The University Of Sheffield.

Models of the Major Evolutionary Transitions

Christopher G Quickfall

Supervised by Prof. James A R Marshall

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

The University of Sheffield
Faculty of Engineering
Department of Computer Science

September 2016

Abstract

This thesis is concerned with the major transitions view of evolution; the idea that general principles operate in the evolution of each new level of the biological hierarchy (Bourke, 2011). We discuss the theoretical background of this field, focussing on inclusive fitness theory and multi-level selection theory, different approaches to analysing the selection of traits. Many of the commonalities between different transitions are dependent on whether they occur within or between species, and whether relatedness is absent ('egalitarian') or present ('fraternal') (Queller, 1997). Altruism underpins fraternal transitions, and mutually beneficial behaviour underpins egalitarian transitions (Bourke, 2011). We focus on several different models relating to this four-way decomposition.

Firstly, we focus on arguments that between-species donation may amount to between-species altruism; this has been a point of contention within the literature (Fletcher and Doebeli, 2009; Gardner et al., 2011; Wyatt et al., 2013). We discuss both deterministic (resting on an assumption of quasi-linkage equilibrium) and stochastic approaches to a simple model of between-species donation, finding that stable donation behaviour can evolve in the presence of assortment across all loci, but is vulnerable to unassorted modifiers. We argue that this behaviour can be interpreted as within-species altruism, using the other species as a vector for altruism, and, further, consider our models in relation to the current literature on greenbeards.

Our second model concerns maternally-transmitted sex-distorting endosymbionts. Many species, particularly insect populations, are infected by sex-distorting parasites such as the bacteria *Wolbachia*, which are maternally-transmitted; thus, distortion of sex ratios towards the production of females may be beneficial to the symbiont. We investigate the potential for a reproductive parasite to transition towards mutualism, laying the foundation for an egalitarian transition between species; in particular, we find that population structure is key to this transition.

Finally, we discuss several potential avenues for future research; in particular, we note that the social group transformation phase of a major transition involves a number of open questions, or ideas open to further investigation.

Declaration

I declare that this thesis is original work, apart from the exceptions indicated by a reference in the text, and that it has not been submitted to any other university for examination.

Acknowledgements

As my only supervisor, I owe a great debt of gratitude to Prof. James Marshall, for taking on a significant responsibility, and for an awful lot of patience, support, expertise, and encouragement. Without James gently prodding me in the right direction at every turn, this thesis would not have been possible.

I'm grateful to the department of Computer Science and the EPSRC, for administrative support and funding.

I'd also like to thank the many people within academia I have spoken to over the course of my PhD, that have shaped the course of my research and maintained my interest in the area, both in Sheffield and elsewhere.

Finally, thanks are also in order for my family (Rachel, Graham, Natasha and Sarah) and friends, old and new, for plenty of support and good times over the last few years.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Chapter Overview

The aim of this thesis is to explore the 'major transitions' view of evolution through computational modelling; this requires some understanding of inclusive fitness theory and multi-level selection theory. Fundamentally, the major transitions view of evolution is concerned with the evolution of the biological hierarchy, and common principles that operate during the evolution of each transition in individuality (Bourke, 2011). Both inclusive fitness theory and multi-level selection theory can be used to analyse the spread of social traits relevant to a major transition, though the two approaches have to be used carefully to provide appropriate causal explanations of evolutionary phenomena (Okasha, 2006, 2015). This chapter is concerned with the introduction of inclusive fitness theory, multi-level selection theory, and the major transitions view of evolution.

1.2 Applications of Social Evolution Theory

Darwin's original proposal of the theory of evolution (Darwin, 1859) introduced the idea of natural selection. There have been many applications since, including (but not limited to) medical science (Williams and Nesse, 1991; Foster, 2005; Aktipis et al., 2015), psychology (Crawford and Krebs, 2013), computation (Fogel et al., 1966), and economics (Friedman, 1998) through game theory, as formalised by Maynard Smith and Price (1973).

The greatest theoretical extension to Darwinian evolutionary theory is inclusive

fitness theory, proposed by Hamilton (1963, 1964a,b). Hamilton's revolutionary idea was that individuals gain fitness through their contribution to the increased reproduction of genetically related individuals, due to shared genes. This is referred to as kin selection theory and shall be discussed in greater depth in Section 1.4. Section 1.5 shall introduce Grafen's 'formal Darwinism project', a philosophical argument that attempts to justify fitness maximisation approaches, such as Hamilton's rule and the Price equation.

Inclusive fitness theory has led to further additions to the application of evolutionary theory to medical science - a development that Foster (2005) refers to as 'Hamiltonian medicine'. For example, a knowledge of genetic relatedness (see Section 1.4.3) in relation to virulence of pathogens can lead to the development of new medical strategies to tackle them. Another important example would be of cancers - these represent conflict between two different levels in the biological hierarchy; the cells within an organism, and the organism itself (Pepper et al., 2007). Cancer cells optimise their own reproduction to the cost of the multicellular host (Aktipis et al., 2015), resulting in what can be understood as a 'tragedy of the commons' scenario (Hardin, 1968); the host dies due to overexploitation by a selfish cell lineage. These simple examples illustrate the importance of understanding multi-level selection, which will be discussed in Section 1.6; this subject is closely tied to our understanding of the major transitions view of evolution (Okasha, 2006). This shall be discussed in Section 1.8. Finally, in Section 1.9, we shall discuss the structure of the thesis in further detail, and we shall also detail how each question tackled fits in with the major transitions view of evolution.

1.3 Semantics

Social Evolution Theory is concerned with understanding the evolution of social behaviours. Intuitively, you would expect that if an individual displays a given social behaviour, the behaviour will have an effect on the individual, and an effect on a set of other individuals. The measurement of an effect of a social behaviour is given by the lifetime consequences of the behaviour, in terms of absolute fitness (West et al., 2007). Fitness can be described, informally, as fecundity, or the reproductive success of an individual (Marshall, 2015). Fitness effects of social behaviours are

often analysed in the context of a social interaction between a focal individual and a social partner, known as a pairwise interaction; examples of such interactions shall be described over the course of this chapter. The fitness effects of a social behaviour can be either negative or positive for both the focal individual and social partners; consequently, Hamilton created a four-way classification of social behaviours according to the signs of the fitness effects (Hamilton, 1964a).

We shall classify social behaviours in pairwise interactions according to West et al. (2007), in which a social action providing a fitness benefit to both a focal individual and their social partner is referred to as 'mutually beneficial'; if the effect on the social partner is instead costly, this social behaviour is classified as 'selfishness'. The other social behaviours, by contrast, involve a fitness cost to the focal individual; if a benefit is conferred to the social partner, this is referred to as 'altruism', and if there is a fitness cost to the social partner, this is instead referred to as 'spite'. These classifications are standard in any discussion of social evolution theory, having been introduced alongside Hamilton's initial discussion of inclusive fitness, and are displayed in Table 1.1, adapted from a similar table by Hamilton (1964a).

Table 1.1: Classification of Social Actions

		Effect on Partner's Fitness	
		Benefit	Cost
Effect on Focal Individual's Fitness	Benefit	Mutual Benefit	Selfishness
	Cost	Altruism	Spite

There is some confusion with regards to the exact terms given in the literature; however, we choose the scheme advocated by West et al. (2007), in an attempt at as much consistency with the literature as possible. The term 'cooperation' in particular has been used to mean a variety of different things; for example, it has often been used to refer to any social behaviour which confers a benefit (i.e. altruistic or mutually beneficial). We shall use the definition of cooperation provided by Birch (2016), a refinement of the definition West et al. (2007). provide in their paper on semantics; 'a cooperative behaviour is one which, in its recent selection history, has

been favoured by selection by virtue of its beneficial effect on the recipient'. The reason for this adjustment to the standard definition is that it fits more closely with our intuition for cooperation, by filtering out any beneficial behaviour which does not result in feedback. Birch's addition to the definition ('in its recent selection history') allows behaviours that have evolved due to feedback from the provided benefit to be part of the definition, even if the behaviour is not currently under selection due to the provided benefit (Birch, 2016).

Many definitions and ideas rely on our intuitive notion of the individual; however, there is no straightforward definition of the 'biological individual'. There have been many proposals, but closer examination of many of these notions can lead to widely varying conclusions on exactly what constitutes an organism, that do not necessarily chime with our intuition (Clarke, 2010). In the following work, we shall talk about individuals as biological units on which natural selection can act, thus attempting to circumvent the debate by avoiding referring to exact properties of 'individuals'. Lewontin proposed that there are three properties of units of selection; phenotypic variation, differential fitness between phenotypes, and heritability of fitness (Lewontin, 1970). The final condition should be adjusted to heritability of phenotypes, since we are concerned with evolutionary changes in phenotype rather than fitness (Okasha, 2006); this corresponds to Darwin's (1859) original vision. For example, assume that there are two different phenotypes, labelled A and B, which yield the same expected fitness. If we simply assume that there is heritability in fitness, A parents could always produce B offspring in our scenario, and the Lewontin conditions would still be satisfied; this scenario violates our understanding of natural selection, showing that including a condition on heritability in fitness does not lead to an equivalency between the Lewontin conditions and biological reality. Thus, heritability in phenotype is the stronger condition, but is necessary for the three conditions to correspond to our intuition.

Further terms shall be introduced in the next section, which describes Hamilton's innovations and the inception of inclusive fitness theory.

1.4 Inclusive Fitness Theory

1.4.1 Game Theory and the Donation Game

One method of formalising social interactions is game theory (Von Neumann and Morgenstern, 1944; Maynard Smith, 1982); whilst this is an expansive field with applications far beyond the study of evolution alone, we briefly introduce one specific game, the donation game, which is central to the study of social behaviour, and Chapter 3 in particular. Evolutionary game theory involves modeling social interactions as a game, involving a number of players (or social partners) to whom a selection of actions are available; actions are chosen according to players' strategies, resulting in fitness payoffs on the basis of actions chosen (Marshall, 2015).

In the specific terms of the pairwise donation game, two players choose from a set of two actions; donation or non-donation. If a player donates, their social partner receives a fitness benefit b', while the donating player incurs some fitness cost c'; it is assumed that $b' > c' > 0$. This a standard assumption given the common practice of studying this game in the context of the evolution of altruism (costly helping behaviour). If a player does not donate, neither their fitness, nor that of the social partner, changes. Individuals maximise their own fitness payoffs by not donating; however, each players' fitness would be maximised if both players donated, since $b' - c' > 0$ (hence the assumption that $b' > c'$; were this not the case, there would be no incentive for mutual donation). Note the use of primes on benefits and costs in this section; this is to ensure distinction between these quantities and the more used and mathematically rigorous versions of fitness benefits and costs introduced and discussed in the subsequent sections.

Donation in the most basic form of this game can be considered altruism, according to the four-way decomposition of social behaviours introduced in the previous section. Many variations of this basic game have been considered in the literature, along with many proposed mechanisms that encourage donation behaviour, allowing the two players to jointly obtain the fitness benefits of mutual donation (Trivers, 1971; Maynard Smith, 1982; Marshall, 2015). However, it does not follow that donation in every variation of this basic game is altruistic; in the next sections, we discuss Hamilton's rule, which can be used to both provide an explanation for the

problem of altruism, and leads to methods for analysing the true classification of social behaviours, such as donation in the donation game.

1.4.2 Hamilton's Rule and Relatedness

One of the significant problems facing evolutionary theory is to explain the evolution of individually costly social traits; to be precise, altruism and spite, as defined in Section 1.3. Hamilton's first foray into answering this problem came in 1963, highlighting the importance of shared genes, and the relative benefits and costs associated with a social behaviour (Hamilton, 1963). Hamilton's rule for the spread of a gene is now written most commonly as:

$$rb > c \tag{1.1}$$

Here, using informal language relating to a single, altruistic, pairwise interaction, r is the coefficient of relatedness between two individuals, b is the fitness benefit conferred to the social partner, and c is the fitness cost incurred by the focal individual. Note that while Hamilton's rule is couched in much of the same language as the donation game, the components of Hamilton's rule take on different, more precise, mathematical meanings. The rule states, in other words, that a gene will receive positive selection if the cost to the focal individual is outweighed by the benefit given to a social partner, weighted by how closely related the social partner is. Note that the benefit, cost, and relatedness can be either positive or negative (even though we use terminology directly pertaining to altruism); thus we can find conditions for stability of any of the four types of social behaviour as defined in section 1.3 (Bourke, 2011). Formal definitions of the benefit, cost and relatedness involve regression or partial regression coefficients (depending on the exact variant of Hamilton's rules considered) (Marshall, 2015; Birch, 2016), and are thus population-level properties.

1.4.3 Relatedness

Initially, Hamilton used Sewall Wright's Coefficient of Relatedness in his first formulation of Hamilton's rule (Wright, 1922; Hamilton, 1963). This is a measure of the likelihood of two individuals possessing the same gene, but only in terms of genes

identical by descent. In reality, genetic similarity between individuals may exist for reasons other than descent (Marshall, 2015). Thus, it has been noted by many authors, including Hamilton himself, that relatedness should take the form of a regression coefficient (Grafen, 1985; Hamilton, 1972). There are many approaches to calculating r, and some of them were summarised in 1980 by Michod and Hamilton (1980). They concluded that each of these formulations of relatedness were equivalent, or deviated from the general case by invoking certain assumptions.

One definition of relatedness, in words, is that it is 'the probability of sharing the focal gene relative to the average probability that two organisms share the gene' (Bourke, 2011). Mathematically, relatedness can take the following form, the regression of focal individuals' genetic value g on social partners' genetic value g', where genetic value is a linear function of underlying genes (Gardner et al., 2011; Orlove and Wood, 1978; Marshall, 2015; Hamilton, 1972):

$$r := \beta_{g,g'} = \frac{\mathrm{Cov}(g, g')}{\mathrm{Var}(g)} \tag{1.2}$$

The numerator measures how the genetic values of paired individuals covary, a measure of statistical association between two random variables. In other words, if individuals with relatively high genetic value are more likely to be paired with social partners with relatively high genetic value, then genetic values between social partners positively covary. We shall use this formulation of relatedness in deriving Hamilton's rule in Section 1.4.4.

Relatedness is a population statistic, which measures the relative probability of genetically similar individuals being social partners in comparison to randomly paired members of the reference population (Birch, 2016). As this term measures the regression of one alike measure on another alike measure, r is always in the range [-1,1] (Grafen, 1985). Positive relatedness indicates that social partners share a greater number of genes than two randomly chosen population members. Negative relatedness is possible, and indicates the opposite. High levels of relatedness are particularly relevant to the evolution of multicellularity, as shall be discussed in due course.

Moving beyond mathematical definitions and technicalities, understanding of relatedness and Hamilton's rule has had numerous impacts on our understanding of social behaviour. For example, relatedness plays a significant role in the level of virulence in pathogens (Foster, 2005). Many bacteria have the potential to cooperate to achieve a common good, such as by releasing enzymes which break down host tissues; these enzymes would be associated with a higher level of virulence. However, individuals may act selfishly, and not participate in the production of public goods; thus, low relatedness is associated with lower virulence in infections, since low relatedness widens the opportunity for selfish behaviour (Foster, 2005; Bourke, 2011). An example of such a public good is the production of siderophores in the bacteria *Pseudomonas aeruginosa*; indeed, underlining the relevance of the theoretical literature on Hamilton's rule, Griffin et al. (2004) use this system to provide evidence for a version of Hamilton's rule which accounts for competition between kin.

1.4.4 Inclusive and Neighbour-Modulated Fitness

Hamilton's rule (see equation 1.1) can be conceptualised as involving two components of fitness. The first, c, pertaining to the fitness cost to a focal individual, is called 'direct fitness', and can be summarised as the component of reproductive success of a focal individual for which it is causally responsible (Birch, 2016). The second, rb, pertaining to the fitness benefit to social partners weighted by relatedness, is termed 'indirect fitness'; it is the relatedness-weighted sum of components of reproductive success of genetic relatives that can be accounted for by the focal individual's effect on those relatives. The two quantities summed together constitute 'inclusive fitness' (Hamilton, 1964a; Marshall, 2015). Formally, inclusive fitness may be expressed as a regression analysis (Marshall, 2015). It takes the following form, where ω refers to relative fitness, ω_0 refers to baseline relative fitness, g refers to the genetic value of focal individuals, and primes denote properties of social partners:

$$\omega = \omega_0 + g\beta_{\omega g \cdot g'} + g'\beta_{g'g}\beta_{\omega' g \cdot g'} \tag{1.3}$$

Since this is a partial regression analysis, residuals as well as intercepts do not

correlate with either individual's genetic value, thus they are omitted. The partial regression terms $\beta_{AB \cdot C}$ refer to the regression of A on B, holding C constant (i.e. 'partialling out'), while the simple regression term β_{AB} refers to the regression of A on B. Here, an individual's inclusive fitness is the sum of baseline fitness, w_0, the direct component, $\beta_{\omega g \cdot g'}$, which comprises of fitness effects on the focal individual caused by the focal individual, and the indirect component, which comprises of fitness on the social partner caused by the focal individual, $\beta_{\omega' g \cdot g'}$, weighted by the relatedness term, $\beta_{g' g}$.

This approach to fitness entails what Birch (2016) refers to as the 'indirect reproduction' view of relatedness; high relatedness means that social partners will give a focal individual's genes better representation in the next generation. An alternative viewpoint is the 'correlated interaction' view of relatedness; high relatedness means that a social partner is more likely to return a fitness benefit to the focal individual. This viewpoint relates to an alternative conception of fitness, known as 'neighbour-modulated fitness'. Formally, neighbour-modulated fitness can, like inclusive fitness, be expressed as a regression analysis (Marshall, 2015):

$$\omega = \omega_0 + g\beta_{\omega g \cdot g'} + g'\beta_{\omega g' \cdot g} \tag{1.4}$$

Under this viewpoint, an individual's fitness is simply the sum of fitness effects on the focal individual caused by the focal individual, and fitness effects on the focal individual caused by social partners. Indeed, this approach, while generating equivalent predictions to the inclusive fitness aproach (Hamilton, 1970; Marshall, 2015; Birch, 2016), relaxes assumptions required by the inclusive fitness approach.

Hamilton (1970) spelt out the formal equivalence of the two approaches, which Birch (2016) expanded, noting that it invoked two assumptions. Firstly, both approaches rely on the assumption of weak additivity; i.e. that fitness effects are not determined by focal individual and social partner genotypes do not correlate with genes. The neighbour-modulated fitness approach is more easily adjusted to cope with deviations from additivity (Birch, 2016). Secondly, the inclusive fitness approach makes the assumption of 'actor's control'; correlations between the focal

individual and social partner's genotypes entirely predict fitness effects. Birch gives the example of an individual that sends an alarm call, which thus causes any benefit to that individual of subsequently receiving an alarm call to be diminished. This would not be accounted for under the inclusive fitness approach since the relatedness term involves regression on the focal individual's genotype. Despite spelling out arguments for the stronger assumptions required by the inclusive fitness approach, Birch does go on to suggest that it nonetheless retains great relevance, since it can lead to more powerful conclusions regarding causation and directly costly traits. Under the inclusive fitness approach, positive selection for an altruistic behaviour is explained by indirect benefits, whereas under the neighbour-modulated fitness approach, it is explained by the behaviour correlating with the return of a fitness benefit (Birch, 2016).

1.4.5 The Price Equation

The Price Equation was introduced by George Price (1970), and allows us to model the evolution of a given character value on the basis of natural selection and transmission. Although underappreciated at the time, it has since been recognised as being of immense importance; it makes minimal biological assumptions, and can be extended to decompose selection at multiple levels of the biological hierarchy (Price, 1972; Hamilton, 1975; Okasha, 2006). It can take the following form, though there are alternative decompositions which we shall describe later on (Price, 1970; Okasha, 2006):

$$\Delta\bar{z} = \mathrm{Cov}(\omega, z) + E_w(\Delta z) \tag{1.5}$$

Here, for the ith member of the population, z_i is the measure of the character of interest, w_i is the absolute fitness, ω_i is the relative fitness (i.e. $\omega_i := w_i/\bar{w}$), z_i' is the mean character value of the ith individual's offspring, and Δz_i is the difference between the mean character of the ith individual's offspring, and the ith individual's character (i.e. $\Delta z_i = z_i' - z_i$). The lack of a subscript indicates a vector of length n, where n is the population size, e.g. $z = (z_1, ..., z_n)$, and means over the population are denoted with a bar, e.g. $\bar{z}$ is the mean character value. Note that, following (Okasha, 2006), we have dropped indices from ω_i, z_i and w_i in equation (1.5) for the

sake of readibility.

We are interested in the change in mean character value, which is given by $\Delta\bar{z}$ on the left-hand side. Importantly, this formula decomposes change in mean character value into two components. The first term, $\mathrm{Cov}(\omega, z)$ is the covariance between the relative fitness and character value. In other words, if individuals with greater character value on average have increased fitness, then this term will be positive. This makes sense intuitively, since if a trait gives its bearer increased fitness, you would expect that trait to spread.

The second term, $E_w(\Delta z)$, can be interpreted as the transmission bias weighted by relative fitness values of population members. If there is no transmission bias (in other words a trait is, on average, passed to offspring perfectly), then $\Delta z = 0$, and this term is zero. However, if a trait has a tendency to be more or less represented in offspring than their parents, this term will be non-zero (respectively; positive or negative). Often, transmission bias is assumed to be zero, thus we arrive at a simplified version of the Price equation:

$$\Delta\bar{z} = \mathrm{Cov}(\omega, z) \tag{1.6}$$

We shall now follow Price's original derivation of the Price equation (Price, 1970). This requires some further notation in addition to that previously provided; note that this proof is specific to genic value (Marshall, 2015), though the formulae are general enough that they apply to genetic value and character value too, provided that character value is a linear function of genetic value (Okasha, 2006). We consider two populations, respectively P_1 and P_2; P_2 is the set of offspring of members of P_1. We denote the number of copies of a gene A within an individual i by A_i, and the total number of copies in their offspring by A'_i. The maximum number of A genes an individual can possess is the gametic ploidy, n_G; thus, we label the genic frequency of A within the ith individual as $a_i := A_i/n_G$. We want to consider the difference in gene frequency between the two populations. Firstly, we note that the gene frequency in P_1 is trivially $\bar{a}$, the mean of the a_i genic frequencies for individuals in the group P_1. Next, we calculate the gene frequency in P_2, denoted $\bar{a}'$, starting

by noting that it is the total number of A genes in offspring of P_1 members, divided by the maximum number of A genes they could possess, which is the gametic ploidy multiplied by the total fitness of P_1 individuals:

$$\begin{aligned}
\bar{a'} &= \frac{\sum_i A'_i}{\sum_i w_i n_G} \\
&= \frac{\sum_i w_i n_G a'_i}{\sum_i w_i n_G} \\
&= \frac{\sum_i w_i a'_i}{n\bar{w}} \\
&= \frac{\sum_i w_i a_i}{n\bar{w}} + \frac{\sum_i w_i \Delta a_i}{n\bar{w}}
\end{aligned}$$

To obtain the second line, we have used the fact that the total number of copies of a gene in the offspring of an individual is the product of the gene frequency in the offspring in that individual, the gametic ploidy, and the fitness of the individual. We then note that n_G is constant, so can be removed, and that the mean fitness in a population is the sum of the fitnesses of population members, divided by the population size (i.e. $\bar{w} = \sum_i w_i/n$). The final expression decomposes the measure of character in P_2 into the original character measure from P_1 and the difference in character measure between the two populations.

We now consider the covariance of w and a, by splitting it up into expectations as follows; we then use the definition of expectation to obtain the following expression:

$$\begin{aligned}
\mathrm{Cov}(w,a) &= \mathrm{E}(wa) - \mathrm{E}(w)\mathrm{E}(a) \\
&= \frac{\sum_i w_i a_i}{n} - \bar{w}\bar{a}
\end{aligned}$$

From this, it requires some trivial algebra to plug this into our expression for $\bar{a'}$, and calculate the change in gene frequency between generations:

$$\begin{aligned}
\bar{a'} &= \bar{a} + \frac{\mathrm{Cov}(w,a)}{\bar{w}} + \frac{\sum_i w_i \Delta a_i}{n\bar{w}} \\
\Rightarrow \Delta\bar{a} &= \bar{a'} - \bar{a} = \frac{\mathrm{Cov}(w,a)}{\bar{w}} + \frac{\sum_i w_i \Delta a_i}{n\bar{w}} \\
\Rightarrow \Delta\bar{a} &= \mathrm{Cov}(\omega, a) + E_w(\Delta a)
\end{aligned}$$

Thus, we arrive at the Price Equation. Note that the covariance term now refers to relative fitness, due to the weighting by $\bar{w}$. We can derive a useful reformulation of this expression by multiplying the penultimate expression through by $\bar{w}$, and replacing genic value a with the more commonly used character value z:

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w, z) + E(w\Delta z) \tag{1.7}$$

We also note that a reduced version of this equation is regularly used, which assumes that there is no transmission bias; note that this formulation is an absolute fitness version of equation (1.6):

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w, z) \tag{1.8}$$

Price made some further notes about the nature of these equations. In particular, he noted the generality of the equation, highlighting only the assumption that gene A ploidy is the same for all members of P_1. He also pointed out that although the derivation above assumed discrete generations, the equation in fact holds for overlapping generations (Price, 1970). One further point that should be made is that there is much flexibility in our choice of character value (rather than being simply gene frequency); Okasha (2006) gave the example of defining the character variable as 1 if an individual has some characteristic of interest, and 0 if not; in this case, Price's equation holds. In fact, Price (1970) pointed out that it could be applied to modelling change more generally; he gave a non-genetical example, in which the difference between student IQs entering a course, and those completing it, is considered.

One useful application of the Price equation involves the use of 'breeding values' (Queller, 1992). These are linear combinations of allelic values which are chosen to 'best explain phenotypic variation in a trait'. By applying the Price equation to breeding values, Queller found four equivalent 'separation conditions' which must be satisfied in order to validly decompose selection into components describing the link between phenotype and genes, and phenotype and fitness. An application of this is found in the 'breeder's equation', which calculates the response to selection (R) as the product of a selection differential (S) and heritability (h^2) (Birch and

Marshall, 2014; Marshall, 2015):

$$R = Sh^2 \tag{1.9}$$

R corresponds to changes in mean breeding value, so, if the separation condition is satisfied, it is possible to perform a least-squares regression of breeding value on phenotype (i.e. readily observable information, unlike genotype), thus finding precise definitions of S and h^2 (Marshall, 2015).

One final insight about the Price equation was suggested by Okasha (2006) again; the equation depicts a statistical decomposition rather than causal one. In other words, a component being non-zero suggests correlation rather than causation. For example, suppose increased height in humans caused a gain in fitness, and that the character of interest is liver size. Then there would be a significant correlation between liver size and fitness (by virtue of the correlation between height and liver size), and hence the first component of the Price equation would be positive, but you would not necessarily expect increased liver size to cause increased fitness.

1.4.6 Derivation and Formulations of Hamilton's Rule

In this section we shall derive Hamilton's rule using the Price equation; this work follows the derivation provided by Queller (1992). It also ties in with the definition of relatedness provided earlier, though we now choose the character of interest to be genetic value, g.

Queller starts with the relative fitness form of the Price equation, and assumes that there is no transmission bias; thus, we start with equation (1.6), using genetic value:

$$\Delta\bar{g} \;=\; \mathrm{Cov}(\omega, g)$$

Next, we introduce the regression analysis form of neighbour-modulated fitness

(equation (1.4)) into the Price equation:

$$\Delta\bar{g} = \beta_{\omega g \cdot g'}\mathrm{Var}(g) + \beta_{\omega g' \cdot g}\mathrm{Cov}(g', g)$$

Note that since ω_0 is a constant, and we have already assumed that the residuals and intercepts are uncorrelated with genetic value, the covariances involving these terms are zero, so we have omitted them. Thus, we obtain a condition for the relevant trait to receive positive selection:

$$\beta_{\omega g \cdot g'} + \frac{\mathrm{Cov}(g', g)}{\mathrm{Var}(g)}\beta_{\omega g' \cdot g} > 0$$

Now, how can we interpret this equation? The first term on the left-hand side is $\beta_{\omega g \cdot g'}$, the effect on the focal individual's fitness of the genetic value of the focal individual, assuming independence between the focal individual and social partner's genetic values. Thus, this corresponds to the direct fitness component, c of inclusive fitness. We next note that relatedness, as shown in equation (1.2), weights the second partial regression coefficient, $\beta_{\omega g' \cdot g}$. This describes the effect on the focal individual's fitness of the social partner's genetic value, g', partialling out g, the focal individual's genetic value. Thus, the second component of the equation corresponds to the indirect fitness component of inclusive fitness, rb. In other words, we have recovered a neighbour-modulated form of Hamilton's rule (Queller, 1992).

Queller's formulation of Hamilton's rule based on partial regression can be compared and contrasted with a number of other, mathematically equivalent, formulations (Marshall, 2015). They involve different interpretations of costs, benefits and relatedness; for example, the cost may be calculated by the negative of either the regression coefficient of genetic value on bearer's fitness, or of the regression coefficient of behaviour on bearer's fitness, in addition to the interpretation already given (Marshall, 2015). Although each generalisation is mathematically correct, different interpretations of r, b, and c may consequently lead to different interpretations of social behaviour; for example, one rule may interpret a 'greenbeard' trait (i.e. a gene which causes the owner to recognise other holders of the 'greenbeard' gene and to donate towards them (Dawkins, 1976)) as a mutually beneficial trait, and another

may interpret it as an altruistic trait (Marshall, 2015).

