12}$)
3) New inversion captures the haplotype $P_0T_1M_1N_1$ in habitat 1 ($\Delta = 10^{-10}$)
4) End ($\Delta = 10^{-10}$)

Scenario 3:

1) Secondary contact
2) New inversion captures the haplotype $P_0T_1M_1N_1$ in habitat 1 (before the introduction of P_1) ($\Delta = 10^{-12}$)
3) P_1 is introduced *within* the inverted region (if the inversion has spread) in habitat 1 ($\Delta = 10^{-10}$)
4) End ($\Delta = 10^{-10}$)

The new mutation (P_1 or inversion) is introduced at a frequency of 0.001 in the given habitat. Migration, selection, and mating happen in that order in each generation in all scenarios.

Since the genetic length between *N* and *]* is set to zero and the inversion captures the N_1 haplotype in all scenarios, haplotypes with N_0 and $]_1$ will never occur. These can therefore be removed from consideration, which significantly reduces the running time of the simulations (the program has a feature that allows user-defined haplotypes or alleles to be removed from individual inter-equilibria steps or the full simulation). This is the main motivation for setting the genetic length between *N* and *]* to zero, though it is also arguably more realistic than the alternative, since the formation of chiasmata is typically more strongly suppressed close to breakpoint boundaries (e.g. Schaeffer and Anderson 2005). The genetic length between *[* and *P* is deliberately not set to zero, so as to explore the effects of recombination of *P* alleles in and out of the different arrangements. Note that setting the genetic length between *L* and *]* to 0 is equivalent to setting the suppression factor for this interval to 0, since *[,]* and @ are always homozygous in homokaryotypes. For the same reason, the positioning of the loci in my model is equivalent to one in which the left breakpoint, *[,* is moved an additional interval to left, if the chiasmata are perfectly suppressed in this interval. Hence, assuming non-zero recombination between *[* and *P* in heterokaryotypes in my model

is not necessarily in contradiction to the finding that chiasmata formation is more strongly suppressed close to the breakpoint boundaries.

The studies in Tricket and Butlin (1994) and Dagilis and Kirkpatrick (2016) are similar to scenario 1, but one important difference is that they both assume that the inversion captures only one of the postzygotic isolation loci. As Dagilis and Kirkpatrick (2016) note, a perfectly chiasma-suppressing inversion that captures a high-fitness haplotype at two or more epistatic or locally adapted loci will spread by avoiding recombination with lower-fitness haplotypes (Charlesworth and Charlesworth 1973, Kirkpatrick and Barton 2006); their objective is to show that this is not a necessary condition for the inversion to spread in a reinforcement setting. My objective is rather to compare the outcome of the three scenarios above when the inversion captures all four loci. In scenario 2, this means that the inversion can spread even though it captures a locally maladapted preference allele (P_0), which generates some interesting dynamics, as we shall see. Another important difference is that Tricket and Butlin (1994) and Dagilis and Kirkpatrick (2016) both assume that the inversion perfectly suppresses recombination and is free of direct selection, which in my terminology is equivalent to a standard inversion with $d_k = 0$ for all intervals $\mathbb{i}_k$. In my study, I will investigate the effect of varying the value of $d = d_{\mathbb{h}} = d_{\mathbb{p}}$, for standard as well as paracentric linear inversions, in the latter case with recombination only in females, as in *Drosophila*. Note that this implies that the paracentric linear inversion is neutral in males, since it neither causes meiotic irregularities, nor suppresses recombination between the loci (since recombination is absent anyway). Since higher values of *d* imply more recombination inside the inversion *and* higher degree of underdominance, both of which lower the fitness of a newly introduced inversion that captures a high-fitness haplotype (Kirkpatrick and Barton 2006), a useful metric is what I will call the *d toleration limit,* i.e. the highest value of $d = d_{\mathbb{h}} = d_{\mathbb{p}}$ that still allow the inversion to spread by selection. This parametrization models recombination and underdominance of the inversion as an indirect consequence of varying the degree to which chiasma generation is suppressed in heterokaryotypes (see chapter 2), which is arguably more realistic than treating recombination and underdominance as independent parameters, as do all other evolutionary models that I am aware of.

I ran simulations with *L = 0.02, 0.06, 0.125, 0.2.* These values were chosen so as to make the total length of the inverted region ($4L$) span the range observed in the Coyne et al. (1993), Navarro and Ruiz (1997) dataset (see figure 2.4, chapter 2). *L = 0.02* correspond approximately to the shortest observed inversion, *L = 0.06* is within the higher range for which $d \approx 0$, *L = 0.125* is within the range for which *d* is significantly larger than zero, and *L = 0.2* corresponds approximately to the largest inversion. For the two-way migration runs, I used $t = t_{01} = t_{10} = 0.001, 0.01, 0.065$. $t = 0.065$ is the approximate highest migration rate for which differentiation between the two habitats was maintained before the introduction of P_1 in all runs. When *t* is further increased, a *tipping point* is reached above which both habitats become fixed for the P_0, T_0, M_0, N_0 alleles. The reason, presumably, is that before the introduction of P_1, T_0 has higher fitness than T_1 when averaged over both habitats because it is preferred by all the (P_0) females. Hence, when migration is sufficiently high, T_0 invades both habitats and M_0 and N_0 increases in frequency in habitat 1 due to linkage disequilibrium with T_0. The *M* and *N* alleles are under positive frequency-dependent selection (common alleles are more likely to be paired with a copy of itself), so when M_0 and N_0 become the more prevalent ones in habitat 1, they are selected in their own right and go to fixation (Servedio and Kirkpatrick 1997 makes a similar point). The exact value of the tipping point depends on the other parameters (Servedio and Kirkpatrick 1997).

4.2 Results

4.2.1 Control scenario

Figure 4.3 (top) displays the differentiation between the habitats at the *P* locus, i.e. the frequency of P_1 in habitat 1 minus the frequency of P_1 in habitat 0, at the final equilibrium in the two-way migration control scenario (i.e. before the introduction of the inversion in scenario 1/2) for the pure counting model, *m = 0, 3, 7, c = 0* (no cost of searching) and *t = 0.065*. All runs with symmetrical two-way migration considered here are *perfectly symmetrical* with regards to preference, selection and migration, meaning that the preference of P_0 for T_0 is mirrored by the preference of P_1 for T_1, the selection on T_0 in habitat 0 and 1 is mirrored by the selection on T_1 in habitat 1 and 0, and the selection on M_0 or N_0 in any given diploid genotype is mirrored by the selection of M_1 or N_1 in the same genotype with the indices reversed. Accordingly, I found that the frequency of P_0 in habitat 0 mirrored the frequency of P_1 in habitat 1 at the final equilibrium whenever the latter spread under these conditions. Consistently with previous results (Felsentein 1981, Servedio and Kirkpatrick 1997, Servedio 2000, Servedio and Sætre 2003), the figure shows that closer linkage implies higher differentiation, i.e. more complete reinforcement. As mentioned in chapter 1, this is because recombination breaks up the linkage disequilibrium between the pre- and postzygotic loci, so as to make the trait allele a less reliable indicator of male fitness (Felsentein 1981, Servedio and Kirkpatrick 1997). The figure also shows that varying chiasma interference has little effect on the degree of differentiation, except that it is slightly lower for stronger interference, presumably because the recombination rate in any given interval is slightly higher for stronger interference when the genetic length is held constant (Foss et al. 1993 and figures 4.3, bottom, and 4.6 in this text). Note that because the equilibrium frequency of P_1 in habitat 1 ($\approx\frac{1+diff.P}{2}$) is not that much higher than 0.5, a randomly occurring inversion is not much more likely to capture P_1 (scenario 1) than P_0 (scenario 2), except when *L* is small.

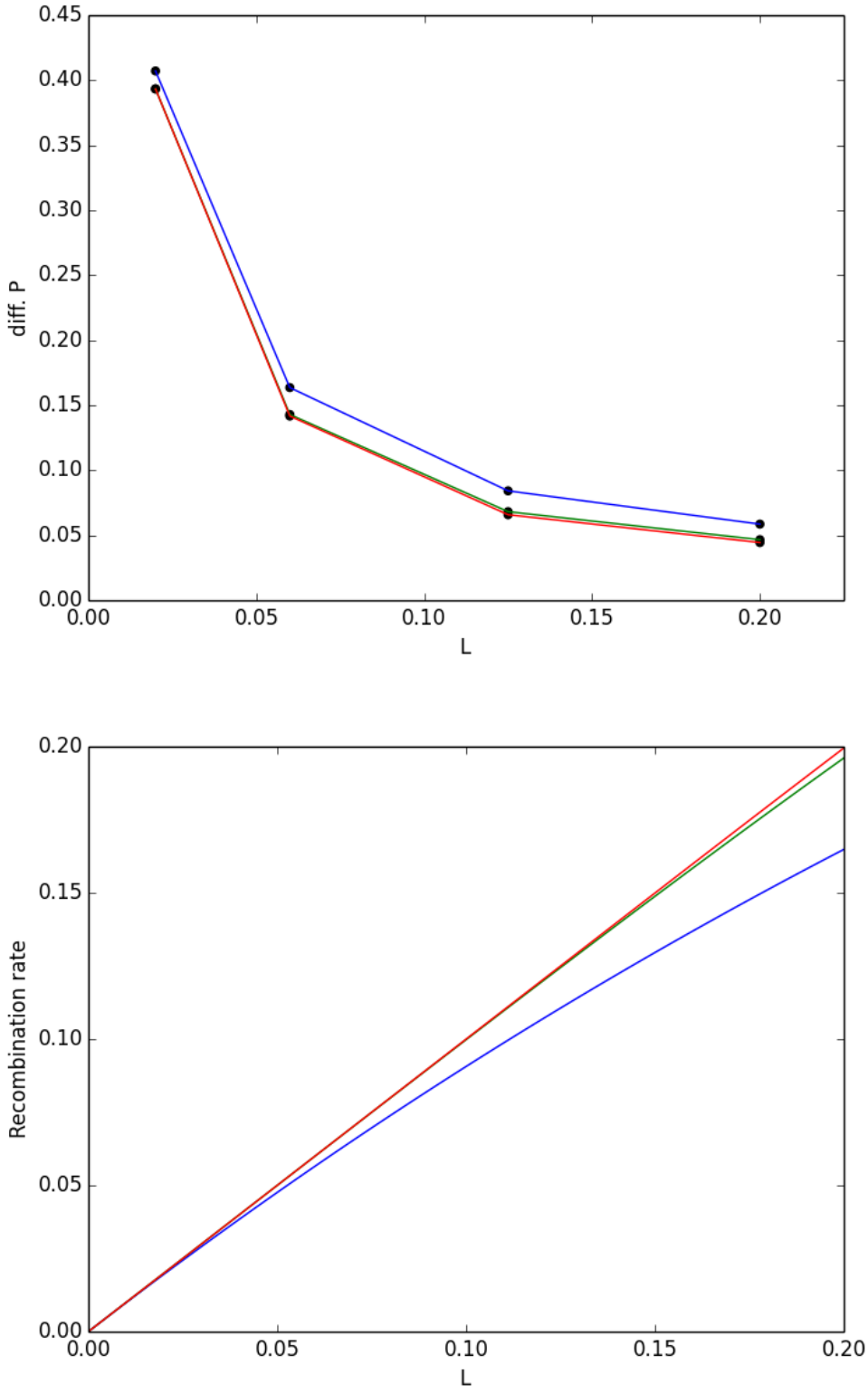


Figure 4.3: Top: The difference between the equilibrium frequencies of P_1 in habitats 1 and 0 in the control scenario with symmetrical two-way migration, t = 0.065, c = 0. Bottom: The recombination rate of a single interval of length L (cf. Figure 1 in Foss et al. 1993, and Figure 4.6 in this text.) Both plots: Pure counting model, blue: m = 0, green: m = 3, red: m = 7. The green and red lines partly coincide in both figures.

4.2.2 Scenario 1

There are two forces influencing whether or not the inversion initially spread in scenario 1. Firstly, it captures an optimal haplotype and reduces recombination with lower-fitness ones, which is favored by selection (Charlesworth and Charlesworth 1973, Tricket and Butlin 1994, Kirkpatrick and Barton 2006, Dagilis and Kirkpatrick 2016). Secondly, if $d > 0$, the inversion is underdominant due to meiotic irregularities in heterokaryotypes (chapter 2), meaning that it will have lower fitness when newly introduced. Whether or not the inversion spread therefore depends on which of these forces initially have the strongest effect; as mentioned above, both depends on d so that the fitness of the inversion is lower for higher values of d. If the derived arrangement spreads and reaches a high frequency in habitat 1, it will still be in strong linkage disequilibrium with P_1, T_1, M_1, and N_1 because of the initial condition and the low (if $d > 0$) or zero ($d = 0$) rate of recombination with haplotypes of the ancestral arrangement. Accordingly, the few P_0, T_0, M_0, and N_0 alleles that might invade it by recombination have low fitness, as they will either be maladapted to the local habitat (T) or the locally prevalent postzygotic isolation alleles (M, N), or cause the bearer to mate with locally maladapted individuals (P, T). Since the runs are once again perfectly symmetrical except for initial condition, the same with the indices inversed will be true for the few P_1, T_1, M_1 and N_1 alleles remaining outside of the derived arrangement. When $d > 0$, the two alternative arrangements additionally experience positive frequency-dependent selection in their respective habitats, due to underdominance. In sum, therefore, I expect the two pure haplotypes, $[_0P_0T_0M_0N_0]_0$ and $[_1P_1T_1M_1N_1]_1$, to be selected in their respective habitats, and any non-pure haplotypes to be selected out faster than they are generated by low recombination in heterokaryotypes. I will say that the inversion *divide the population* when the population as a whole at equilibrium consist almost exclusively of the two pure haplotypes in approximately equal proportions; this was the case in all the runs in which the inversion spread. To be more precise, the total frequency of pure haplotypes in the population as a whole at equilibrium depended on the recombination rate in heterokaryotypes (higher for higher L and higher d) and ranged from over 0.99999999 to about 0.96 for different runs (not shown).

Figure 4.4a (top) compares the differentiation at the P locus before and after the introduction of a standard inversion for two-way migration, $t = 0.065$, $d = 0.001$, $m = 3$, $c = 0$. The imperfect final differentiation (≈0.8 rather than ≈1) is primarily due to exchanges of pure genotypes, as indicated in figure 4.4a (bottom), which plots the normalized linkage disequilibrium ($D/D_{\max}$; Lewontin 1964) between the P and T loci in the population as a whole (both habitats combined) before and after the introduction of the inversion. This measure is 1 when all haplotypes are pure (P_0T_0 or P_1T_1), which is approximately the case at the final equilibrium. The differentiation is higher for lower migration rates (figure 4.4b), as expected. A typical progress for an inversion that spread in scenario 1 is shown in figure 4.5. Note that when the inversion spreads, the frequency of P_1 is significantly reduced in habitat 0 and the frequency of pure adapted haplotypes is increased in both habitats, for the reasons discussed. Also note that P_1 spreads at a much lower rate than the inversion, a point to which I will return.

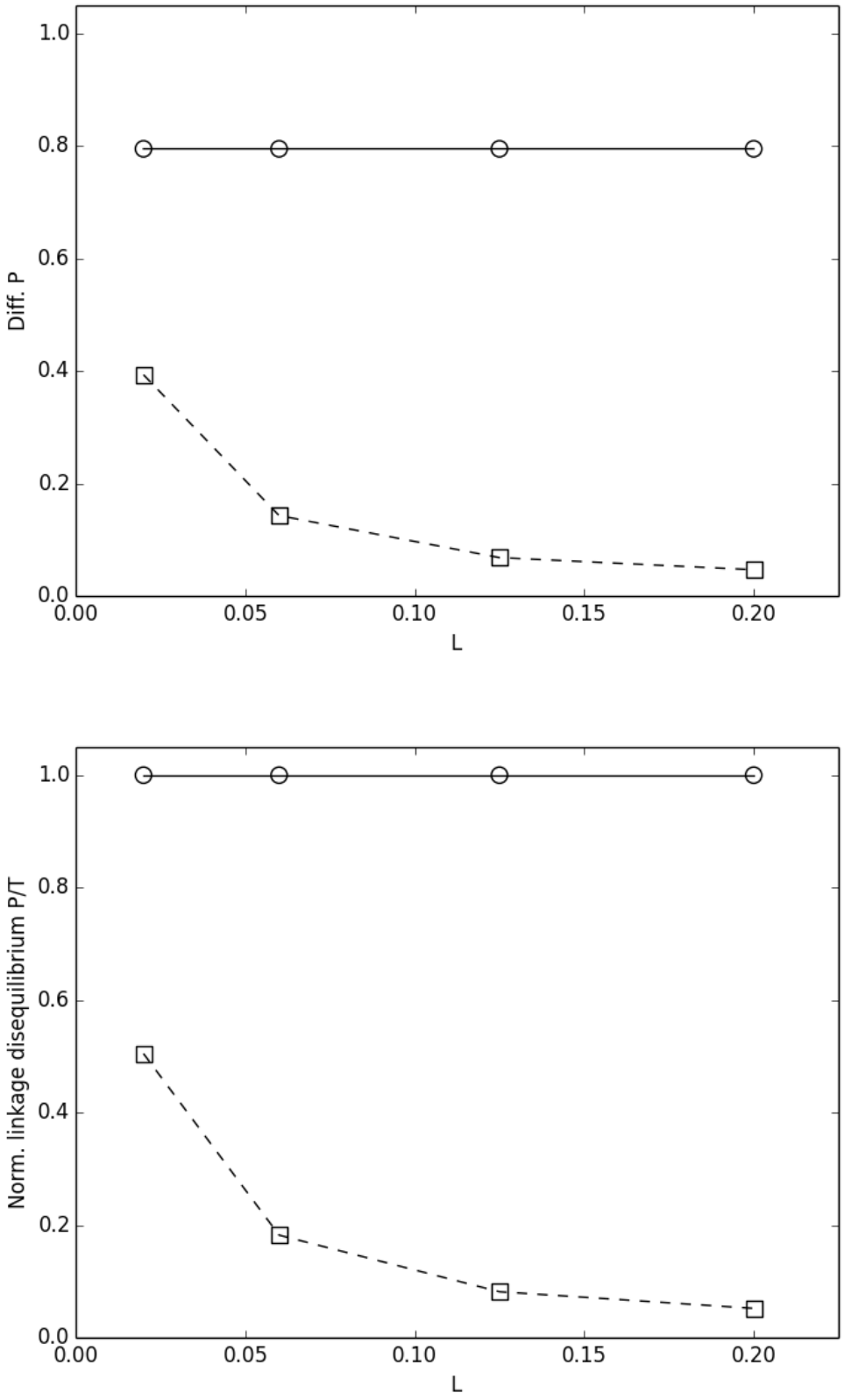


Figure 4.4a: Top: The difference between the equilibrium frequencies of P_1 in habitats 1 and 0 in scenario 1 with symmetrical two-way migration before (squares, dashed) and after (circles, solid) the introduction of the inversion (note the change of axis from figure 4.3, top). Bottom: The normalized linkage disequilibrium between P and T in the pooled population before (squares, dashed) and after (circles, solid) the introduction of the inversion. t = 0.065, c = 0, m = 3, d = 0.001.

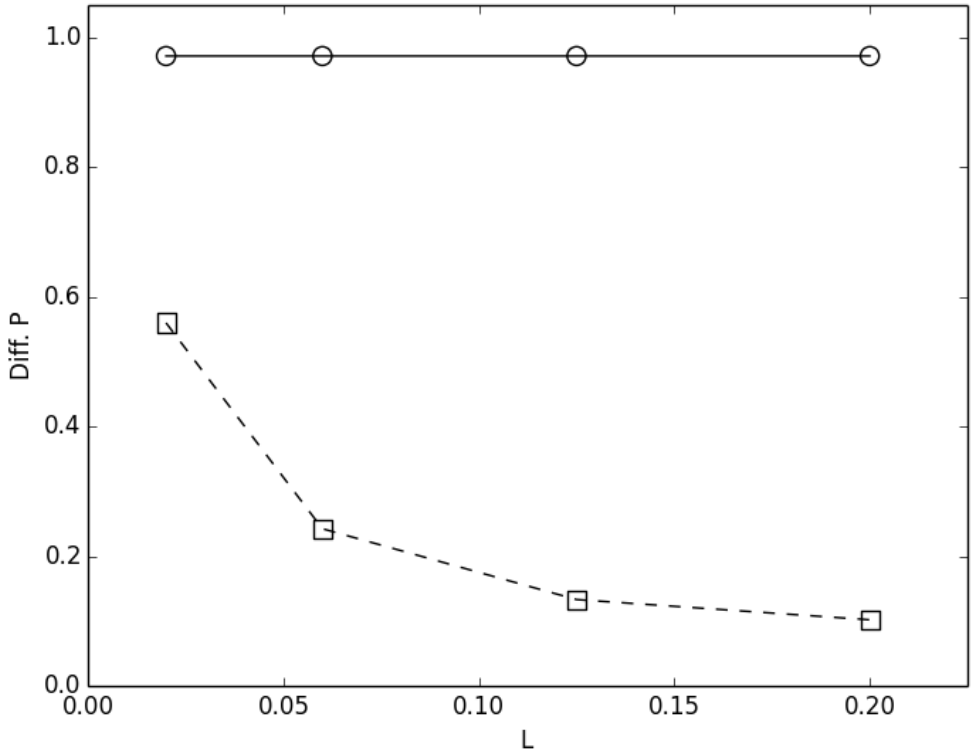


Figure 4.4b: Same as figure 4.4a (top), but for t = 0.01.

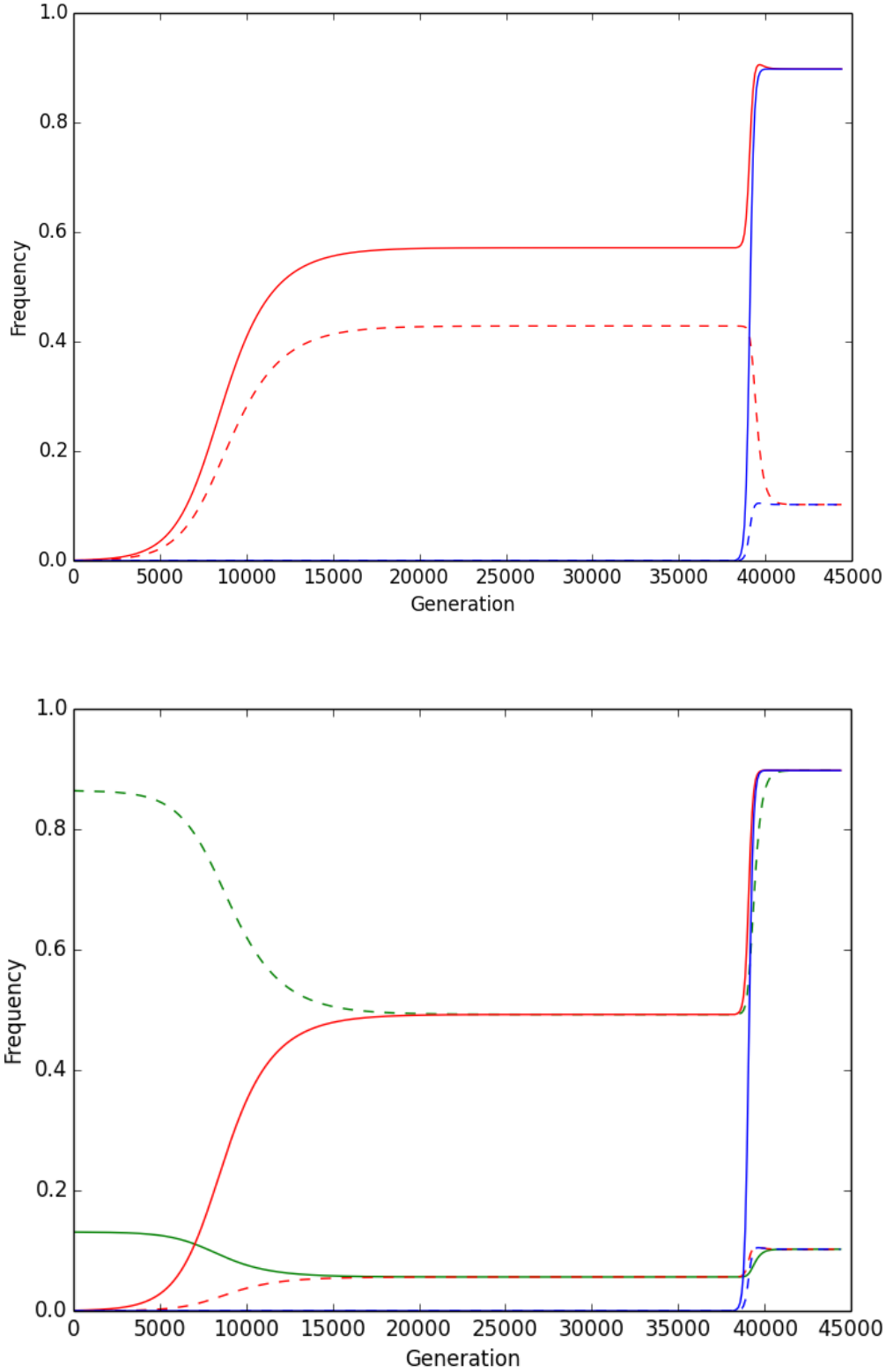


Figure 4.5: Progress example for scenario 1 in habitat 0 (dashed) and habitat 1 (solid). Top: The frequencies of the P_1 allele (red) and the inversion (blue). Bottom: the frequencies of the pure $P_1T_1M_1N_1$ haplotype (red), the pure $P_0T_0M_0N_0$ haplotype (green) and the inversion (blue). The two plots represent the same run. The P_1 allele is introduced at generation 0, and the inversion at generation 39,100. Standard inversion, symmetrical two-way migration, pure counting model, $t = 0.065$, $m = 3$, $d = 0.001$, $L = 0.06$, $c = 0$.

Table 4.2 displays whether or not the standard inversion spread in scenario 1 for *d = 0, 0.001, 0.01, 0.05* and *t = 0.001, 0.01, 0.065.* Interestingly, those results hold for all tested degrees of linkage and interference (*L = 0.02, 0.06, 0.125, 0.2, m = 0, 3, 7)*. An inversion capturing a high-fitness haplotype with more closely linked loci will, all else being equal, spread at a lower rate because if the loci are closely linked then the recombination rates between them are low even in homokaryotypes, so the presence of an inversion makes less of a difference (Kirkpatrick and Barton 2006, Dagilis and Kirkpatrick 2016). On the other hand, as figure 4.6 shows, a long (in genetic length) standard inversion is generally more strongly underdominant (the sterility is generally higher for stronger interference, but for the relevant values of *d* the difference is small). Hence, in my parametrization, these two forces work to the opposite effect when varying *L*, and at the granularity tested here, they approximately cancel out.

	d=0	***d=0.001***	***d=0.01***	***d=0.05***
t=0.001	yes	no	no	no
t=0.01	yes	yes	no	no
t=0.065	yes	yes	yes	no

Table 4.2: Displays whether the standard inversion spread (yes) or not (no) in scenario 1. Valid for L = 0.02, 0.06, 0.125, 0.2; m = 0, 3, 7; c = 0

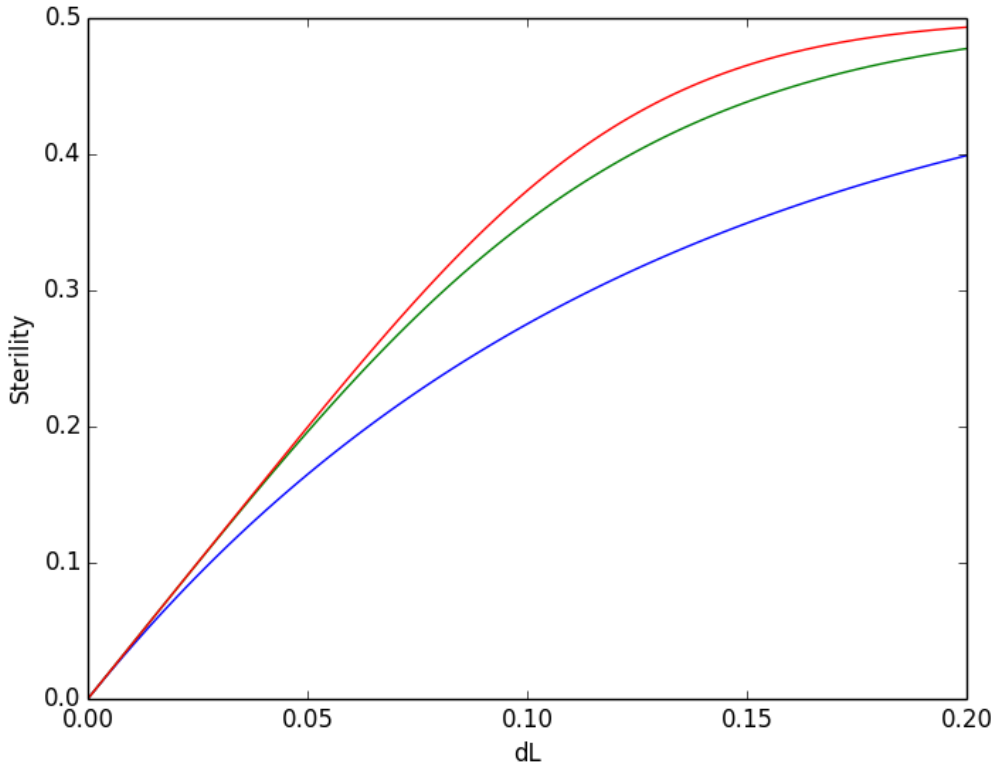


Figure 4.6: The sterility of a standard inversion as a function of dL (the parameters d and L multiplied) for different strengths of interference. Pure counting model, m = 0 (blue, bottom line), m = 3 (green), m = 7 (red, top line). Note that the whole inverted region has genetic length 4dL in heterokaryotypes. The figure also gives the recombination rate for an interval of length 4dL (cf. figure 1 in Foss et al. 1993 and figure 4.3, bottom, in this text).