We can also introduce our expression for neighbour-modulated fitness (1.4) into the reduced absolute fitness version of the Price equation (equation (1.8)) to obtain an expression for the change in mean genetic value, where $\bar{w}$ is simply a normalising constant (Okasha, 2006):

$$\begin{aligned} \bar{w}\Delta\bar{g} &= \mathrm{Cov}(w, g) \\ \Rightarrow \bar{w}\Delta\bar{g} &= \beta_{wg\cdot g'}\mathrm{Var}(g) + \beta_{wg'\cdot g}\mathrm{Cov}(g', g) \\ &= (\beta_{wg\cdot g'} + \beta_{wg'\cdot g}\beta_{g'g})\mathrm{Var}(g) \\ \Rightarrow \bar{w}\Delta\bar{g} &= (-c + rb)\mathrm{Var}(g) \end{aligned} \tag{1.10}$$

This neighbour-modulated decomposition of change in mean genetic value is useful for comparison with two approaches introduced in Section 1.6, where it shall be taken to represent the kin selection approach in general, following Okasha's (2015) analysis utilising causal graphs. We shall introduce a 'multi-level' version of the Price equation, and the contextual analysis approach, which can be regarded as a hybrid of the kin selection and Price equation approaches (Okasha, 2004). However, we start by discussing an important philosophical point underlying the use of Hamilton's rule and the Price equation in modelling.

1.5 Grafen's Formal Darwinism Project

We now turn to Grafen's 'formal Darwinism' project, an attempt to reconcile population genetic models with the argument that individuals are designed objects, and the associated explanatory tool of them being 'maximising agents' (otherwise known as the 'individual as maximising agent', or IMA, analogy) (Grafen, 1999, 2014a; Gardner, 2009). Population genetic models are designed according to the principle that genotype frequencies change over time; as Grafen (2014a) notes, this is an accurate description of the process of natural selection. Fitness maximisation principles are represented by the Price equation and Hamilton's rule. In particular, inclusive fitness provides a candidate maximand; it is possible to consider a focal individual's genotype and ask when it maximises its own fitness. This is in contrast

to the neighbour-modulated approach, which involves a component not directly dependent on a focal individual's genes (Birch, 2016). These approaches are widely used in modelling, invoking the 'phenotypic gambit', the assumption that it is possible to study the selective benefit of a trait in an organism without knowledge of the precise underlying genetic architecture (Grafen, 1991, 2014a). It is precisely this assumption which Grafen aims to explore and justify in his formal Darwinism project.

The IMA approach is not firmly grounded, as fitness maximisation does not always take place. Grafen (2014a) gives several examples in his 2014 outline of the project, including the 'overdominance' scenario (in which AB heterozygotes have higher fitness than AA or BB homozygotes, but the equilibrium frequencies include AA and BB types), the 'all heterozygotes' scenario (in which there is a population of only heterozygotes, in the same scenario - this is not stable, due to Mendelian genetics), and the 'separate sexes' scenario (in which AA and BB homozygotes have the same fitness, which is strictly greater than that of AB heterozygotes, all females are AA type, and all males are BB type; in the next generation, all individuals will be AB). However, application of the fitness-maximisation principle has been the foundation for much work; in particular, the phenotypic gambit is implicitly invoked by field-workers, modellers, and in meta-analyses. Grafen's aim is to provide a formal justification for this work, and the original argument by Darwin for the appearance of design.

Specifically, Grafen's approach is to establish formal links between population genetics models, and an optimisation program, which Grafen (2014a) constructs in order to represent the fitness-maximisation approach. The links he specifies are as follows:

> '1. If all individuals in the population solve the optimisation program, then the expected change in every gene frequency equals zero, and there is no possible phenotype which, if produced by a rare dominant mutant, would initially invade the population.
>
> 2. If all individuals attain the same value of the maximand but do not solve the optimisation program, then the expected change in every gene frequency equals zero, but there is a possible phenotype which, if produced by a rare dominant mutant, would initially invade the population.

3. If individuals attain different values of the maximand, then the change in every gene frequency equals its covariance across individuals with those attained values.

4. If the expected change in every gene frequency equals zero, and if there is no possible phenotype which, if produced by a rare dominant mutant, would initially invade the population, then every individual in the population solves the optimisation program.'

The point of these links is to show that for every population genetic model, a corresponding optimisation program can be constructed in such a way that satisfying the links will ensure that the maximand in the optimisation program will be uniquely defined, and will represent fitness (Grafen, 2014a). Grafen uses the three problem scenarios previously introduced as test cases for these links. For example, consider the 'separate sexes' scenario described above. In this case, link one holds, as gene frequencies will on average not change, even though the genotypes will become suboptimal due to recombination when individuals of the two types breed with one another. The 'if' statement (or antecedents) is not satisfied for link two, as the entire population is acting optimally, and nor is it satisfied for link three, as every individual has the same fitness. Thus, Grafen concludes that these two links 'hold trivially'. Link four is also satisfied - any rare mutants would be either AA or BB type and so could not invade.

Grafen's papers have mathematically proved links between population genetics and IMA-based optimisation programs under more and more plausible assumptions. His 2002 paper assumed no social interactions or frequency dependence (Grafen, 2002); however, his subsequent 2006 paper replaces the assumption of no social interaction with the assumption that social effects are additive (Grafen, 2006). The central point that Grafen makes in his papers is that while gene frequencies change according to inclusive fitness, the same does not apply to genotype frequencies (Grafen, 2006).

Grafen's 'formal Darwinism' project was the subject of a special issue of 'Biology and Philosophy' in 2014, in which he outlined the project. A wide mixture of responses to his project (Okasha and Paternotte, 2014b) were published, and we now briefly summarise some of the least technical of the perceived limitations and criticisms brought up in the responses.

Firstly, Grafen focuses only on a single level in the biological hierarchy, specifically treating multicellular organisms as an 'ideal'; this was heavily criticised as being biased by Shelton and Michod (2014b), while Bourke (2014b) argued that Grafen's lack of acknowledgement of the major transitions view (the subject of Section 1.6) of evolution is a limitation. Grafen's (2014b) response was to suggest that the extension to further levels of the biological hierarchy is too technically challenging.

Elsewhere, Birch (2014) argued that Grafen's wish to demonstrate the 'appearance of design' is beyond the scope of mathematical equations, since it is not itself equivalent to the IMA analogy. Grafen's (2014b) response is to allow that his project may have been oversold, leaving open the question of how much biological reality is covered by his project .

Birch (2016) has also noted that Grafen actually aims to link a fitness maximisation program to points in a dynamical system at which there is 'neither scope nor potential for selection'. He argues that these points do not necessarily constitute equilibrium points from the perspective of population genetics; thus Grafen's argument is not sufficient to prove legitimate links. Unfortunately, Birch's work is yet to be published, so a response from Grafen is unavailable.

However, despite these arguments and reservations, and a number of more detailed points (Grafen, 2014b), some researchers still find value in the aim of the project, to bridge a gap in Darwin's argument between natural selection and the appearance of design (Gardner, 2014; Birch, 2014; Bourke, 2014b; Huneman, 2014; Okasha and Paternotte, 2014a). Alternatively, it has been suggested that the project is unnecessary to illustrate the importance of social evolution theory; so long as the IMA analogy leads to correct conclusions about biological reality, it retains value (Birch, 2016).

1.6 Multi-Level Selection Theory

1.6.1 The Levels of Selection Problem

Up until now we have only considered selection at one level. However, in nature, selection often acts at multiple levels at the same time. In Section 1, we briefly discussed the example of cancers, in which selection at the level of the cell conflicts with selection at the level of the organism (Foster, 2005). Thus, we need to consider the effects of selection on multiple levels. For simplicity, in the following, we shall consider selection at two different levels. Following Okasha (2006), we refer to individuals at the lower level as 'particles', and individuals at the higher level as 'collectives'. Collectives are assumed to be non-overlapping groups of particles.

We want to consider the selection of characters at both of these levels, but this is difficult, because we also need to consider the relation between particle and collective characters. There is a distinction to be made between different types of collective character. Collective characters which are directly dependent on the characters of particles within the collective are known as 'aggregate' characters. For example, if the relevant particle character is height, then average particle height within the collective would be an aggregate character. Collective characters which are indirectly dependent on characters of particles within the collective, instead pertaining to relations between the characters, are known as 'emergent' characters. For example, variance in particle height within the collective would be an emergent character (Okasha, 2006).

We must consider two alternative perspectives of multi-level selection, that differ in their treatment of the 'focal' level of selection - these are referred to as multi-level selection 1 (MLS1) and multi-level selection 2 (MLS2) (Damuth and Heisler, 1988; Okasha, 2006). In MLS1, the focal individuals are particles. Particles are organised into collectives, and any character based on collective membership is referred to as a 'contextual' character of particles within the group. Thus, the examples of aggregate and emergent character given in the previous paragraph would be considered contextual characters of particles under the MLS1 perspective. Character and fitness are properties of particles, and the fitness of collectives is simply defined to be the mean fitness of the constituent particles. Selection between collectives

affects particles, but no inferences can be made about frequencies of different types of collectives within the population.

In MLS2, by contrast, both particles and collectives are the focal individuals; in other words, they both possess the properties of character and fitness. It follows then, that collective fitness is treated independently of particle fitness. In this scenario, we are interested in the evolution of different kinds of collectives as opposed to particles (Damuth and Heisler, 1988; Okasha, 2006).

There are several things to note here. Firstly, the choice between modelling MLS1 and MLS2 is not a matter of taste, it is a matter of suitability. The question of when exactly we should apply MLS1 or MLS2 is not always straightforward, and we shall return to this later on. Secondly, 'particles' and 'collectives' simply denote individuals on two levels of the biological hierarchy, with the relations previously described - they aren't meant to denote specific levels. Finally, one of the problems that we face is quantifying selection at different levels; in the next three sections, we discuss approaches to understanding this problem.

1.6.2 The Price Approach to MLS1

It takes a simple extension to expand the Price Equation from the single-level scenario to the multi-level scenario; this process is described by Okasha (2006), more rigorously described by Wade (1985), and originally introduced by Price (1972). We start by considering selection at the level of the particle, and assume that there is no transmission bias - this can be reintroduced into our final equation if the assumption does not hold, but for simplicity we start with equation (1.8):

$$\bar{w}\Delta\bar{z} = \mathrm{Cov}(w, z)$$

Corresponding to our previous notation, $\bar{w}$ refers to mean absolute particle fitness, $\bar{z}$ refers to mean particle character, w is the vector of absolute particle fitnesses w_i, and z is the vector of particle characters z_i. We now make use of the MLS1 approach, thus taking collective fitness, W_i to be the mean of particle fitness within the ith particle's collective, and also taking collective character, Z_i to be the mean of

particle character within the ith particle's collective; note that this latter assumption is not necessary under the MLS1 approach, but it allows us to simplify equations. It also means that the character of interest is an aggregate character. Note also that we use upper cases to refer to properties of collectives, and lower cases to refer to properties of particles. In addition, the subscript i denotes the particles to which a property belongs; since this is an MLS1 perspective, properties of the collective to which a particle belongs are ascribed to the particle.

Under these assumptions, we can apply a decomposition to the right-hand side and arrive at the following expression (Wade, 1985; Okasha, 2006), in which we have included the subscript j to make the point that the second term is the expectation of the particle-level covariance between character and fitness, calculated across all collectives j:

$$\mathrm{Cov}(w, z) = \mathrm{Cov}(W, Z) + \mathrm{E}(\mathrm{Cov}_j(w, z))$$

We therefore obtain a partition of selection into two different levels:

$$\bar{w}\Delta\bar{z} = \mathrm{E}(\mathrm{Cov}_j(w, z)) + \mathrm{Cov}(W, Z) \tag{1.11}$$

The first component of the right-hand side can be interpreted as the particle-level selection, whereas the second component can be interpreted as the collective-level selection. Of course, this equation only applies to MLS1 and aggregate characters, as it relies on the assumptions of collective character being mean particle character, and collective fitness being mean particle fitness. Under MLS2, we can do little more than apply the single-level Price equations at the collective level, and also at the level of the particle, for each collective (Okasha, 2006).

1.6.3 The Contextual Analysis Approach to MLS1

An alternative decomposition of selection on multiple levels is provided by the contextual analysis approach; once again, this is relevant only to MLS1. The essential idea of contextual analysis is simple - to perform a regression of particle fitness w on both particle character z, and collective character Z, in the simplest case

(Heisler and Damuth, 1987). In this way, the relative strengths of within-group and between-group selection are evaluated. The collective character Z could be either an aggregate or emergent character, but, either way, it is the particles which are considered to possess both particle and collective character, since this is only relevant to MLS1; this motivated Heisler and Damuth to refer to collective characters as 'contextual' characters.

Thus, the following linear regression model of particle fitness is created; we follow Okasha (2006) as opposed to the more complex terminology of Heisler and Damuth, though unlike Okasha we fully display partial correlation coefficients in order to more explicitly convey their meaning:

$$w = \beta_{wz \cdot Z} z + \beta_{wZ \cdot z} Z \tag{1.12}$$

In reality, we would need to include an intercept and errors as well as, depending on our model assumptions, other possible explanatory variables and interaction terms, but for simplicity they are omitted (Okasha, 2006). Following standard partial regression notation, the coefficient $\beta_{wz \cdot Z}$ is the effect of particle character on particle fitness, partialling out collective character. Similarly, $\beta_{wZ \cdot z}$ is the effect of collective character on particle fitness, partialling out particle character. It is then simple to insert this into the reduced form of the Price equation (1.8), giving the following:

$$\begin{aligned} \bar{w}\Delta\bar{z} &= \mathrm{Cov}(\beta_{wz \cdot Z} z + \beta_{wZ \cdot z} Z, z) \\ &= \beta_{wz \cdot Z}\mathrm{Var}(z) + \beta_{wZ \cdot z}\mathrm{Var}(Z) \end{aligned} \tag{1.13}$$

As with the Price Equation form of multi-level selection, the first component of the right-hand side can be interpreted as the particle-level selection, whereas the second component can be interpreted as the collective-level selection. Note that the intercept will never covary with z, the errors are assumed not to, and we have assumed that there are no other possible explanatory variables or interaction terms. By definition, $\mathrm{Var}(z) = \mathrm{Cov}(z, z)$, and we also note that $\mathrm{Var}(Z) = \mathrm{Cov}(z, Z)$, if

we again make the assumption that Z is an aggregate character. We interpret the first term in equation (1.10) as accounting for the direct causal influence of particle character on particle fitness, and the second term as accounting for the by-product of selection at the level of the collective (Okasha, 2006). This is known as a 'cross-level by-product' in the downwards direction. In other words, this component of collective-fitness covariance at the level of the particle is detected as being a by-product of collective-fitness covariance at the level of the collective; this link is not necessarily causal. Cross-level by-products are also possible in the upwards direction - this can be easily envisaged, since it is very possible that particle character has a causal effect on particle fitness, which in turn may have a causal effect on collective fitness, and collective character; thus there will be a spurious character-fitness covariance at the level of the collective (Okasha, 2006).

Thus, under the same assumptions as the Price approach, we have obtained an alternative decomposition of selection, which is mathematically equivalent, but not causally equivalent in interpretation.

We can also note the similarity between the contextual analysis and kin selection approaches (equation (1.10); note that genetic values g and g' here could be replaced with phenotypic values, z and z' (Okasha, 2004)); essentially, the contextual analysis approach regresses on collective character instead of neighbour character. Indeed, the two components of the kin selection approach can be considered to be relevant to particle-level selection and collective-level selection, respectively. Okasha (2004) derives a simple mathematical relation between social partner character, individual character, and collective character, and notes that the component of neighbour-modulated fitness derived from social partners is conceptually similar to the group selection component of the contextual analysis approach. Thus, the three approaches to partitioning fitness relevant to multi-level selection, captured by equations (1.10), (1.11), and (1.13), bear comparison; this is the subject of the next subsection.

1.6.4 Problems with the Price, Contextual and Kin Selection Approaches in MLS1

We have established that the Price, contextual, and kin selection approaches are all mathematically valid, but there can be problems in interpretation, as they do not provide the same causal explanations. In this section, we shall apply Okasha's (2015) causal graph approach to illustrate some examples of when each approach fails . He describes a model in which possible variables are particle character z_i, average character of social partners is z_i' (note that social actions are not necessarily pairwise), collective character is Z_i, particle fitness is w_i, average fitness of social partners is w_i', and collective fitness is W_i. As before, upper cases denote properties of collectives, lower cases denote properties of particles, and subscripts are used to denote the particles to which a property is considered to belong.

Okasha (2015) notes that for each analysis to provide a causally valid decomposition of selection, any statistical associations detected must be wholly the result of direct causal influence between variables. His idea is to make this explicit, displaying causal graphs for certain example situations which make clear the applicability of each approach. For the Price approach to be causally valid, by examination of equation (1.11), we can see that the only direct causal link to W_i must be from Z_i, and the only direct causal link to w_i must be from z_i. Likewise, for the contextual analysis approach to be causally valid, by examination of equation (1.13), we see that the only direct causal links to w_i must be from z_i and Z_i. Finally, by examining the kin selection approach, represented by equation (1.10), we see that, after substituting character values for genetic values, the only direct causal links to w_i must be from z_i and z_i'.

The much-discussed 'soft selection' scenario (Goodnight et al., 1992; Okasha, 2006) is a scenario in which all collectives have the same fitness regardless of their composition; thus, there is no selection at the level of the collective, according to the Lewontin conditions (see section 1.2). However, the fitness of a particle is dependent on its relative size (for example, size could be replaced by any other method of ranking particles on the basis of phenotype) within the collective. To illustrate this in terms of Okasha's causal graph approach, we simply note that there are causal

links from particle character and partner character to collective character, and that particle (partner) fitness is determined by particle (partner) character, collective character, while collective fitness is constant. This causal graph is displayed in Figure 1.1.

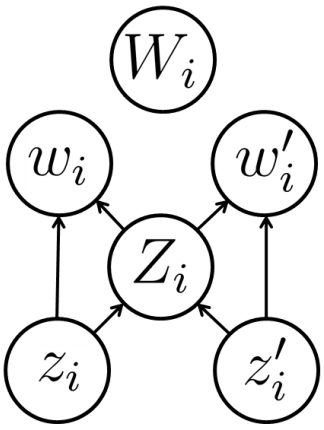

Figure 1.1: Causal links in the 'soft selection' scenario.

In the case of the Price approach, since there is no causal link from Z_i to W_i, no component of collective-level selection is detected. This corroborates with the Lewontin conditions, and thus the Price approach provides a correct causal decomposition of selection in this scenario. Next, we consider the contextual approach. The soft selection scenario has frequently been touted as a prime example of a failing of the contextual approach, and the reason for this can be seen by observing that both direct and indirect causal links from z_i to w_i and z_i' to w_i' are present. As pointed out by Goodnight et al. (1992) and Okasha (2006), contextual analysis wrongly detects a component of collective-level selection, since there are collective effects on particle fitness. The kin selection approach also fails to reflect causality here, since particle character indirectly affects particle fitness, and partner character indirectly affects partner fitness.

On the other hand, we can consider an alternative scenario, the 'non-social trait case', in which a particle's fitness is determined solely by its character. Particles are sorted into groups at random, independently of character. A correct causal decomposition of selection here will not attribute any selection to the collective level, since collectives are entirely arbitrary. However, collectives will have varying mean particle character and fitness, and thus a spurious covariance between collective character and fitness will be detected by the Price approach - this is a cross-level by-product, arising from the covariance between character and fitness at the level of the particle. This situation is described and discussed by Heisler and Damuth (1987) and Okasha (2006, 2015); Figure 1.2 depicts the causal graph for this scenario, adapted from a similar figure in Okasha's (2015) causal graph paper. It is pointed out by Okasha that the kin selection approach applies here, since there are only direct causal links

from z_i to w_i and z_i' to w_i'. There is no correlation between partner character and particle fitness, so no collective-level component of selection is detected. Similarly, the contextual approach also applies, as noted by Heisler and Damuth (1987) and Okasha (2006) - this is because the contextual approach correctly accounts for the cross-level by-product. As shown in Figure 1.2, there are direct causal links from z_i to Z_i and from w_i to W_i.

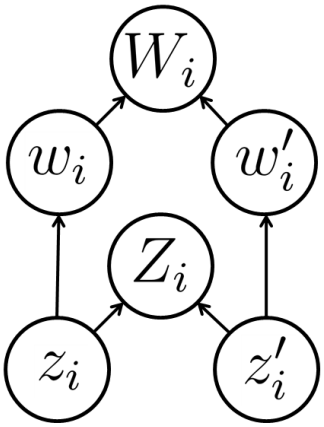

Figure 1.2: Causal links in the 'non-social trait' scenario (adapted from Okasha (2015)).

Elsewhere, Okasha (2004) has noted that the contrasting results in these two scenarios come from every decomposition having to balance different types of error. On the one hand, we only want selection at the level of the collective to be detected if there is variance in collective fitness. The Price equation makes no mistake here, since the collective-level selection component is a product of $\mathrm{Cov}(W, Z)$. The kin selection and contextual analysis approaches fail, as illustrated by the soft selection scenario.

On the other hand, we also want selection at the level of the collective to only be detected if individual fitnesses are group-dependent; if this is not satisfied, false positives are possible in detection of selection at the collective level. The Price approach does not satisfy this condition, as illustrated by the non-social trait group scenario. The contextual analysis approach was created to solve this problem (Heisler and Damuth, 1987; Okasha, 2004), by introducing the partial regression of fitness on collective character; the kin selection approach satisfies the condition by conducting a partial regression of particle fitness on social partner character.

Other examples can be found of scenarios in which one or more of the three approaches to decomposing selection fail to provide an adequate causal explanation, but the point here is that not every approach is applicable all of the time. Illustrating where causal links are in a given situation can help to choose one of the three approaches.

1.7 The Inclusive Fitness Controversy

Recent debate has erupted over the status of inclusive fitness theory, particularly in relation to multi-level selection theory (Nowak et al., 2010; Abbot et al., 2011; Marshall, 2011; Gardner et al., 2011). Charges levelled at inclusive fitness theory by Nowak et al. (2010) include that it does not extend beyond 'standard natural selection theory', that there is a lack of evidence, that it has not made a meaningful contribution, and that it requires many assumptions, including weak selection and additive, pairwise interactions. While inclusive fitness theory underlies specific understanding of certain systems, such as the evolution of eusociality (social groups of multicellular organisms in which a major transition - defined in the next section - has occurred), Nowak et al. (2010) pay particular attention to the haplodiploidy hypothesis, which describes how haplodiploidy (a type of breeding system) predisposes species to eusociality (Hamilton, 1964b, 1972); they present an alternative theory of eusociality, based loosely around multi-level selection.

Swift responses were delivered to Nowak et al. (2010); over one hundred academics co-authored a rebuttal (Abbot et al., 2011) which addressed each of the points raise by Nowak et al. (2010). Firstly, Abbot et. al. argue that Nowak et. al. misunderstand inclusive fitness theory in relation to 'standard natural selection theory'; Nowak et. al. seem to view 'standard natural selection theory' as simply involving analysis of neighbour-modulated fitness, incorrectly relabelled as direct fitness, instead of inclusive fitness (Marshall, 2011). Abbot et. al., by contrast, consider the role of natural selection differently; rather than involving maximisation of any quantity, natural selection is the mechanism by which individuals maximise their inclusive fitness (though this statement rather ignores neighbour-modulated fitness, and the work of Grafen (1999, 2014a), described in Section 1.5). Abbot et al. (2011) also state, straightforwardly, that none of the assumptions Nowak et al. (2010) view as necessary for inclusive fitness are actually required. Further, Abbot et. al. provide examples of a large number of behavioural phenomena which have been understood with the aid of inclusive fitness theory, and a further number of examples of behaviours within the field of eusocial insects about which ift has made successful predictions. They regard Nowak et. al.'s targeting of the haplodiploidy hypothesis a straw man argument, since the importance of haplodiploidy to the evolution

of eusociality has since been understood to be much smaller than initially thought (Abbot et al., 2011).

More in-depth critiques have been provided of these points, and a number of other criticisms of inclusive fitness theory in the literature (Marshall, 2011; Bourke, 2011; Gardner et al., 2011; Birch and Okasha, 2014; Birch and Marshall, 2014). Birch and Marshall (2014) argues that a charitable reading of Nowak et al. (2010) is that they are in fact arguing against a particular form of Hamilton's rule, and that poor presentation and misinterpretation of their argument has led to a continuing debate. However, Birch and Marshall (2014), along with many other authors (Abbot et al., 2011; Bourke, 2011; Gardner et al., 2011), refute claims against the generality of inclusive fitness; discussion of the actual level of generality of inclusive fitness, particularly in comparison to neighbour-modulated fitness, is located in Section 1.4.4. Marshall (2011) specifically focuses on Nowak et. al.'s (2010) preference for a multilevel, or group, selection approach, seeing this as a resurgence of an old strain of arguments arguing for the primacy of one approach or the other; in fact, as Marshall (2011) shows and as we have mentioned, the kin selection and multilevel selection approaches are mathematically equivalent.

In response to the claim by Nowak et al. (2010) that there is a lack of evidence for inclusive fitness theory, Bourke (2014a) has provided a thorough review of a number of studies that provide explicit tests of Hamilton's rule. One such study is that of Hatchwell et al. (2014); their long-term study of the cooperatively breeding long-tailed tits *Aegithalos caudatus* has yielded data which allows individual estimates of the components of Hamilton's rule, r, b, and c. This is specifically possible in this species since complete life-histories of individual birds are easily obtainable due to short life-spans and there are a limited set of actions within a relatively simple breeding system; if an individual fails to breed in the breeding season, it helps to rear offspring of related individuals. Hatchwell et al. (2014) analyse the relatedness of helping individuals to the breed parents whose offspring they help to rear, the benefits of helping through increased survival of young and a reduced burden on the male parent, and the costs of helping in terms of future productivity. Thus, Hatchwell et al. (2014) obtain values of r and b relating to each benefit, and a single value of c. Together, these components show that Hamilton's rule is satisfied for

the helping behaviour, and that it primarily rests on the benefit of brood productivity rather than the reduced burden on the male parent. This is, of course, just one example of a study which aims to quantify and check Hamilton's rule (Bourke, 2014a).

While Nowak et. al.'s (2010) paper resulted in a great deal of attention, subsequent papers criticising inclusive fitness theory have provoked less response. Given how thoroughly each point has been refuted (Marshall, 2011; Bourke, 2011; Gardner et al., 2011; Birch and Okasha, 2014; Birch and Marshall, 2014), the status of inclusive fitness theory as being part of the status quo does not remain validly questioned; thus, this topic remains a side-note, and the rest of this thesis will continue to use kin selection and inclusive fitness theory unimpeded.

1.8 The Major Transitions

The central idea of the major transitions view of evolution is that all organisms can be organised into a biological hierarchy, in which individuals at one level evolved through particles at the lower level cooperating to such an extent that collectives attain individuality at a higher level (Buss, 1987; Maynard Smith and Szathmáry, 1995; Bourke, 2011; West et al., 2015). The general principles involved in each of the 'evolutionary transitions in individuality' (ETIs) are the subject of the major transitions view of evolution.

Bourke (2011) defines major transitions as 'transitions that involve the evolution of individuality, or at least groupings that are regarded as being candidates for individuals'. This refers back to the philosophical uncertainty of the exact definition of an individual. Bourke provides a definition ('some stable, physically discrete entity that is composed of interdependent parts acting in a coordinated manner to achieve common goals and is typified by the very property of lacking a high degree of within-individual conflict', with the qualification that 'physically discrete' means 'parts of the individual are either physically joined or tend to remain in close proximity'), though recognises that it is not precise, and is possibly wider than that considered elsewhere in the literature - again, this echoes the problems highlighted elsewhere

(Clarke, 2010). The concept of individuality, relabelled 'organismality' by Queller and Strassmann (2009) shall be returned to in Chapter 5. Bourke's definition ensures that each of the major transitions considered is structurally similar. Maynard Smith and Szathmary considered a slightly wider definition of the major transitions, suggesting that they could involve transitions in 'the language whereby information is transmitted'; thus, they considered the evolution of RNA, and human language, as major transitions. Clearly, these do not fit under Bourke's definition, as they do not involve the grouping of individuals (Maynard Smith and Szathmáry, 1995). We list Bourke's (2011) more recent classification of six major transitions in table 1.2, adapting the information from his table 3.1 :

Table 1.2: Classification of Major Transitions

Transition	Egalitarian/Fraternal?	Between/Within Species?
Separate replicators (genes) ⇒ Cell enclosing genome	Both	Within
Separate unicells ⇒ Symbiotic unicells	Egalitarian	Between
Asexual unicells ⇒ Sexual unicells	Egalitarian	Within
Unicells ⇒ Multicellular organisms	Fraternal	Within
Multicellular organisms ⇒ Eusocial societies	Fraternal	Within
Separate species ⇒ Interspecific mutualisms	Egalitarian	Between

We have included some details about each transition that require further explanation. Two decompositions of types of major transitions shall be identified. The first of these is the distinction between transitions that occur between or within species. The second decomposition was suggested by David Queller (1997) when reviewing Maynard Smith and Szathmary's book; he took inspiration from the French revolutionary motto, 'Liberté, Egalité, Fraternité' to propose that there are two types of major transition; 'egalitarian', and 'fraternal', alternatives to the 'libertarian' way of life. Egalitarian transitions occur when social groups of non-relatives come together to form an individual at a higher level of the biological heirarchy. Non-relatives could be conspecifics or non-conspecifics; either way, the only social behaviours that can occur are mutually beneficial. Thus, during an egalitarian transition, these individuals must retain their own reproductive capabilities. Fraternal transitions occur between genetically-related individuals, and thus can only occur within species. At the end of a fraternal transition, a reproductive division of labour will occur; this is possible since genetic relatedness allows altruism (Bourke, 2011;

Queller, 1997, 2000).

Bourke's (2011) recent work, 'Principles of Social Evolution', decomposed the major transitions into three components - social group formation, maintenance, and transformation . Social group formation consists of the initial spread of social behaviours in a population; social group maintenance refers to the stage at which processes for the control of conflicts are developed, and social group transformation refers to the processes that transform a stable group into an individual in its own right. The order of these stages may not be entirely strict; aspects of one stage may actually occur before or after aspects of another (McShea and Changizi, 2003). Note that each step in a major transition will not necessarily arise after the previous step; evolution is not necessarily progressive (Maynard Smith and Szathmáry, 1995; Bourke, 2011).

1.9 Thesis Structure

The aim of this thesis is to discuss questions relating to the major transitions in evolution, through the lens of social evolution theory. We now briefly discuss the structure of the thesis, and how each question fit into the major transitions framework. The models we present involve between-species interactions. Recall, as discussed in Section 1.8, that fraternal transitions, involving relatedness between interacting individuals, can involve altruism. Egalitarian transitions, similarly introduced in Section 1.8, do not involve relatedness between interacting individuals, and require mutually beneficial social behaviour. The four types of social behaviour, altruism, mutual benefits, spite, and selfishness, were introduced in Section 1.3.

While inclusive fitness theory and multilevel selection theory, discussed in Sections 1.4-7, are explicitly invoked to varying degrees over the following chapters, they both underlie a wider understanding of analysis of social behaviours. For example, relatedness and the categorisation of social behaviours are given particular weight in Chapter 3, while selection at the level of the group is an important aspect of the model considered in Chapter 4. Chapter 2 gives more detailed introductions, and literature reviews, of topics relating to the specific modelling contexts of Chapters 3 and 4, supplementing the theoretical background we have introduced in this chapter.

Chapter 3 discusses models of pairwise between-species donation, using a modification of the donation game discussed in Section 1.4.1; under what conditions does between-species donation behaviour spread, what parallels exist to biological reality, and how can donation behaviour be classified according to the four-way decomposition of social behaviours provided in Section 1.3? It has been proposed that between-species donation amounts to between-species altruism, which would imply the possibility of the existence of fraternal transitions between species, which are notably absent from Table 1.2.

In Chapter 4, we discuss a question related to egalitarian transitions between species, introducing a model of maternally-transmitted symbionts. We are interested in when reproductive parasites can become mutualists, engaging in mutually-beneficial behaviour between species. This is a necessary condition for an egalitarian transition between species.

Finally, in Chapter 5 we discuss further theories regarding the major transitions, focusing in particular on fraternal transitions within species. This is fertile ground for future research, and in particular relates to Bourke's (2011) social group transformation stage of an ETI .

Chapter 2

Between-Species Interactions: Donation, Symbiosis, and Sex-Ratio Distortion

2.1 Introduction

In Chapter 1, we introduced the idea of the major transitions in evolution, and several related topics, including inclusive fitness theory and multi-level selection theory. We now turn to specific areas within this wide field of study, starting with some discussion of models of between-species donation and proposed between-species altruism (Section 2.2) and introducing the concept of greenbeard genes (Section 2.3), establishing the context for the model we analyse in Chapter 3. In addressing proposed models of between-species altruism, this chapter focuses on fraternal transitions between species.

We also introduce some further background relating to endosymbiosis and sex-distortion, and look at models of sex-distortion in the context of parasite invasions and transitions from parasitism to mutualism (Section 2.4). This will be relevant to Chapter 4, in which we discuss a model of maternally-transmitted sex-distorting symbionts in a host population. Since mutualism is necessary for an egalitarian major transition, this chapter focuses on egalitarian transitions between species.

2.2 Between-Species Donation

Fraternal transitions in individuality require altruism; examples of fraternal transitions are those from unicellularity to multicellularity, and from multicellularity to eusociality (Bourke, 2011). Both of these transitions are within species, since altruism requires relatedness. However, there have been attempts to generalise the concept of relatedness to between-species interactions (Frank, 1994; Fletcher and Doebeli, 2009), thus allowing the argument that altruism is possible between members of different species. Were this possible, theory surrounding major transitions would require some reconsideration, since fraternal transitions could in principle be possible between different species. Therefore, claims that between-species altruism is possible require careful examination; various models have been proposed arguing this, which are examined in the next section.

Chapter 3 discusses models of between-species donation, intended to test whether or not stable altruism can occur between species. We now introduce and discuss aspects of the literature relating to between-species donation.

2.2.1 Previous Models

An early discussion of altruism in the context of between-species interactions was provided by Trivers (1971), who coined the term 'reciprocal altruism' to describe altruism that occurs without relatedness. He argued that if reciprocal altruists are non-randomly paired with other reciprocal altruists, their expected fitness may be greater than non-reciprocal altruists. However, since social behaviours are defined by expected lifetime fitness, the behaviour Trivers calls 'reciprocal altruism' is actually correctly defined as mutually beneficial rather than altruistic (Hamilton, 1996). Recall that both altruistic and mutually beneficial behaviour involve social partners receiving some fitness benefit; however, altruism involves incurring an expected lifetime fitness cost, as opposed to the expected lifetime fitness benefit associated with mutually beneficial behaviour (see Table 1.1).

Frank (1994) considers a model of donation between groups of hosts and mutualists, using it to show that selection can create positive correlations between donation alleles in two species. He compares two classifications of social behaviour recognised

in the literature to the behaviour expressed in this model. He argues that the donation behaviour is not altruism due to kin selection, since donators do not directly donate to related individuals of the same species. He also argues that it is not what Trivers terms 'reciprocal altruism', since reciprocity here involves correlations rather than direct return of fitness benefits. Instead, Frank summarises the behaviour as altruism resting on a genetic correlation between species.

Fletcher and Zwick (2006) also focus on proposed altruism between species, by attempting to unify Queller's (1985) formulation of Hamilton's rule and reciprocal 'altruism'. They consider an iterated prisoner's dilemma, a donation game in which two individuals are repeatedly given the chance to either cooperate, donating some fitness benefit b to the social partner while incurring a fitness cost c, or defect, in which case no fitness changes are made. Possible strategies are to always defect, always cooperate, or act conditionally on the social partner's behaviour. When Axelrod and Hamilton (1981) considered this game, they discovered that the relatively simple tit-for-tat strategy (specifying cooperation if the social partner cooperated last round and defection if not) was an evolutionarily stable strategy (ESS: in a pairwise scenario, a strategy is an ESS if and only if the ESS yields greater fitness to an individual than any other strategy, in interactions with other ESS individuals (Maynard Smith and Price, 1973)). Note that we consider a non-iterated version of the prisoner's dilemma in Chapter 3.

Fletcher and Zwick firstly consider an asymmetric version of the iterated prisoner's dilemma; that is, benefits and costs differ across species, though the cost to a given individual is always less than the benefit gained from a cooperating member of the other species. In this case, the tit-for-tat strategy is stable for both species. They also find that it is possible for cooperative behaviours to be stable in the symmetric (i.e. benefits and costs relating to cooperation and defection do not vary across species) version of the iterated prisoner's dilemma when different potentially-cooperative strategies are considered across species. However, they note that the reason for the social behaviour reaching fixation in both populations is in fact that cooperators have greater expected fitness than defectors. Thus, although their analysis is framed as a discussion of altruism, in fact, by once again applying the definitions described in Chapter 1, this social behaviour is instead classified as

mutually beneficial.

Foster and Wenseleers (2006) analyse two-species interactions with a general model focusing on the evolution of mutualism. Their model assumes that fitness benefits donated from an individual in a group to member(s) of an associated group of another species are returned, and that this feedback has several causes; cooperator association, partner-fidelity feedback, and partner choice. Cooperator association refers to the tendency of individuals with genotypes that specify cooperation to be paired, partner-fidelity feedback refers to cooperation causing passive phenotypic effects in the partner species that provide feedback benefits (e.g. increases in group size or survival), and partner choice refers to cooperation causing active phenotypic effects in the partner species, such that partners discriminate between cooperators and non-cooperators. These latter two effects correspond, respectively, to what Queller (2011) labels 'kith' and 'kind' selection; he derives an extended version of Hamilton's rule that includes terms relating kin, kith and kind selection.

Foster and Wenseleers (2006) derive a condition for the spread of the donation behaviour, which resembles Hamilton's rule; however, while the cost and benefit terms relate to the social behaviour in the model, the relatedness term is replaced by the sum of three components, each relating to the effects described above.

Each of the three components of fitness feedback are analysed in turn. Cooperator association is shown to be a form of genotype assortment across species; the mathematical form found by Foster and Wenseleers (2006) is in fact equivalent to what Frank (1994) terms between-species relatedness. However, this effect requires that mutualist genotypes remain associated across generations; Foster and Wenseleers consider this a limiting constraint. They emphasise the importance of partner-fidelity feedback in relation to cooperator association; indeed, in this model, partner-fidelity feedback tends to have a greater effect than cooperator association, since phenotypic changes can occur much more quickly than genetic changes.

Foster and Wenseleers (2006) conclude that cooperator association mechanisms tends to be disproportionately represented in the literature. Partner-choice mechanisms, on the other hand, are potentially very effectual since they can involve high

within-species relatedness and between-species fidelity (the tendency of different species to remain associated over time); however, they are usually found in only one species in a mutualism, thus are not sufficient to entirely explain the evolution of a given mutualism. Throughout this model, Foster and Wenseleers identify within-species relatedness as having particular importance to the evolution of mutualism; they decompose cooperator assortment in such a way that it is dependent on within-species relatedness. The importance of within-species relatedness to between-species donation with cooperator association has been noted by other authors, and may apply to the model of between-species donation we discuss in Chapter 3.

A contrasting viewpoint is apparent in a more recent paper by Fletcher and Doebeli (2009); they view altruism between species as possible, with assortment between species as fundamental to that. They consider a two-locus model, in which A and B genes encode for donation to a public good, whereas recessive a and b genes do not; AB individuals are not viable. Assortment is applied such that Ab and aB individuals are paired up as much as possible; under this assortment, the A and B alleles go to fixation. They argue that the donation behaviour, which they label altruism, reaches fixation due to assortment, which they consider to be a more general form of relatedness. Fletcher and Doebeli also consider an extension to this model, in which Ab and aB individuals suicidally donate with probability q. They conclude that this scenario is possible for some parameter values, and take it as an ultimate example of altruism between non-relatives. Therefore, they conclude, assortment is a fundamental mechanism, and relatedness is simply a particular case of this mechanism; thus, altruism can evolve in the absence of genetic relatedness.

Gardner et al. (2011) apply a kin selection analysis to Fletcher and Doebeli's thought experiment, specifically focusing on the extension involving suicidal donation. They use a regression analysis to obtain the cost, benefit and relatedness according to Hamilton's rule, recovering Fletcher and Doebeli's condition for the spread of donation. However, unlike Fletcher and Doebeli, Gardner et. al. conclude that the coefficient of relatedness is in fact 1 rather than 0, citing the fact that relatedness is genetic rather than genotypic (see also Table 7.1 in Marshall (2015)). Thus, Fletcher and Doebeli's statement that assortment is more fundamental than relatedness is unfounded.

A further discussion of the possibility of altruism between species was carried out by Wyatt et al. (2013). They consider an infinite one-dimensional stepping-stone model, in which each patch holds an individual of each species who play a single-shot prisoner's dilemma. As before, donators confer b fitness to the social partner, and incur c loss of fitness. However, under this model all individuals are replaced with clonal offspring every generation, apart from a proportion of patches in which competition occurs between adjacent individuals of the same species for reproduction in those patches. Wyatt et. al. consider the invasion of donators into a population of non-donators, observing that two paired helpers of different species may give rise to an expanding chain of adjacent helpers. Indeed, they derive a mathematical condition for donators to receive positive selection.

The most significant aspect of Wyatt et. al.'s work is their analysis of this social behaviour. Is it truly between-species altruism or not? They consider a regression analysis of fitness, observing that the choice of predictors is important. For example, an individual's fitness is dependent on both their genes, their social partner's genes, the genes of the potential competition for reproduction (i.e. same-species individuals two patches away), and the genes in the corresponding two-step neighbours of the other species. Using these predictors, a regression on fitness concludes that donation is directly costly, with benefits dependent on the other species; thus, the behaviour is classified as between-species altruism.

An alternative approach is to consider only genetical predictors from members of the same species, in which case only genes of same-species individuals of previous generations are used as predictors, as the genes of members of the other species are dependent on these. Under this regression analysis, the behaviour is diagnosed as within-species altruism, since it is directly costly and any indirect fitness benefits are due to member of the same species. This reflects the work of Foster and Wenseleers (2006), who considered that within-species relatedness plays an important role in stable social behaviour between species.

Ultimately, Wyatt et. al. leave an open conclusion, stating only that between-species altruism is not an exclusive interpretation of a social behaviour, since benefits

can be traced back to member of the same species.

The model we consider in Chapter 3 is an adaptation of the model that Fletcher and Doebeli (2009) discuss, which involves a single-shot prisoner's dilemma between assorted, but allegedly unrelated, members of the same species. The simplest form of the model we consider can be considered as analogous to the model of Fletcher and Doebeli, except that interactions are now specifically between-species; thus, we avoid the ambiguities of relatedness and assortment which complicate interpretation of their model. The simplicity of our between-species model allows for consideration of different types of assortment and possible mechanisms which allow between-species donation to remain stable.

2.3 Greenbeards

An application of the model considered in Chapter 3 is to the study of greenbeard traits; we briefly introduce and summarise these here. Greenbeard genes are genes which satisfy several conditions; they cause bearers of the greenbeard gene to recognise other bearers of the greenbeard gene and behave differently towards them. The idea, and name, for greenbeard genes arose as a result of a thought experiment, in which the recognition component of the greenberad gene behaviour took the form of literal green beards (Dawkins, 1976). Greenbeards may be 'facultative', in which case greenbeard individuals adjust their behaviour in response to other greenbeard individuals, or 'obligate', in which case greenbeards individuals perform a fixed behaviour which differentially affects greenbeard individuals and non-greenbeard individuals (Gardner and West, 2010).

The social ameoba *D. discoideum* forage for bacteria in forest soil, but if food is scarce, some cells will altruistically sacrifice themselves to allow related cells to form clusters of spores which can relocate to find better food sources. The *csA* gene allows the formation of these spores through cellular adhesion, thus qualifying as an obligate greenbeard gene since non-csA individuals cannot be part of these spores as they lack the adhesion property (Queller et al., 2003).

A second example of a greenbeard gene has been identified in the red fire ant *Solenopsis invicta* (Keller and Ross, 1998). They find that the $Gp\text{-}9^b$ allele causes workers who possess it to attack queens that initiate reproduction without possessing the allele; said workers identify queens without the allele through odour. Thus, this is an example of a facultative greenbeard gene.

Greenbeard traits are relevant to our model; they are vulnerable to genes which suppress the donation behaviour but retain the mark of recognition for receiving fitness benefits. We investigate genes arising on unassorted loci that undermine stable donation behaviour in Chapter 3; this is analogous scenario.

2.4 Maternally-Transmitted Symbionts

A symbiosis is a close physical association between two species; symbioses can be split into three types (Bronstein, 1994). Parasitism involves one species (the parasite) gaining a fitness benefit to the detriment of the other species (the host); the cost to the host is often termed 'virulence'. Commensalism involves one species (the commensalist) gaining a fitness benefit, while the fitness of the other species (the host) is unaffected. A mutualism, on the other hand, involves fitness benefits going to both species. Mutualisms are required for an egalitarian major transition to occur between species (Bourke, 2011); thus, modelling potential pathways towards the formation of mutualism is an area of interest. One possible pathway is that from an established parasitism towards mutualism; other origins may be of enslavement of a symbiont by a host followed by coevolution towards mutualism, commensalism between two species followed by coevolution towards mutualism, or changes in ecological conditions resulting in an environment capable of supporting a specific mutualism (Aanen and Bisseling, 2014)

The method of transmission of a symbiont is also relevant to our understanding of this field. Transmission of the symbiont from parent to offspring (either paternally or maternally) is labelled vertical transmission; all other types of between-individual transmission of symbionts are labelled horizontal (Fine, 1975). Horizontal transmission often occurs at the expense of a host, depending on the exact mechanism, whereas vertical transmission requires host reproduction; therefore a trade-off be-

tween horizontal and vertical transmission is often assumed in modeling, or arises out of other assumptions, such as positive correlation between horizontal transmission and virulence (Ferdy and Godelle, 2005).

Chapter 4 discusses a model of maternally-transmitted endosymbionts, particularly in relation to a potential egalitarian transition in individuality. Specifically, we focus on a maternally-transmitted symbiont with the potential to distort sex ratios; we investigate what mechanisms could cause a host and symbiont to peacefully coexist without the symbiont distorting sex ratios. In this way, we model the transition from parasitism to mutualism, which, as noted above, is required for an egalitarian transition between species.

This section deals with specific aspects of the literature relevant to this model, both expounding on topics mentioned in passing here, and a summary of previous models which relate to this topic.

2.4.1 Sex Ratio Theory

Fisher (1930) argued that parents will invest equal resources into male and female offspring. His argument starts by supposing that males and females require equal amounts of investment to produce. Suppose also that parents tend to produce more females. Therefore, males have greater fitness than females as they give rise to more offspring; thus, parents that produce more males than the average will also have more grandchildren on average. Genes for greater male-production will then spread; this effect continues until the mean sex ratio is 0.5. At this point, there is no better strategy than to give birth to equal numbers of male and female offspring. This argument has been re-stated in the literature multiple times, since Fisher's original formulation was extremely concise (Fisher, 1930; Hamilton, 1967; West, 2009). The sex ratio of 0.5 is known as the 'Fisherian' sex ratio.

However, there are exceptions to this sex ratio being optimal. Hamilton (1967) gave several examples. One example involves the gene controlling sex ratio being paternally-inherited, for example due to being located on the Y-chromosome in humans. In this case, a male-biased sex ratio will be favoured by selection, since paternal inheritance means that only males contribute directly to the spread of the

gene, so mutations favouring the production of males will spread. Hamilton discusses extreme theoretical examples of paternally and maternally inherited sex-distorting genes, noting the possibility of extinction due to excessive distortion of a population's sex ratio. We shall return to the subject of sex distortion due to maternal inheritance in Chapter 4.

Another example Hamilton (1967) discusses is that of local mate competition (LMC). Under LMC, females move between food sources inhabited by males, who remain. Thus, males compete with related males for reproduction. The ESS here involves a sex ratio biased towards production of females. There are two reasons for this; producing fewer males results in less competition between related males, and more daughters means disproportionately more mates for sons (Taylor, 1981; West, 2009). In general, ESSs can be formulated for sex ratios which depend on the degree of matings between siblings ('sibmatings'), for given mating systems (Taylor, 1993; West, 2009); resulting predictions have been tested and validated in many species (see West (2009) for a long list).

Trivers and Hare (1976) noted that relatedness asymmetries may also lead to deviations from Fisher's theory; for example, in Hymenoptera, workers are three times more related to sisters than brothers, whereas queens are equally related to sons and daughters (Trivers and Hare, 1976; Bourke, 2015). Thus, conflict arises between workers and queens over sex ratio. Trivers and Hare (1976) survey several other species, outlining how observed sex ratios match with predictions on the basis of relatedness asymmetries; for example, they test the prediction that sex ratios in monogynous ants (i.e. those with single queen colonies) will approach a 3:1 female to male ratio since workers control the rearing of young. They utilised data on 21 species of monogynous ants to test this prediction, finding that it is on average borne out by the data. While contradictory theories have been proposed, extensions to Fisher's sex ratio theory remain consistent with sex ratios based on relatedness asymmetries (Bourke, 2015).

Trivers and Willard (1973) discuss another mechanism that causes deviations from the Fisherian sex ratio: conditional sex allocation. They give the example of a situation in which females vary in condition, and offspring inherit that condition;

however, the reproductive success of male offspring is more affected by condition than the reproductive success of female offspring. Females in poor condition will then have more grandchildren if they have a bias towards production of female offspring, while females in good condition will have more grandchildren if they have a bias towards producing male offspring. This is because the downside of being in poor condition can be mitigated by producing offspring less affected by that condition; the converse applies.

Conditional sex allocation may result in a deviation in the population-level sex ratio from 0.5; since it involves a non-linear relationship between parental investment and reproductive output, Charnov's non-linear model applies, which results in potential deviations from equal sex allocation (Charnov, 1979; Frank, 1987, 1990). The theory holds more widely than the simple scenario posited above; for example, an analagous circumstance in parasitoid wasps has been reported, in which sex ratios of offspring, which are laid in batches in different hosts, vary depending on the size of the host (West, 2009). As with LMC theory, there have been many additions to this theory, and tests of its predictions (West, 2009).

2.4.2 Endosymbionts

'Endosymbiosis' refers to a symbiosis in which one of the two organisms lives inside the other; the smaller organism is referred to as an 'endosymbiont'. Endosymbiosis has played a key role in the biological hierarchy; eukaryotic cells are the result of the endosymbiosis of prokaryotic cells and ancestral forms of plastids such as mitochondria and chloroplasts (Margulis, 1981; Maynard Smith and Szathmáry, 1995; Bourke, 2011). In Chapter 1, we noted that one of the major transitions was from separate unicells to symbiotic unicells; the transition from prokaryotes to eukaryotes is an instance of this type of between-species egalitarian transition (Bourke, 2011). Such transitions still occur today; Marin et al. (2005) found evidence that an amoeba has recently obtained photosynthetic organelles through endosymbiosis.

Maynard Smith and Szathmáry (1995) discuss the acquisition of mitochondria and chloroplasts in some detail. They first note that mitochondria and chloroplasts retain properties associated with free-living prokaryotes, but that their genes have been transferred to the cell nucleus since becoming symbionts. One interest-

ing hypothesis that they discuss is that of 'slavery'. Under this alternative idea, the endosymbionts do not necessarily gain a benefit from the association; the argument hinges on the fact that the association may not have been initially beneficial to the endosymbionts. They propose that initially, mitochondria and chloroplasts were subject to controlled exploitation - in other words, they were forced to remain within the host cell on the basis that the host could gain a fitness benefit from their presence. They also discuss details of mechanisms which allowed the host to reap more elaborate benefits from the endosymbiont, but propose that these could only have evolved after the association had been established. In reality, it is worth investigating the possibility of immediate direct fitness benefits to the host, since, as shall be discussed, it has been shown that they can occur with the endosymbiotic bacteria *Wolbachia*; as a result, the 'slavery' hypothesis is worthy of investigation.

Endosymbionts are present in many types of organism, including amoebae, flagellates, ciliates, plant cells, and invertebrates (Jeon, 2011). We now focus on those endosymbionts within sexually-reproducing organisms that cause sex ratio distortion.

2.4.3 Sex-Ratio Distorters

Sex-ratio distorters can be either maternally (through the cytoplasm) or paternally (through the nucleus) inherited; we focus here only on maternally-inherited sex-distorters. Maternally-inherited (i.e. cytoplasmically inherited) sex-ratio distorters can cause various phenotypic effects in their host and have been summarised extensively in the literature (Werren, 1997; Hurst and Werren, 2001; West, 2009); we discuss four of the most common phenotypes below. However, each phenotype reflects a general principle; endosymbionts which are maternally transmitted have greater fitness when they cause female-biased distortion in sex ratios for the host species. If infected females have more infected female offspring than uninfected females have uninfected female offspring, selection favours the spread of the endosymbiont. Males are effectively an evolutionary dead-end for maternally-transmitted endosymbionts.

In addition to benefiting their own spread through sex-ratio distortion, endosymbionts may provide fitness benefits to the host, though there may also be inherent fitness costs associated with sex-ratio distortion. Many maternally-transmitted

symbionts confer resistance against parasites; for example, *Spiroplasma* protects *Drosophila neotestacea* against parasitic nematodes (Jaenike et al., 2010), and a cytoplasmic-incompatibility-inducing form of *Wolbachia* can successfully invade a species of mosquitoes, conferring resistance against RNA viruses such as Dengue fever (Walker et al., 2011). It has also been proposed that the subsequent development of increases in fitness of *Drosophila melanogaster* as a result of *Wolbachia* infection could mean that some parasitic forms of *Wolbachia* are part-way through a development in their relationship towards mutualism (Fry and Rand, 2002). Since genetic changes resulting in reduction of costs are more likely than the addition of new benefits, a potential pathway towards mutualism would be parasites that cause a net cost, in spite of conferring fitness benefits such as infection resistance to a host, reducing the associated fitness costs to the host (Ewald, 1987).

Horizontal transfer of sex-ratio-distorting symbionts can be compared to new mutations; they result in some change in mean fitness to the new host, and occur at a certain rate (Jaenike, 2012). It has been hypothesised that the distribution of fitness changes resulting from a horizontally-transferred symbiont is less biased towards fitness costs than that of fitness changes resulting from mutations, since sex-ratio-distorting symbionts are more likely to have been transferred from host species which they have succesfully invaded, which is more likely to occur when the symbiont provides a fitness benefit (Jaenike, 2012). However, the exact rate of horizontal transfer is dependent on the biological specifics of the two host species and the symbiont. Beneficial symbionts may also experience an increasing probability of transmission as they spread through communities of potential hosts (Jaenike, 2012).

We now discuss the four types of sex-ratio distortion caused by endosymbionts; while none of these are specifically modelled in Chapter 4, the more general principle of a symbiont inducing a biased sex ratio in a host is. Each of these types of sex-distortion can be understood as reflecting this general principle; bias in favour of female production as a means to spread the maternally-transmitted endosymbiont. In Chapter 4, we briefly discuss how specific types of sex-distortion can be considered by our model through reparameterising the general form of sex-distortion that we model.

Cytoplasmic Incompatibility

Cytoplasmic incompatibility (CI) refers to the inability of female members of many insect species to reproduce with males infected with certain strains of *Wolbachia* (O'Neill et al., 1992; Rousset et al., 1992; Yen and Barr, 1973; Hoffmann et al., 1996; Famah Sourassou et al., 2014) or *Cardinium* bacteria (Gotoh et al., 2007; Famah Sourassou et al., 2014). CI-inducing bacteria cause sperm to be 'modified' in some way such that only eggs from individuals infected by the same strain of CI-inducing bacteria can be fertilised by these sperm. This is referred to as a 'modification-rescue' system (Werren, 1997). The result of the modification without rescue is a haploid zygote due to failed karyogamy (O'Neill et al., 1992), which causes male-biased sex ratios in hymenoptera (West, 2009; O'Neill et al., 1992).

Thus, aside from the special case of haplodiploidy, there are two types of compatibility. Under 'unidirectional' compatibility, which occurs when there is one CI-inducing bacteria present in a population, infected males can not breed with uninfected females, since the modified sperm are not rescued; however, sperm are viable in every other pairing. Under 'bidirectional' compatibility, which occurs when there are two CI-inducing bacteria present in a population (we label them types 'A' and 'B'), A-males cannot breed with B-females, and B-males cannot breed with A-females; every other male-female pairing is compatible (Yen and Barr, 1973).

The advantage of CI to bacterial endosymbionts such as *Wolbachia* and *Cardinium* is that (speaking in the language of unidirectional incompatibility, though the same logic holds for the bidirectional case), by ensuring that infected males cannot mate with uninfected females, they are reducing the fitness of uninfected females in comparison to infected females. Thus, they aid their own spread, since they are concerned with the fitness of infected females. This primarily occurs when CI-inducing bacteria are not rare; however if CI-inducing bacteria are rare, they can still experience positive selection if there is local resource competition (West, 2009).

Thelytokous Parthenogenesis

Parthenogenesis refers to reproduction in which unfertilised eggs give rise to offspring. In haplodiploids such as Hymenoptera, males develop from unfertilised eggs

('arrhenotoky'); however, *Wolbachia* can induce a form of parthenogenesis in which females develop from unfertilised eggs ('thelytoky') (Hamilton, 1967; Arakaki et al., 2001; Weeks and Breeuwer, 2001). Thelytokous parthenogenesis (TP) takes two forms. 'Apomictic' TP entails offspring being genetically identical to their mother. 'Automictic' TP, by contrast, involves genetic recombination; this covers several different mechanisms (Rabeling and Kronauer, 2011). Three types of bacteria have been shown to induce TP; *Wolbachia* have been found to do so in non-eusocial Hymenoptera and species of thrips and mites (Huigens et al., 2000; Arakaki et al., 2001; Weeks and Breeuwer, 2001; Rabeling and Kronauer, 2011), *Cardinium* have been found in spider mites and parasitoid wasps (Zchori-Fein et al., 2001; Gotoh et al., 2007), and *Rickettsia* do so in species of eulophids and parasitoid wasps (Hagimori et al., 2006; Giorgini et al., 2010). All of these species are haplodiploid, though the mechanism of *Wolbachia*-induced TP in mites is of an apomictic type, which suggests that it is also possible in non-haplodiploid species, though it may be difficult to detect (Weeks and Breeuwer, 2001). In addition to being inherited vertically, parthenogenesis-inducing *Wolbachia* can also infect new hosts horizontally; this allows another source of genetic variation (Huigens et al., 2000; Watanabe et al., 2013).

TP is not necessarily deleterious to the host species; in fact, it has evolved in some ants, bees, and termites, in the absence of *Wolbachia* or *Cardinium* (Wenseleers and Van Oystaeyen, 2011). In several cases, it is a successful part of a reproductive system involving mixed modes of reproduction. For example, queens may reproduce asexually, while workers reproduce sexually (Wenseleers and Van Oystaeyen, 2011; Rabeling and Kronauer, 2011).

Another proposed consequence of TP is that it could work to suppress within-group conflict; if it is obligate, a group may be composed of genetically identical individuals (Rabeling and Kronauer, 2011). This could be considered to satisfy the condition for a major transition to occur illustrated by Boomsma's (2009) monogamy hypothesis of individuals being equally related to siblings and offspring; this shall be introduced and discussed in Chapter 5. However, worker policing may still be present in obligate-TP groups; whether this is due to existing social conflicts through multiple clonal lineages or remains due to recent ancestry appears undetermined (Hartmann et al., 2003; Kellner and Heinze, 2011).

Feminisation

Wolbachia-induced feminisation involves the development of genetic males into females; the process is not entirely reliable as it needs to act continuously through development, sometimes resulting in intersex individuals (Narita et al., 2007; Werren et al., 2008). This is in contrast to CI and TP, which act at a very early stage of development. *Wolbachia* has been found to cause feminisation in species of butterfly (Narita et al., 2007) and isopods (Rigaud and Juchault, 1995; Bouchon et al., 1998), while *Cardinium* causes feminisation in arthropods (Weeks and Breeuwer, 2001; Zchori-Fein and Perlman, 2004).