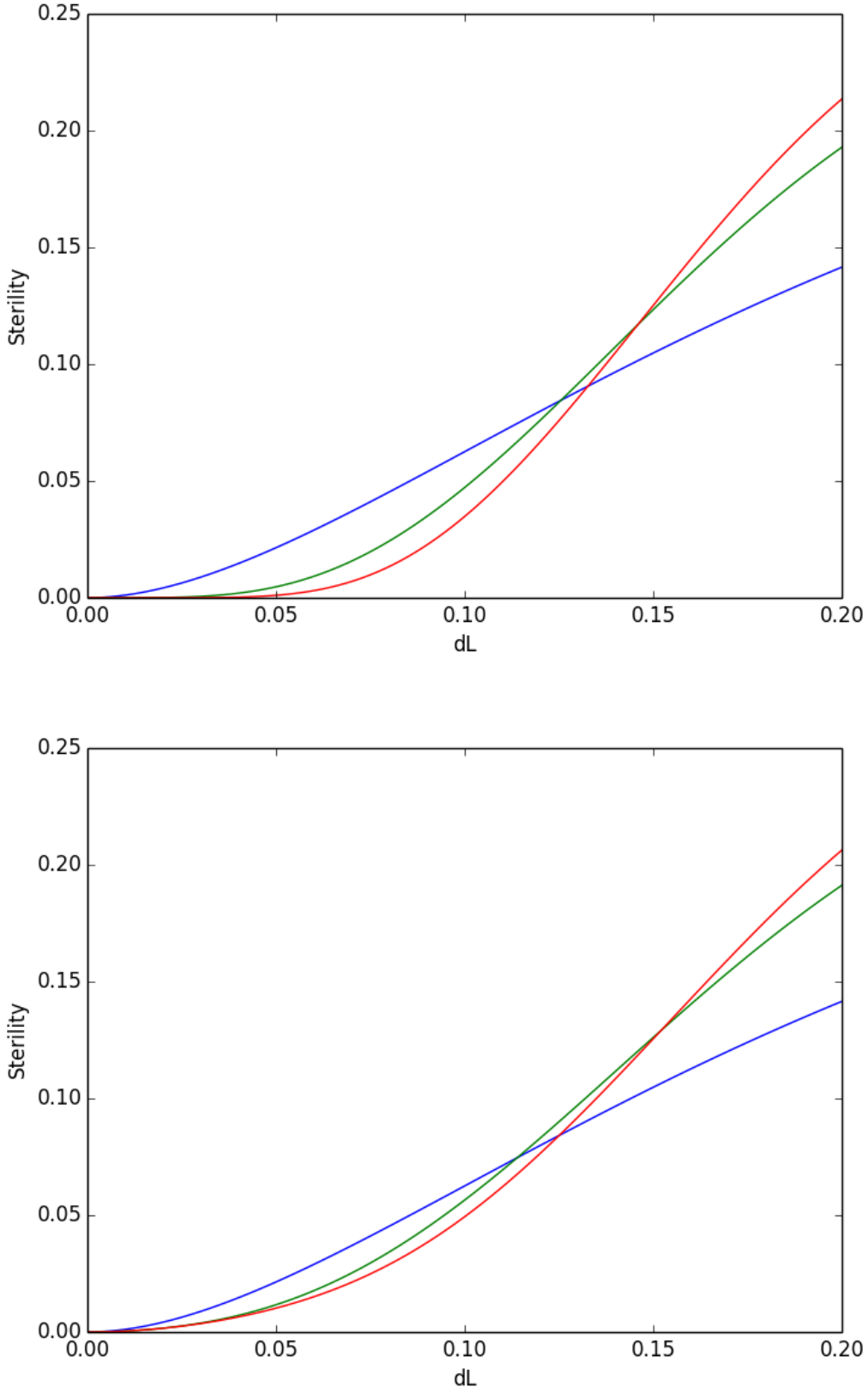


Figure 4.7: the sterility of paracentric linear inversions with $\alpha = 1$ (top) and $\alpha = 0$ (bottom) as a function of dL for m = 0 (blue), m = 3 (green), and m = 7 (red). The inverted and proximal regions have lengths 4dL and dL, respectively. Note that the two blue curves are identical. Also note the change of axis from that in figure 4.6.

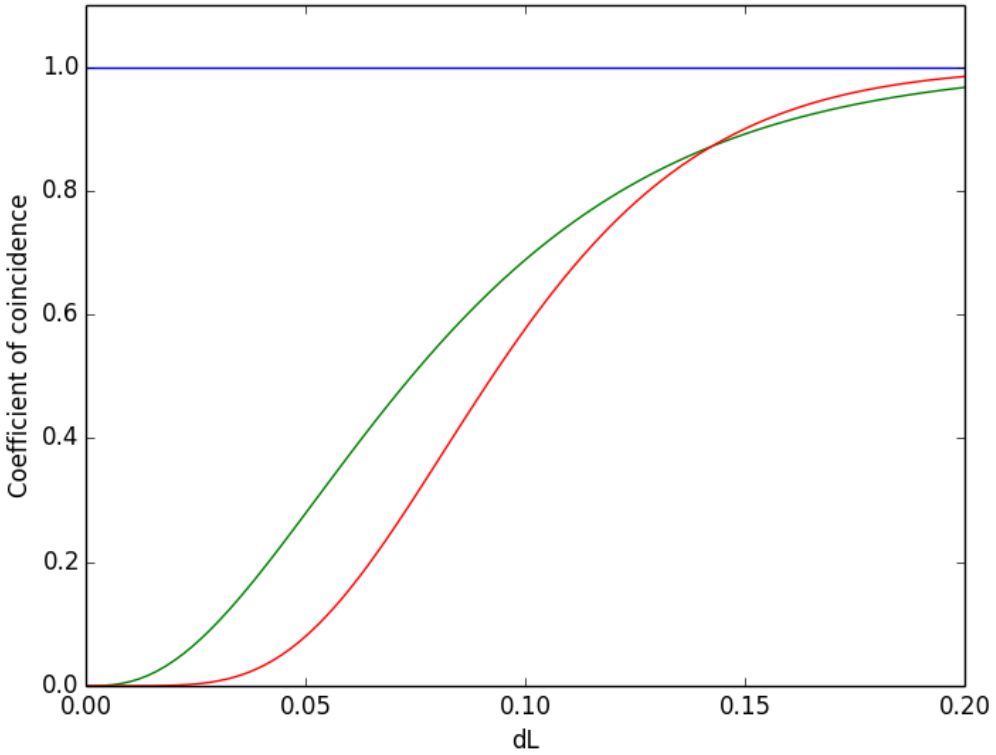


Figure 4.8: The coefficient of coincidence for the two intervals comprising the inverted (4dL) and the proximal (dL) regions, as a function of dL. Pure counting model, m = 0 (blue), m = 3 (green), and m = 7 (red).

Also evident from table 4.2 is the observation that the *d* toleration limit is generally higher for higher migration rates (*t*). This is because when the migration rate is high, there will be more unfit migrants to recombine with, and hence stronger selection for recombination suppressors that makes recombination less likely (Kirkpatrick and Barton 2006, Dagilis and Kirkpatrick 2016). The *d* toleration limit for standard inversions is nevertheless fairly low even for the high migration runs (*t = 0.065)*; in no runs did an inversion with *d = 0.05* spread. In contrast, figure 4.9 plots the *d* toleration limits for a paracentric linear inversion with $\alpha = 1$ (interference across breakpoint boundaries) and $\alpha = 0$ (no interference across breakpoint boundaries) for *m = 0, 3, 7*; *L = 0.02, 0.06, 0.125, 0.2; t = 0.001, 0.065*, showing *d* toleration limits as high as 1.0 (no chiasma inhibition; values higher than 1 were not tested) for short inversions with strong interference. These values were found by gradually increasing the *d* values in steps of 0.05 until the inversion no longer spread. Note that *d = 1.0* imply no chiasma inhibition, but for paracentric linear inversions there is still significant reduction of recombination in heterokaryotypes compared to homokaryotypes, since the unbalanced anaphase 1 chromatids that are retained in the polar bodies in the former case would have been balanced recombinant gametes in the latter case. As the figure shows, the limit for this type of inversion depends strongly on *m* and *L.* We can understand why this is so by comparing the heterokaryotype sterility for standard inversions and paracentric linear inversions (in females) with $\alpha = 1$ and $\alpha = 0$, as shown in figures 4.6 and 4.7. Three things stand out. Firstly, standard inversions are overall much more strongly underdominant than paracentric linear ones. This is expected, since for a paracentric linear inversion heterokaryotype to produce unbalanced gametes, there must be at least one chiasma event in the inverted region *and* one or more additional chiasma event in the inverted region and/or the proximal region, whereas a standard inversion heterokaryotype always produces 50% unbalanced gametes when there is at least one chiasma event in the inverted region (chapter 2;

the latter follows from Mather's equation).

Secondly, the degree of interference makes little difference for the underdominance of short standard inversions, but it makes a large difference for short paracentric linear inversions, in particular when $\alpha = 1$. This is again an intuitive result: the underdominance of standard inversion depends only the probability of observing zero chiasma events in the inverted region (Mather's equation), which for short inversions does not vary much with the degree of interference. For paracentric linear inversions, on the other hand, there must be at least two chiasma events in close proximity for there to be any unbalanced gametes at all, and this is much less likely with strong positive interference since the first chiasma event will inhibit a second one from forming in its vicinity. This effect is even more stark when there is interference across the breakpoint boundaries ($\alpha = 1$), because then a chiasma event in the inverted region interfere not only with other chiasma events in the inverted region, but also with chiasma events in the proximal region. The result is that for strong interference and $\alpha = 1$, a short paracentric linear inversion is almost selectively neutral (compare the curves for *m = 0, m = 3* and *m = 7* in figure 4.7 (top)). Figure 4.8 plots the coefficient of coincidence (see chapter 2) for the inverted and proximal regions, indicating that allowing interference across the breakpoint boundary has a large effect for small L but matters less as L gets larger. For *m = 0,* there is no interference anyway, so the curve is flat at one. Thirdly, and consistently with the results in Navarro et al. (1997), when the inversions are sufficiently long, the sterility increases rather than decreases with stronger interference. This is not relevant for the *d* toleration limits in our case, however, since no inversions above the necessary heterokaryotype length spread in any of the runs. A comparison between the top and bottom plots in 4.9 also show the effect of migration rate, with generally higher *d* toleration limits for higher rates, as discussed above.

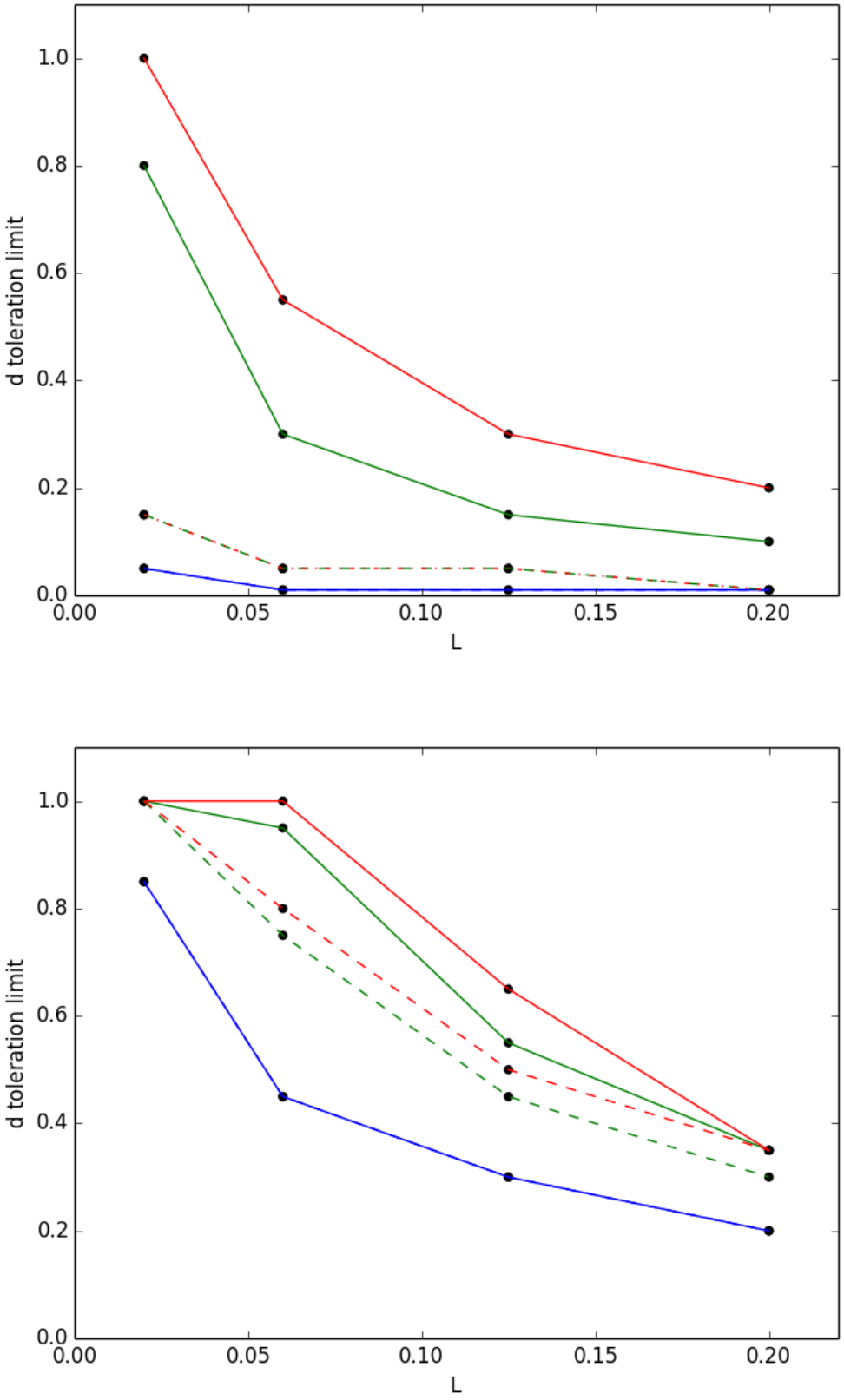


Figure 4.9: The d toleration limits in scenario 1 for a paracentric linear inversion when t = 0.001 (top) and t = 0.065 (bottom). Values larger than 1 were not considered. Two-ways symmetrical migration, pure counting model, m = 0 (blue), m = 3 (green), m = 7 (red), $\alpha = 1$ (solid) and $\alpha = 0$ (dashed). c = 0. The dashed and solid blue lines coincide in both plots. The dashed red and dashed green lines coincide in the top plot.

4.2.3 Scenario 2

As in scenario 1, the newly introduced inversion in scenario 2 experience the opposing evolutionary forces of underdominance due to meiotic irregularities (if $d > 0$) and positive selection for reducing recombination of a directly selected high-fitness haplotype ($T_1M_1N_1$). What differs in scenario 2 is that the P_0 rather than the P_1 allele is initially in close linkage with the inversion, which causes bearers of the new arrangement to prefer mating with T_0 males. Since the T_0 allele is poorly adapted to habitat 1, and is in linkage disequilibrium with the M_0 and N_0 alleles, which are incompatible with the M_1 and N_1 alleles inside the inversion, this introduces an additional force working against the spreading of the inversion (though note that the T_0 allele is rare in habitat 1 and that the preference is by assumption rather weak.)

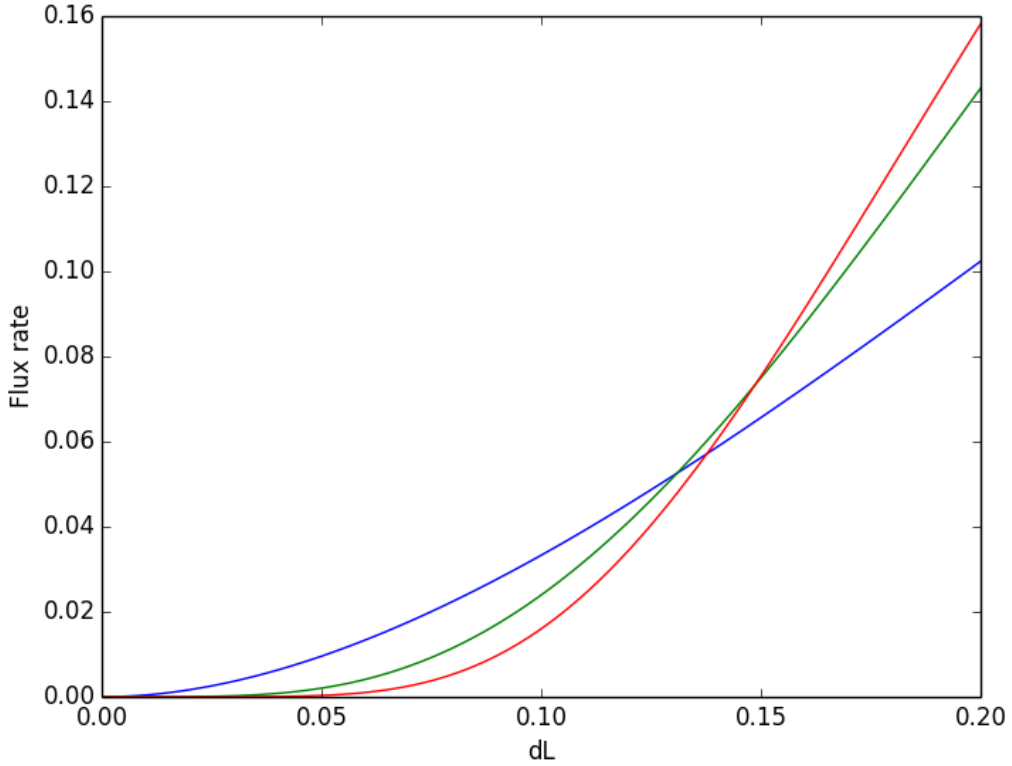
Unlike in scenario 1, I found that the standard inversion did not initially spread in scenario 2 for $L = 0.02$ under any of the tested combinations of migration rates and interference strengths, meaning that it is sometimes positively selected when it captures P_1 but not when it captures P_0. In some of these runs, the inversion accordingly rebounded and spread to divide the population once the P_1 allele had invaded it through recombination, though this implies that its frequency initially decreased to a value lower than the introduction frequency, which is biologically implausible. For $L = 0.06, 0.125, 0.2,$ the standard inversion spread in scenario 2 for the same combinations of t and d as in scenario 1 (table 4.2), regardless of m. The paracentric linear inversion once again showed much higher d toleration limits, though I did not perform systematic investigations to find the exact values for this type of inversion in scenario 2. When an inversion of either type did spread in scenario 2, one of two things happened. Either the inversion went to fixation with the P_0 allele in habitat 1, meaning that the P_1 allele was replaced and lost from the population; or the inversion initially increased with the P_0 allele, before being invaded through recombination by the P_1 allele, so that the pure $[_1P_1T_1M_1N_1]_1$ haplotype replaces the lower-fitness $[_1P_0T_1M_1N_1]_1$ haplotype and the population eventually equilibrated at approximately the same genotype frequencies as in the corresponding run in scenario 1.

Since the P_1 allele decreases in frequency before rebounding once it invades the inversion, its overall lowest frequency – and accordingly the probability that it would be lost in a finite population – depends on the flux rate at the P locus in heterokaryotypes, i.e. the proportion of balanced gametes that show recombination in the interval between $[$ and P *and* the interval between P and $]$. Figures 4.10 plots the flux rate at the P locus in heterokaryotypes as a function of genetic distance (dL) with my parameter settings, for standard inversions and paracentric linear inversions with $\alpha = 1$ and $\alpha = 0$. The shapes of these curves are all similar to each other and to the shape of the sterility curves for paracentric linear inversions (figure 4.7); this is because in all these cases at least two closely spaced chiasma events are required. Consistently with Navarro et al.'s (1997) results, flux rates are lower for stronger interference when inversions are short, again because a chiasma event in the interval between $[$ and P interferes with the generation of the necessary additional chiasma event in the interval between P and $]$. We should therefore expect the P_1 allele to rebound earlier, and at a higher lowest frequency, when interference is weak. I found this to be the case for all runs tested (standard inversion with $\alpha = 1$, $t = 0.001, 0.01, 0.065$; $m = 0, 3, 7$; $d=0, 0.001, 0.01$; $L = 0.06, 0.125, 0.2$; all combinations for which the inversion spread).

Figure 4.11 compares the progress of the P_1 allele in habitat 1 when all parameters except m are held constant at. $L = 0.06$, $t = 0.065$, $d = 0.01$, $c = 0$, showing the expected pattern for both types of inversion (I only tested $\alpha = 1$ for paracentric linear inversions). The figure also show that the minimum frequency of P_1 is lower for standard than for paracentric linear inversion for all values of m. I suggest that this is because the inversion spread at a faster rate when it captures a haplotype with more high-fitness loci (Kirkpatrick and Barton 2006), and the $P_0T_1M_1N_1$ haplotype has higher

fitness in males than in females. Recall that for paracentric inversions, I assume that there is no recombination at all in males (as in *Drosophila*). Therefore the paracentric linear inversion with the $P_0T_1M_1N_1$ allele is selected only in females, for whom P_0 is maladaptive and T_1 is not directly selected, whereas the standard inversion is also selected in males, for whom the P_0 allele is neutral and T_1 is positively selected. Hence, I expect the inversion to spread more slowly, and the P_1 allele to be replaced more quickly in the latter case**.** Figure 4.12 plots the progress the inversion in the same runs as those in figure 4.11**,** indicating that the standard inversions initially spread relatively fast to a high frequency before increasing somewhat further once invaded by P_1, whereas the paracentric linear inversions initially spread relatively slowly before accelerating once invaded. This is seen more clearly in figure 4.13, which shows the progress of the P_1 allele and the $[_1P_1T_1M_1N_1]_1$ and $[_1P_0T_1M_1N_1]_1$ haplotypes in a single run, for both types of inversions. The figures also show that even though the paracentric linear inversion initially spread at slower rate, the total number of generations for the two runs is not necessarily much different.

When interference is strong and the inversion is short (table 4.4), or when $d = 0$ (tables 4.3 and 4.4), recombination in heterokaryotypes is respectively very low (figure 4.10) or zero, so that the P_1 allele does not invade the inversion at all, and is lost (though note that as long as the recombination is non-zero and the population is infinite, the P_1 allele will invade sooner or later if Δ is set to a low enough value).



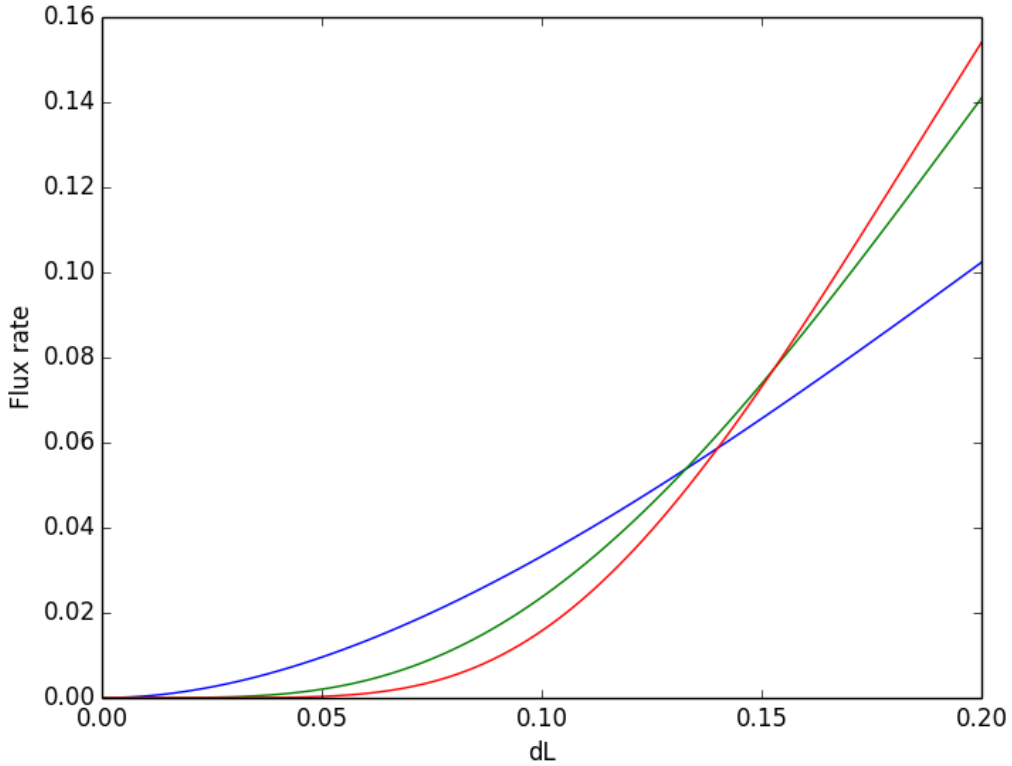


Figure 4.10a: The flux rates at the P locus for a paracentric linear inversion with $\alpha = 1$ *(top)* and $\alpha = 0$ *(bottom), as a function of dL.*

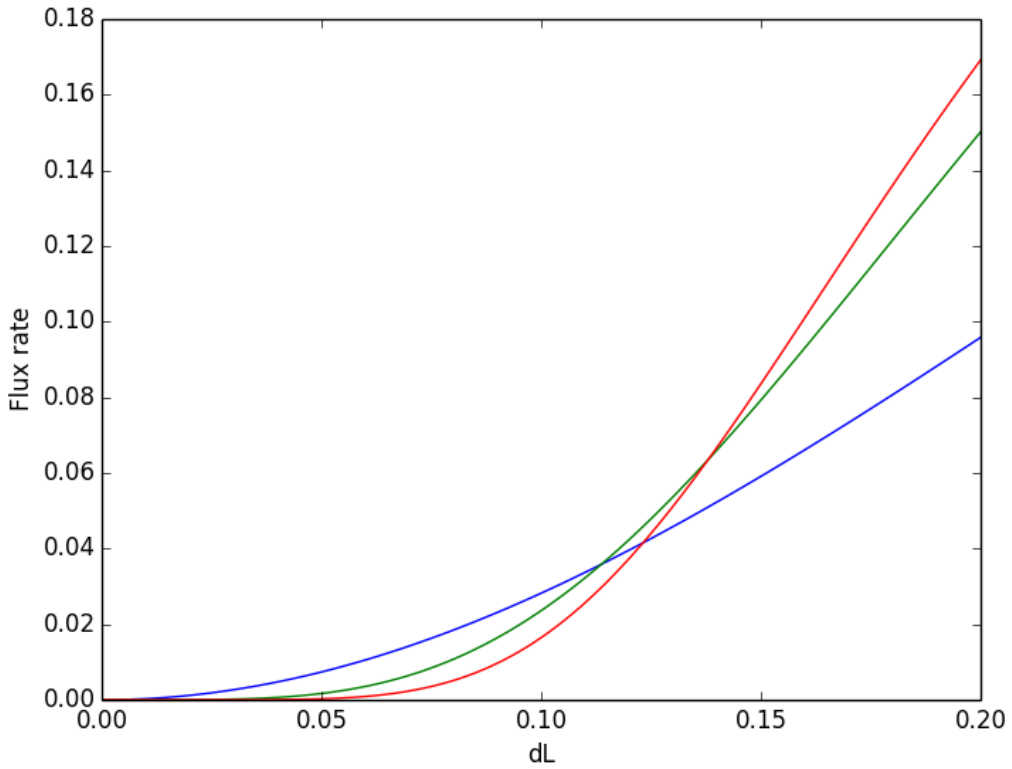


Figure 4.10b: The flux rates at the P locus for a standard inversion as a function of dL.

	d=0	***d=0.001***	***d=0.01***	***d=0.05***
t=0.001	yes (P_0)	no	no	no
t=0.01	yes (P_0)	yes (P_1)	no	no
t=0.065	yes (P_0)	yes (P_1)	yes (P_1)	no

Table 4.3: Displays whether the standard inversion spread (yes) or not (no) in scenario 2, and which allele at the P locus ended up fixed in habitat 1. Valid for L = 0.06, 0.125, 0.2; m = 0, 3

	d=0	***d=0.001***	***d=0.01***	***d=0.05***
t=0.001	yes (P_0)	no	no	no
t=0.01	yes (P_0)	yes (P_0)	no	no
t=0.065	yes (P_0)	yes (P_0)	yes (P_1)	no

Table 4.4: Displays whether the standard inversion spread (yes) or not (no) in scenario 2, and which allele at the P locus ended up fixed in habitat 1. Valid for L = 0.06, 0.125, 0.2; m = 7

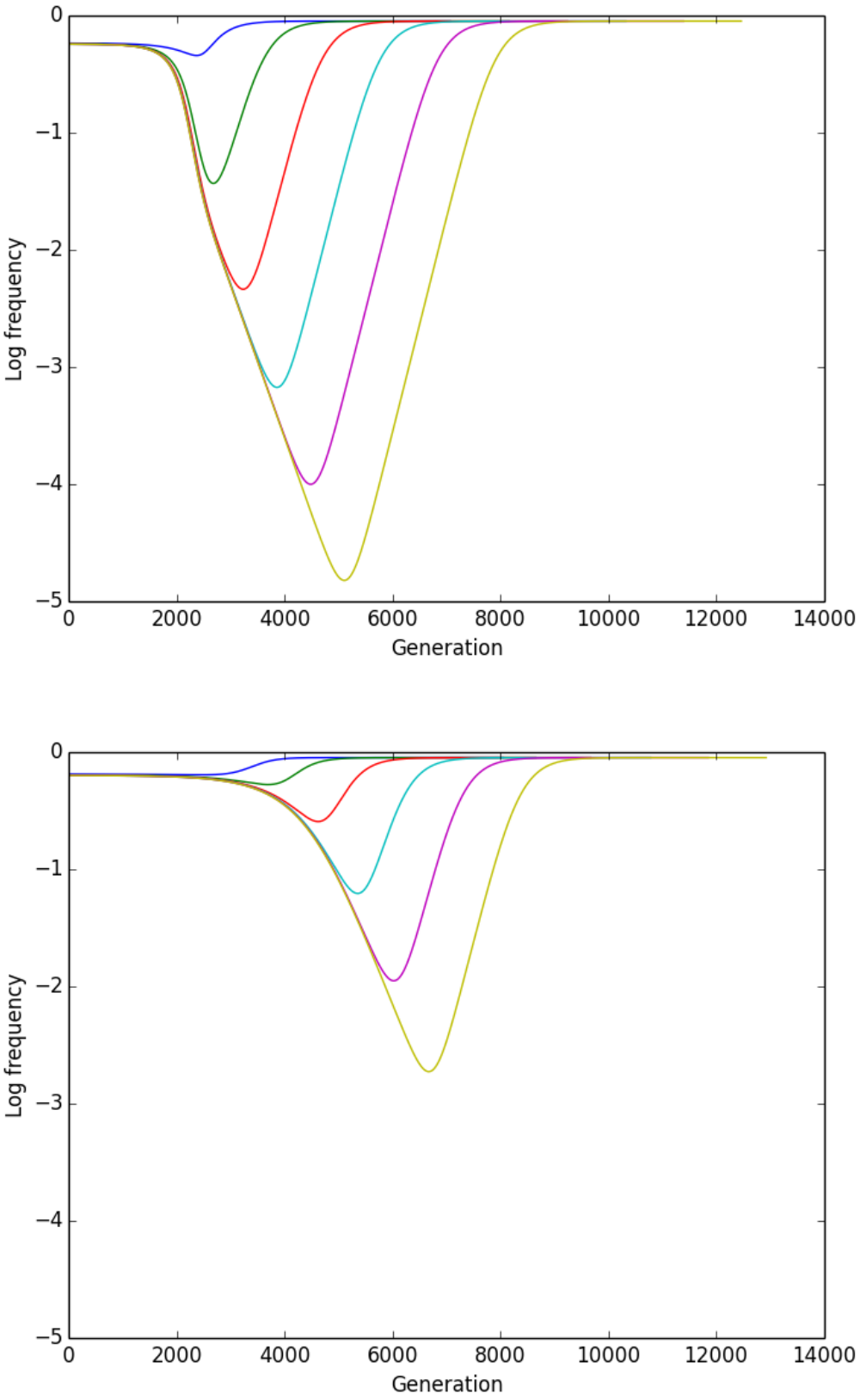


Figure 4.11: The progress of the base 10 logarithm of the frequency of P_1 in habitat 1, scenario 2 for m = 0, 1, 2, 3, 4, 5 (top to bottom curves). Top: standard inversion. Bottom: paracentric linear inversion, $\alpha = 1$. All runs, both plots: L = 0.06, t = 0.065, d = 0.01, c = 0