Male-killing

Male-killing involves the deaths of males in development (early) or at a later stage in order to maximise horizontal transmission (late) (Hurst, 1991; West, 2009). It is caused by *Spiroplasma*, *Rickettsia* and *Wolbachia*, among others (Werren et al., 1994; Hurst et al., 1999a,b; Hurst and Jiggins, 2000), and primarily occurs in insects (Hurst and Jiggins, 2000; West, 2009); the precise mechanism causing death varies widely (Duplouy et al., 2013). Of the four commonly recognised methods of sex distortion by bacterial symbionts, this is the only one which *Cardinium* have not been found to exhibit (Martin et al., 2013). It may provide a fitness benefit to infected females since they face less between-sibling conflict for resources and the likelihood of inbreeding is decreased; alternatively, defense against parasites may also be conferred (Xie et al., 2014). In addition, male death may result in horizontal transmission of the symbiont (Hurst and Majerus, 1993; Hurst and Jiggins, 2000; Hurst et al., 2000).

An interesting case of co-evolution between a male-killing symbiont and its host is that of the *Wolbachia* strain *w*Bol1-b in the butterfly *Hypolimnas bolina* (Duplouy et al., 2013; Hornett et al., 2014). The strain *w*Bol1-b has recently emerged, rapidly diversifying due to horizontal gene transfer from other eukaryotes (Duplouy et al., 2013). This male-killing symbiont resulted in a 100:1 female to male sex ratio, before a suppression locus arose in the butterfly which allowed a return to the Fisherian sex ratio (Hornett et al., 2014).

2.4.4 Previous Models

While there is a vast literature discussing aspects of symbiosis, in particular regarding parasitism and virulence, we focus instead on models pertinent to the formation of mutualisms, in keeping with our aim of investigating a potential egalitarian major transition between species. Mathematical models of symbiosis necessarily simplify reality; for example, they typically deal only in expected outcomes of between-species interactions, whereas in reality, conditional outcomes may occur, with host and symbiont behaviour dependent on changes in environment, population size, or individual state (Bronstein, 1994).

Early work modelling the evolution of mutualism often focused on vertical transmission, in both the context of the formation of symbioses, and potential changes in host-symbiont relationships from parasitism to mutualism (Roughgarden, 1975; Ewald, 1987).

Fine (1975) derived the 'fundamental equation of vertical transmission', which could be used to find the stable equilibrium frequency of a purely vertically-transmitted symbiont in a sexually-reproducing host population. A sufficient condition for this kind of symbiosis was derived:

$$vf(\alpha_f + \alpha_m) > 1 \tag{2.1}$$

Here, f is the relative reproductive fitness of infected individuals in comparison to uninfected individuals (a.k.a. fecundity), v is the relative survival to reproduction of infected individuals in comparison to uninfected individuals (a.k.a. viability), α_f is the maternal vertical transmission rate, and α_m is the paternal vertical transmission rate.

This equation captures two important forces that determine whether or not a symbiont will be present in a host population at equilibrium. If the symbiont imposes fitness costs on the host through lower fecundity and/or viability, it must be highly efficient at vertical transmission to remain in the host population; in this scenario, the symbiont is a parasite. However, if the symbiont does not transmit

itself well, it can counteract this selective disadvantage by providing fitness benefits to the host; in this scenario, the symbiont is a mutualist. Note that Fine's model involves a host population of constant size, thus the possibility of a parasite driving a host population to extinction, as may happen if host fitness is decreased by a parasite, is not considered.

Lipsitch et al. (1995) consider a model of mixed vertical and horizontal transmission, splitting up the 'basic reproductive rate' of a symbiont, R_0, into two additive components, one due to vertical transmission and one due to horizontal transmission. Thus, a symbiont with inefficient vertical transmission requires horizontal transmission to persist, and vice versa. When a symbiont is entirely vertically transmitting, Fine's condition (2.1) is retrieved. Since the vertical transmission component accounts for fitness effects on the host, Lipsitch et. al.'s condition states that parasites require some degree of horizontal transmission to persist in a host population; the more harmful a parasite, the greater the importance of horizontal transmission. However, Lipsitch et. al. note that once a parasite reaches a high frequency in a host population, a small increase in vertical transmission will have a relatively much greater effect on equilibrium prevalence of the symbiont than a small increase in horizontal transmission, since the number of susceptible hosts decreases as parasite frequency increases. Mutualists, on the other hand, simply need to have sufficient vertical transmission to persist in a host population.

However, neither of these models consider selective pressures acting on vertical transmission, nor how selection on hosts and selection on symbionts may result in competing selective pressures.

Yamamura (1993) addresses this final point, modeling mixed vertical and horizontal transmission of parasites. He argued that conflict over vertical transmission will exist, since selection on hosts will favour a decrease in the vertical transmission rate, as parasites give hosts a fitness disadvantage, whereas selection on parasites will favour an increase in the vertical transmission rate. Yamamura specifically wished to investigate the progression of symbioses from parasitism towards mutualism; he argued that if the vertical transmission rate crosses a certain threshold, host and parasite fitness interests are aligned, leading to mutualism with a high

vertical transmission rate. This threshold value of the vertical transmission rate is dependent on baseline parasite and host fitnesses during a symbiosis, and the fitness effects of further exploitation by the parasite.

None of the models considered so far have invoked sex-ratio distortion. We now turn to models that include this feature, having established the importance of horizontal transmission to the establishment of parasitism, the importance of vertical transmission to the evolution of mutualism, and the potential for conflict between hosts and symbionts over the evolution of vertical transmission.

Werren (1987) created a general model of sex-ratio distortion, considering the conflict between cytoplasmic sex-ratio genes (CSR) and autosomal sex-ratio genes (ASRs). While CSRs are maternally-inherited, ASRs are inherited by either gender, so there may be conflict between the two genes over sex ratio. Werren (1987) looked at the coevolution of the two genes. Note that the conflict that Werren describes could manifest itself in conflict in preferred sex ratio between a sex-distorting endosymbiont (with sex-ratio strategy analagous to a CSR) and a host (with sex-ratio strategy analagous to an ASR).

Under perfect vertical transmission, a CSR will go to fixation if it can invade initially, which occurs when the product of the relative fitness of CSR individuals to ASR individuals and the sex-ratio of CSRs is greater than 1/2. This suggests that a CSR which involves a fitness cost can nonetheless spread if it distorts the sex-ratio in favour of females to a sufficient extent. In this case, there would be no compensatory selection on ASRs, for whom the ESS would remain 0.5, since ESR females will be in a separate subpopulation from ASR females (Werren, 1987).

However, under imperfect vertical transmission, a CSR will invade a population if the product of its relative fitness, vertical transmission, and sex ratio is greater than 1/2. In the scenario, a polymorphic equilibrium will occur, and ASRs will evolve to compensate, producing only males ('monogeny'). Selection on CSRs causes them to produce only females at equilibrium. Thus, in the absence of selection on vertical transmission, there is no coincidence of fitness interests between the two genes and their effects are entirely opposed.

Both situations have the potential to cause the host population to go extinct. Werren notes several possible mechanisms that may avoid this possibility: resistance genes that suppress the CSR phenotype, population structure (an idea explored more fully in Chapter 4), and frequency dependent changes in parameters.

Hatcher and Dunn (1995) consider a model of feminising symbionts in a host population, adjusting Werren's (1987) framework . Under Hatcher and Dunn's reformulation of Werren's model, feminisation occurs in infected individuals prior to sex determination of offspring, which is subject to the sex ratio of the host. Thus, the actual sex ratio of infected individuals is a function of host sex ratio, vertical transmission, and the degree of feminisation.

Hatcher and Dunn found that feminising symbionts may invade a host population even if they are deleterious to host fitness (i.e. parasites), so long as vertical transmission and feminisation are sufficiently high. As with Werren's (1987) model, Hatcher and Dunn found that a successful feminising symbiont could lead to host monogeny, producing only males in order to compensate for feminisation. An intermediate alternative that they discuss is that the feminising symbiont does not completely invade the population, and the host sex ratio is simply biased in favour of males; this can occur if feminisation is not perfect. This alternative did not arise out of Werren's model; Hatcher and Dunn suggest that it occurs since, if feminisation fails, infected offspring are subject to host sex ratios.

Hatcher et al. (2000) further considered how population structure affects a feminising parasite. Population structure refers to the organisation of a population in such a way that certain sets of individuals interact more with one another, or within themselves, than others. For example, Hatcher et al. (2000) consider a patch structure, so that the population is arranged into a grid of subpopulations between which individuals migrate at a certain stage of the life cycle. They use a stochastic individual-based model to analyse the same system Hatcher and Dunn (1995) previously discussed, which itself was an adaption of Werren's (1987) model; Hatcher et al. (2000) also include patch structure with dispersal each generation. The invading parasite feminises all hosts that inherit the parasite, which occurs according

to the vertical transmission parameter; uninfected host offspring are subject to the standard Fisherian sex ratio, 0.5. Distortion of patch sex ratio potentially causes extinction of hosts in the patch since males are assumed to only be able to mate with a limited number of females. However, recolonisation allows extinct patches to recover.

Hatcher and Dunn (1995) found that feminising parasites could not reach the frequency they did in unstructured populations, thus allowing the parasite to persist at intermediate frequencies; however, with too low or too high dispersal, the patches, or metapopulation, respectively, acted like single populations and both host and parasite were driven to extinction.

The model we present in Chapter 4 contains features that reflect those of models discussed here. Our focus is on how a transition can occurs from parasitism to mutualism, in the context or maternally-transmitted endosymbionts. Thus, we consider selection on vertical transmission, since as noted by Yamamura (1993), this is important in deciding whether there is a coincidence of fitness interests between the host and symbiont. We also adapt Werren's (1987) simple model of conflict between host and symbiont sex ratio; however, we significantly extend his analysis, considering population structure and evolution of vertical transmission, and focus on coincidence of fitness interests between hosts and symbionts. Our model also bears comparison to models of feminising parasites discussed by Hatcher and Dunn (1995) and Hatcher et al. (1999, 2000). Whereas they focus on compensatory host sex ratios (Hatcher and Dunn, 1995), and extend their model to focus on extinction of host populations and its dependence on male mating limit (Hatcher et al., 1999), and maintenance of parasites in patch-structured populations (Hatcher et al., 2000), our focus is on the effects of a patch-based population structure on the invasion of parasites and the development of that relationship towards, ultimately, mutualism. Specifically, the amount of sex-distortion determines where a symbiont is located on the parasite-commensalist-mutualist scale; by allowing the strategy to be subject to selection, we investigate coincidences in fitness interests and the transition towards mutualism.

Chapter 3

Modelling Between-Species Donation

Abstract

In this chapter, we analyse and extend previous models of between-species donation (Fletcher and Doebeli, 2009). Two primary versions of our basic model are analysed; in model (a), donation behaviours is determined by two loci, only one of which is subject to assortment, whereas in model (b), assortment is genome-wide.

A deterministic approach, utilising an assumption of quasi-linkage equilibrium (Kimura, 1965), is used to carry out an initial analysis. Stochastic simulations, which drops the assumption of weak selection while simultaneously introducing stochasticity, are used to test the robustness of these results.

We find that stable donation behaviour is possible under genome-wide assortment, but is vulnerable to suppressing modifiers arising on unassorted loci.

We question previous interpretation of the donation behaviour as between-species altruism, which was previously argued for by Fletcher and Doebeli (2009); we introduce, but do not test, an alternative possibility. We also contextualise our series of models, arguing that there are direct parallels to greenbeard behaviours, both within and between species.

3.1 Introduction

In Chapter 2, we discussed the importance of the question of whether or not altruism can evolve between species, specifically highlighting the relevance of this question to the major transitions view of evolution. Using a simple model (and a number of slight variations on this model) of between-species donation, we analyse this question, implementing and discussing deterministic and stochastic approaches. Ultimately, we conclude that between-species donation is transient due to the potential for suppressing modifiers to arise at unassorted loci; in addition, in the presence of assortment acting across all relevant loci, between-species donation may be interpretable as within-species altruism, whereby altruists use the other species as a vector for fitness benefits provided to genetic relatives. Similar arguments have been stated elsewhere (Foster, 2009; Bourke, 2011; Wyatt et al., 2013). We also discuss our models in relation to greenbeard traits, which offer some biological parallels to out model.

We use two approaches. Firstly, a deterministic approach, which invokes the assumption of quasi-linkage equilibrium; this approach involves numerically updating allele and genotype frequencies in an infinite population until an equilibrium is attained. Secondly, we consider a stochastic simulation, which drops the entailed assumption of weak selection whilst allowing randomness in mutation, breeding, and recombination; this approach involves a finite population of individuals, for whom fitness, breeding, mutation and recombination are individually calculated. We provide model details and results from the deterministic approach first (Sections 3.2, 3.3), before doing the same for the stochastic approach (Sections 3.4,3.5); finally, we discuss these results (Section 3.6).

3.2 Deterministic Approach: Specification

We consider a two-locus, biallelic model with two haploid populations of equal size. The two populations represent two different species; all interactions are interspecific, and individuals act as both potential donors and potential recipients. The first locus represents donation and non-donation and is referred to as the donation locus, and the second locus represents suppression and non-suppression, and is referred to as

the suppression locus. Individuals donate, and are labelled donators, if and only if they possess the donation and non-suppression alleles. All other individuals do not donate. For example, if an individual possesses the donation and suppression alleles, the individual will not be a donator since the suppression allele will act to suppress the donation behaviour. The types of individuals present in this model are summarised in Table 3.1, alongside genotype notation, which shall be introduced in due course.

Table 3.1: Genotypes and Phenotypes in the Between-Species Donation Model

Genotype	Donation or Non-Donation Allele?	Suppression or Non-Suppression Allele?	Phenotype
(0,0)	Non-Donation Allele	Suppression Allele	Non-Donator
(0,1)	Non-Donation Allele	Non-Suppression Allele	Non-Donator
(1,0)	Donation Allele	Suppression Allele	Non-Donator
(1,1)	Donation Allele	Non-Suppression Allele	Donator

Note our use of the term 'donator', as opposed to the more common term 'donor'; this is to distinguish between individuals engaged in an act of donation (donors) and individuals genetically predisposed towards donation, but whose predisposition may be disrupted by conditionality (Marshall, 2015). For clarity, individuals that only hold the donation allele shall be referred to as bearers of the donation allele. The cost associated with the donation behaviour is c, and the benefit conferred to the social partner is b; we constrain the two so that $0 < c < b$ and $c < w_0$, where w_0 is the baseline absolute fitness. The former is a standard constraint in any study of the donation game, as noted in Section 1.4.1; setting the cost so that it is strictly less than baseline fitness is necessary to avoid individuals being able to have negative fitness. We set w_0 to be 1; we lose no generality, since c and b are both scaled to w_0.

The life cycle of individuals within this model may be detailed as follows. Haploid individuals are paired up between species according to assortment. Fitnesses are calculated on the basis of whether or not an individual and their social partner are donators; each individual donates only once, to their social partner, if they are donators. A number of breeding partners are then determined at random within each species, with the probability of breeding directly proportional to fitness. Breeding partners then breed, with recombination taking place since offspring go through

a diploid stage in their life cycle; this is a simplification for modelling purposes with some biological and modelling precedent (De Massy et al., 1994; Zeng and Charlesworth, 2011). The original population of individuals then dies, replaced by their offspring.

Our deterministic approach, involving an assumption of 'quasi-linkage equilibrium' (or QLE; this is introduced and explained in Section 3.2.1), uses numerical methods to calculate changes in allele and genotype frequencies from one generation to the next, following Kirkpatrick et al. (2002). Each process is initiated with a set of initial allele and genotype frequencies, and terminated either once convergence has been detected, or manually if appropriate. The specifics of our model are all incorporated into 'selection coefficients', which measure how selection acts on a particular set of loci. Mean fitnesses of each genotype, which depend on assortment of each genotype and the associated phenotype, are used to find expressions for the selection coefficients.

We explore two primary versions of the basic model, which vary in terms of which individuals are more likely to be paired with one another; this is termed assortment, and more fully explained later in this section. In model (a), assortment (represented by the assortment parameter α) is applied to only the donation locus, whereas in model (b), it is applied to the whole genome. The main conclusions of this work come from models (a) and (b), though we briefly discuss four further models which address additional questions; we avoid precise details for the sake of brevity. In models (a) and (b), individuals act as both potential donors and potential recipients; that is, if an individual is a donator, they donate, losing c fitness, and if their social partner is a donator, they receive b fitness. Model (c) involves genome-wide assortment, but, in contrast to models (a) and (b), individuals are uniformly at random assigned a single role in their lifetime, of potential donor or potential recipient. If an individual is a potential donor, with probability 0.5, they cannot receive fitness benefits from a social partner, however, if they are a potential recipient, they cannot incur the fitness cost of donation.

Model (d) is a single-locus model, in which individuals are donators if and only if they possess the donation allele; therefore, assortment takes place on the single

remaining locus. Model (e) considers a single population with genome-wide assortment. Model (f) involves genome-wide assortment, but donators no longer donate unconditionally; instead they only donate if their social partner is also a donator.

We now discuss the details relevant to the construction of models (a) and (b) only, with brief notes on how these details extend to cover models (c-f); further notes on the construction of models (c-f) can be found in the Appendix, along with sample results relating to each of these additional models, and both deterministic and stochastic approaches.

First, we define some notation; the subscript i refers to the population index of the focal individual (or population), and the subscript i' refers to the index of the other population; note that necessarily, $i \neq i'$. We let the frequency of the donation allele in population i be $p_{1,i}$, and the frequency of the non-suppression allele in population i be $p_{2,i}$. The frequencies of the non-donation and suppression alleles are $q_{1,i} := 1 - p_{1,i}$ and $q_{2,i} := 1 - p_{2,i}$ respectively. The allelic value at the donation locus for an individual is 1 if the individual bears the donation allele, and 0 if they bear the non-donation allele. Similarly, the allelic value at the suppression locus is 1 if the individual bears the non-suppression allele, or 0 if the individual bears the suppression allele. Note that allelic values for each locus are within the set $\{0,1\}$ since each locus is biallelic. Thus, the set of genotypes is G=$\{(0,0),(0,1),(1,0),(1,1)\}$, where donators are (1,1) individuals. We denote the frequency of individuals of genotype $u \in G$ in population i by $f_i(u)$. An individual of arbitrary genotype u has allele u_1 on the donation locus, and allele u_2 on the suppression locus. Note that Table 3.1, displayed at the start of this section, clarifies how genotype notation matches phenotypes, and possession of alleles.

A major aspect of our model is assortment; we require a general method of pairing up individuals from each population, that can be extended to each of our models. We begin by constructing a function $P_i(u, v)$ which gives the probability that a focal individual with genotype $u \in G$ from population i will be matched up with a social partner with genotype $v \in G$ from population i'. We need to constrain function P so that it satisfies the following two conditions:

$$\sum_{v \in G} P_i(u, v) = 1 \quad \forall u \in G \tag{3.1}$$

$$\sum_{u \in G} f_i(u) P_i(u, v) = f_{i'}(v) \quad \forall v \in G \tag{3.2}$$

Condition (3.1) states that the sum of the matching probabilities for a genotype u is 1. Condition (3.2) states that the sum of the frequencies of each genotype multiplied by the probability that they are matched with a given genotype, is the frequency of the given genotype in the other population. This must be true, otherwise the pairings for each population do not correspond to one another.

We shall construct P in several stages. Firstly, we create a 'bias' function, which measures the overall frequency of pairings due to assortment between individuals of genotype u in population i with individuals of genotype v in population i'; this shall be denoted by $\beta_i(\alpha, u, v)$, but since the individuals subject to assortment vary between the different models, this shall be specified for each model. Assortment is non-negative, so $\beta_i(\alpha, u, v) \geq 0$ for all $i \in \{1, 2\}$ and $u, v \in G$. Alike individuals (i.e. those with the same alleles on whichever loci are subject to assortment) are paired up according to the bias function, which is in turn dependent on the assortment parameter α, before the remaining individuals are paired up uniformly at random.

Therefore, in the following models the expressions for unmatched genotype frequencies after assortment are as follows:

$$g_i(u) = f_i(u) - \sum_{v \in G} \beta_i(\alpha, u, v). \tag{3.3}$$

Necessarily, $\sum_{u \in G} g_1(u) = \sum_{v \in G} g_2(v)$, since the population sizes are assumed to be the same, as are the numbers of individuals paired by assortment. We only have a proportion $\sum_{u \in G} g_1(u)$ of individuals left to match up after assortment, so the frequency of pairings of u-types from population i, with v-types from population i', due to post-assortment matching, is $g_i(u) g_{i'}(v) / \sum_{u \in G} g_1(u)$. Thus, we reach the final expression for the P function:

$$P_i(u,v) = \frac{\beta_i(\alpha,u,v)}{f_i(u)} + \frac{g_i(u)g_{i'}(v)}{f_i(u)\sum_{t\in G} g_1(t)} \tag{3.4}$$

We can prove that this choice of P satisfies condition (3.1) for P to be valid as follows:df

$$\begin{aligned}
\sum_{v\in G} P_i(u,v) &= \frac{\sum_{v\in G}\beta_i(\alpha,u,v)}{f_i(u)} + \frac{g_i(u)}{f_i(u)} \\
&= \frac{f_i(u) - g_i(u)}{f_i(u)} + \frac{g_i(u)}{f_i(u)} \\
&= 1
\end{aligned}$$

The first step here sums equation (3.4) over all partner genotypes v, and notes that $\sum_{v\in G} g_{i'}(v) = \sum_{t\in G} g_1(t)$, thus cancelling out the two expressions where they occur in the second component. The second step subsititutes equation (3.3), and the final step is trivial. We can also prove condition (3.2) by substituting equation (3.4) into the left-hand side of condition (3.2):

$$\begin{aligned}
\sum_{u\in G} f_i(u)P_i(u,v) &= \sum_{u\in G}\left(\beta_i(\alpha,u,v) + \frac{g_i(u)g_{i'}(v)}{\sum_{t\in G} g_i(t)}\right) \\
&= \sum_{u\in G}\beta_i(\alpha,u,v) + g_{i'}(v) \\
&= f_{i'}(v) - g_{i'}(v) + g_{i'}(v) \\
&= f_{i'}(v)
\end{aligned}$$

We use similar techniques in this proof, once again noting that $\sum_{v\in G} g_{i'}(v) = \sum_{t\in G} g_1(t)$ to reach the second step, and introducing equation (3.3) to reach the third step; again, the final step is trivial.

We shall now construct the β functions for the two primary models, and discuss how it is constructed for the remaining four minor models. Note that the P functions discussed above are general; differences between each model are mostly subsumed into different constructions of the bias function, β.

Model (a): Single-locus Assortment

The construction of the β function in the case where only the donation locus is subject to assortment starts by noting that the maximum frequency of paired individuals with allele u_1 on the donation locus is:

$$\min\left(\sum_{t:t_1=u_1} f_i(t), \sum_{t:t_1=u_1} f_{i'}(t)\right) \tag{3.5}$$

This is because we cannot match more bearers of a given allele with one another than exist in either population. However, we need to further decompose this expression to find the maximum frequencies of pairs of individuals of given genotypes matched by assortment. The proportion of pairings by assortment between individuals of genotype u in population i, and individuals of genotype v in population i', of the total matchings by assortment for individuals with allele u_1 on the donation locus is:

$$\frac{f_i(u)f_{i'}(v)}{(\sum_{t:t_1=u_1} f_i(t))(\sum_{t:t_1=u_1} f_{i'}(t))} \tag{3.6}$$

By taking the product of expressions (3.5) and (3.6), we obtain the maximum frequency of pairings given focal and partner genotypes; in order to construct the β function, we weight this linearly by the assortment parameter α, which takes a value in the range [0,1], and an indicator function, which takes the value 1 if the focal individual and social partner have the same allele on the donation locus, and 0 otherwise:

$$\beta_i(\alpha, u, v) = \alpha I_{\{u_1=v_1\}} \frac{f_i(u)f_{i'}(v)}{\max\left(\sum_{t:t_1=u_1} f_i(t), \sum_{t:t_1=u_1} f_{i'}(t)\right)} \tag{3.7}$$

Model (b): Genome-Wide Assortment

In the case of genome-wide assortment, the β function is as follows:

$$\beta_i(\alpha, u, v) = \alpha I_{\{u=v\}} \min(f_i(u), f_{i'}(u)) \tag{3.8}$$

This requires slightly less construction than in model (a); we simply have to note that the maximum proportion of pairings by assortment for a genotype u is the minimum frequency of genotype u in either population. This is weighted once again by α, and the indicator function which only allows assortment between alike genotypes. This β function is then simply inserted into expressions (3.3) and (3.4) to obtain our P functions required for the deterministic simulation and mathematical analysis.

Further Models

We now discuss how assortment works in each of the four minor models (see Appendix for further details). Models (c) and (f) utilise the same β equation (3.8) as model (b), and the standard assortment equation (3.4), since they both involve genome-wide assortment between two species with two loci. Model (d) involves a single locus, so has a smaller set of genotypes. Model (e) involves a single population, so population indices are dropped from the assortment equations (3.3,3.4,3.8) and the fitness equations.

3.2.1 Quasi-Linkage Equilibrium

The model we consider in this chapter is multi-locus, thus the possibility of linkage and recombination arises. When offspring inherit genetic information from parents, they inherit a section of one parent's chromosome, with random crossover to the other parent's chromosome occurring with some probability; this is referred to as recombination. However, associations between alleles may arise, in such a way that they are inherited non-randomly, i.e. they are linked. The term 'linkage disequilibrium' describes this occurrence. The state of quasi-linkage equilibrium (QLE) was originally investigated by Kimura (1965). This states that under certain assumptions, linkage disequilibrium reduces to a low value which depends on allele frequencies and selection, but not previous values of linkage disequilibrium (Rice, 2004). The utility of QLE is that it allows us to neglect higher orders of selection coefficients and linkage disequilibrium, meaning that the mathematical recursions we consider in the deterministic simulation are significantly reduced (Kirkpatrick et al., 2002).

Specifically, the assumptions are of weak selection and weak linkage disequilibrium in comparison to selection; these are formulated mathematically as follows. Firstly, we must have that $a \ll 1$, where a is the greatest of the selection coefficients as defined by Kirkpatrick et al. (2002). Secondly, linkage disequilibrium must always be of the order a.

Assuming two loci, three selection coefficients are defined for each population i, which are $a_{1,i}$, $a_{2,i}$, and $a_{\{1,2\},i}$; we also denote $a := \max(a_{1,i}, a_{2,i}, a_{\{1,2\},i})$. Kirkpatrick constructed selection coefficients so that they have a high degree of generality, and represent selection acting on the relevant loci. They are defined according to the following equation (adapted from equation (7) in Kirkpatrick et al. (2002)):

$$W_i(z) = \bar{W}_i(1 + \sum_U a_{U,i}(Y_{U,i,z} - D_{U,i})) \tag{3.9}$$

Here, $\bar{W}_i$ is the mean fitness for population i, the possible sets of loci are $U := \{1, 2, \{1, 2\}\}$, and $Y_{U,i,z}$ is the product of deviations of the allelic value from the mean allelic value over the set of loci U in population i, for allelic values $z = (z_1, z_2)$, where z_1 is the allelic value on the donation locus, and z_2 is the allelic value on the non-suppression locus. For example, if $U = \{1, 2\}$, $i = 1$, and $z = (0, 1)$ (i.e. individuals of genotype z possess the non-donation and non-suppression alleles) then $Y_{U,i,z} = -p_{1,1}(1 - p_{2,1})$. The measure of linkage disequilibrium in population i will be labelled as $D_i := D_{\{1,2\},i}$, since $D_{1,i} = D_{2,i} = 0$ by definition. Consequently, to satisfy the second condition of QLE, we must simply show that D_i is of order a.

An example of the implementation of equation (3.9) would be as follows, where we evaluate the fitness of a donator (i.e. a (1,1) individual):

$$W_i(1, 1) = \bar{W}_i(1 + a_1(1 - p_1) + a_2(1 - p_2) + a_{\{1,2\}}(1 - p_1)(1 - p_2))$$

The first term in the parentheses corresponds to the average fitness, the second corresponds to the adjustment for possessing the donation allele, the third corresponds to the adjustment for possessing the non-donation-suppression allele, and

the final term corresponds to the adjustment from possessing both alleles; these are all weighted by the mean population fitness. Since every single model involves either different fitness equations, or different methods of assortment, each model must be considered individually. As before, we discuss the finer details of the first two models, and briefly mention the remaining four models.

Model (a): Single-locus Assortment

We begin by noting that in order to calculate selection coefficients, we must have equations for the average fitness of individuals of each genotype. Given some baseline fitness w_0, we find that the average fitnesses of individuals in the ith population are as follows:

$$W_i(0,0) = w_0 + P_i((0,0),(1,1))b \tag{3.10}$$

$$W_i(0,1) = w_0 + P_i((0,1),(1,1))b \tag{3.11}$$

$$W_i(1,0) = w_0 + P_i((1,0),(1,1))b \tag{3.12}$$

$$W_i(1,1) = w_0 + P_i((1,1),(1,1))b - c \tag{3.13}$$

Consequently, using equations (3.3,4,7,9,10-13), we find that the selection coefficients for the ith species in this case are as follows (where we have set $\tilde{c} = c/w_0$ and $\tilde{b} = b/w_0$):

$$a_{1,i} = \frac{\tilde{b}(P_i((1,0),(1,1)) - P_i((0,0),(1,1))) - p_{2,i}\tilde{c}}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} \tag{3.14}$$

$$a_{2,i} = \frac{-p_{1,i}\tilde{c}}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} \tag{3.15}$$

$$a_{\{1,2\},i} = \frac{-\tilde{c}}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} \tag{3.16}$$

The maximum absolute value of the numerator of equation (3.14) is $\tilde{b}$, since, using equations (3.4) and (3.7), we can show that $1 \geq (P_i((1,0),(1,1)) - P_i((0,0),(1,1))) \geq 0$. This is the highest value any of the numerators can take. The minimum absolute value of each of the denominators is $1 - \tilde{c}$, since $1 \geq f_i(u) \geq 0$. Since $a := \max(a_{1,i}, a_{2,i}, a_{\{1,2\},i})$, and $a_{1,i}$ can take the maximum value of equations (3.14-16), we find the following expression to be true:

$$a < \frac{\tilde{b}}{1-\tilde{c}} \tag{3.17}$$

Thus, in order to satisfy the first assumption of QLE, that $a \ll 1$, it is sufficient to choose b and c such that $\tilde{b}/1-\tilde{c} \ll 1$, or equivalently:

$$\tilde{b} \ll 1-\tilde{c} \tag{3.18}$$

This shall be satisfied by our parameter choices.