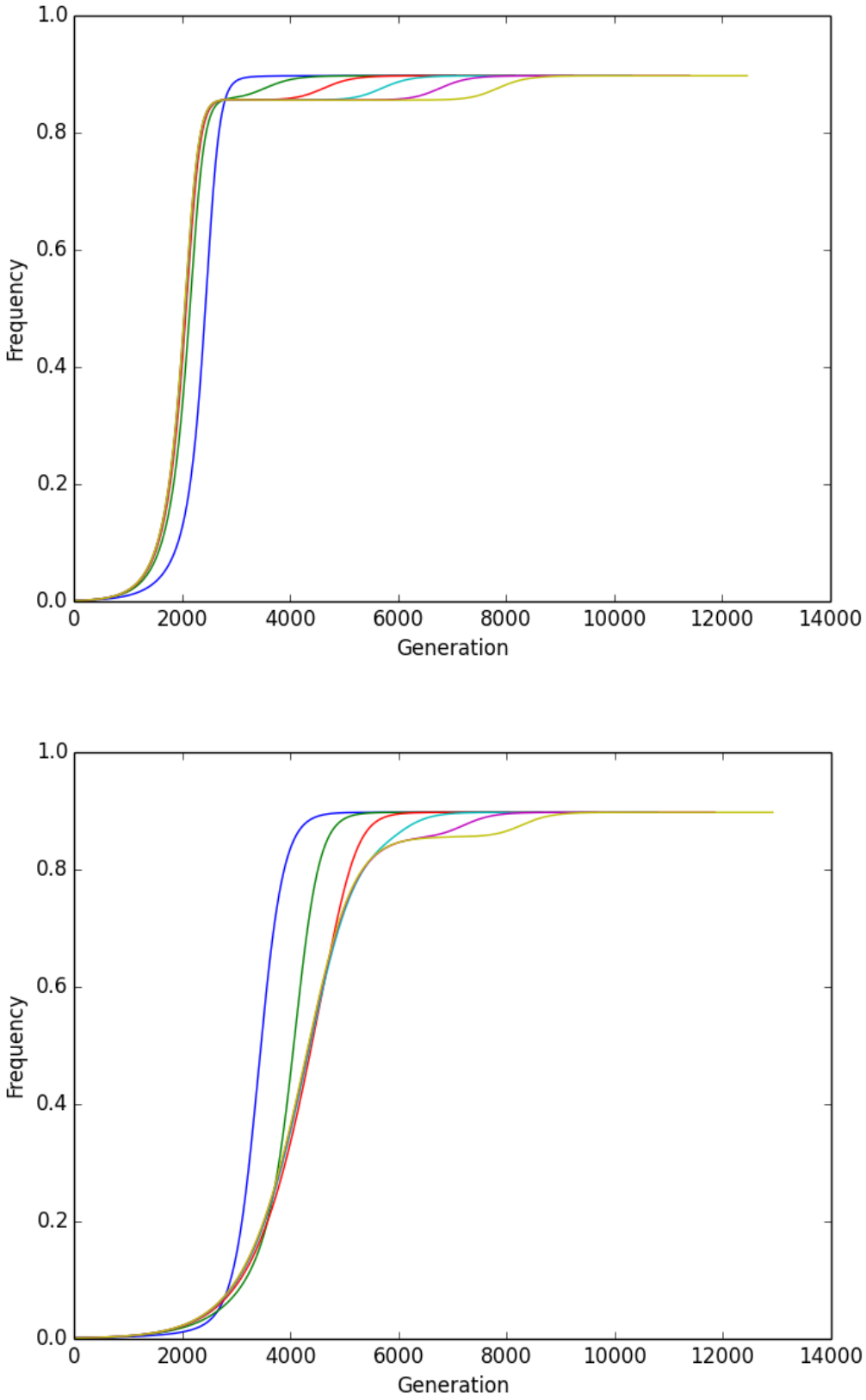


Figure 4.12: The progress of the inversion in habitat 1, scenario 2, for m = 0, 1, 2, 3, 4, 5 (same colors as in figure 4.11). Top: standard inversion. Bottom: paracentric linear inversion, $\alpha = 1$. L = 0.06, t = 0.065, d = 0.01, c = 0.

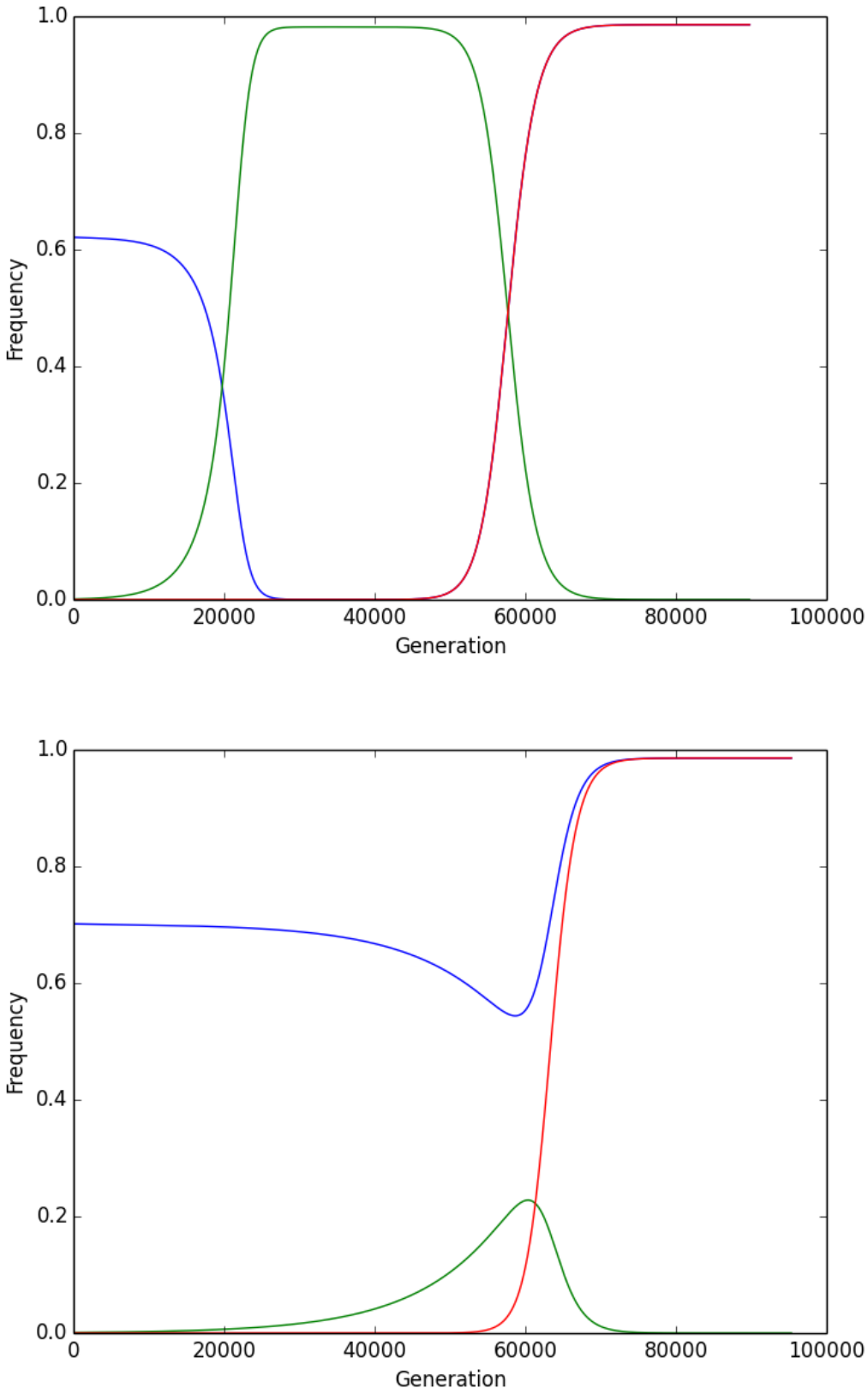


Figure 4.13: The progress of the P_1 allele (blue), the $[_1P_0T_1M_1N_1]_1$ haplotype (green) and the $[_1P_1T_1M_1N_1]_1$ haplotype (red) for a standard inversion (top) and paracentric linear inversion (bottom). The blue and red curves coincide in the rightmost half of the top plot. Both runs: $t = 0.01$, $d = 0.001$, $m = 3$, $L = 0.06$, $c = 0$.

4.2.4 Scenario 3

In scenario 3, the P_1 allele is introduced inside the already (nearly) fixed inversion, meaning that the high-fitness $P_1T_1M_1N_1$ haplotype mostly avoids recombining with locally maladapted 0-index alleles, so that the trait allele (T_1) remains a reliable indicator of male fitness. Since recombination is the major obstacle to reinforcement (Felsenstein 1981) and the initial spreading of the P_1 allele seems to be the limiting step in scenario 1 (see figure 4.5), scenario 3 tests the idea that reversing the order of the equilibrium actions can accelerate the process. Figure 4.14 plots the total number of generations needed to reach the final equilibrium in the control scenario, scenario 1 and scenario 3 for *t = 0.001, 0.065, d = 0, m = 3,* showing that scenario 3 reaches equilibrium significantly faster than both the control scenario and scenario 1 when *L > 0.02.* The results for *d > 0* are almost identical (figure 4.15). That the inversion spread slightly faster for higher *L* in scenario 3 is consistent with Kirkpatrick and Barton's (2006) results showing that an inversion with $d_{\mathbb{h}} = 0$ that capture a high-fitness haplotype spread at a lower rate when recombination in homokaryotypes is lower; as mentioned above, this is because the difference between the recombination rates in homokaryotypes and heterokaryotypes is then less significant. Figure 4.15 shows that this effect is also present when $d = d_{\mathbb{h}} = 0.001$. Scenarios 1 and 2 show the opposite pattern, because in those cases the limiting step is the initial spreading of the P_1, which is impeded when there is high recombination between pre- and post-zygotic loci (since the trait allele is then no longer a reliable indicator of male fitness). Also note the substantial effect of varying *t*: for *t = 0.065* the longest run lasted about 133,000 generations, whereas for *t = 0.001* the number is over 3,750,000. This is presumably because higher migration brings more maladapted individuals to habitat 1, which increases the selection on both the inversion (Kirkpatrick and Barton 2006, Dagilis and Kirkpatrick 2016) and the preference allele. Figure 4.16 plots the progress of the P_1 allele and the inversion in scenarios 1 and 3 for two different settings of *L*, showing that some, but not all, of the difference is due to P_1 allele taking longer to settle on an equilibrium after reaching a high frequency in scenario 1.

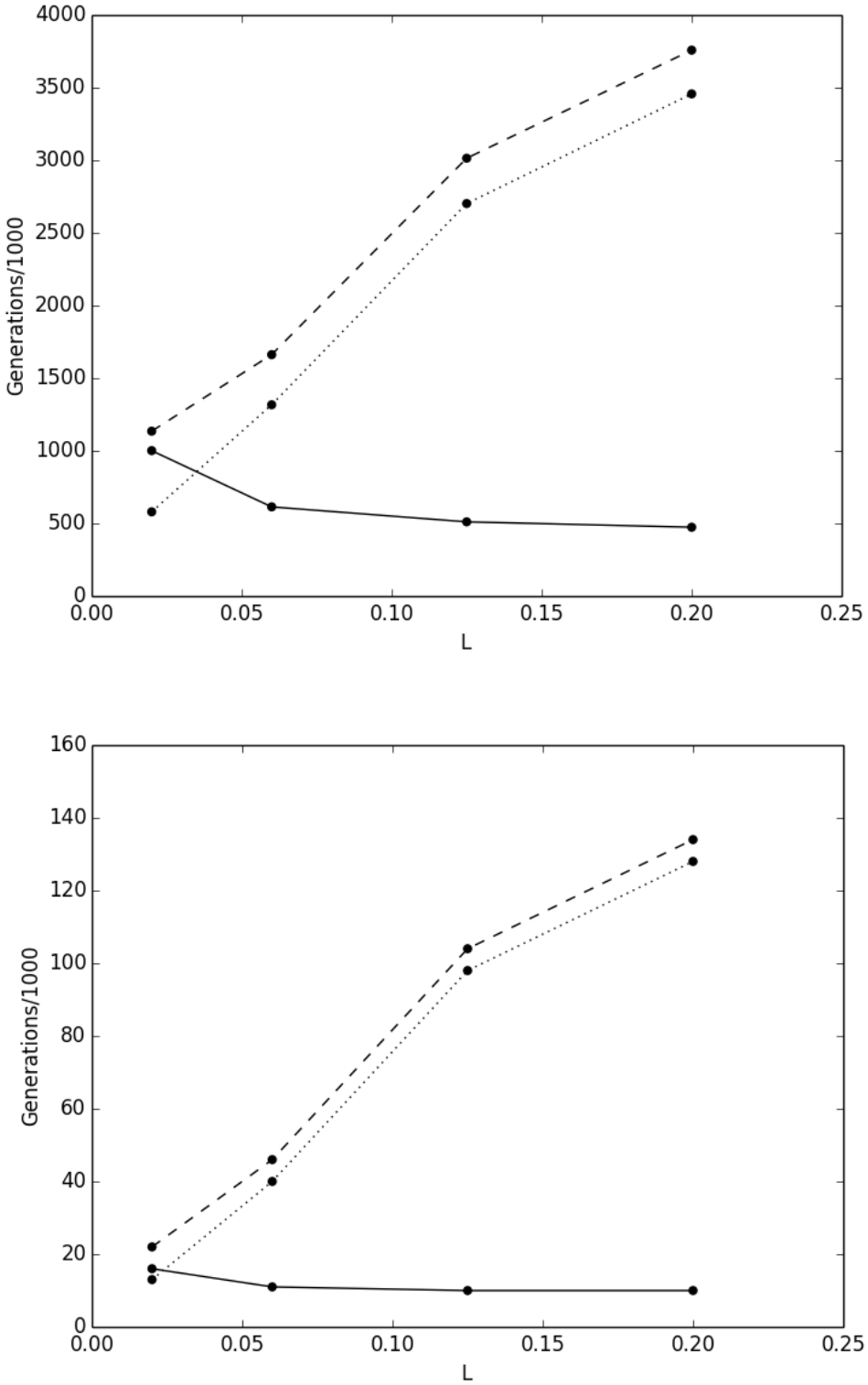


Figure 4.14: The total number of generations (in thousands) needed to reach the final equilibrium in the control scenario (dotted), scenario 1 (dashed), and scenario 3 (solid). Top: t = 0.001. Bottom: t = 0.065. Both figures: d = 0, m = 3, c = 0. All generations are rounded up to the nearest thousand. Note that it necessarily takes more generations to reach equilibrium in scenario 1 than in the control, because the former include the latter plus one extra step.

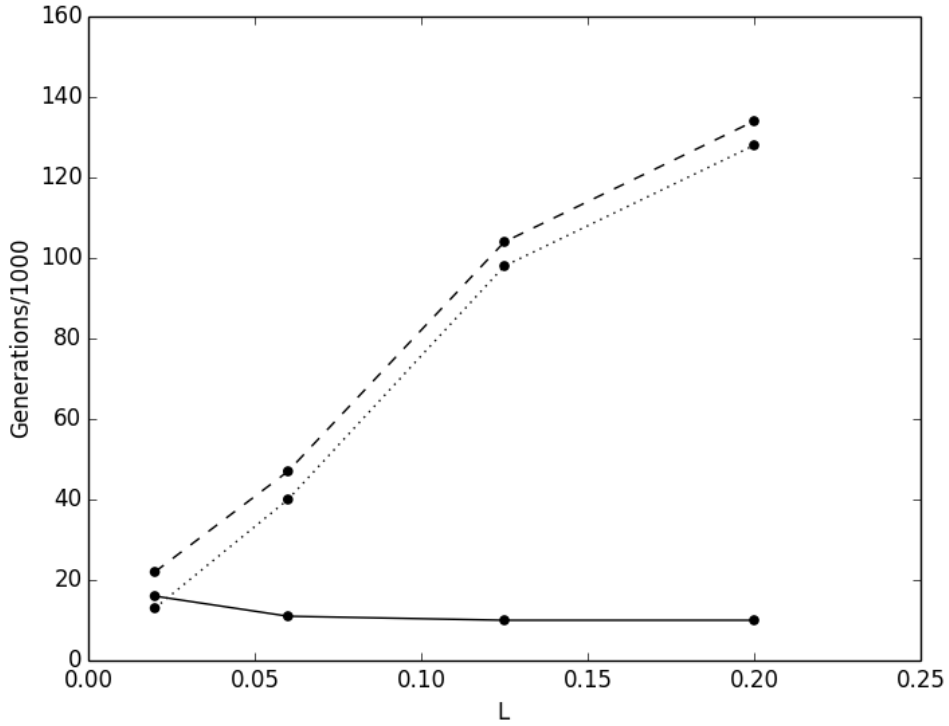


Figure 4.15: Same as figure 4.14, bottom, except d = 0.001

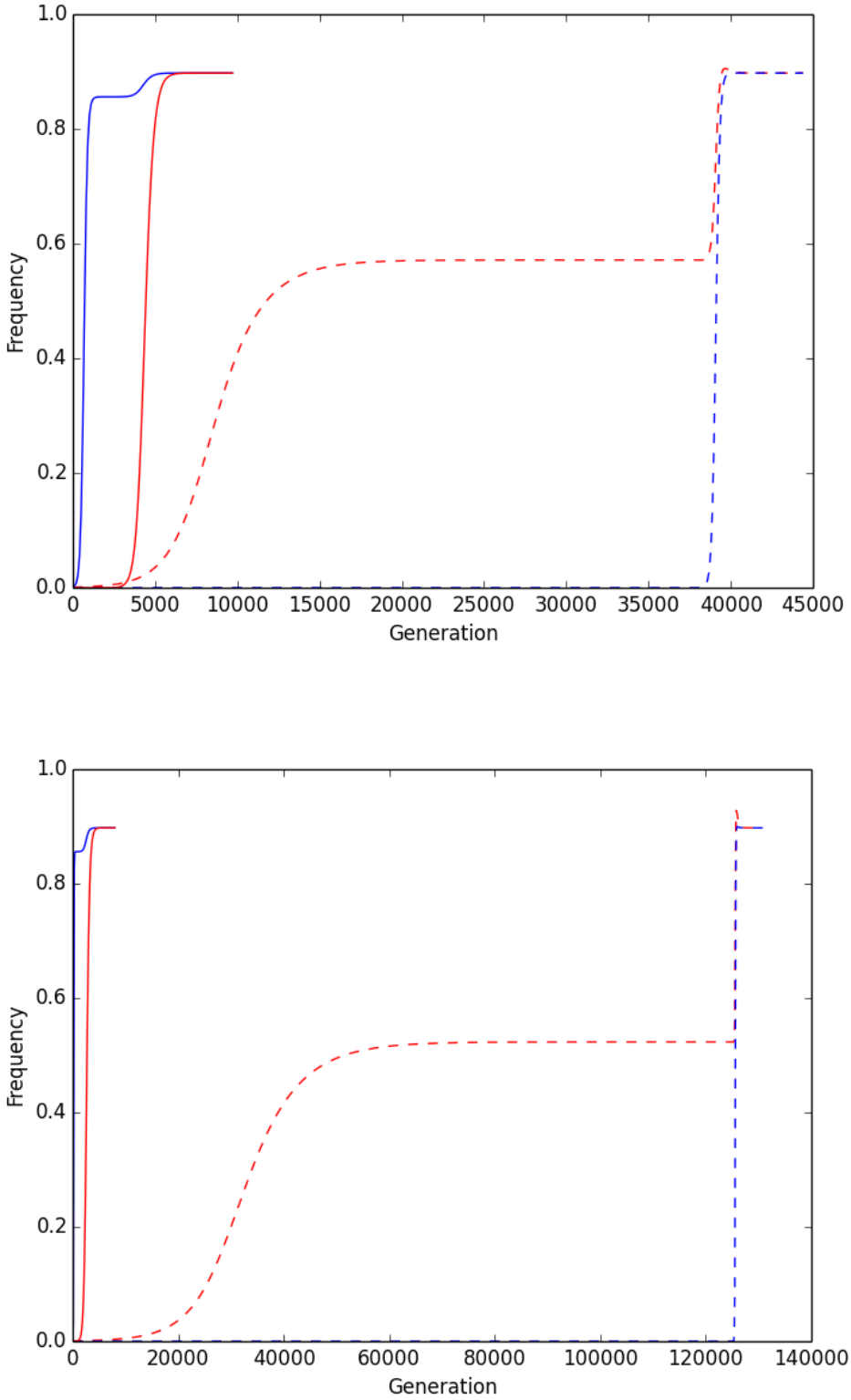


Figure 4.16: Example progress plots for scenario 1 (dashed) and scenario 3 (solid) for equivalent parameter settings. Blue: Inversion. Red: P_1. Top: L = 0.06, Bottom: L = 0.2. All runs: standard inversion. Note that, unlike in figures 4.14 and 4.15, the progress before the introduction of P_1 and the inversion is not included. The solid blue curve in the bottom plot rises so sharply that it almost coincides with the y-axis. All runs: t = 0.065, m = 3, d = 0.001.

I also tested the idea that an initially (nearly) fixed inversion can make reinforcement happen more easily when the P_1 allele is introduced on the island in a continent-island model (one-way migration). One-way migration is generally less conducive to reinforcement than two-way migration because the direct effect of migration of P_0 individuals from the continent can swamp the indirect effect of selection on P_1 on the island (Servedio and Kirkpatrick 1997). Figure 4.17 compares the final equilibrium frequency of P_1 on the island in the control scenario for $c = 0, 0.005, 0.01, 0.05, t = 0.01, m = 3$, and scenario 3 for the same settings except $c = 0$, showing that for $c = 0$, the P_1 allele does not spread when $L > 0.02$ (meaning that scenario 1 would be impossible), but it *does* spread almost to fixation for all values of L when an inversion is introduced as intermediate step (i.e. in scenario 3). The figure also shows the effect of varying c, the cost of searching for mates, in the control scenario. As I showed in chapter 3, when $c > 0$, females favoring rare males are disadvantaged, and increasingly so with increasing c. Since P_0 females favor T_0 males, which are rare on the island, and vice versa for P_1, we should expect reinforcement to happen more readily for higher values of c, which the figure shows is the case. Nevertheless, the equilibrium frequencies in the control scenario for the highest value of c tested (0.05) is still not as high as the ones for the corresponding runs in scenario 3 with $c = 0$.

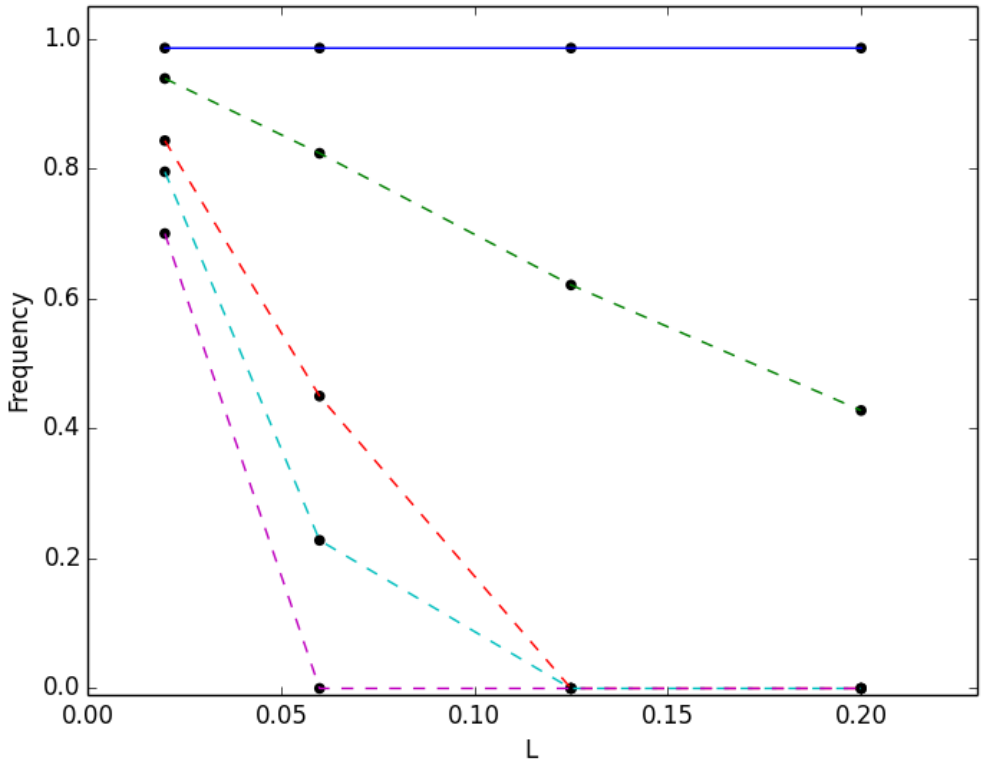


Figure 4.17: The equilibrium frequency of the P_1 allele in habitat 1 for one-way migration (continent-island model). Solid line: scenario 3, $c = 0$, $d = 0.001$ (no runs with $c > 0$ were executed). Dashed lines: Control scenario, $\tau = 1$, $c = 0.05, 0.01, 0.005, 0$ (top to bottom lines). All runs: $t = 0.01$, $m = 3$.

5 Discussion

Previous studies have investigated whether or not a perfectly chiasma-suppressing inversion ($d_{\mathbb{h}} = 0$ in my terminology) can spread and facilitate reinforcement when it captures an adapted set of pre- and postzygotic isolation loci (Tricket and Butlin 1994, Dagilis and Kirkpatrick 2016). However, the effects of recombination and underdominance, of capturing a maladapted preference allele, and of reversing the order in which the preference allele and inversion appears, have not yet been examined. The results in the previous chapter indicate that a chromosomal inversion can enhance differentiation at the preference loci and cause the two habitats to be fixed for alternative optimal haplotypes (scenario 1), even when initially capturing a maladapted preference allele (scenario 2), and that the process is significantly accelerated when the preference allele is introduced after the inversion (scenario 3). Furthermore, although my simulations only examine the initial stages of reinforcement, I will suggest that an inversion can enhance differentiation further in later stages for four reasons.

Firstly, my results indicate that the inversions can spread and divide the population into the two pure haplotypes even when underdominant ($d > 0$). This means that almost all individuals with co-adapted index 1 alleles in habitat 1 will also have the derived arrangement ($[_1$ and $]_1$), and vice versa for index 0, so that the underdominance of the inversion enhances the postzygotic barrier to gene flow. I stress that the underdominant inversions in my model are favored by selection, which is in contrast to classical chromosomal rearrangement speciation models (e.g. White 1978) in which the inversion spread by drift against a selection gradient in allopatry.

Secondly, as Navarro and Barton (2003) show, new pairs of universally favored two-allele incompatibilities (see figure 4.2) that appear successively in parapatry are more likely to cause differentiation of the two subpopulations when located inside an inversion than when located in a colinear region. This is because when a new allele at a locus M (allele M_0) starts spreading in habitat 0, the reduced gene flow between the two arrangements means that there is more time for an incompatible allele N_1 at locus N to appear and spread in habitat 1 before this habitat is also invaded by M_0 (see figure 5.1). They furthermore show that each such pair of fixed incompatibility alleles that differentiates the two arrangements will make the recruitment of additional incompatibility alleles more likely, because of the further reduction in gene flow, and because higher differentiation increases the probability that new mutations will be incompatible with at least some alleles in the other habitat. Hence, a snowball effect of increasing postzygotic barriers can be initiated.

Thirdly, the enhanced postzygotic barriers will in turn create stronger selection for enhanced prezygotic barriers, since the disadvantage of mating with individuals from the other deme will be larger. Since stronger prezygotic barriers will reduce gene flow further by making cross-species mating less likely, it will accelerate the Navarro-Barton postzygotic snowball effect and initiate a positive feedback loop between the recruitment of new pre- and postzygotic isolation alleles.

Fourthly, additional preference alleles will spread faster when the inversion is present, for the reasons discussed in the previous chapter (scenario 3).

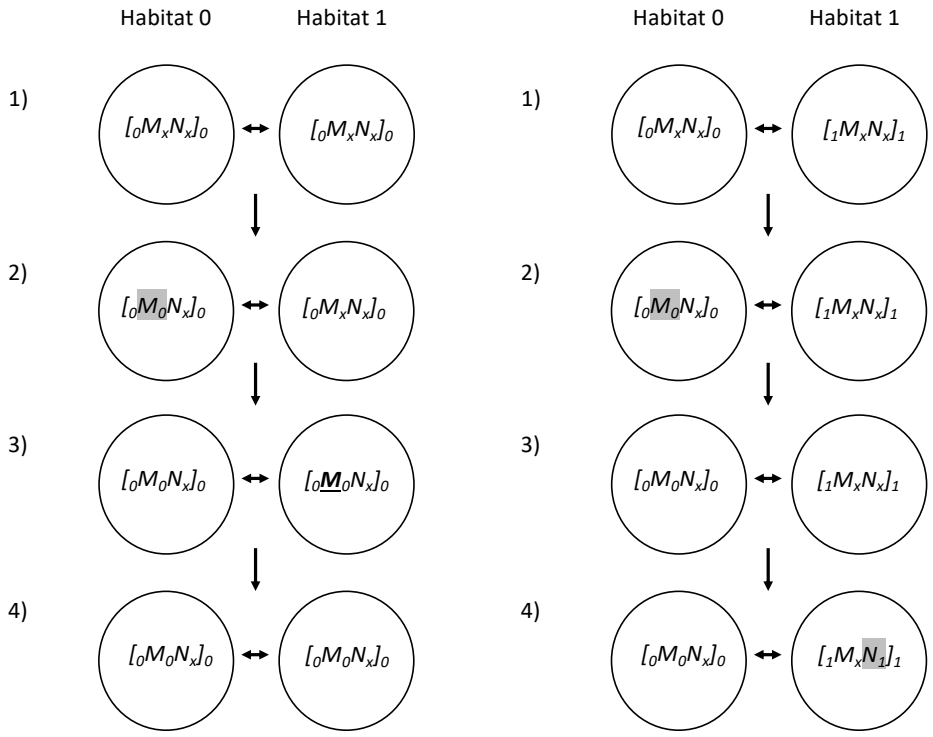


Figure 5.1: Navarro and Barton's (2003) model of the differentiation of two-allele incompatibilities in parapatry. Left: No inversion polymorphism. Right: Inversion polymorphism. M_0 and N_1 are universally selected and incompatible. Grey indicates new fixed mutation, bold underline indicates new fixed migrant. The numbers represent time, as follows. Left: 1) The subpopulations starts out the same. 2) The M_0 allele appear in habitat 0. 3) The M_0 allele, being universally selected, quickly spread into habitat 1. 4) The N_1 allele appear as a mutation in habitat 1, but is inhibited from spreading by the presence of the incompatible M_0 allele. The two habitats ends up the same. Right: 1) The two subpopulations starts fixed for the M_xN_x haplotype, but differentiated by an inversion polymorphism (maintained by some sort of selection). 2) M_0 appear in habitat 0. 3) The invasion of M_0 into habitat 1 is delayed because of low recombination in inversion heterokaryotypes. 4) N_1 appear as a mutation in habitat 1 and spreads, hence preventing M_0 from invading. The two subpopulations end up differentiated at all four loci.