We now turn to the second assumption of QLE, that D_i is of order a. The maximum absolute value linkage disequilibrium can take is 0.25 (Lewontin, 1964), and, since we have already found that $a \leq \tilde{b}/1-\tilde{c}$, we can deduce that:

$$D_i < \frac{4a(1-\tilde{c})}{\tilde{b}}$$

Here, D_i is of order a. Thus, we have satisfied the second requirement of QLE. By combining the two requirements for the population to be in a state of QLE, we can ignore higher powers of D_i and a, which lets us reduce the original equations specified in (Kirkpatrick et al., 2002) for the associations D_i. We obtain the following approximation for D_i, where x_i is the rate of recombination between the two loci in the ith species:

$$D_i = \frac{a_{\{1,2\},i} p_{1i} p_{2i} q_{1i} q_{2i}}{x_i} \tag{3.19}$$

Now that the conditions for QLE have been satisfied, we can take a closer look at the selection coefficients themselves. From equation (3.9), we can see that an individual with genotype $z = (z_1, z_2)$ will gain the following fitness if they possess the non-suppression allele ($z_2 = 1$) instead of the suppression allele ($z_2 = 0$):

$$\begin{aligned} W_i(z_1,1) - W_i(z_1,0) &= \bar{W}_i(a_{2,i} + a_{\{1,2\},i}(z_1 - p_{1,i})) \\ &= \bar{W}_i(p_{1,i}a_{\{1,2\},i} + a_{\{1,2\},i}(z_1 - p_{1,i})) \\ &= \bar{W}_i a_{\{1,2\},i} z_1 \end{aligned}$$

Here, we have noted in obtaining the second term that, from equations (3.15) and (3.16), $a_{2,i} = p_{1,i}a_{\{1,2\},i}$. The third term is reached through simple cancelling out. Thus, we obtain a condition for bearers of the non-suppression allele to have strictly greater fitness than bearers of the suppression allele, given possession of allele z_1 on the donation locus:

$$\bar{W}_i a_{\{1,2\},i} z_1 > 0 \tag{3.20}$$

Since $\bar{W}_i$ is strictly positive, z_1 is non-negative, and by examination of equation (3.16), we can see that $a_{\{1,2\},i}$ is strictly negative (since the numerator is strictly positive, and the denominator strictly negative), this term is non-positive. Therefore, the non-suppression allele will always receive either neutral or negative selection.

We can carry out a similar analysis on the donation locus. Consequently, using equations (3.9), (3.14), and (3.16), we can see that an individual with genotype $z = (z_1, z_2)$ will gain the following fitness if they possess the donation allele ($z_1 = 1$) instead of the non-donation allele ($z_1 = 0$):

$$\begin{aligned} W_i(1,z_2) - W_i(0,z_2) &= \bar{W}_i(a_{1,i} + a_{\{1,2\},i}(z_2 - p_2)) \\ &= \frac{\tilde{b}(P_i((1,0),(1,1)) - P_i((0,0),(1,1))) - z_2\tilde{c}}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} \end{aligned}$$

In other words, we can see that bearers of the donation allele will have strictly greater fitness than bearers of the non-donation allele if the following condition is satisfied:

$$P_i((1,0),(1,1)) - P_i((0,0),(1,1)) > \frac{z_2 c}{b} \tag{3.21}$$

Thus, when an individual possesses the suppression allele (i.e. $z_2 = 0$), we can see that, since the left-hand side is greater than or equal to 0, they have either greater or equal fitness when they possess the donation allele. If an individual possesses the suppression allele, they will never donate; the only difference the donation allele will make is an increase in the chance of being paired with a donator, in the event that there are donators present in the other population. If not, then the donation allele will make no difference. In behavioural terms, individuals that possess the donation and suppression allele may be considered cheaters under this form of assortment; they are assorted as if they are donators, but suppress this behaviour, avoiding the costs of donating themselves.

When an individual possesses the non-suppression allele, (i.e. $z_2 = 1$), then the condition (3.21) is sometimes satisfied, and sometimes not satisfied. This is dependent on the c/b ratio, α, and genotype frequencies within the two population. Unfortunately, the condition is fairly intractable, but we can conclude that the evolution of donation is aided by a low c/b ratio.

Model (b): Genome-Wide Assortment

The fitness equations (3.10-13) utilised in the single-locus assortment apply to the genome-wide assortment case, since the only difference between the two models is expressed in terms of the β function (3.8). Thus, using equations (3.3,4,7,8,10-13), we find the following selection coefficients:

$$a_{1,i} = \frac{\tilde{b}(P_i((1,0),(1,1)) - P_i((0,0),(1,1)))}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} + p_{2,i}a_{\{1,2\},i} \tag{3.22}$$

$$a_{2,i} = \frac{\tilde{b}(P_i((0,1),(1,1)) - P_i((0,0),(1,1)))}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} + p_{1,i}a_{\{1,2\},i} \tag{3.23}$$

$$a_{\{1,2\},i} = \frac{\tilde{b}\big(P_i((1,1),(1,1)) + P_i((0,0),(1,1))\big)}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} + \frac{\tilde{b}\big(-P_i((1,0),(1,1)) - P_i((0,1),(1,1))\big) - \tilde{c}}{1 + \tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} \tag{3.24}$$

We must again check that the two assumptions of QLE are satisfied. Firstly, we note that $\max|a_{1,i}| < \max|a_{\{1,2\},i}|$ and $\max|a_{2,i}| < \max|a_{\{1,2\},i}|$. By noting that each P_i expression falls between 0 and 1, we find:

$$|a_{\{1,2\},i}| \leq \frac{2\tilde{b} - \tilde{c}}{1 - \tilde{c}} \tag{3.25}$$

Thus, the first assumption of QLE, that $a \ll 1$, is satisfied if $\tilde{b} \ll 1/2$. As before, we shall satisfy this condition through our choices of paramater values in the modelling section.

We now turn to the second assumption of QLE; that is, that all values of D_i are of order a. Since $a < 2\tilde{b} - \tilde{c}/1 - \tilde{c}$, and necessarily, $D_i \leq 1/4$, we can see that:

$$D_i < \frac{4a(1-\tilde{c})}{2\tilde{b} - \tilde{c}}$$

We can once again find conditions for positive selection on the donation and non-suppression alleles, by using equations (3.9) and (3.22-24). Both conditions involve the donation or non-suppression alleles receiving positive selection if some function of genotype frequencies and α is greater than c/b. This underlines the importance of the c/b ratio, and the unimportance of the baseline fitness w_0. Unfortunately, the conditions are so mathematically intractable that they do not allow us to make any meaningful observations about the genotype frequencies or α. However, we can note that in the special case in which allele and genotype frequencies are equal between populations, then $\alpha > c/b$ is the condition for the donation and non-suppression

alleles to positive selection. This is analogous to Hamilton's rule for the spread of a social behaviour within species.

Further Models

The same techniques shown above for calculating selection coefficients and satisfying the assumptions of QLE can be followed for each of the models (c-f) (see Appendix for further details and sample numerical examples); thus, for brevity the exact calculations and selection coefficients are omitted. We now discuss the conclusions we can draw from each set of selection coefficients.

Under model (c), in which each individual is randomly assigned the role of either potential donor or potential recipient, fitness equations are altered from those in model (b), since the model entails a weight on all fitness outcomes due to social interactions of 1/2. This results in different selection coefficients. By scaling $\tilde{b}$ and $\tilde{c}$ by 2, we obtain the same selection coefficients, and the same conclusions. Essentially, models (b) and (c) are equivalent, except that model (c) involves one social interaction per individual per generation rather than two, thus halving the selection strength.

Under model (d), only a single selection coefficient is calculated, since there is only a single locus. As with model (b), we reach the conclusion that donation will evolve if the allele frequencies in each population are the same, and $\alpha > c/b$; again, this is analogous to Hamilton's rule. However, when allele frequencies are unequal between populations, the selection coefficient is mathematically intractable, and no further conclusions can be drawn.

Model (e) involves genome-wide assortment within a single population. This is mathematically equivalent to the scenario briefly mentioned at the end of Section 3.2, in which we noted that donation evolves to fixation if allele and genotype frequencies are equal in both populations. Here, individuals are paired up within the same population, so frequencies match up as they did in the between-species scenario, thus we reach the same conclusion that donation will reach fixation when $\alpha > c/b$.

Model (f), in which donators to refuse to donate to non-donators, fitness equations reflect the fact that social interactions only occur between donators; when donators are paired, they each gain $b - c$ fitness. This results in three strictly positive selection coefficients, implying that both the donation allele and non-suppression allele always receive positive selection. Thus, donation always reaches fixation.

3.2.2 Implementing the Deterministic Simulation

We now discuss the implementation of the deterministic analysis, utilising the simplifying equations than QLE permits; note that all of the code used in this chapter is located online (Quickfall, 2016), and runs in R (R Core Team, 2016).

We start with some initial allele frequencies, $p_{1,1}$, $p_{1,2}$, $p_{2,1}$, and $p_{2,2}$, which define the other allele frequencies $q_{1,1} = 1 - p_{1,1}$, $q_{1,2} = 1 - p_{1,2}$, $q_{2,1} = 1 - p_{2,1}$ and $q_{2,2} = 1 - p_{2,2}$. We are interested in whether donators can invade a population when donation and suppression alleles are initially rare, so we choose $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$. We have examined more extreme initial frequencies (e.g. 0.01 initial donation and suppression), which always resulted in the same endpoints. However, we present results from the more moderate scenarios as they are more easily interpretable visually, and are computed more quickly.

The parameters of the system that we choose in every numerical example are the recombination rate, x (assumed to be the same in both populations, hence the dropping of the subscript), the baseline fitness w_0, and the benefits and costs associated with donation, respectively, b and c. We vary these initial conditions and parameters to investigate the system; aspects of the sensitivity analysis which result in no change to the outcome of a given experiment are mentioned, but not shown.

In every iteration, the selection coefficients $a_{1,i}, a_{2,i}$ and $a_{\{1,2\},i}$ are calculated. Selection coefficients, and allele and genotype frequencies are then used to update allele frequencies in the next iteration, using the following equations, taken from equation (10) in Kirkpatrick et al. (2002):

$$\Delta p_{1,i} = a_{1,i} p_{1,i} q_{1,i} + a_{2,i} D_i + a_{12,i}(1 - 2p_{2,i}) D_i \tag{3.26}$$

$$\Delta p_{2,i} = a_{2,i} p_{2,i} q_{2,i} + a_{1,i} D_i + a_{12,i}(1 - 2p_{1,i}) D_i \tag{3.27}$$

Finally, the new values of D_i need to be calculated at the end of every generation; these take into account the effects of selection in the previous generation, and the new allele frequencies. The following expression finds the new values of D_i, derived using equation (19) in Kirkpatrick et al. (2002):

$$D_i = \frac{a_{12,i} p_{1,i} p_{2,i} q_{1,i} q_{2,i}}{x_i} \tag{3.28}$$

The process is terminated once the sum of the absolute changes in genotype frequencies between two successive generations falls below a certain threshold, or alternatively after a specified number of generations in order to aid visual comparison of different results.

In the following section, we discuss results of the deterministic simulation for the models (a) and (b). We omit discussion of models (c-f) (though see Appendix for further details), since our findings simply confirm the results of the analysis of selection coefficients that were discussed in Section 3.2.1.

3.3 Deterministic Approach: Results

3.3.1 Model (a): Single-locus Assortment

In Section 3.2.1, an analysis of the selection coefficients revealed that the non-suppression allele will never receive positive selection, and that a lower c/b ratio favours the donation allele. We shall now provide numerical examples of these conclusions by using the deterministic simulation formulated in Section 3.2.2.

We choose initial parameter values of $x = 0.05, c = 0.005, b = 0.01$, and $w_0 = 1$. These are valid parameters given the constraints found in section 3.2.1; preliminary testing shows that the scale of c and b only affects the number of generations until convergence, thus we choose these values so that they are small enough to satisfy the condition of weak selection. Our first analysis focuses on the effect of α, thus we do not specify a default value of α initially.

We choose initial allele frequencies of $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$. This

means that both populations contain an initially low frequency of the donation and non-suppression alleles, and, consequently, a low initial frequency of donators. We choose these parameters as they reflect the most biologically interesting question, of whether donation can invade, and if so, whether suppression can then invade in turn. Since this simulation is deterministic, initial allele frequencies are the same in each population (as are parameters), and by symmetry of the simulation, allele and genotype frequencies will remain the same in each population over the course of the simulation. Thus, we display only one set of allele and genotype frequencies for each numerical example in which we do not deviate from these default initial allele frequencies.

Our parameter choices and initial allele frequencies are informed by a preliminary sensitivity analysis. The results are omitted for brevity, but we note some of the conclusions. Firstly, the amount of recombination makes no difference to the evolutionary outcomes or dynamics; this is a consequence of very little linkage disequilibrium arising out of the model. Secondly, we tried different amounts of selection strength (i.e. scaling w_0, or alternatively, c and b), while retaining the same c/b ratio. Under weak selection, an increase in computing time was the only consequence; evolutionary outcomes and dynamics remain the same. Of course, both c and b must take low values to satisfy condition (3.18), so the values of 0.005 and 0.01 respectively reflect a compromise between satisfying the QLE assumptions and a desire for swift computation.

We also note that initial allele frequencies make no substantive changes to the outcome of each analysis; the frequency of donators at equilibrium is always the same, even if the dynamics are different. This also holds when initial allele frequencies vary between populations.

We now discuss the effects of the remaining parameters, by turn focusing on assortment α and the c/b ratio.

The Assortment Parameter α

Firstly, we consider the sensitivity of our model to the assortment parameter α; results are displayed in Figure 3.1. Analytically, we were unable to quantify the

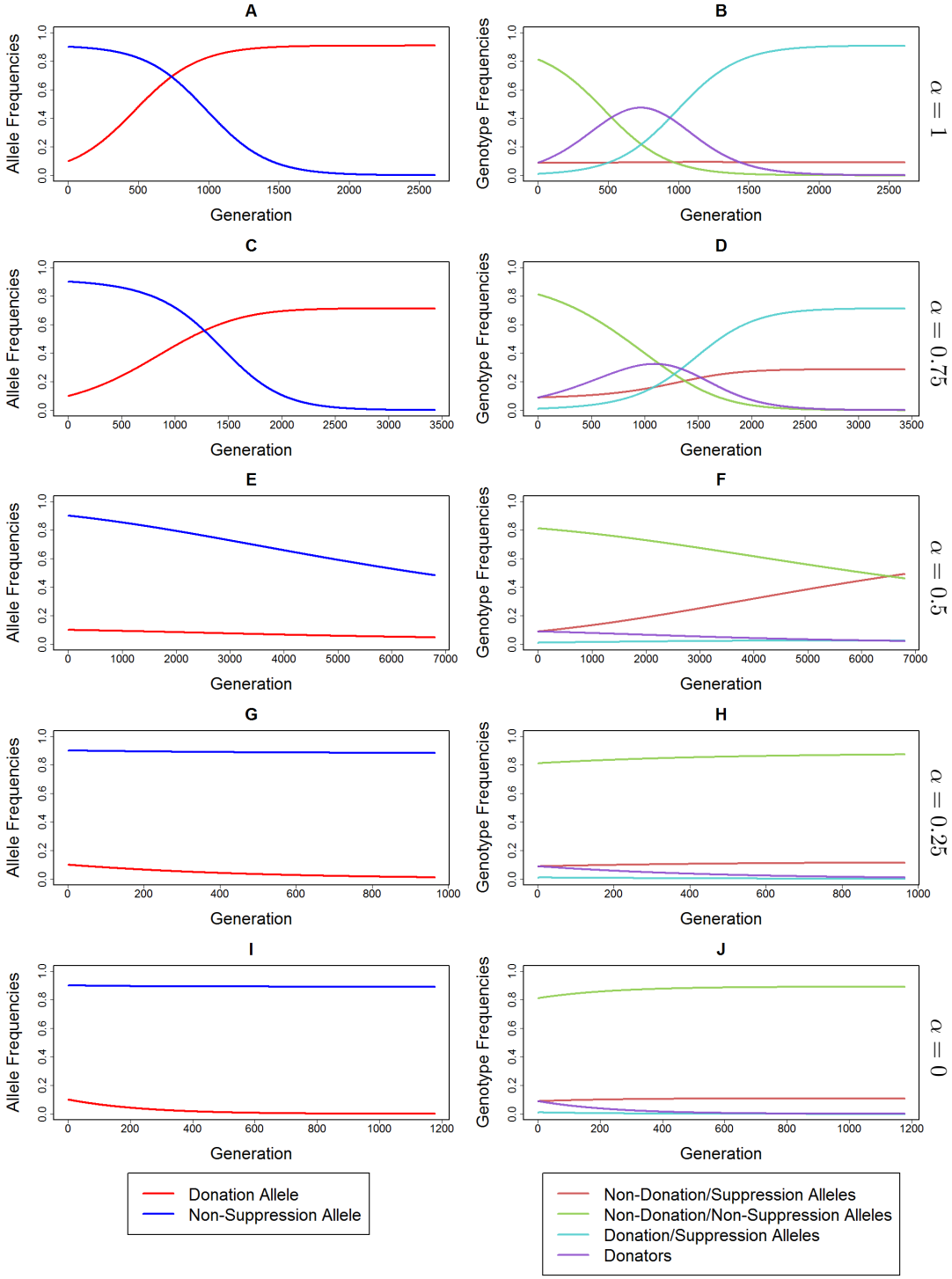

Figure 3.1: Evolution of allele frequencies (A/C/E/G/I), genotype frequencies (B/D/F/H/J) (using the deterministic simulation with donation assortment); initial frequencies are $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1$. In A/B, C/D, E/F, G/H and I/J, α is, respectively, 1, 0.75, 0.5, 0.25 and 0. Non-donation and suppression allele frequencies are not displayed as they can be trivially derived from, respectively, the displayed donation and non-suppression allele frequencies. The greater the level of assortment, the higher the frequency of the donation allele at convergence; donation remains transient.

effects of α due to the intractable nature of our condition for positive selection on the donation allele; however, we discovered that the suppression allele always receives positive selection. The latter result is corroborated by our results, though we can note that the donation allele is more favoured by selection initially as α increases. These observations reflect a verbal argument as follows; the donation allele receives positive selection initially when α is high, since donators are more likely to be paired with other donators. Since there is no assortment on the suppression locus, there is no penalty to bearers of the donation allele suppressing the donation behaviour; they still receive the fitness benefits associated with assortment, though they do not donate themselves, and thus do not lose c fitness. As suppression spreads, fewer individuals are donators, so the benefit of bearing the donation allele and thus being more likely to be paired with other bearers of the donation allele decreases. Meanwhile, the suppression allele reaches fixation, at which point the allele on the donation locus is irrelevant since no donation occurs, so the donation allele receives neutral selection. In summary, the donation behaviour is costly when assortment is independent of alleles on the suppression locus; indeed, in each numerical example displayed in Figure 3.1, we can see that it is always transient. The effect of an increase in α appears to be to increase selection for the donation allele, which in turn increases selection for suppression; the overall effect is that donation reaches a higher frequency, but remains transient, since the suppression allele always reached fixation.

The c/b Ratio

Next, we look at the effect of c/b. In Figure 3.2, we provide numerical examples for three different values of c/b. In every case, donation is transient once again, as predicted. It appears that a decrease in c/b promotes the evolution of the donation allele, thus having a similar effect to an increase in α.

This makes sense in light of the findings of Section 3.2.1. In equation (3.21), we presented a condition for selection for the donation allele; this stated that if some function of allele and genotype frequencies and α exceeded a linear function of c/b, then the donation allele would receive positive selection.

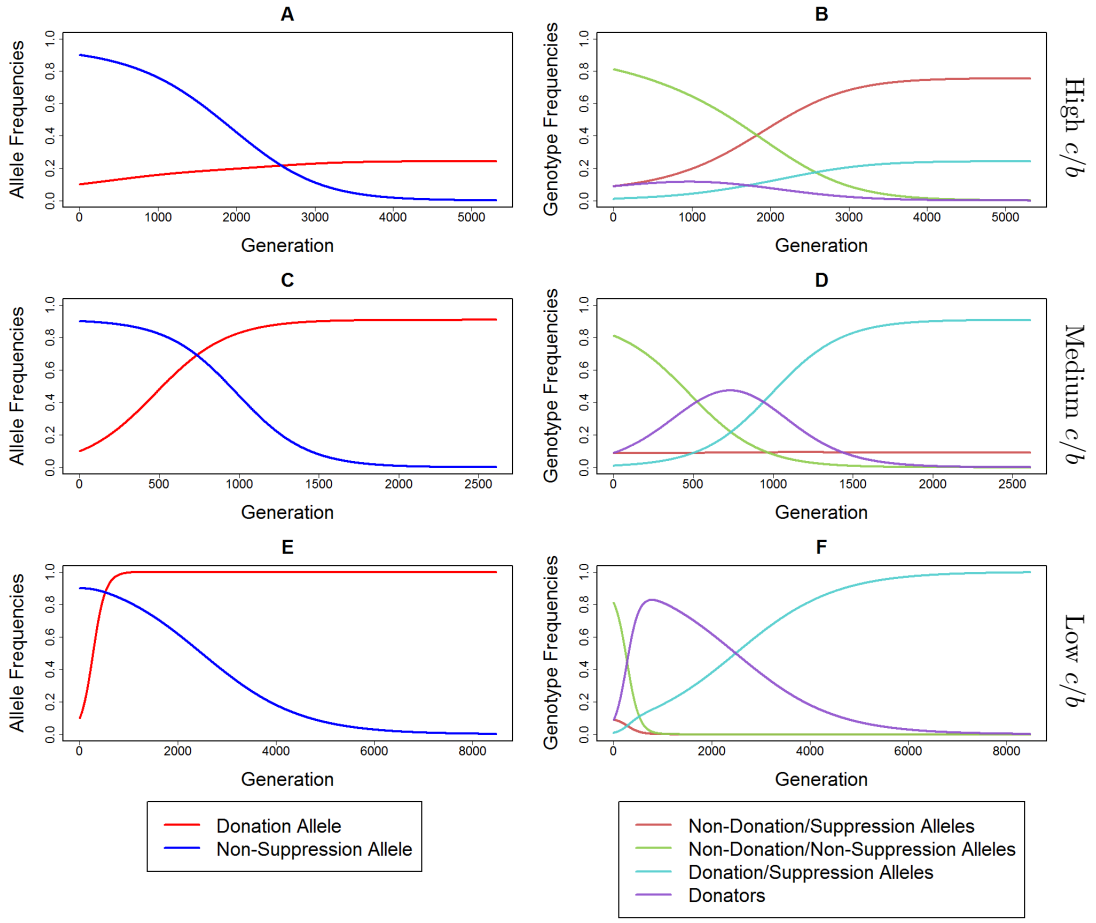

Figure 3.2: Evolution of allele frequencies (A/C) and genotype frequencies (B/D) (using the deterministic simulation with donation assortment); initial frequencies are $p_{1,1} = p_{1,2} = 0.9, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, w_0 = 1, \alpha = 1$. In A/B, C/D, and E/F, c is, respectively, 0.009, 0.005 and 0.001. A smaller c/b ratio causes the transient frequencies of donators to be higher.

Conclusion

We have carried out a sensitivity analysis of this model, in particular highlighting the effects of α and c/b. However, even when both of these parameters are set in such a way that the donation allele receives significant positive selection, we can find no case in which the donation behaviour is anything other than transient, since suppression always receives positive selection. This was also shown by the mathematical analysis discussed in Section 3.2.1.

3.3.2 Model (b): Genome-Wide Assortment

At the end of Section 3.2.1, which discussed our mathematical analysis of this model, we concluded that equal allele and genotype frequencies between populations meant that the donation and non-suppression alleles were tied together in terms of the kind of selection they received. We discovered that both alleles would receive positive selection if $\alpha > c/b$; however, we could draw no conclusions when allele and genotype frequencies were different between populations.

Therefore, in this analysis, we choose the following initial allele frequencies: $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901$, and $p_{2,2} = 0.9$. The other parameters remain the default parameters chosen for model (a), specified in Section 3.3.1. Figure 3.3 depicts results of this approach; the striking result is that frequencies of donation are cyclic and intermediate in both populations.

We can see from Figure 3.3 that in this case, the non-suppression allele reaches fixation in the first population, and the donation allele reaches fixation in the second population. Thus, whether individuals are donators or not is entirely a result of the frequency of the donation allele in the first population, and the non-suppression allele in the second. The frequencies of donators cycle. They increase at similar rates and frequencies, before the frequency of donators in the population with fewer donators stops increasing and starts to fall. After a brief period, the frequency of donators in the other population does the same. They decrease at a similar rate, but at different frequencies. The population with fewer donators begins to increase again first, leading to the period of increase in donator frequencies in both populations.

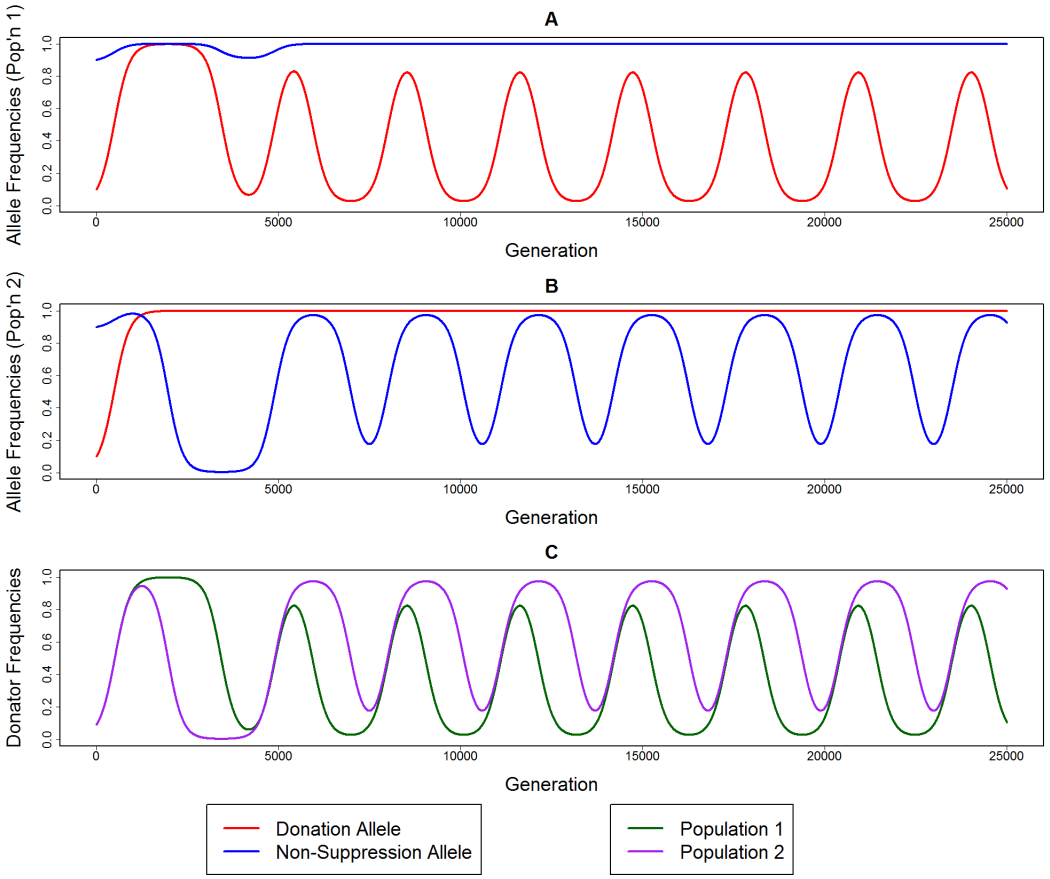

Figure 3.3: Evolution of allele frequencies (A/B), and donator frequencies (C) (using the deterministic simulation with genome-wide assortment); initial frequencies are $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1, \alpha = 1$. Maximal assortment leads to cyclic donator frequencies and no stable equilibria.

We can derive a condition for donators to have greater fitness, which can be used to explain this cyclicity in more mathematical terms. We start by expressing the fitness of donators compared to the average fitness of non-donators. In the following, the phenotype of donators is denoted by a 1, and the phenotype of non-donators is denoted by a 0. Note that in the first population, the only non-donators present are (0,1) genotypes, and in the second population, the only non-donators present are (1,0) genotypes. We denote the probability of an individual of phenotype s in population i being paired with an individual of phenotype t by $P'_i(s,t)$. These terms could be re-formulated using the P functions found earlier, though this form is intuitive for our current use. We derive the following condition for donators to have greater fitness in population i:

$$bP'_i(1,1) - c > bP'_i(0,1) \tag{3.29}$$

Assume arbitrarily that population 1 has a lower frequency of donation. Then, $P_1'(1,1) = 1$, and $P_1'(0,1) > 0$, since donators will always be paired with donators, as $\alpha = 1$, and the unmatched donators from the other population will be paired with non-donators. However, in population 2, $P_2'(1,1) < 1$, and $P_2'(0,1) = 0$.

$P_1'(0,1)$ and $P_2'(1,1)$ are determined by the ratio of unmatched donators in population 2 to non-donators in population 1. When this ratio is high, which occurs when there is a large difference in donator frequencies between the two populations, or when there are few non-donators in population 1, then $P_1'(0,1)$ is relatively high, and $P_2'(1,1)$ relatively low, so non-donators are more likely to have greater fitness. The converse is true; donators have higher fitness when there is a small difference in donator frequencies, and when donators themselves are at low frequencies. Each of these properties is reflected in the dynamics of donator frequencies displayed in part (c) of figure 3.3.