The previous paragraphs are reminiscent of Servedio and Sætre's (2003) suggestion of a positive feedback loop between pre- and post-zygotic isolation genes in collinear (i.e. non-inverted) regions. According to this hypothesis, the frequency of post-zygotic isolation alleles (M_1 and N_1 in a four-allele incompatibility model) can increase in habitat 1 by hitchhiking with the newly introduced preference allele (P_1). Their simulations (and mine, results not shown) show that this does indeed happen, though the effect is small, even for closely linked loci (in Servedio and Sætre's simulations the frequency of the M_1N_1 haplotype increases by around e^{-8} when $L \approx 0.01$ and around e^{-13} to e^{-14} when $L > 0.1$ for autosomes, and by slightly more when M and N recombine freely and/or the loci are sexlinked). The increased frequency of the M_1 and N_1 alleles in habitat 1 will, the hypothesis goes, further increase the selection pressure on new prezygotic isolation alleles, hence initiating the feedback loop. It is not intuitively obvious to me that this latter effect is real, as one could make the opposite argument that selection against mating with maladapted (M_0 and N_0) individuals is lower rather than higher when there are fewer maladapted individuals around, and to my knowledge it is yet to be confirmed and, if it is real, quantified. Either way, since the increase in the frequency of the M_1N_1 haplotype is small, the effect on the selection pressure on new prezygotic alleles will presumably be small as well.

The simulations of scenario 2 show that the inversion can cause the fixation of the P_1 allele in habitat 1 even when it does not initially capture it, though it can also cause it to become lost if the flux rates are too low and/or the inversion spread to fixation too fast. I will conjecture that this finding is quite general, or more precisely that an inversions that (1) spread by capturing a differentially adaptive haplotype with high but suboptimal fitness in parapatry, and (2) show some recombination in heterokaryotypes, will tend to either absorb remaining differentially adaptive alleles through recombination and end up fixed with an optimal or close to optimal haplotype, or cause the non-included high-fitness alleles to become lost, depending on linkage, *d*, interference, and the speed at which the inversion spreads. Hence, higher recombination in heterokaryotypes can facilitate differentiation when the inversion does not capture the full optimal haplotype. Note that this effect works only one way, since low-fitness alleles that invade the inversion through recombination are not selected.

My results might at first seem to contradict the ones in Feder and Nosil (2009), which are sometimes cited as a potential reason to doubt the effect of inversions on speciation (e.g. Faria and Navarro 2010). In that study, the authors ran simulations of different models with two postzygotic isolation loci in two habitats initially fixed for alternative chromosomal arrangements (loci order *[MN]*), and found that in some runs the differentiation between the two habitats quickly disappeared whenever there is some recombination in heterokaryotypes. In all those runs, however, they assumed that the inversion captures a haplotype with low fitness in both demes, and that the alternative haplotype with highest fitness is not differentially adaptive. For example, in their model 3 the *M* and *N* loci are standard Dobzhansky-Muller two-allele incompatibilities (see figure 4.2), and the two habitats start out fixed for the haplotypes $[_0M_0N_x]_0$ and $[_1M_xN_1]_1$, respectively, where M_x and N_x are the ancestral compatible alleles, and M_0 and N_1 are derived and incompatible but not positively selected. Accordingly, M_x and N_x are universally compatible, and when allowed to recombine, they take over both habitats. The only evolutionary force influencing the inversion will then be underdominance from meiotic irregularities in heterokaryotypes, meaning that the more common one will soon become universally fixed. This does not happen when the inversion captures a differentially adaptive haplotype, like the M_1N_1 haplotype in my four-allele incompatibility model, or the M_xN_1 haplotype in a two-allele model where N_1 and M_0 are positively selected (so that the fitness of M_xN_1 is higher than that of M_xN_x and M_0N_1 in deme 1, but lower than that of the prevalent M_0N_x in deme 0), as in Feder and Nosil's model 4, and in the later (Time 4 in figure 5.1, right) stages of Navarro and Barton's (2003) model (the latter explicitly assume some sort of selection maintaining the initial differentiation of the alternative arrangements at time 1). Hence, the key factor determining whether or not differentiation is maintained in spite of recombination in Feder and Nosil's study is simply whether or not the differentiation is adaptive. In other words, while recombination erodes neutral and maladaptive differentiation (Feder and Nosil's models 1, 3, and 5), it does not necessarily erode *adaptive* differentiation, and in some cases it actually facilitates it (this study, scenario 2). Furthermore, Feder and Nosils' models assume that the habitats are initially fixed for the alternative arrangements without explaining how this came to be; if the inversion does not capture a high-fitness haplotype then it does not spread by selection in parapatry, and will not become fixed in the first place unless one invokes drift against the selection gradient in allopatry (assuming that the inversion, being subject to recombination, is also underdominant). There is therefore reason to doubt whether such scenarios are representative of naturally occurring inversions.

In sum, I suggest that in order to have a stable inversion polymorphism between two subpopulation, there needs to be some kind of selection to establish and maintain it. Such selection can either be the result of reduced recombination of local high-fitness (but not necessarily optimal) haplotypes (Charlesworth and Charlesorth 1973, Tricket and Butlin 1994, Kirkpatrick and Barton 2006, Dagilis and Kirkpatrick 2014, this text), or simply from direct selection on the inversion's

effects on gene expression (Avelar et al. 2013). Once such a polymorphism is in place, further differentiation can, I suggest, occur through the processes discussed here and in Navarro and Barton (2003), even when there is some recombination in heterokaryotypes.

As expected, I found a stark difference between the *d* toleration limit for standard and paracentric linear inversions. In the former case, the inversion did not spread for *d=0.05*, even under the most favorable conditions (scenario 1, highest possible migrations rate). In the dataset in Coyne et al. (1993) and Navarro and Ruiz (1997), only inversions shorter than about 0.25 Morgan have *d* values in this range (see figure 2.4, this text); if this is a general result, it could impose an upper limit on the length of standard inversions that can spread in scenarios similar to the ones considered here. This is a testable hypothesis that can be addressed in future empirical studies. Paracentric linear inversions, on the other hand, are only weakly underdominant even when *d* is quite large, and accordingly they did spread in my simulations for a wide range of *d* values, even for the maximum value of 1 when conditions are ideal. As many people have noted before me (e.g. Coyne et al. 1993), this is probably the reason why paracentric inversion polymorphisms are much more common than pericentric ones in *Drosophila* (Stone 1955), which is one of the groups with linear meiosis. My results from scenario 2 furthermore indicate that a paracentric linear inversion that does *not* initially capture the adapted preference allele is less likely to cause it to get lost from the population, compared to a standard inversion, presumably because of the assumption of no recombination in males.

Future research should investigate whether the *Drosophila d* values in Navarro and Ruiz (1997) are representative for other species. It would also be interesting to systematically compare the *d* values of laboratory-induced inversions (as in the Coyne/Navarro and Ruiz-dataset) with those of naturally occuring inversion polymorphisms. Naturally occuring inversion polymorphisms are sometimes found to not be underdominant at all (Nachman and Myers 1989, Coyne et al. 1991) implying $d \approx 0$, though it is not clear whether this is because more underdominant inversions did not spread in the first place or because the inversions spread when underdominant and subsequently underwent selection to further suppress chiasma formation.

All theories and assumptions, however well established, should and must be open to further criticism (Popper 1934/2002). I will therefore end this thesis with a critical look at some of mine. Firstly, I have throughout this text disregarded the effect of gene conversions. Although the rate of flux from gene conversion is typically small compared to that from crossing over in homokaryotypes, this need not necessarily be the case in heterokaryotypes, especially when the inversion is short and interference is strong (Navarro et al. 1997). My simulations therefore probably underestimate the degree of flux in heterokaryotypes for a given value of *d*. This might be considered a conservative assumption for scenario 2, in that higher gene flux rates would make it easier for the P_1 allele to invade, but not necessarily so for the other scenarios. Allowing for gene conversions would not affect the degree of underdominance of the inversion and would increase flux in homokaryotypes as well as heterokaryotypes, so I would not expect it to have a large effect in scenarios 1 and 3, except perhaps that the *d* toleration might be slightly lowered. Another interesting implication of gene conversions is that a short paracentric linear inversions with strong interference and $\alpha = 1$ would be nearly selectively neutral (i.e. not underdominant) while still allowing some degree of gene flux. I plan to include gene conversions in a future version of my program.

Secondly, while there is plenty of evidence that interference depends on genetic, as opposed to physical, distance in homokaryotypes (chapter 2), evidence either way is to my knowledge lacking in heterokaryotypes. My assumption in this text is that the intermediate events are suppressed by a factor *d*, and that the counting process works on this depleted set in the same manner that it does in homokaryotypes. This might seem like a straightforward implication of the counting models, but, as I mentioned in chapter 2, much work remains to determine which aspects of these models are

physically real, and which are just useful mathematical abstractions. In essence, the problem is that the counting models lack a *good explanation* (Deutsch 2011) for why they work so well. I here use the term *good explanation* in Deutsch's non-standard meaning of "an explanation that is hard to vary while still accounting for what it purports to account for". As he explains, the ancient Greek myth that seasonal variation in weather is caused by the mood swings of the goddess Demeter is not a good explanation in this regard, as one can easily vary the myth to account for any weather pattern, or indeed anything at all. By contrast, the theory that seasons are due to the earth's axial tilt *is* a good explanation: the theory makes clear, non-trivial and unchangeable predictions – *risky* predictions in Popper's (1963) terminology – that allow us to unequivocally reject it if proven false. A good explanation, furthermore, has *reach,* in Deutch's (2011) sense of being able to address problems beyond those which it was designed to solve. Hence, just as a good explanation of seasons informs us about seasons in other locations and on other planets, so a good explanation of how interference works in homokaryotypes might tell us, by implication, how it works within an inverted region and across breakpoint boundaries, how and why it varies across species, and all sorts of other things we are yet to consider. At the very least, a good explanation will tell us where to look, and what to test. While the counting models have proved adept at modelling interference in a wide range of species, the theory of why they work so well is at present not a good explanation in the sense used here; as long as the mechanism of the hypothetical "machine that can count" (Foss and Stahl 1995) remains as elusive as Demeter's mood swings, the inferences we can draw from one study to the next will remain limited. For this reason, a good, hard to vary, long-reaching, and provably true explanation of how interference actually works, will be infinitely more valuable than any set of data points.

Finally, I recognize that the simulations in the previous chapter explore only a small region of the vast parameter space in which reinforcement and chromosomal evolution can occur. This is mostly for reasons of simplicity, though I do not consider it a major shortcoming. The theories presented in this thesis do not *consist* of the outputs of the model designed to test them; the former, unlike the latter, have reach that stretches well into hitherto unexplored areas of parameter space, in which they can be further tested and eventually refuted or improved. I intend to take up this task.. In particular, I plan to investigate how the Navarro-Barton postzygotic snowball effect interacts with the evolution of new preference alleles in later stages of divergence, and how the whole process is affected by chiasma interference and recombination, with and without a chromosomal inversion. Other studies could follow in this path, or focus on testing the influence of parameters that I have here investigated only briefly or not at all, such as the cost of searching, the symmetry and degree of mating preferences, or the type and strength of selection on the postzygotic isolation loci.

References

Anton, E., Blanco, J., Egozcue, J., & Vidal, F. (2005). Sperm studies in heterozygote inversion carriers: a review. *Cytogenetic and genome research*, 111(3-4), 297-304.

Avelar, A. T., Perfeito, L., Gordo, I., & Ferreira, M. G. (2013). Genome architecture is a selectable trait that can be maintained by antagonistic pleiotropy. *Nature communications*, 4, 2235.

Ayala, D., Guerrero, R. F., & Kirkpatrick, M. (2013). Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. *Evolution: International Journal of Organic Evolution*, 67(4), 946-958.

Berchowitz, L. E., & Copenhaver, G. P. (2010). Genetic interference: don't stand so close to me. *Current genomics*, 11(2), 91-102.

Carson, H. L. (1946). The selective elimination of inversion dicentric chromatids during meiosis in the eggs of *Sciara impatiens*. *Genetics*, 31(1), 95-113.

Charlesworth, B., & Charlesworth, D. (1973). Selection of new inversions in multi-locus genetic systems. *Genetics Research*, 21(2), 167-183.

Cobbs, G. (1978). Renewal process approach to the theory of genetic linkage: the case of no chromatid interference. *Genetics*, 89(3), 563-581.

Colombo, P. C., & Jones, G. H. (1997). Chiasma interference is blind to centromeres. *Heredity*, 79(2), 214.

Copenhaver, G. P., Housworth, E. A., & Stahl, F. W. (2002). Crossover interference in Arabidopsis. *Genetics*, 160(4), 1631-1639.

Coyne, J. A., Aulard, S., & Berry, A. (1991). Lack of underdominance in a naturally occurring pericentric inversion in Drosophila melanogaster and its implications for chromosome evolution. *Genetics*, 129(3), 791-802.

Coyne, J. A., Meyers, W., Crittenden, A. P., & Sniegowski, P. (1993). The fertility effects of pericentric inversions in Drosophila melanogaster. *Genetics*, 134(2), 487-496.

Coyne, J. A., & Orr, H. A. (2004). *Speciation.* Sunderland, MA: Sinauer.

Dagilis, A. J., & Kirkpatrick, M. (2016). Prezygotic isolation, mating preferences, and the evolution of chromosomal inversions. *Evolution*, 70(7), 1465-1472.

Deutsch, D. (2011). *The Beginning of Infinity*. London: Penguin

Dobzhansky, T. (1937). *Genetics and the Origin of Species.* New York: Columbia University Press.

Eiben, A. E. & Smith, J.E. (2015). *Introduction to Evolutionary Computing.* Springer

Erdelyi, A. (Ed.) (1955). *Higher Transcendental Functions, vol. 3*. New York: McGraw-Hill.

Etges, W. J., & Levitan, M. (2004). Palaeoclimatic variation, adaptation and biogeography of inversion polymorphisms in natural populations of Drosophila robusta. *Biological Journal of the Linnean Society*, 81(3), 395-411.

Faria, R., & Navarro, A. (2010). Chromosomal speciation revisited: rearranging theory with pieces of

evidence. *Trends in Ecology & Evolution*, 25(11), 660-669.

Feder, J. L., Roethele, J. B., Filchak, K., Niedbalski, J., & Romero-Severson, J. (2003). Evidence for inversion polymorphism related to sympatric host race formation in the apple maggot fly, *Rhagoletis pomonella. Genetics*, 163(3), 939-953.

Feder, J. L., & Nosil, P. (2009). Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions?. *Evolution: International Journal of Organic Evolution*, 63(12), 3061-3075.

Felsenstein, J. (1981). Skepticism towards Santa Rosalia, or why are there so few kinds of animals?. *Evolution*, 35(1), 124-138.

Foss, E., Lande, R., Stahl, F. W., & Steinberg, C. M. (1993). Chiasma interference as a function of genetic distance. *Genetics*, 133(3), 681-691.

Foss, E. J., & Stahl, F. W. (1995). A test of a counting model for chiasma interference. *Genetics*, 139(3), 1201-1209.

Futuyma, D. J. (2013). *Evolution* (3rd ed.). Sunderland, MA: Sinauer.

Gethmann, R. C. (1988). Crossing over in males of higher Diptera (Brachycera). *Journal of Heredity*, 79(5), 344-350.

Gomulkiewicz, R. S., & Hastings, A. (1990). Ploidy and evolution by sexual selection: a comparison of haploid and diploid female choice models near fixation equilibria. *Evolution*, 44(4), 757-770.

Gorlov, I. P., & Borodin, P. M. (1995). Recombination in single and double heterozygotes for two partially overlapping inversions in chromosome 1 of the house mouse. *Heredity*, 75(2), 113.

Guillaume, F., & Rougemont, J. (2006). Nemo: an evolutionary and population genetics programming framework. *Bioinformatics*, 22(20), 2556-2557.

Haldane, J. B. S. (1919). The combination of linkage values and the calculation of distances between the loci of linked factors. *J. Genet*, 8(29), 299-309.

Hillers, K. J. (2004). Crossover interference. Current Biology, 14(24), R1036-R1037.

Houle, D., & Kondrashov, A. S. (2002). Coevolution of costly mate choice and condition-dependent display of good genes. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1486), 97-104.

Housworth, E. A., & Stahl, F. W. (2003). Crossover interference in humans. *The American Journal of Human Genetics*, 73(1), 188-197.

Jaarola, M., Martin, R. H., & Ashley, T. (1998). Direct evidence for suppression of recombination within two pericentric inversions in humans: a new sperm-FISH technique. *The American Journal of Human Genetics*, 63(1), 218-224.

Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., ... & Wilkinson, P. A. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature, 477(7363), 203.

King, J. S., & Mortimer, R. K. (1990). A polymerization model of chiasma interference and

corresponding computer simulation. *Genetics*, 126(4), 1127-1138.

Kirkpatrick, M. (1982). Sexual selection and the evolution of female choice. *Evolution*, 36(1), 1-12.

Kirkpatrick, M., & Ravigné, V. (2002). Speciation by natural and sexual selection: models and experiments. the american naturalist, 159(S3), S22-S35.

Kirkpatrick, M., & Barton, N. (2006). Chromosome inversions, local adaptation and speciation. *Genetics*, 173(1), 419-434.

Kleckner, N., Zickler, D., Jones, G. H., Dekker, J., Padmore, R., Henle, J., & Hutchinson, J. (2004). A mechanical basis for chromosome function. *Proceedings of the National Academy of Sciences of the United States of America*, 101(34), 12592-12597.

Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D. S., Hoeppner, M. P., ... & Chen, W. (2016). Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). *Nature Genetics*, 48(1), 84.

Lande, R., & Stahl, F. W. (1993). Chiasma interference and the distribution of exchanges in Drosophila melanogaster. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 58, pp. 543-552). Cold Spring Harbor Laboratory Press.

Lange, K., Zhao, H., & Speed, T. P. (1997). The Poisson-skip model of crossing-over. *The Annals of Applied Probability*, 299-313.

Langtangen, H. P. (2008). *Python Scripting for Computational Science.* Springer

Lewontin, R. C. (1964). The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics*, 49(1), 49.

Malkova, A., Swanson, J., German, M., McCusker, J. H., Housworth, E. A., Stahl, F. W., & Haber, J. E. (2004). Gene conversion and crossing over along the 405-kb left arm of Saccharomyces cerevisiae chromosome VII. *Genetics*, 168(1), 49-63.

Mary, N., Barasc, H., Ferchaud, S., Priet, A., Calgaro, A., Loustau-Dudez, A. M., ... & Pinton, A. (2016). Meiotic recombination analyses in pigs carrying different balanced structural chromosomal rearrangements. *PloS one*, 11(4), e0154635

Mather, K. (1938). Crossing-over. *Biological Reviews*, 13(3), 252-292.

McPeek, M. S., & Speed, T. P. (1995). Modeling interference in genetic recombination. *Genetics*, 139(2), 1031-1044.

Muller, H. J. (1916). The mechanism of crossing-over. *The American Naturalist*, 50(592), 193-221.

Munz, P. (1994). An analysis of interference in the fission yeast Schizosaccharomyces pombe. *Genetics*, 137(3), 701-707.

Nachman, M. W., & Myers, P. (1989). Exceptional chromosomal mutations in a rodent population are not strongly underdominant. *Proceedings of the National Academy of Sciences*, 86(17), 6666-6670.

Navarro, A., Betrán, E., Barbadilla, A., & Ruiz, A. (1997). Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics*, 146(2), 695-709.

Navarro, A., & Ruiz, A. (1997). On the fertility effects of pericentric inversions. *Genetics*, 147(2), 931-

933.

Navarro, A., & Barton, N. H. (2003). Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. *Evolution*, 57(3), 447-459.

Nolan, J. P. (2017). Erlang renewal models for genetic recombination. *Journal of Statistical Distributions and Applications*, 4(1), 10.

Noor, M. A. (1999). Reinforcement and other consequences of sympatry. *Heredity*, 83(5), 503.

Noor, M. A., Grams, K. L., Bertucci, L. A., Almendarez, Y., Reiland, J., & Smith, K. R. (2001). The genetics of reproductive isolation and the potential for gene exchange between Drosophila pseudoobscura and D. persimilis via backcross hybrid males. *Evolution*, 55(3), 512-521.

Novitski, E., & Braver, G. (1954). An analysis of crossing over within a heterozygous inversion in Drosophila melanogaster. *Genetics*, 39(2), 197-209.

Pegueroles, C., Ordóñez, V., Mestres, F., & Pascual, M. (2010). Recombination and selection in the maintenance of the adaptive value of inversions. *Journal of evolutionary biology*, 23(12), 2709-2717.

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Müller, I., Baglione,V., Unneberg, P., Wikelski, M., Grabherr, M. G. & Wolf, J. B. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, 344(6190), 1410-1414.

Popper, K. (1963). *Conjectures and Refutations.* London: Routledge

Popper, K. (2002). *The Logic of Scientific Discovery.* London: Routledge (Original work published in German 1934, in English 1959)

del Priore, L., & Pigozzi, M. I. (2015). Heterologous synapsis and crossover suppression in heterozygotes for a pericentric inversion in the Zebra Finch. Cytogenetic and genome research, 147(2-3), 154-160.

Roberts, P. A. (1976). The genetics of chromosome aberration. In M. Ashburner and E. Novitski (Eds.) *The Genetics and Biology of Drosophila* (67-184). Academic Press: London

Ross, S. M. (2014). *Introduction to probability models*. London: Academic press.

Schaeffer, S. W., & Anderson, W. W. (2005). Mechanisms of genetic exchange within the chromosomal inversions of D. pseudoobscura. *Genetics*, 171. 1729-1739.

Servedio, M. R., & Kirkpatrick, M. (1997). The effects of gene flow on reinforcement. *Evolution*, 51(6), 1764-1772.

Servedio, M. R. (2000). Reinforcement and the genetics of nonrandom mating. *Evolution*, 54(1), 21-29.

Servedio, M. R., & Noor, M. A. (2003). The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, 34(1), 339-364.

Servedio, M. R., & Sætre, G. P. (2003). Speciation as a positive feedback loop between postzygotic and prezygotic barriers to gene flow. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1523), 1473-1479.

Speed, T. P. (1996). What is a genetic map function?. In *Genetic Mapping and DNA sequencing* (pp. 65-

88). New York, NY: Springer.

Sinclair-Waters, M., Bradbury, I. R., Morris, C. J., Lien, S., Kent, M. P., & Bentzen, P. (2018). Ancient chromosomal rearrangement associated with local adaptation of a postglacially colonized population of Atlantic Cod in the northwest Atlantic. *Molecular ecology*, 27(2), 339-351.

Stahl, F. W., Foss, H. M., Young, L. S., Borts, R. H., Abdullah, M. F., & Copenhaver, G. P. (2004). Does crossover interference count in Saccharomyces cerevisiae?. *Genetics*, 168(1), 35-48.

Stam, P. (1979). Interference in genetic crossing over and chromosome mapping. *Genetics*, 92(2), 573-594.

Stone, W. S. (1955) Genetic and chromosomal variability in Drosophila. *Cold Spring Harbor Symp. Quant. Biol,* 20, 256-270.

Sturtevant, A. H., & Beadle, G. W. (1936). The relations of inversions in the X chromosome of Drosophila melanogaster to crossing over and disjunction. *Genetics*, 21(5), 554-604.

Trickett, A. J., & Butlin, R. K. (1994). Recombination suppressors and the evolution of new species. *Heredity*, 73(4), 339.

Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., ... & Gonser, R. A. (2016). Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology*, 26(3), 344-350.

Wang, J., Wurm, Y., Nipitwattanaphon, M., Riba-Grognuz, O., Huang, Y. C., Shoemaker, D., & Keller, L. (2013). A Y-like social chromosome causes alternative colony organization in fire ants. *Nature*, 493(7434), 664.

Wang, S., Zickler, D., Kleckner, N., & Zhang, L. (2015). Meiotic crossover patterns: obligatory crossover, interference and homeostasis in a single process. Cell Cycle, 14(3), 305-314.

White M. (1978). *Modes of Speciation*. San Francisco: W.H. Freeman

Zhao, H., Speed, T. P., & McPeek, M. S. (1995a). Statistical analysis of crossover interference using the chi-square model. *Genetics*, 139(2), 1045-1056.

Zhao, H., McPeek, M. S., & Speed, T. P. (1995b). Statistical analysis of chromatid interference. *Genetics*, 139(2), 1057-1065.

Appendix A: A note on matrix notation

All vectors and matrices are denoted in bold. A matrix or vector followed by square brackets (*[]*) denotes the particular element of that matrix or vector, so that e.g.

$$\boldsymbol{M}[i,j] = i + j, \quad for\ i,j = 0,1,2 \ldots n-1$$

indicate that $\boldsymbol{M}$ is a zero-indexed matrix of size *n,n* with element *i,j* (row, column) equal to *i+j*. All vectors are row-vectors unless otherwise indicated, so that e.g.

$$\boldsymbol{v}[j] = j, \quad for\ j = 0,1,2 \ldots n-1$$

means that

$$\boldsymbol{v} = (0 \quad 1 \quad 2 \quad \ldots \quad n-1)$$

When referring to a column-vector, I either do so by writing out the vector in full, like this:

$$\boldsymbol{w} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ \ldots \\ n-1 \end{pmatrix}$$

or by first defining the vector elements and then indicate in subsequent expression that the vector is transposed, like this:

$$\boldsymbol{w}[j] = j, \quad for\ j = 0,1,2 \ldots n-1$$
$$a = \boldsymbol{v}\boldsymbol{w}^T$$

Two vectors or matrices placed side by side always indicate matrix multiplication. When matrix multiplication operations are to be performed on a given number of matrices with increasing indices from left to right, I sometimes write it in the following short form

$$\prod_{i=0}^{n-1} \boldsymbol{M}_i = \boldsymbol{M}_0\boldsymbol{M}_1\boldsymbol{M}_2 \ldots \boldsymbol{M}_{n-1}$$

The multiplication of *x* instances of the same matrix ($\boldsymbol{M}$) is denoted $\boldsymbol{M}^x$, so that

$$\boldsymbol{M}^x = \prod_{i=0}^{x-1} \boldsymbol{M} = \boldsymbol{M}\boldsymbol{M}\boldsymbol{M} \ldots \boldsymbol{M}$$

By definition,

$$\boldsymbol{M}^0 = \boldsymbol{I}$$

i.e. a matrix to the zeroth power is always equal to the corresponding identity matrix.

Multiplication of a matrix with a scalar, or addition of two matrices with the same dimensions, indicate that each element in the matrix is multiplied or added. For example, if

$$\boldsymbol{M}_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

$$\boldsymbol{M}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

then

$$2\boldsymbol{M}_1 = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$$

and

$$\boldsymbol{M}_1 + \boldsymbol{M}_2 = \begin{pmatrix} 2 & 2 \\ 3 & 5 \end{pmatrix}$$

The *one-vector,* denoted $\mathbf{1}$, is a vector of only ones, i.e. $\mathbf{1} = (1 \quad 1 \quad 1 \quad \dots \quad 1)$. The number of elements in the one-vector is determined by context.

Appendix B: Example input file with comments

The following is a version of the input file '*example.txt*' (online appendix) with brief comments (indicated with *//*). Note that the comments interfere with the reading of the input, so use the comment-free version if you want to run the simulation.

The key settings for this simulation are: *Standard inversion; L = 0.125; d = 0.001; pure counting model, m = 3; symmetrical migration, t = 0.065.*

```
# chromosome 1                          // create an autosomal chromosome
loci = [PTMN]                           // loci keys in the order they appear on the chromosome
lambda = 1.0, 1.0, 1.0, 1.0, 0.0        // lambda values for each interval
Mu = 0.0                                // mu values (if only a single 0 is given, it means 0 for all
                                        intervals)
d = 0.001, 0.001, 0.001, 0.001, 0.001   // d values

# chromosome sex                        // create the sex chromosome
loci = $                                // $ is the sex determination locus
heterogametic = female



# Equilibrium 1                         // input for first equilibrium
check = 500                             // indicate how often to check if equilibrium is reached

% remove                                // alleles/haplotypes to remove from this equilibrium step
alleles = P1, [1

% condition                             // delta
delta = 1.0e-12

% follow                                // which alleles/haplotypes to follow
alleles = T1, M1, N1
haplotypes = T1&M1&N1, T0&M0&N0
screen = true

% do mutate                             // introduce the inversion
[00P00T11M11N11]00 to [01P00T11M11N11]01
frequency = 0.002
deme = 1


# Equilibrium 2                         // input for second equilibrium
check = 1000
```

% remove
alleles = P1

% condition
delta = 1.0e-10

% follow
alleles = [1
haplotypes = [1&T1&M1&N1, T0&M0&N0, T1&M1&N1
screen = true

% do mutate *// introduce P_1*
[11P00T11M11N11]11 to [11P01T11M11N11]11
frequency = 0.002
deme = 1

 # Equilibrium 3 *// input for third equilibrium*

check = 100

% condition
delta = 1.0e-10

% follow
alleles = P0, P1, [1
haplotypes = [1&P1&T1&M1&N1&]1, [1&P0&T1&M1&N1&]1,[0&P0&T0&M0&N0&]0, P1&T1&M1&N1, P0&T0&M0&N0, T1&M1&N1, T0&M0&N0
screen = true
file = progress *// store progress as Numpy array*

% do end *// end simulation*

 # Population *// population settings*
Gamma = 0.0, 0.0, 0.0, 1.0 *// gamma values*
c = 0.0 *// cost of searching*

% remove *// remove globally*
haplotypes =]1&N0

% migration *// migration matrix*
0.935 0.065
0.065 0.935

% Interactions
*fitness = $&T*M&N* *// multiplicative fitness interaction*
preference = PxT

% Fitness

```
Incompatibilities = M,N, 0.5, 0.5        // shortcut for setting incompatibilities with the
                                         parametrization used in the thesis (can also be set manually
                                         to other parametrizations)

% Mating
quickset = P, T, 0.1, 0.1, 1.0           //shortcut for setting mating preferences with
                                         parametrization used in the thesis


# Deme 0                                 // input for habitat 0

static = false

% Allele frequencies
P|T|M|N0=1.0
P|T|M|N1=0.0
[|]0=1.0
[|]1=0.0

%Fitness                                 // fitness specific to habitat 0 ({m} indicate a male)
{m}&T00 = 1.2
{m}&T01 = 1.0
{m}&T11 = 0.833333333


# Deme 1                                 // input for habitat 1

% Allele frequencies
T|M|N0=0.0
T|M|N1=1.0
P0=1.0
P1=0.0
[|]0=1.0
[|]1=0.0

%Fitness
{m}&T11 = 1.2
{m}&T01 = 1.0
{m}&T00 = 0.833333333


# Report                                 //settings for the final report
alleles = P1, T1, M1, N1, [1, ]1
haplotypes = [0&P0&T0&M0&N0&]0, [1&P1&T1&M1&N1&]1, P0&T0&M0&N0, P1&T1&M1&N1,
T0&M0&N0, T1&M1&N1, M1&N1, M0&N0
equilibria = true
file = +_output.txt
screen = true
```