Further investigation of the c/b parameter under perfect assortment showed that a decrease in c/b increases the period of the cyclicity; we do not display these relatively uninformative results here. However, since we have assumed perfect assortment, an unrealistic condition which led to cyclicity due to the frequencies of certain non-donator genotypes, we now investigate imperfect assortment. Figure 4 shows a comparison of the dynamics of donation frequencies for several different values of α; the remaining parameters are once again unchanged. Allele and genotype frequencies are not presented, since they follow the same pattern as in figure 3; the donation allele goes to fixation in one population, while the non-suppression allele goes to fixation in the other.

Figure 3.4 shows that intermediate frequencies of α can lead to stable intermediate frequencies of donation. Three types of equilibrium are shown here; with the lowest level of assortment, $\alpha = 0.4$, there are no donators at equilibrium. When $\alpha = 0.6$, the same frequency of donators is present in each population at equilibrium. Finally, when $\alpha = 0.8$, stable frequencies of donators are even higher; however, these are different between the two populations.

A more complete analysis is required to pick out a general trend; the triangle

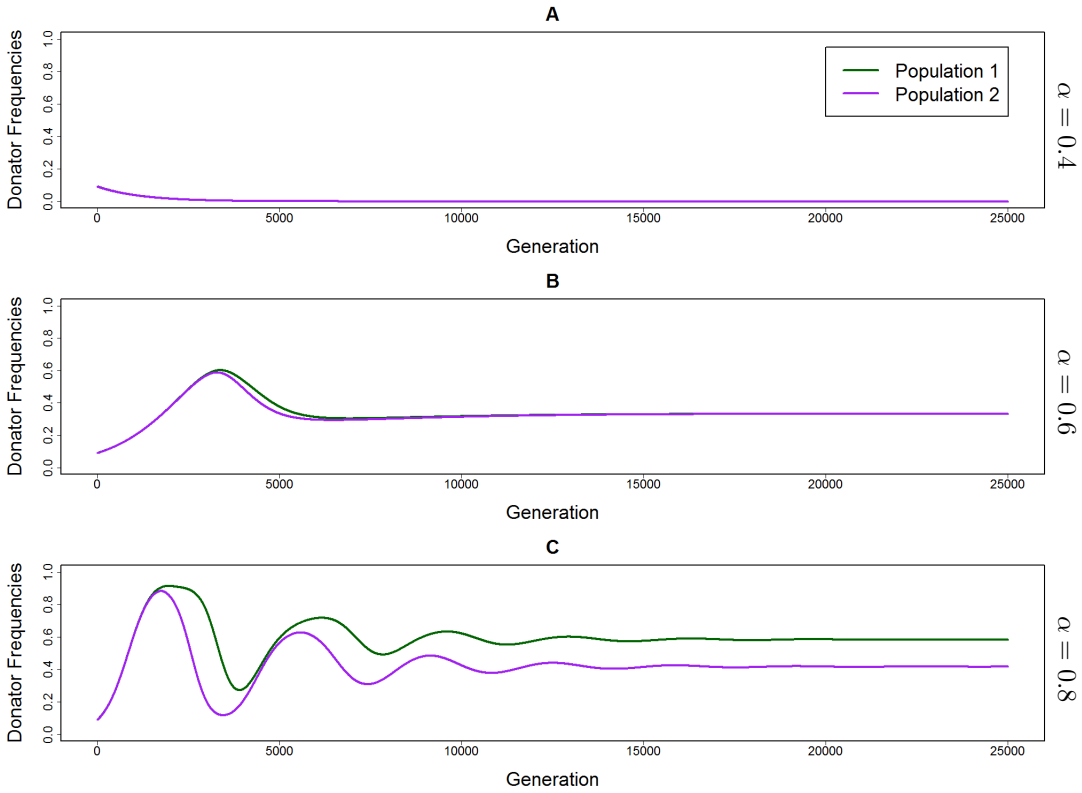

Figure 3.4: Evolution of donation (using the deterministic simulation with genome-wide assortment); initial frequencies are $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1, \alpha = 0.4, 0.6$ and 0.8 in A, B, and C respectively. Intermediate values of α result in stable equilibria with intermediate donator frequencies. Note that in 3.4A and B, donator frequencies are equal in the two populations at equilibrium.

points in Figure 3.5 show the stable frequencies of donation over the entire parameter space of α, except for the cases where there is no assortment ($\alpha = 0$) or maximal assortment ($\alpha = 1$).

We can see that the three distinct types of equilibrium illustrated in Figure 3.4 are shown here. Once again, further numerical examples are omitted, but confirm that the donation allele is in fixation in one population, and the non-suppression allele in fixation in the other, at every equiliibrium involving stable frequencies of donation. We also note that these equilibria are not sensitive to changes in the initial allele and genotype frequencies, apart from the special case where they are equal between populations; here, this model reduces to the within-species model (e).

Observations about the types of equilibria can be used to find analytic versions of these equilibria. We first assume that $p_{1,1} = p_{2,2} = 1$; that is, the donation allele is in fixation in population 1, and the non-suppression allele is in fixation in population 2. A consequence of this assumption is that $p_{2,1} = f_1(1, 1)$ and $p_{1,2} = f_2(1, 1)$. By

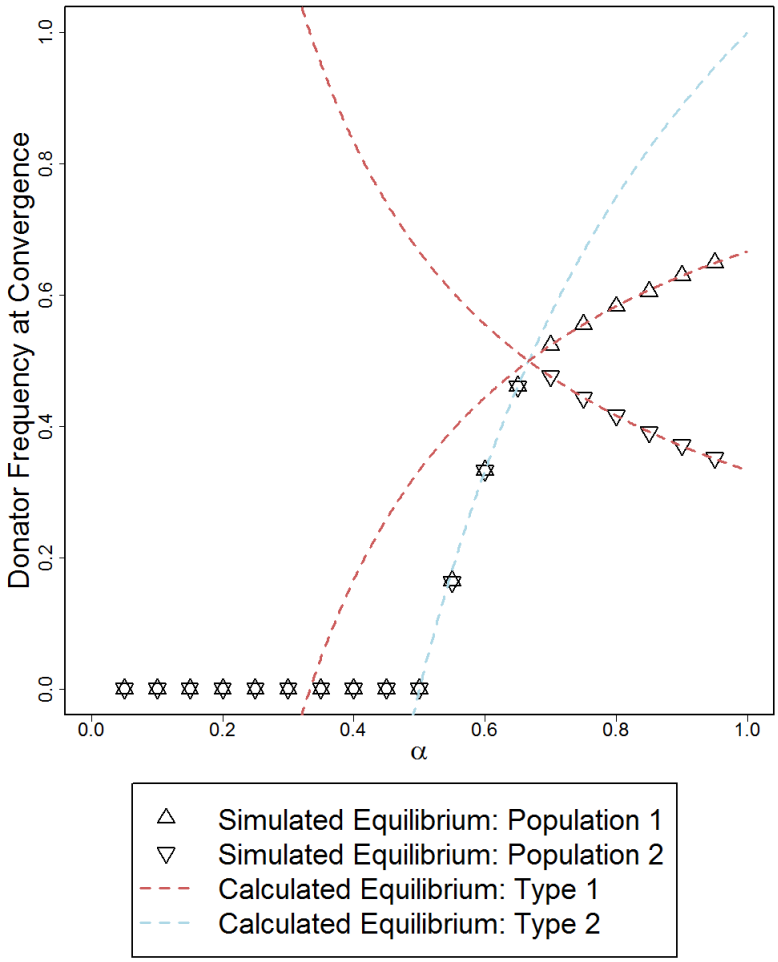

Figure 3.5: Stable equilibria of donation (using the deterministic simulation with genome-wide assortment) are shown through triangles; initial frequencies are $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005$, and $w_0 = 1$. Stable donation is present with sufficient α. Non-zero equilibria calculated and presented in equations (3.30-32) are presented as light red and blue dashed lines.

introducing these assumptions into equations (3.3), (3.4), (3.8), and (3.29), we can find analytic expressions for each equilibrium involving positive donator frequencies:

$$p_{2,1} = p_{1,2} = \frac{\alpha - c/b}{\alpha(1 - c/b)} \tag{3.30}$$

$$\text{or} \quad \text{(a)} \quad p_{2,1} = \frac{c/b}{\alpha(1 + c/b)}, \text{(b)} \;\; p_{1,2} = \frac{\alpha(1 + c/b) - c/b}{\alpha(1 + c/b)} \tag{3.31}$$

$$\text{or} \quad \text{(a)} \quad p_{1,2} = \frac{c/b}{\alpha(1 + c/b)}, \text{(b)} \;\; p_{2,1} = \frac{\alpha(1 + c/b) - c/b}{\alpha(1 + c/b)} \tag{3.32}$$

These equilibria have been plotted in Figure 3.5; the blue dotted line shows the equilibria calculated by equation (3.30), while the red dotted lines show the equilibria represented by equations (3.31a,b) and (3.32a,b).

For the equilibrium represented by equation (3.30), $\alpha > c/b$ is necessary for there to be strictly positive frequencies of donation. However, when $\alpha > (c/b)/(1+c/b)$, we find that the equilibria represented by equations (3.31a,b) and (3.32a,b) are stable, and (3.30) is unstable. The converse is true: when $\alpha < (c/b)/(1 + c/b)$, we find that the equilibria represented by equations (3.31a,b) and (3.32a,b) are unstable, and (3.30) is stable. Equations (3.31a,b) and (3.32a,b) represent the case where donation is more represented in one population than the other. This may seem initially counterintuitive, but in fact it is a result of individuals' fitnesses only being compared to the fitnesses of others individuals within the same population. In this model, a population can be at equilibrium even if its members have on average less fitness than individuals in the other population.

Conclusion

Overall, we conclude that in the deterministic case, donation can evolve to some intermediate level when there is genome-wide assortment. It always reaches an evolutionarily stable intermediate level when $\alpha \in (c/b, 1)$. However, this raises the question of what type of behaviour this is - perhaps we have found a case of stable between-species altruism? We shall return to this question in the discussion.

3.4 Stochastic Approach: Specification

We now turn to the stochastic approach. We consider this approach for two reasons; firstly, to drop the assumptions of QLE, which entail determinism, infinite populations, and weak selection, and secondly, to either confirm the results of the QLE analysis, or understand what causes any differences. Thus, we are able to extend parameter values beyond those previously tested, and check that previous results are robust to the joint introduction of both stochasticity and strong selection. Note that, once again, all of the code used here is located online (Quickfall, 2016), and runs in R (R Core Team, 2016).

While it would perhaps be preferable to introduce these two changes to our modelling approach individually, the forces of stochasticity and selection can be manipulated through parameter choices, allowing us to study the effects of each change in relative isolation. Thus, we give active consideration to the relative forces of

stochasticity and selection in the following section.

As shall become clear, it is difficult to consider very low amounts of stochasticity under this approach, since this requires large population sizes, which entail long computation times. If, on the other hand, the relative force of stochasticity is large in relation to that of selection strength, then due to mutation on biallelic loci, stochasticity results in negatively frequency-dependent selection of all alleles, which outweighs the selection strength. The results we display aim for a balance between two competing desires; one to study strong selection in relative isolation (i.e. very low stochasticity), and one to study the robustness of our results to stochasticity itself.

We start with two randomly generated finite populations, as opposed to the monomorphic, infinite populations that we were working with previously. We are interested in the within-population changes in genotype frequency, so the populations are constrained to have some constant size N.

Individuals are paired up with alike individuals in the other population according to the assortment parameter α and bias functions β defined in Section 3.2; remaining unpaired individuals are paired up uniformly at random between populations. Next, expected fitnesses are evaluated for every individual, and individuals are selected to be part of N breeding pairs per population, with probability directly proportionate to their expected fitness. We use stochastic universal sampling (SUS) (Baker, 1987), which chooses many individuals to breed on the basis of one random number, while retaining the property that the probability of being part of a breeding pair is proportionate to expected fitness. This technique reduces computational complexity since it requires less computationally demanding random number generation; it also reduces sampling variance.

Next, individuals arising from breeding pairs are created through uniformly at random choosing one parent's genotype and applying recombination with probability x, the recombination rate. Mutations are also applied, with p_m being the fixed probability per locus of a mutation to the other allelic value. Thus, we have created two new populations of individuals of size N. This process continues for an

appropriate number of generations, which is chosen to aid visual analysis of each test.

3.5 Stochastic Approach: Results

3.5.1 Model (a): Single-locus Assortment

In addition to all of the default parameters we discussed and chose in Section 3.3.1, we must now specify three extra parameters pertaining to the stochastic approach. Firstly, the population size, N is set to a default of 10000; the associated SUS parameter, n_{SUS}, specifies the number of individuals chosen for breeding per random number generation. This must be constructed so that individuals cannot breed with themselves. Thus, we calculate the maximum possible ratio of an individual's fitness to that of the total of the rest of their population:

$$\frac{w_0 + b}{Nw_0 + b - (N-1)c} \tag{3.33}$$

This is the maximum probability of an individual being selected. If $1/n_{SUS}$ is smaller than this, an individual could be selected twice by SUS. Thus, n_{SUS} must be constrained such that:

$$n_{SUS} < \frac{N(w_0 - c) + b + c}{w_0 + b} \tag{3.34}$$

We must also choose n_{SUS} to be even, and a divisor of the population size N. For consistency, we choose $n_{SUS} = 2000$; in each analysis, this satisfies the above constraint. It is desirable for this parameter to be constant since, while this method of choosing breeding partners means that fecundity reflects expected fitnesses in an unbiased manner, a decrease in n_{SUS} will result in higher variance, since an individual's maximum number of offspring will be N/n_{SUS}. When N is deviated from 10000, we also adjust n_{SUS} proportionately, such that $N/n_{SUS} = 5$.

Finally, we choose $p_m = 0.001$; this was settled on through some preliminary testing, which showed that this value has the desirable quality that mutations are

rare, whilst still allowing extinct genotypes another chance on occasion.

Previously, under the assumptions of QLE, we were forced to set our parameters in such a way that $\tilde{b} \ll 1 - \tilde{c}$. This ensures that both c and b must be very low, thus QLE entails weak selection. With the introduction of the stochastic approach, we are able to drop this constraint since the stochastic approach does not invoke QLE. However, we start by presenting a comparison of an experiment performed in Figure 3.1A/B using the deterministic simulation (reproduced in Figure 3.6A/B), with the same experiment using the stochastic simulation (Figure 3.6C-F).

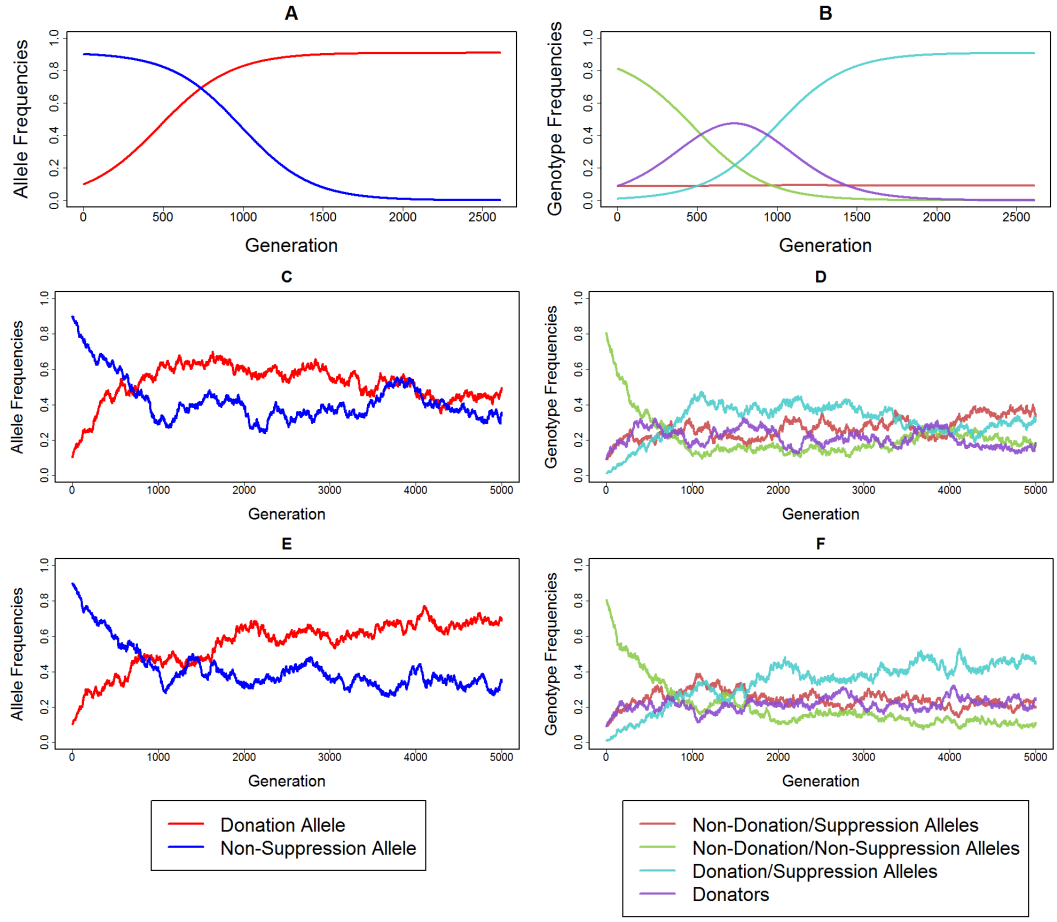

Figure 3.6: A comparison of the deterministic (A/B) and stochastic simulations (C-F), using model (a). Evolution of allele frequencies (A/C/E) and gene frequencies (B/D/F); initial frequencies are $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, c = 0.005, b = 0.01, w_0 = 1, \alpha = 1$. Additionally, in (C-F), $p_m = 0.001, N = 10000, n_{SUS} = 2000$. Population 1: C/D; Population 2: E/F.

Note that the initial allele frequencies are the same for both populations, in both tests; thus, in the deterministic simulation, allele and genotype frequencies are the same for both populations for the entirety of the process. By contrast, in the stochastic simulation, allele and genotype frequencies are not constrained to remain

the same due to the introduction of stochasticity. Figure 3.6C-F only shows the results from the first 5000 generations, due to the computational intensity of the simulation with such a large value of N. However, even after 5000 generations, allele and genotype frequencies have not converged.

The reason for this is likely to be that the effects of stochasticity outweigh the weak selection strength; this was a requirement of the deterministic approach, given the QLE assumptions. In addition, mutations have the side-effect of giving a very small selective benefit to alleles at low frequency. Thus, while it is evident that the patterns of donation and suppression on average receiving positive selection are reproduced successfully, suppression has not reached as high a frequency as it did under the deterministic simulation. This illustrates the hazards of having high levels of stochasticity relative to selection strength.

In Figure 3.7, we relax the assumption of weak selection for the stochastic simulation, instead choosing $c = 0.5$, and $b = 1$; the two parameters have been scaled up by a factor of 100, while retaining the same c/b ratio.

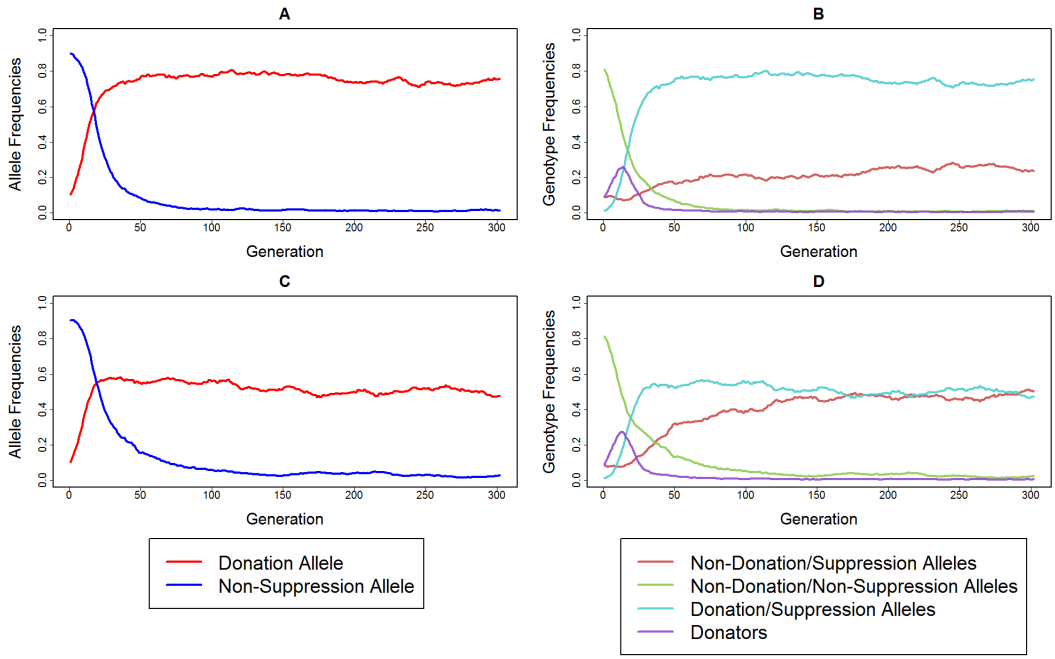

Figure 3.7: Results of the stochastic simulation under stronger selection strength. Evolution of allele frequencies (A/C) and gene frequencies (B/D); initial frequencies are $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, c = 0.5, b = 1, w_0 = 1, \alpha = 1$. Additionally, in (C-F), $p_m = 0.001, N = 10000, n_{SUS} = 2000$. Population 1: A/B; Population 2: C/D.

In Section 3.3.1, we noted that under the deterministic simulation, the scale of c

and b changes only the number of generations before convergence. However, we can see comparison of Figures 3.6 and 3.7 that in this case, the increased scale of c and b is important to ensure that the effects of selection outweigh the effects of stochasticity. We observe that allele and genotype frequencies have stabilised at equilibrium, to the point that the suppression allele has almost entirely invaded both populations and selection on the donation allele has halted, leaving it at intermediate values in the two populations. Since so many individuals possess the suppression allele, the frequency of the donation allele makes little difference since possession of the allele very rarely results in the donator phenotype. Thus, the results of model (a) are robust to the joint introduction of stochasticity and strong selection; we have reproduced the experiment that most favoured the evolution of donation, and found once again, backing up our analytic results, that the suppression allele will invade both populations, and that donation is transient.

3.5.2 Model (b): Genome-Wide Assortment

In the experiments investigating assortment on both loci with the deterministic simulation, we found that when assortment is maximised, there are cyclic fluctuations in allele and genotype frequencies. The reasons for cyclicity were discussed at length in Section 3.3.2; in summary, the relative fitnesses of donators and non-donators were dependent on the difference in donator frequencies between the two populations, which ensured that both populations did not reach equilibrium at the same time. We also noted that each population only had one type of non-donator genotype; donator frequencies were determined by the frequency of the donation allele in one population, and the non-suppression allele in the other. The original experiment is shown in Figure 3.3; we have replicated this in Figure 3.8, using the same parameters, slightly different initial frequencies, and introducing the appropriate parameters for the stochastic simulation, entailing the introduction of strong selection and stochasticity. Note that the parameters relating to selection strength and stochasticity are unchanged from those used for Figure 3.7, which achieves a reasonable balance of the forces of selection and stochasticity, as noted above. Each of these changes was motivated and discussed in Section 3.4.

A significant difference between the outcomes using the two approaches is that it

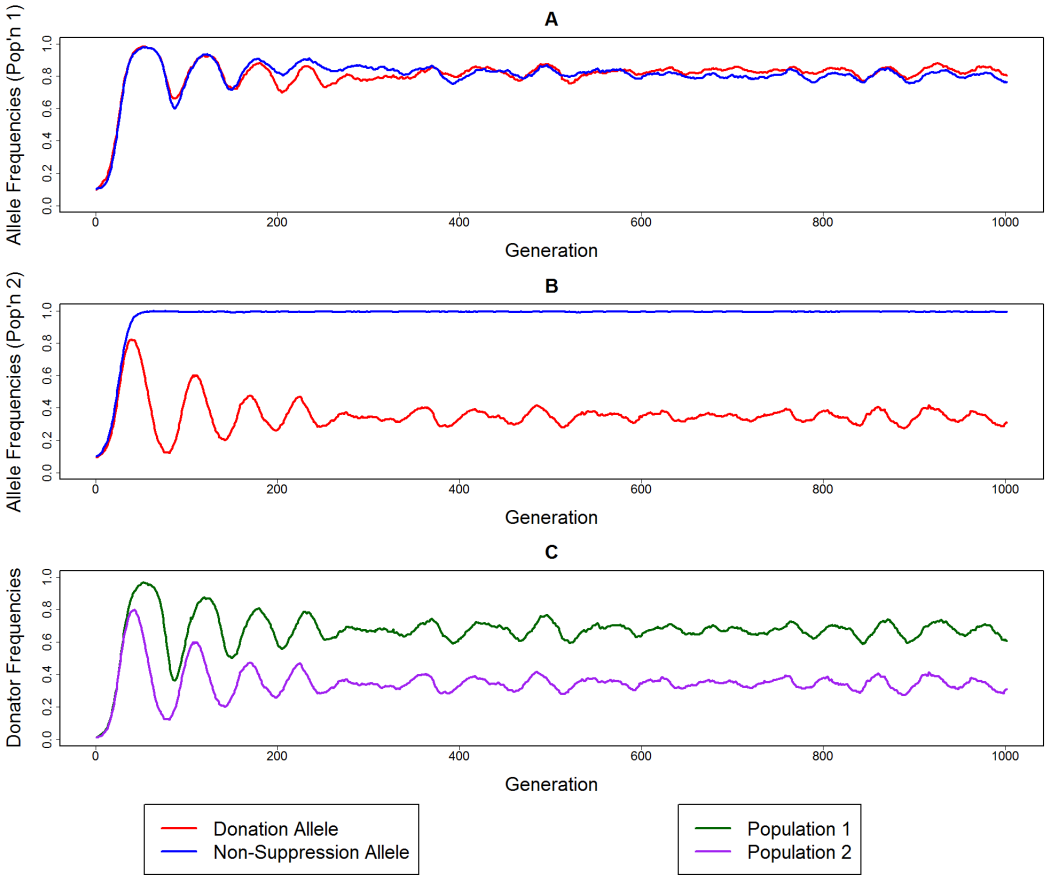

Figure 3.8: Evolution of allele frequencies (A/B), and donator frequencies (C) (using the stochastic simulation with genome-wide assortment); initial frequencies are $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 1, c = 0.5, w_0 = 1, \alpha = 1$. Additionally, $p_m = 0.001, N = 10000, n_{SUS} = 2000$. Intermediate frequencies of donation evolve, though cyclicity is moderated by the introduction of stochasticity. Population 1: A; Population 2: B.

is now not the case that the donation allele reaches fixation in one population, and the non-suppression allele reaches fixation in the other. Here, the non-suppression allele reaches fixation in population 2, and the donation and non-suppression alleles are at high levels in population 1. Initially, the cyclicity of genotype frequencies is once again in evidence, though this breaks down to a certain extent. A stable equilibrium is never reached, as donator frequencies still cycle between the two populations. This makes sense, since previously the cyclicity was based on precise differences in genotype frequencies, whereas in this case, donators may receive positive selection over a period, but find that, due to randomness, this is interrupted when the difference in donator frequencies between the two populations deviates from the expected change.

In fact, even if allele frequencies do not quite act as we expected, donator frequencies do appear to follow the pattern established in Figure 3.5. There, high

levels of assortment led to a difference in donator frequencies between populations at equilibrium. We found that two types of stable equilibria involving donation were possible, if $\alpha > c/b$. We now extend our replication of previous results to consider a range of intermediate α. Figure 3.9 displays results using the stochastic simulation that correspond to results displayed in Figure 3.4.

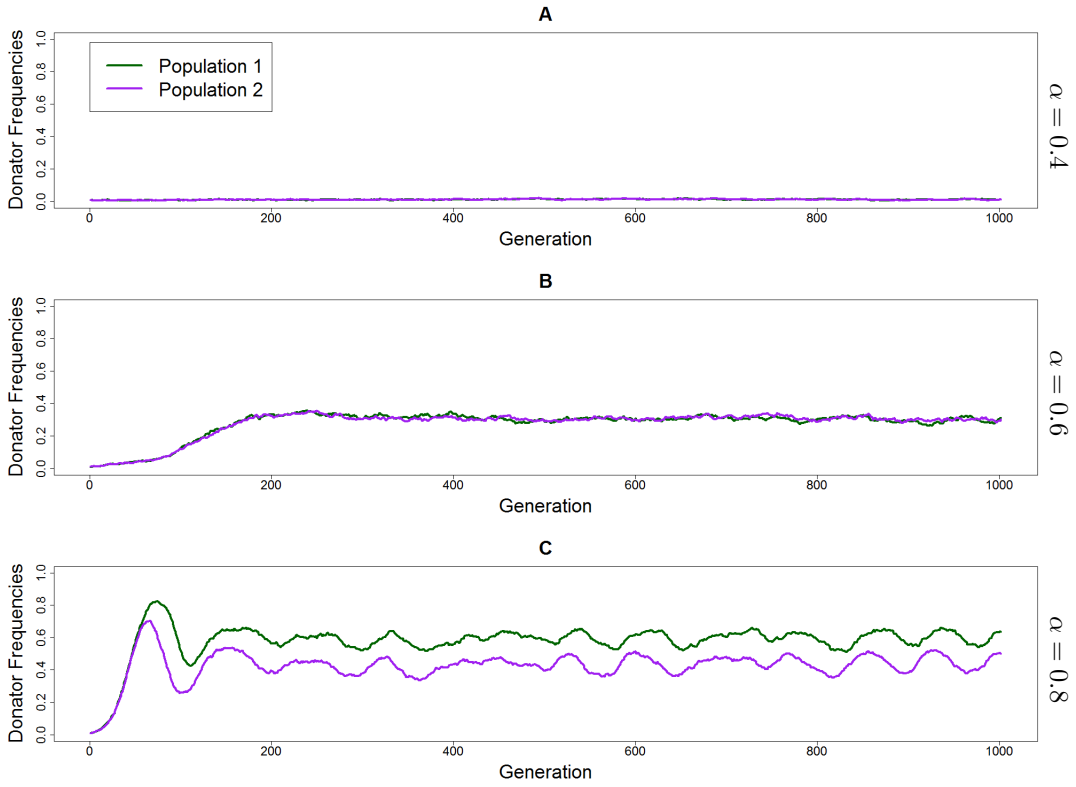

Figure 3.9: Evolution of donation (using the stochastic simulation with genome-wide assortment); initial frequencies are $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 1, c = 0.5, w_0 = 1, \alpha = 0.4, 0.6$ and 0.8 in A, B, and C respectively. Additionally, $p_m = 0.001, N = 10000, n_{SUS} = 2000$. Once again, intermediate values of α result in stable equilibria with intermediate donator frequencies; however, a small degree of cyclicity is evident when α is high.

We can see from Figure 3.9 that the stochastic simulation closely reproduces results of the deterministic approach. Specifically, we have once again found that when α is low, donation becomes extinct; when α is intermediate, donation reaches low equal frequencies in the two populations, and finally, when α is high, the frequencies of donators in the two populations are separated, while nonetheless remaining stable. One small difference here is that some degree of cyclicity is evident in Figure 3.9C. Under the deterministic approach, this settled down; however, with stochasticity, small deviations can occur. For example, imagine that the frequency of donators increased in the population 1, which we assume without loss of generality is the

population with more donators, from the stable equilibrium frequency. Then, using equation (3.29), we can see that, since there are now more unmatched donators in population 1, $P_1(1,1)$ will be decreased, and $P_2(0,1)$ will be increased. Thus, in both populations donators will receive positive selection, encouraging a cycle according to the dynamics explained in Section 3.3.2.

We now extend the tests we carried out in Figure 3.9, plotting stable equilibria of donation frequencies against α, essentially replicating the same test shown in Figure 3.5. One difference is that, due to stochasticity, donator frequencies fluctuate. Thus, we plot the mean donator frequencies in each population over the final 500 generations of each 1000-generation experiment, omitting the first 500 generations due to some initial cyclicity, as illustrated by Figure 3.9C. We have checked visually that for each α, populations do not swap between having the most and having the least donators in any of the experiments, as this would invalidate our results. Figure 3.10 shows the results of this test.

While Figure 3.10 does show slightly different results in comparison to those displayed in Figure 3.5, it is not difficult to argue that the effects of stochasticity account for the changes. Firstly, the two types of stable equilibria are clearly in evidence. Previously, we discovered that when $c/b < \alpha \leq (c/b)/(1 + c/b)$, equal but low frequencies of donators between populations occurred at equilibrium. Note that in this case, we have chosen $c/b = 1/2$, so equal but low frequencies of donators should be present in each population in the interval [1/2,2/3]. We can see that this is the case, though there appear to be low frequencies of donation in the two populations when $\alpha = 0.45$ and 0.5 too; this occurs since selection against donation at low frequencies is not strong, so low frequencies of donators in the two populations that arise through mutation can be temporarily sustained by chance. The second type of equilibrium is also apparent; this occurs when $\alpha > (c/b)/(1 + c/b)$. As discussed in relation to Figure 3.9, stochasticity again accounts for fluctuations since a small deviation from equilibrium can launch a feedback process that exacerbates this deviation.

Thus, we have concluded that all of the findings related to our deterministic model are robust to the joint introduction of stochasticity and strong selection.

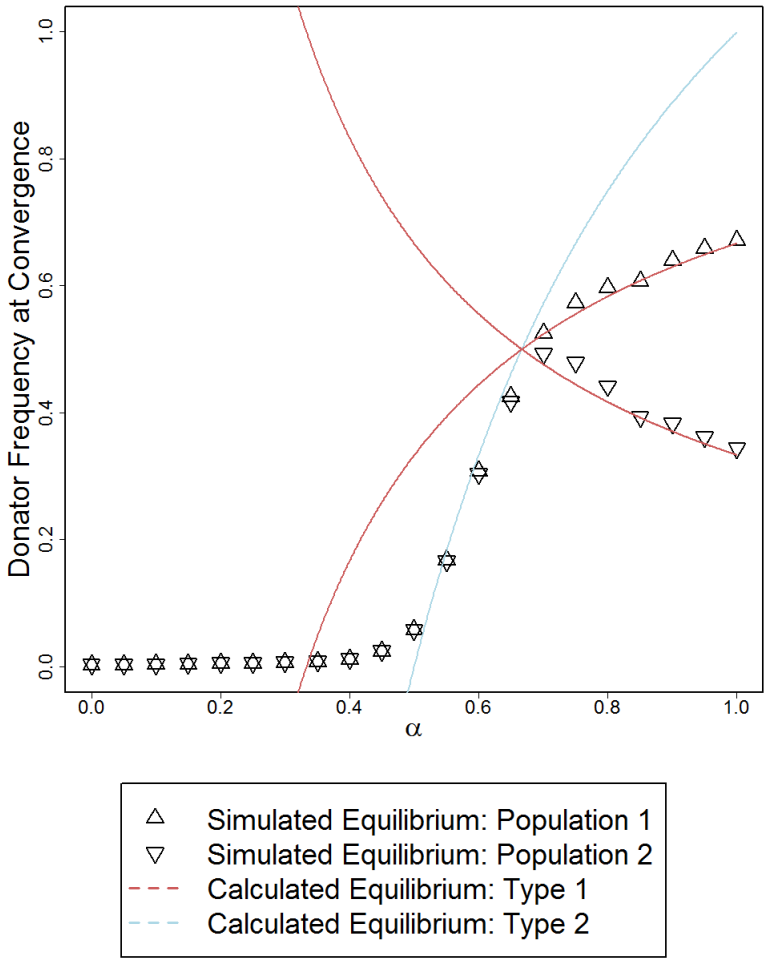

Figure 3.10: Stable equilibria of donation (using the stochastic simulation with genome-wide assortment) are shown through triangles; initial frequencies are $p_{1,1} = p_{1,2} = 0.1, p_{2,1} = p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 1, c = 0.5$, and $w_0 = 1$. Additionally, $p_m = 0.001, N = 10000, n_{SUS} = 2000$. Stable donation is present with sufficient α. Non-zero equilibria calculated and presented in equations (3.30-32) are presented as light red and blue dashed lines.

Indeed, this is also the case for the results of each of the minor models, the details of which are located in the Appendix.

3.6 Discussion

In this chapter, we have presented results for two different models of donation between species; results relating to a further four can be found in the Appendix. We have utilised deterministic and stochastic approaches.

In Chapter 2, we discussed the model of Fletcher and Doebeli (2009). Their model involves two loci; alleles on the first locus are labelled a and A, whereas alleles on the second locus are labelled b and B. AB individuals are not viable, Ab and aB individuals are cooperators in a pairwise prisoner's dilemma game (of

which the donation game introduced in Section 1.4.1 is a special case, involving additive interactions (Marshall, 2015)), and ab individuals are defectors. Fletcher and Doebeli impose assortment in such a way that Ab individuals are always paired with aB individuals. They then argue that donation is altruistic, on the basis of a assortment, which they distinguish from relatedness. Responses to this model were discussed in Section 2.2.

Model (a) is intended as a variation of the model of Fletcher and Doebeli (2009). The donation locus in the first species corresponds to the aA locus in Fletcher and Doebeli's model, while the donation locus in the second species is analogous to the bB locus. Similarly, we apply assortment on the donation locus. However, where our model differs from that of Fletcher and Doebeli is our introduction of the suppression locus; individuals that possess the suppression allele do not donate, even if they possess the donation allele. We find that the suppression allele always receives positive selection, thus the donation behaviour is vulnerable to suppression by modifiers. The donation behaviour is analogous to a within-species obligate greenbeard trait; donators performed a fixed behaviour, which is disproportionately aimed at other donators. Notably, greenbeard traits are similarly vulnerable to suppression (Biernaskie et al., 2011).

Model (b) applies genome-wide assortment across both loci, and is intended to correspond to within-species donation. We find that donation reaches a stable level when something like Hamilton's rule applies ($\alpha > c/b$). However, the donation in this scenario may also be vulnerable to suppression, if a further unassorted locus arises to suppress the donation behaviour. For example, model (d) considers donation determined by a single, assorted locus, which receives positive selection when $\alpha > c/b$, but model (a) introduces an unassorted suppression locus which results in the loss of the donation behaviour; model (d) is a closer analogy to Fletcher and Doebeli's (2009) model.

Our model could be used to inform study of greenbeard associations with parallels to our models. One example is that of between-species plant-pollinator associations, in which stable associations may be disrupted by 'nectar-robbing' species which benefit from receiving nectar without pollinating themselves; this is of course analogous

to individuals with the donation and suppression alleles in model (a) taking advantage of being assorted with donators, but suppressing their own donation behaviour (Roubik, 1982; Marshall, 2015). While in general greenbeard genes are vulnerable to modifier genes which retain the recognition component relating to the greenbeard gene, but suppress costly donation, an example of a greenbeard association which may not suffer from this vulnerability may be the FLO1 allele in *S.cerevisae* (Smukalla et al., 2008), which avoid exploitation by cheating cells through self/non-self recognition. Thus, this gene is potentially resistant to suppression genes arising at unassorted loci. Similar behaviour may be exhibited in multi-species associations such as biofilms (Marshall, 2015).

We also considered a facultative greenbeard scenario; model (f) considered genome-wide assortment, in which donators only donate to other donators. In this case, donators are facultative greenbeards since they recognise other donators and behave differentially towards them, adjusting their own social behaviour to provide help (Gardner and West, 2010). In effect, these greenbeards account for the suppression locus as well as the original donation locus, thus these greenbeard genes are better described as gene complexes (Gardner and West, 2010). In this scenario, the donation behaviour always receives positive selection, regardless of the level of assortment. However, once again, these greenbeards may be vulnerable to new modifiers arising on loci not subject to assortment.

The rule for stable donation being possible in models (d) and (b), $\alpha > c/b$, also holds for model (e), in which we considered the evolution of donation under genome-wide assortment in a single population. This rule reflects Hamilton's rule, with α being analogous to relatedness (see Section 1.4.2).

We have also checked that stable donation in these models does not rest on the dependence of donation on returned benefits, by noting that all results found by model (b) were insensitive to individuals only being placed in one role (potential donor or recipient) in a lifetime (model (c)) rather than two (model (b)). Thus, donators may incur lifetime personal fitness costs. This does not have implications for the classification of the donation behaviour under an inclusive fitness analysis (see Table 1.1), since any such classification relies on expected lifetime fitness.

A final point may be made regarding the interpretation of the donation behaviour. Fletcher and Doebeli (2009) regard the behaviour as between-species altruism, arguing that cooperative behaviour in a pairwise prisoner's dilemma is altruism by definition. However, cooperative behaviour in a prisoner's dilemma may be mutually beneficial, if social behaviour leads to an expectation of increased personal fitness on average. Alternatively, the donation behaviour may be within-species altruism; Wyatt et al. (2013) show that fitness can be regressed on genetic predictors within species, despite interactions occurring between species, leading to potential diagnoses of donation between species as within-species altruism. Given our lack of awareness of any evidence of between-species altruism occurring in nature, arguments for social behaviours to be interpreted in such a way while alternative, credible interpretations are possible should be viewed with scepticism.

We propose a causal link from donated fitness benefits in the two-locus genome-wide assortment model (b), to received fitness benefits of related individuals in subsequent generations. There is a causal link between donated fitness benefits and the likelihood of a social partner being a donator. If an individual is a donator that receives fitness benefits, that individual is more likely to have more offspring on the basis of higher fitness, who are more likely to themselves be donators, and send fitness benefits back to members of the first species. Thus, donated fitness benefits may be causally linked to received fitness benefits by related members of subsequent generations, and the donation behaviour may have a valid interpretation as within-species altruism. For the donation behaviour to be properly analysed, a formal Hamilton's rule analysis is necessary, however, beyond the scope of this work.

Chapter 4

Maternally-transmitted Symbionts and Egalitarian Transitions

Abstract

In this chapter, we present and analyse a model of hosts infected with maternally-transmitted symbionts.

We present a deterministic approach, which numerically solves differential equations to reach a solution, before conducting a bifurcation analysis of this model. We find that coincidence in fitness interests between a host and symbiont increase with the level of vertical transmission of the symbiont.

We also present stochastic simulations, which allow for evolution of symbiont strategies. In the absence of population structure, a sex-distorting symbiont will force a host to extinction. However, in the presence of population structure, the symbiont may completely invade a host population, achieving a stable equilibrium in which it does not distort sex-ratios. This novel finding contrasts with previous models of sex-distorting symbionts invoking population structure (Hatcher et al., 2000), and relies on our introduction of the evolution of symbiont strategies. We propose that this presents a potential pathway from sex-distorting parasitism to mutualism; consequently, this may be the foundation of an egalitarian major transition.

4.1 Introduction

We wish to model, in the simplest way possible, the interactions between maternally-transmitted, sex-distorting symbionts (such as the endosymbiotic bacteria, *Wolbachia* and *Cardinium* (Werren, 1997; Gotoh et al., 2007; West, 2009); specific methods of sex-distortion were discussed in Chapter 2), and their hosts. The aim of this chapter is to consider how a coincidence of fitness interests between two different species with a host-symbiont relationship can be achieved; this is necessary for an 'egalitarian' transition in individuality (Bourke, 2011). The wider literature relating to endosymbionts, egalitarian transitions, and sex-ratio theory was discussed in Chapter 2.

4.2 Model Specification

We start by considering a population of hosts, of which some are infected by a symbiont. The symbiont is maternally-transmitted, so it is necessary that we consider male and female hosts. Consequently, we have four types of hosts; uninfected females, uninfected males, infected females, and infected males; the absolute numbers of each are labelled, respectively, U_f, U_m, I_f, and I_m.

Since we are interested in modelling the conflict between host and symbiont sex ratios, we introduce strategies for the host and symbiont, which represent the proportions of female offspring of uninfected and infected individuals; these are labelled ρ_U and ρ_I respectively. We have assumed here that the gender of infected host offspring is under control of the symbiont. However, it may be the case that transmission of the symbiont from a mother to her offspring is not perfect (Ewald, 1987); thus, we include a vertical transmission parameter, α, which represents the proportion of infected offspring of an infected mother.

The life cycle of the host in this model is simple: males and females give birth at certain rates dependent on their status as uninfected or infected, and die at a constant rate independent of their infection status. Since we are interested in the maternal inheritance of the symbiont, the status of the offspring as male or female, and infected or uninfected, are determined by either the sex ratio and vertical trans-

mission strategies of the symbiont if the mother is infected, or the sex ratio strategy of the mother if she is uninfected. Host offspring are then immediately able to reproduce at the rate afforded by their infection status and gender, and die at the same constant rate. The life cycle of the symbiont simply entails birth into infected host offspring, and death when their hosts die.

We are primarily interested in the formation of mutualisms, therefore we must also consider the fitness effects a symbiont will have on a host, such as resistance to infection (Jaenike et al., 2010; Walker et al., 2011) (see Chapter 2). For simplicity, we consider only a differential effect in birth rate; we label the birth rate of hosts infected with the symbiont b_I, and the birth rate of hosts uninfected with the symbiont b_U. The death rate is labelled d, and remains unchanged in the presence or absence of the symbiont. We always choose $d = 1$, as we are only interested in the relative values of the birth rates to one another and the death rate. However, d is explicitly included in the differential equations that describe our system in order to make clear the causes of gains and losses in the numbers of each host type:

$$\frac{dU_f}{dt} = B(b_f, b_m)\big(U_f b_U \rho_U + I_f b_I \rho_I (1-\alpha)\big) - dU_f \tag{4.1}$$

$$\frac{dU_m}{dt} = B(b_f, b_m)\big(U_f b_U (1-\rho_U) + I_f b_I (1-\rho_I)(1-\alpha)\big) - dU_m \tag{4.2}$$

$$\frac{dI_f}{dt} = B(b_f, b_m)\big(I_f b_I \rho_I \alpha\big) - dI_f \tag{4.3}$$

$$\frac{dI_m}{dt} = B(b_f, b_m)\big(I_f b_I (1-\rho_I)\alpha\big) - dI_m \tag{4.4}$$

$$\begin{aligned} \text{where} \quad & b_f := U_f b_U + I_f b_I \\ & b_m := U_m b_U + I_m b_I \\ & B(b_f, b_m) := 2\min(b_f, b_m)/b_f \end{aligned} \tag{4.5}$$

We shall now walk through the meaning of the differential equation (4.1); analogous reasoning holds for the construction of equations (4.2-4). Equation (4.1) shows the change in numbers of uninfected females. Uninfected females are born to either uninfected or infected females. There are U_f uninfected females at a given point in time, that give birth at a rate proportional to b_U. Offspring of an uninfected female will be uninfected, and female with probability ρ_U. Thus, the product of

these terms, $U_f b_U \rho_U$ appear in equation (4.1), alongside the term $I_f b_I \rho_I (1-\alpha)$. This latter term arises for similar reasons; there are I_f infected females, that give birth at rate proportional to B_I, to individuals that are uninfected with probability $1-\alpha$, and female with probability ρ_I.

However, note that the terms described above are only proportional to birth rates; this is because we need another term to account for the effects of biased sex ratios and the different birth rates of infected and uninfected individuals. Without any further weighting, the total birth rate would be $b_f := U_f b_U + I_f b_I$. In words, this is the sum of the number of infected females multiplied by the infected female birth rate and the number of uninfected females multiplied by the uninfected female birth rate. A similar term could be derived for males: $b_m := U_m b_U + I_m b_I$. We label these terms the capacity for births according to females and males, respectively.

We now wish to use b_f and b_m to construct a term that will constrain overall birth rates on the basis of population sex ratio. We choose to set the total birth rate as the minimum of b_f and b_m; in other words, the total birth rate is the lowest limiting factor out of the capacities for births according to males and females. Since, as noted above, the total birth rate without weighting would be $b_f := U_f b_U + I_f b_I$, we divide through by b_f, then multiply by $\min(b_f, b_m)$. Since we will later choose b_U and b_I in relation a baseline of 1, we must also weight by 2, accounting for the fact that each offspring has two parents. Thus, we introduce a term multiplying all of these factors, $B(b_f, b_m) := 2\min(b_f, b_m)/b_f$, which weights each of the birth rate terms.

Finally, the death rate term is simply the base death rate per individual, d, multiplied by the number of uninfected females, U_f.

Unfortunately, the system of equations described above is non-linear, and therefore difficult to solve; it is necessary that we consider an alternative approach. Thus, we initially analyse the model using a deterministic, iterative scheme in order to understand the dynamics of system and investigate the effects of host and symbiont strategies on endpoints of the system. We then utilise a bifurcation analysis, using numerical approaches to focus on the different types of equilibria in the system

and their stability. Finally, we use an individual-based simulation to investigate the evolutionary end-points given that symbiont sex ratio and vertical transmission strategies are allowed to evolve through breeding and mutation.

4.3 Deterministic Model

4.3.1 Method

Here, we use the 'deSolve' package (Soetaert et al., 2010) in R (R Core Team, 2016) to program our system of differential equations; relevant code is located online (Quickfall, 2016). We analyse changes in population size and within-population frequencies of each of the four host types (male or female; symbiont present or absent). We use time-steps of length 0.1 (note that time is dimensionless in this model, and rates are defined relative to time), and analyse the population for convergence every ten time-steps; these are choices that balance the competing needs for computational complexity and accuracy of the process. Convergence is detected when the absolute change in within-population frequencies falls below a certain threshold. Evolutionary outcomes tested over a small selection of parameter choices were not sensitive to changes in the time-step size, convergence test intervals, or the convergence threshold, given reasonable choices of each parameter.

Our remaining parameters are as follows: b_U, the birth rate for uninfected hosts; b_I, the birth rate for infected hosts; ρ_U, the sex ratio strategy of the host; ρ_I, the sex ratio strategy of hosts infected with the symbiont; d, the death rate; α, the probability of successful vertical transmission of the symbiont, and finally, the initial frequencies of each of the four host types. Some of these we shall not test; Table 4.1 summarises our investigation of model parameters. Initial frequencies are chosen such that the symbiont is initially rare (1%) in every test, since we are primarily interested in invasion of the symbiont. We also choose the initial sex ratio to be the Fisherian sex ratio of 1/2. Birth rates are set by default to 1.2 for infected hosts and 1.1 for uninfected hosts, since, if the symbiont confers a fitness benefit through increased birth rate, it essentially has an 'option' to be either parasitic or mutualistic, depending on the fitness cost associated with any sex ratio distortion. The death rate is set to 1, as discussed in the previous section. Note that $b_U > d$; this is so

that there is some leeway in terms of endosymbionts causing detrimental effects on host fitness; population extinction is detected only if this effect is sufficiently high. The remaining parameters, ρ_U, ρ_I, and α are given no default since we test each of these extensively.

Parameter	Status	Interval
b_I	Constant	1.2
b_U	Constant	1.1
d	Constant	1
ρ_I	Variable	[0,1]
ρ_U	Variable	[0,1]
α	Variable	[0,1]

Table 4.1: Investigation of Model Parameters

For each run of the model with a given set of parameters, we analyse certain data relating to the model. Firstly, we track frequencies of each host type, as these can help to provide insight into the mechanics of the process. We also look at the average host fitness at equilibrium, (defined to be the growth rate of the overall population) and the average symbiont fitness at equilibrium (defined to be the growth rate of the population of infected hosts). Notably, for every single run using the deterministic approach, equilibrium of host frequencies was reached, although the population size was not constant at equilibrium.

4.3.2 Results

We start our analysis with an example, in order to illustrate the dynamics of the system; this is depicted in Figure 4.1. Along with the above-specified parameters, we choose $\alpha = 0.95$; in words, the symbiont is almost perfectly transmitted from mother to offspring. We also choose the host to have the standard Fisherian sex ratio, $\rho_U = 0.5$, but the symbiont to distort this sex ratio slightly; $\rho_I = 0.55$.

Firstly, we note that symbionts infect a significant proportion of the host population at equilibrium. This is because symbionts confer a fitness benefit to their hosts, and transmit themselves successfully due to a combination of high vertical

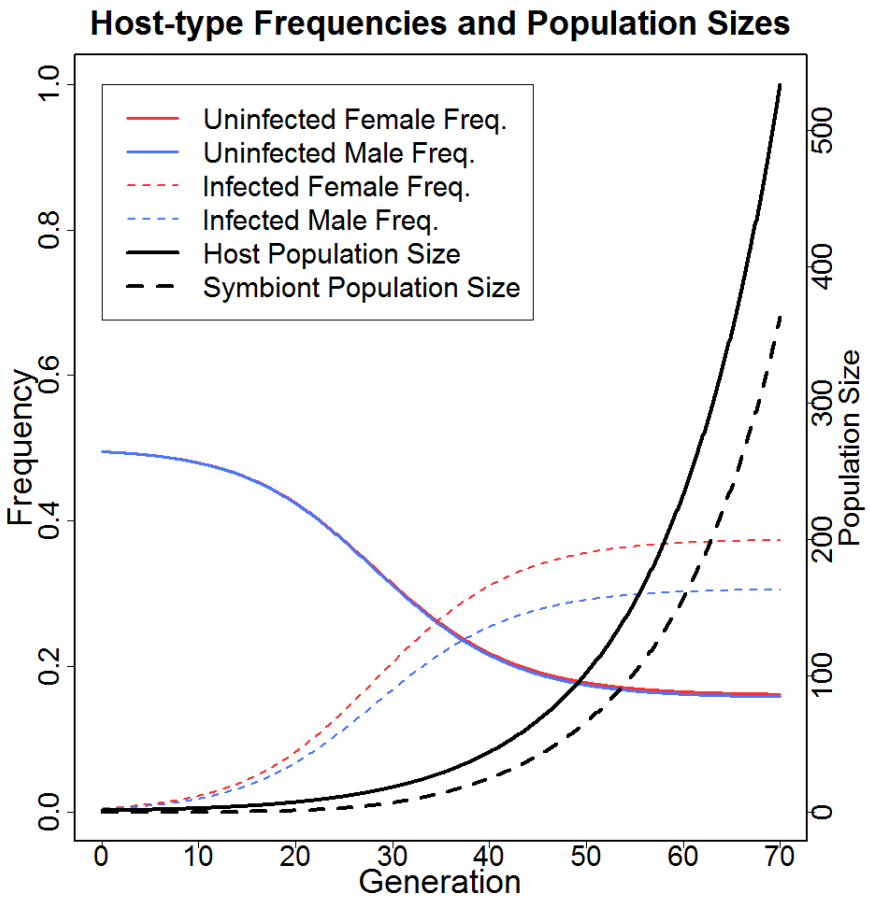

Figure 4.1: Evolution of host-type frequencies and population size. Here, $\alpha = 0.95, \rho_I = 0.55$, $\rho_U = 0.5$, $b_I = 1.2, b_U = 1.1$, and $d = 1$. At equilibrium, the population is increasing, symbionts are present at a high frequency, and there is a biased sex ratio in favour of females.

transmission and sex ratio distortion. They do not have such high sex ratio distortion and frequency within the population that the sex ratio is excessively distorted, and the host goes extinct.

This is one simple example of the dynamics of this system. We shall now conduct a fuller investigation by considering the equilibrium host-type frequencies and population size status over the entire parameter space of the sex ratio parameters ρ_I and ρ_U. These parameters represent the sex ratios for infected and uninfected individuals respectively, and thus take values in the range [0,1], where a sex ratio of 1 entails the production of female offspring only. Once again, we assume that vertical transmission is nearly complete; $\alpha = 0.95$. The results of this experiment are depicted in Figure 4.2.

As shown by the fitness contours in the plot of host fitnesses, Figure 4.2A, the

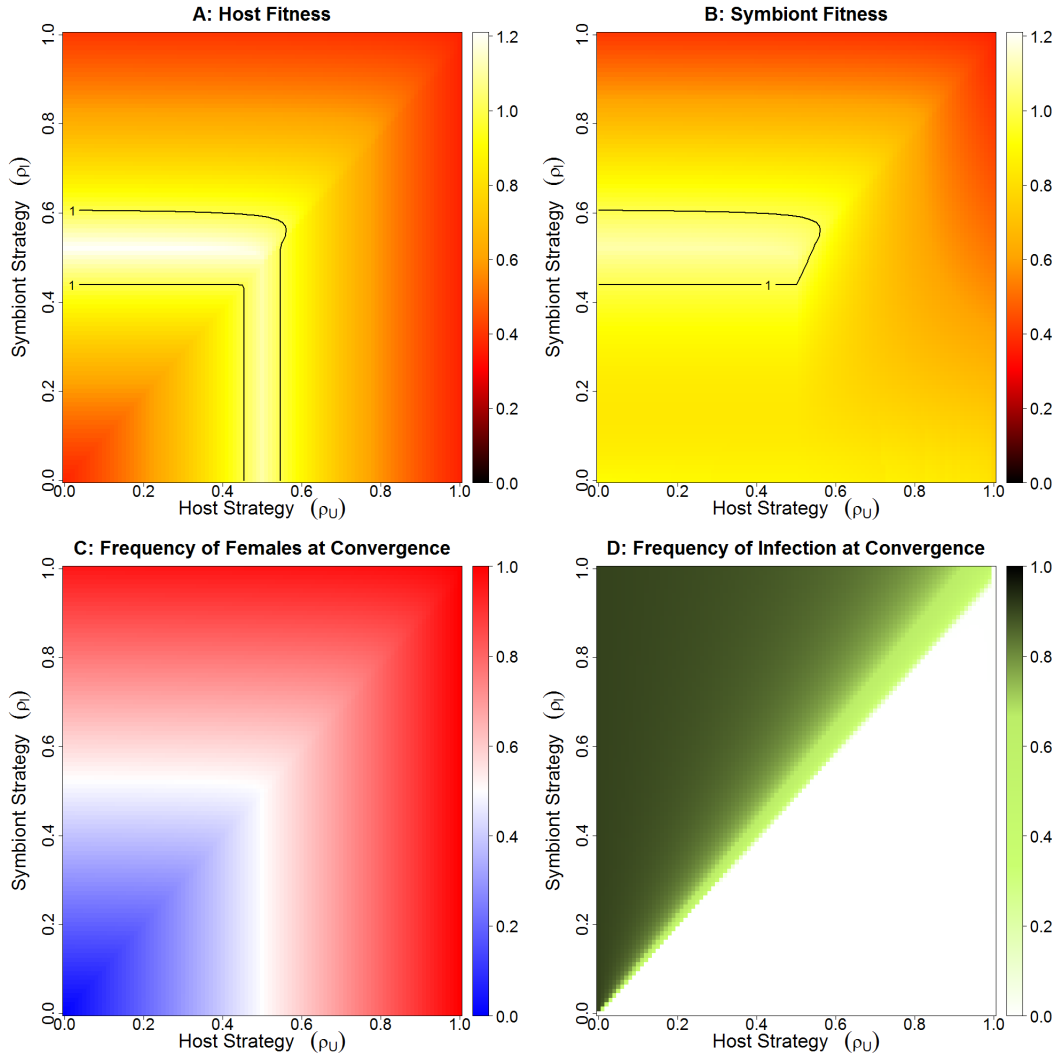

Figure 4.2: Equilibrium fitnesses of hosts (top-left) and symbionts (top-right), and frequencies of females (bottom-left) and symbionts (bottom-right), over the ranges [0,1] for the host and symbiont sex ratio strategies, ρ_U and ρ_I. We also set $\alpha = 0.95, b_I = 1.2, b_U = 1.1$, and $d = 1$. The host population does not crash (i.e. fitness at equilibrium is 1 or greater) only when $\rho_U \approx 0.5$ or $\rho_I \approx 0.5$. In the former case, the symbiont is entirely absent, and in the latter, the symbiont is entirely present. There is a small crossover region with intermediate symbiont frequencies.

host population is only viable over a small range of the parameter space. The failed populations come under three categories. If $\rho_U \gg \rho_I$, then infected females give birth to less infected females than uninfected females give birth to uninfected females, since the fitness benefit to infected hosts is not large enough to overcome the symbiont's failure to sufficiently distort sex ratios towards production of females. Thus, at equilibrium, there are no symbionts, and whether or not the population grows is dependent on the host strategy ρ_U. If this is biased far enough one way or the other from the Fisherian sex ratio of 1/2, then the population crashes as a result.

The second type of outcome occurs when $\rho_I \gg \rho_U$. In this scenario, the symbiont successfully transmits itself, and thus at equilibrium the population sex ratio is entirely dependent on ρ_I. Therefore, if ρ_I is significantly distorted away from 0.5, the population crashes. In this case, there is a little more lee-way in terms of how much the symbiont can distort the population sex ratio, since it confers a fitness benefit through increased birth rate to infected hosts.

Finally, if $\rho_I \approx \rho_U$, then the population crashes if both sex ratios are significantly distorted away from 0.5. If both sex ratio strategies are around 0.5, then an intermediate frequency of symbionts can persist at equilibrium.

A general rule regarding all of these possible equilibria is that a population will crash if the effects of sex ratio distortion at equilibrium outweigh any excess birth rate over death rate. This is why we have chosen the birth rates of uninfected and infected individuals to be strictly greater than 1.

We can also note that if the host deviates too far from the standard Fisherian sex ratio of 0.5 (see Chapter 2 for further discussion), the population will crash; in other words, compensatory sex ratios lead to population extinction in our model. This contrasts with Werren (1987) and Hatcher and Dunn (1995), who, using similar models, found that compensatory sex ratios could evolve; however, they did not model population extinction.

We now consider results from the model if we vary both symbiont strategies; the vertical transmission parameter, α, which was not varied in the previous experiment,

and the symbiont sex ratio, ρ_I. As previously discussed, the host sex ratio is taken to be $\rho_U = 0.5$. All other parameters remain unchanged. The results are depicted in Figure 4.3.

For a large portion of the parameter space, the host fitness is greater than 1, and so the population is expanding at equilibrium; only when both α and ρ_I are sufficiently large do symbionts reach a sufficiently high frequency and distort sex ratios sufficiently to cause the population to crash. However, by analysing symbiont fitnesses, we can see that for much of the parameter space in which the host population persists, the symbionts do not transmit themselves efficiently enough to gain a foothold within the population. Only for a narrow band of α and ρ_I strategies do the symbionts successfully infiltrate the host population without causing extinction. As α increases past 0.5, the required ρ_I strategy for symbionts to both reach a high frequency and not cause extinction decreases to 0.5. In other words, higher transmission fidelity causes fitness interests of the symbiont and host to align, as shown by the optimal ρ_I converging to ρ_U.

This is an important result, since a confluence of fitness interests is required for an egalitarian transition in individuality (Bourke, 2011). However, similar results have been reached in the past. Notably, Hatcher and Dunn (1995) have created a similar model, of host sex ratio, and symbiont transmission and feminisation, and analysed a host's response to the presence of a symbiont. They found that equilibrium presence of infection increases with transmission, and that high transmission and feminisation can lead to overexploitation of the host population, corroborating our findings. However, they also found that compensatory sex ratio evolution in the host is possible, contrasting the results of our model.

Given fixed host sex ratio, we can choose ρ_I and α such that our model and Hatcher and Dunn's (1995) model are equivalent. However, their host sex ratio acts in a slightly different way, since it can influence the gender of infected offspring. Thus, while many of their results are similar to ours, it makes sense that their model results differ from ours when they consider changes in the host sex ratio parameter.

A further, very significant, similarity between our analysis and that of Hatcher

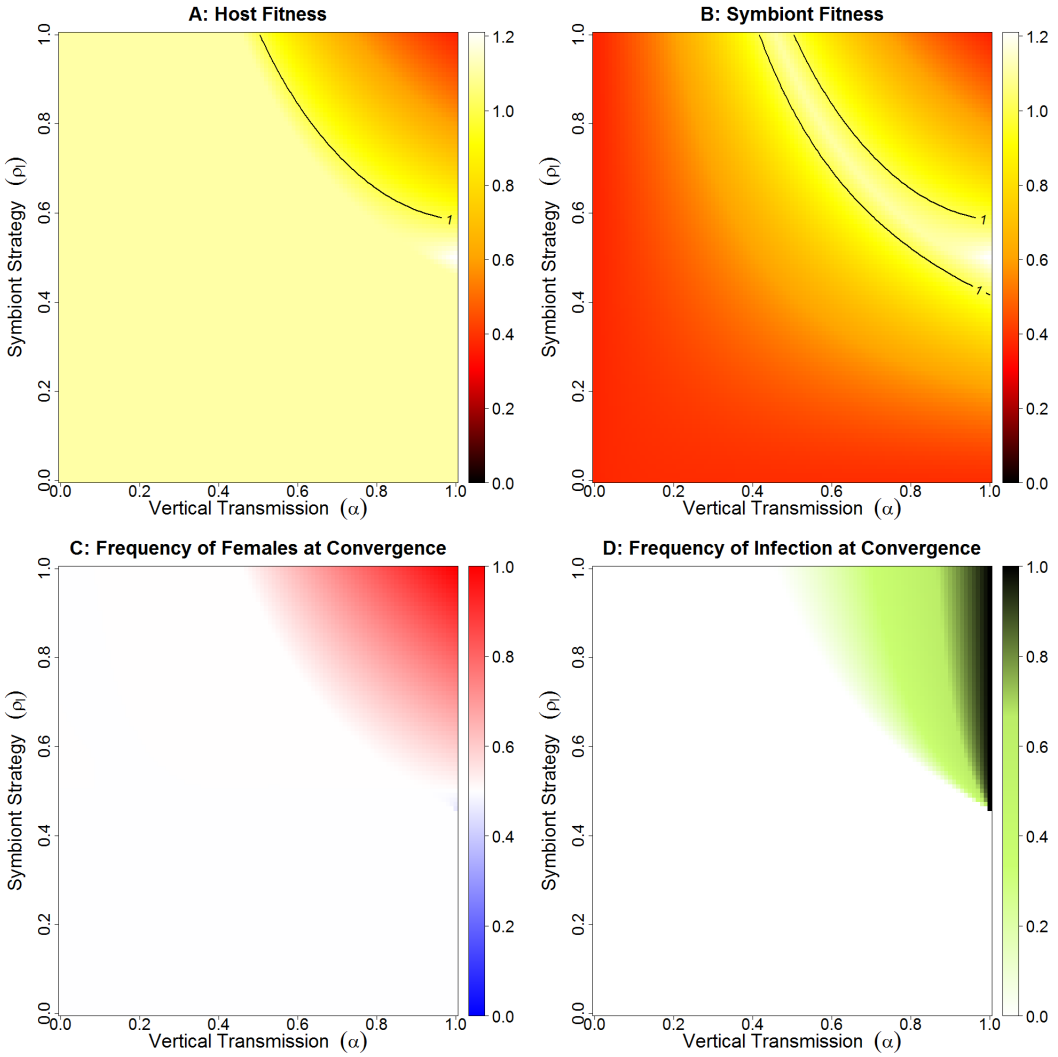

Figure 4.3: Equilibrium fitnesses of hosts (top-left) and symbionts (top-right), and frequencies of females (bottom-left) and symbionts (bottom-right), over the ranges [0,1] for the symbiont's vertical transmission and sex ratio strategies, α and ρ_I. We also set $\rho_U = 0.5, b_I = 1.2, b_U = 1.1$, and $d = 1$. The host population survives (i.e. fitness at equilibrium is 1 or greater) if the symbiont does not over-exploit the population; as α increases, fitness interests, and therefore sex ratio strategies, align.

and Dunn so far is that the feminisation and transmission parameters which can lead to an intermediate frequency of symbionts at equilibrium in Hatcher and Dunn's (1995) model is very similar to the set of ρ_I and α parameters that lead to intermediate symbiont frequencies in our analysis, given that $f = 0$ in their model (no feminisation) is analogous to $\rho_I = 0.5$ in our model (unbiased symbiont sex ratios), when $\rho_U = 0.5$ (see Figure 5, Hatcher and Dunn (1995)).

In Section 2.3.4, we discussed Hatcher and Dunn's (1995) analysis of feminising symbionts in a host population; they found that the degree of feminisation of an invading symbiont could drive a host population to develop compensatory sex ratios, possibly producing only males. This reflected Werren's (1987) model, which reached a similar conclusion, though in the absence of host sex ratios determining some infected offspring's gender, intermediate compensatory sex ratios were not detected . Our model has shown that in developing compensatory sex ratios, a population may be destabilised; in the absence of frequency-dependent fitness effects, the growth rate at equilibrium is shown to be less than 1, since there is an overall bias in sex-ratios. A limitation of these analyses was that they did not consider how symbiont strategies may evolve; this shall be considered in Section 4.5.

However, the next section considers the same model, utilising an alternative numerical approach. This allows us to check the robustness of our results, explore the potential for alternative equilibria, and investigate the stability of the equilibria detected in the deterministic model.

4.4 Bifurcation Analysis

We now use the MATLAB (The Mathworks Inc., 2010) package MatCont (Dhooge et al., 2003) to analyse the stable and unstable equilibria of this system. The model itself can be located online (Quickfall, 2016). This uses numerical methods to obtain equilibria; numerical continuation can then be used to find how a given equilibrium changes with a certain parameter. Unfortunately, our system is four-dimensional, so plotting every equilibrium for all four types of individual is unwieldy; thus, we focus only on the numbers of infected females at equilibrium, and report verbally the other components of the host population where informative.

Equilibria detected using the deterministic approach were only equilibria in terms of within-population frequencies of host types. We now apply population regulation, adjusting birth rates according to how far the population size is from 100. The birth rate discounting function is labelled R; for a given population size N, every individual's birth rate is adjusted by the following factor:

$$R := \frac{200 - N}{100} \tag{4.6}$$

When the population size is 100, no fitness adjustment occurs; fitness adjustment linearly decreases with N, from a factor of 2 when the population size is 0 (of course, this factor is never used) to a factor of 0 when the population size reaches 200 (likewise, this factor is never reached). Note that any population smaller than 100 is comparable to populations decreasing in size in the previous analysis, so the detection of extinction is mitigated here by negatively-frequency-dependent fitness effects. Due to the numerical methods used, we were forced to avoid any population regulation function which invoked use of exponential functions; this explains why there is deviation from the method of population regulation used in the individual-based model presented in Section 4.5.

We were also forced to consider an alternative version of the birth discounting function, since MatCont could not handle a discontinuous version; this remains symmetrical, with mode 0.5 and maximum value 1:

$$B(b_f, b_m) := 4\frac{b_f b_m}{b_f + b_m} \tag{4.7}$$

We begin by investigating how equilibria change with the vertical transmission rate, α; these results are displayed in Figure 4.4. We use the same parameters used in Figure 4.1, choosing the symbiont to confer some fitness benefit ($b_I > b_U$), while distorting sex ratio ($\rho_I = 0.55$) away from that of the host ($\rho_U = 0.5$).

For low α, symbionts are not present at the only equilibrium, which is stable and involves equal numbers of uninfected males and uninfected females. However, once

α crosses a certain threshold, a bifurcation point occurs, and two types of equilibrium are present. Firstly, there is an unstable equilibrium entailing equal number of uninfected males and uninfected females, and a total absence of symbionts. Secondly, there is a stable equilibrium, polymorphic for infection. Symbionts overcome their inefficient transmission by providing fitness benefits to hosts and distorting sex ratios. As vertical transmission increases even higher, the amount of infected hosts further increases. Note that there is no population crash here, since symbionts do not distort sex ratios to a great extent. This corroborates with results found in Figures 4.3A-D.

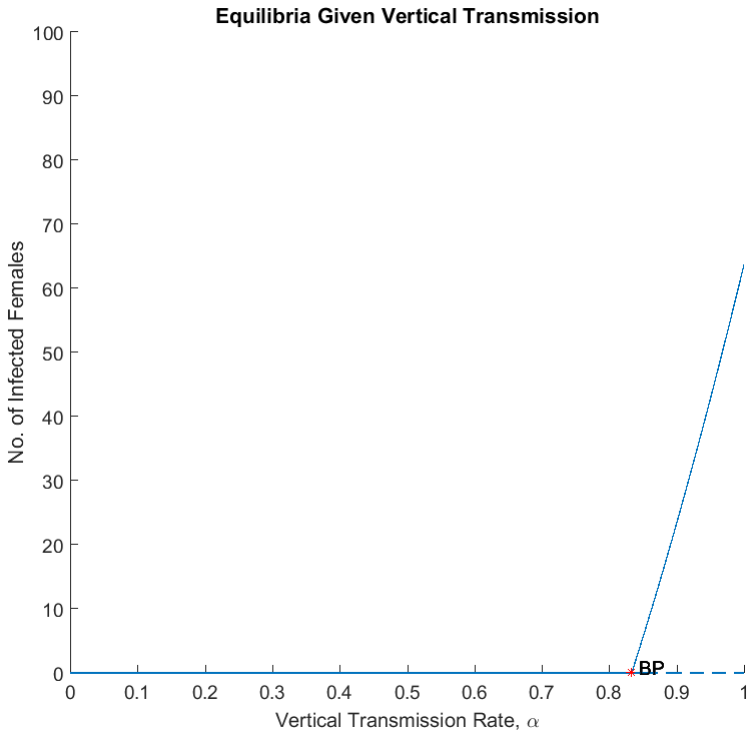

Figure 4.4: Equilibria given changes in vertical transmission. Here, $\rho_I = 0.55$, $\rho_U = 0.5$, $b_I = 1.2, b_U = 1.1$, and $d = 1$. Stable equilibrium (undashed line) frequencies of infected females are possible once α crosses a threshold value, and increase with α from then on. A trivial unstable equilibrium (dashed line) is possible in which no infection is present, for all α.

We now investigate how equilibria change with the symbiont sex ratio strategy, ρ_I; these results are displayed in Figure 4.5. When symbionts distort sex ratios in favour of males, they do not aid their own spread, and the stable equilibrium involves equal numbers of uninfected males and females. Since vertical transmission is perfect and the symbiont provides a fitness benefit, the symbiont spreads through the population when ρ_I reaches close to 0.5. Here, expected numbers of infected

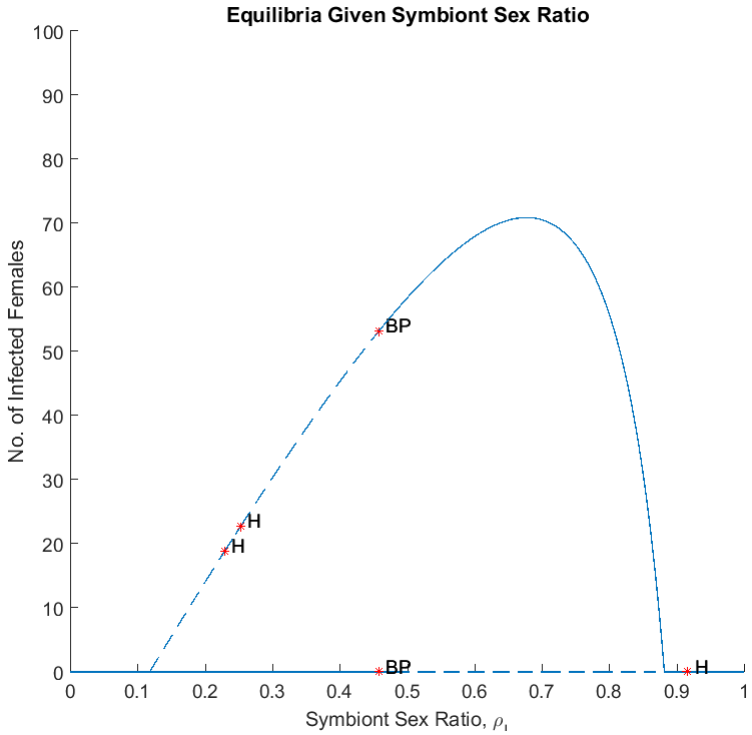

Figure 4.5: Equilibria given changes in symbiont sex ratio. Here, $\alpha = 1$, $\rho_U = 0.5$, $b_I = 1.2, b_U = 1.1$, and $d = 1$. Infected females persist at a stable equilibrium so long as they do not exhibit significantly male-biased sex-ratios; the bifurcation point (BP) occurs at the point where infected hosts transmit more efficiently than uninfected hosts. At the stable equilibria past this level of ρ_I, no uninfected individuals are present in the population. Too high ρ_I causes the population to crash. A trivial unstable equilibrium (dashed line) is possible in which no infection is present, for all ρ_I. Neutral saddles (H) are detected but do not entail bifurcations.

females offspring have eclipsed expected numbers of uninfected female offspring, so the symbiont spreads through the entire population. Since vertical transmission is perfect, this is not a polymorphic equilibrium. Females increase in number with ρ_I, until the sex-ratio distortion has such a negative effect on birth rates that the population begins to fall, going extinct with high enough distortion. Note that the introduction of population regulation has maintained the population far beyond the point at which it crashes in the previous analysis (see Figure 4.3). We can also note the presence of unstable equilibria; due to complete vertical transmission, there is an unstable equilibrium in which only infected individuals are present, which occurs when the symbiont sex ratio is female-biased enough, in addition to the unstable equilibrium in which only uninfected individuals are present, which occurs when symbiont sex ratio is intermediate to significantly female-biased.

The results of this figure can be contrasted to the case where $\alpha = 0.9$; these results are displayed in Figure 4.6. Stable equilibria remain similar to the previous

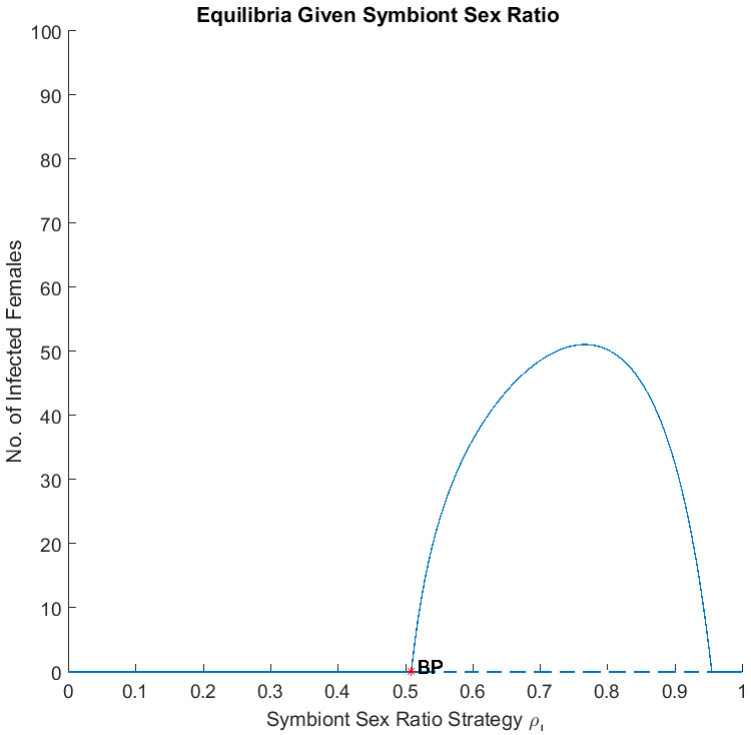

Figure 4.6: Equilibria given changes in symbiont sex ratio. Here, $\alpha = 0.9$, $\rho_U = 0.5$, $b_I = 1.2, b_U = 1.1$, and $d = 1$. Infected females persist at a stable equilibrium so long as they exhibit sufficiently female-biased sex-ratios; the bifurcation point occurs where infected hosts begin to transmit more efficiently than uninfected hosts with increasing ρ_I. At the stable equilibria past this level of ρ_I, uninfected individuals are present in the population but fall in number with ρ_I. Too high ρ_I causes the population to crash. A trivial unstable equilibrium (dashed line) is possible in which no infection is present, for all ρ_I.

case; however, note that the bifurcation point occurs at a slightly greater value of ρ_I. This is because the symbiont sex ratio must be higher for the symbionts to invade, in order to make up for inefficient vertical transmission. Since vertical transmission is not complete, the stable equilibria are polymorphic for infection. In contrast to Figure 4.5, we can see that the only unstable equilibria involve uninfected individuals; a wholly infected population with male-biased sex ratio is impossible because uninfected individuals arise out of incomplete vertical transmission. Once again, these results can be compared with Figure 4.3. A more complete analysis of how the bifurcation point shown in Figures 4.5 and 4.6 responds to α would be desirable; unfortunately, this was not possible for numerical reasons.

We also consider, but do not display here, alternative versions of the birth discounting function $B(b_f, b_m)$, which model how a population copes with distorted sex ratios. We considered two additional functions which, while retaining the properties of having a mode of 0.5 and maximum value of 1, modelled differential effects on

birth rates when males or females had lower birth capacities. Results derived from this work showed little qualitative difference from those displayed in the case of a symmetric birth discounting function, thus are not displayed here; code is provided online (Quickfall, 2016).

We now turn to the extension of this model promised at the end of Section 4.3, using an individual-based model to understand the evolution of symbiont strategies.

4.5 Individual-based Model

4.5.1 Method

We now consider an individual-based model, returning to use of R (R Core Team, 2016); our code is once again found online (Quickfall, 2016). We are interested in the potential invasion of a population of hosts by symbionts, thus we once again choose 1% of the host population to possess symbionts at the start of the process. These symbionts are randomly allocated starting strategies, using a truncated normal distribution (across the possible set of α and ρ_I strategies, [0,1]); the initial mean strategies will be a parameter of the system under investigation. The host sex ratio, ρ_U, is not under selection, and is chosen to be 0.5 for every individual. Life cycles remain mostly the same as those in Section 4.2, though this approach involves discrete generations, so instead of individuals breeding and dying randomly, the entire population is replaced every generation.

Since male hosts do not pass on any information to offspring, as the symbiont is maternally-transmitted, we consider only reproduction from the perspective of females. Fitnesses are determined by whether or not an individual possesses the symbiont, and discounting of the birth rate according to the sex ratio. This is calculated in line with the differential equations discussed in Section 4.2, with one slight alteration in the form of population regulation.

We introduce an exponential function as an alternative to the linear discounting of population fitness with sex-ratio distortion that was forced by the use of a numerical approach. Here, we use the following population regulation function:

$$R := \exp\{-2\log(r)\frac{N - 10000}{10000}\} \tag{4.8}$$

In words, this function states that for every 5000 individuals over or below the neutral group size of 10000, every individual's fitness will be, respectively, divided or multiplied by the regulation parameter r. This shall be set to 1.5 initially; however, this will be subject to investigation later on. This function is plotted in Figure 4.7 in order to illustrate how it works more clearly.

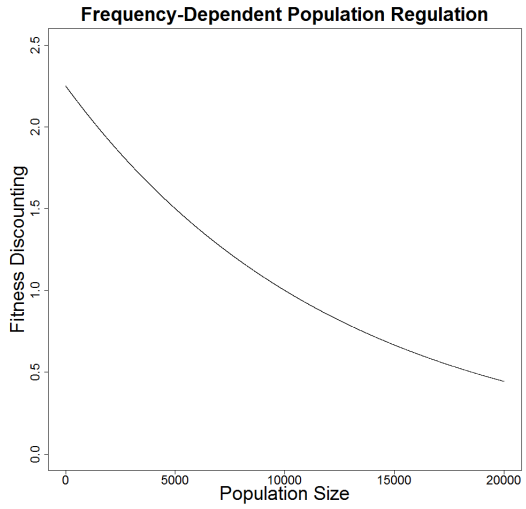

Figure 4.7: Frequency-Dependent Population Regulation: for every 5000 extra members of a population, fitness is discounted by a further 1/3.

Once every female's fitness has been evaluated, their fecundity is determined randomly, using the Poisson distribution with parameter chosen to be fitness; thus fecundity is expected to match fitness. Gender and infection status of offspring is randomly chosen according to the sex ratio strategy of the mother, and the vertical transmission strategy of the mother if they are infected. Infected offspring's symbionts inherit the mother's symbiont's strategies, with mutations occurring on both loci according to a truncated normal distribution, with means being the mother's strategies and variance being 0.01.

Convergence is determined to have occurred when the mean ρ_I and α strategies have registered little absolute change in comparison to the previous five generation of mean ρ_I and α strategies. Alternatively, the process ends in either the symbiont

going extinct, or the host population size falls by a factor of 10.

We initially consider an unstructured population. This shall be contrasted to a population which include a patch-structure; this choice shall be briefly motivated and discussed at the start of Section 4.4.3.

4.5.2 Unstructured Population Results

First, we consider a population initiated with optimal symbiont strategies according to the previous analysis. Can a perfectly vertically-transmitting, non-sex-ratio-distorting symbiont invade a host population if the symbiont strategies ρ_I and α are allowed to evolve? Figure 4.8 shows an experiment which starts with $\rho_I = 0.5$ and $\alpha = 1$.

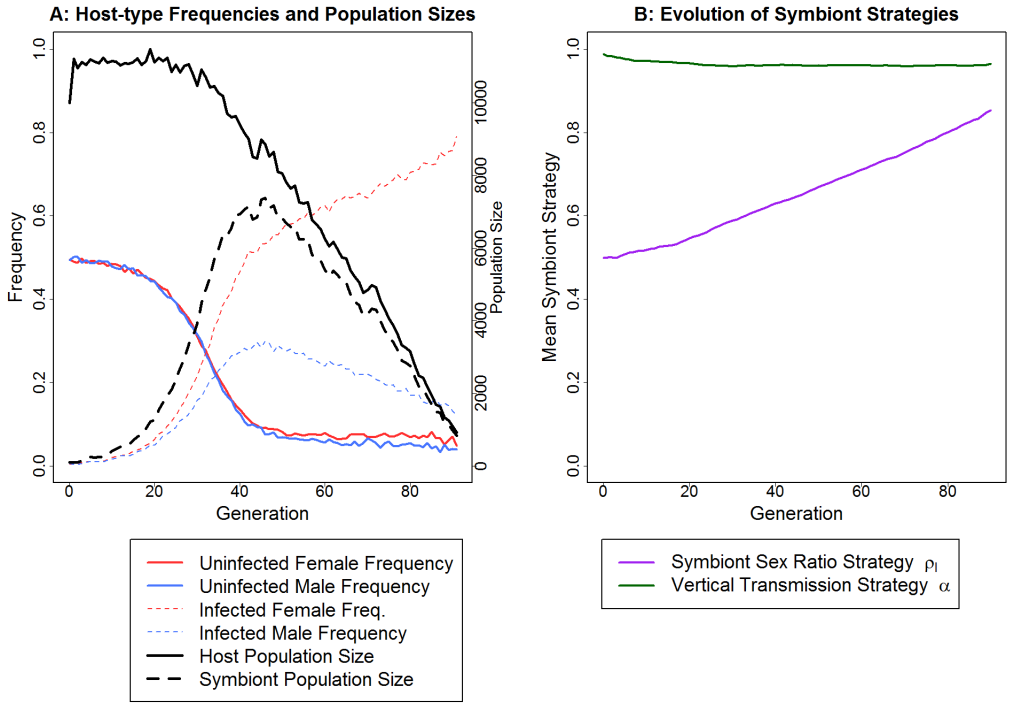

Figure 4.8: Evolution of host-type frequencies and population size (A), and symbiont strategies (B). Initially, symbionts are rare and possess strategies $\alpha = 1$ and $\rho_I = 0.5$. We also set $\rho_U = 0.5, b_I = 1.2, b_U = 1.1$, and $d = 1$. Symbionts invade the population, then evolve more and more exploitative sex ratio strategies, causing the host population to crash due to an excessively biased sex ratio.

Initially, the population increases in size and maintains a size above 10000, since symbionts are rare and not distorting the sex ratio, and $b_U > d$. Symbionts increase in frequency, since they are close to perfectly vertically-transmitting (due to the mutation scheme, the mean strategy drifts away from 1), and $b_I > b_U$. However,

the mean ρ_I strategy of the symbiont is always increasing, since an increase in the proportion of female offspring for infected hosts will ensure that more infected female offspring are produced. As the symbiont frequency and mean ρ_I increase, the overall sex ratio of the host population becomes more and more distorted, meaning average host fitness linearly decreases as the symbiont sex ratio increases.

Figure 4.8 is just one example of this process; we can repeat this experiment for the entire parameter space of initial ρ_I and α, with 100 repeats of each experiment, and plot the trajectories of mean symbiont strategies. These results are depicted in Figure 4.9.

Firstly, we can note from Figure 4.9 that when α and ρ_I are both initially small, the symbiont goes extinct very swiftly, since the symbiont is rarely transmitted to offspring, who are mostly male and thus cannot transmit the symbiont any further in any case; this scenario is illustrated by blue crosses at the ends of evolutionary trajectories in Figure 4.9. However, if α and ρ_I are sufficiently high that the symbiont does not go extinct within the first few generations, there is scope for symbiont strategies to evolve. This typically occurs when both strategies are greater than 0.5, with the chances of the symbiont persisting increasing as α and ρ_I increase further. When the symbiont persists, both α and ρ_I receive positive selection, since symbionts that vertically transmit more efficiently will have greater fitness as more offspring will be infected, and symbionts that bias sex ratios in favour of females will also have greater fitness since males are evolutionary dead ends for the symbiont. However, when α and ρ_I increase too far, they bring about a population crash, symbolised by red crosses in Figure 4.9, and discussed in relation to Figure 4.8.

These results can be compared to the host and symbiont fitnesses depicted in Figure 4.3 (A and B respectively); note that the two experiments utilise the same parameter values. When the symbiont starts with strategies far away from the lower fitness isocline, the symbiont goes extinct, whereas the host persists, since host fitness is on average greater than 1. Both the host and symbiont persist when symbiont strategies are such that symbiont and host fitnesses are greater than 1; however, when symbiont strategies evolve to the space outside the upper fitness isocline, the host and symbiont fitnesses are aligned, but smaller than 1, so the pop-

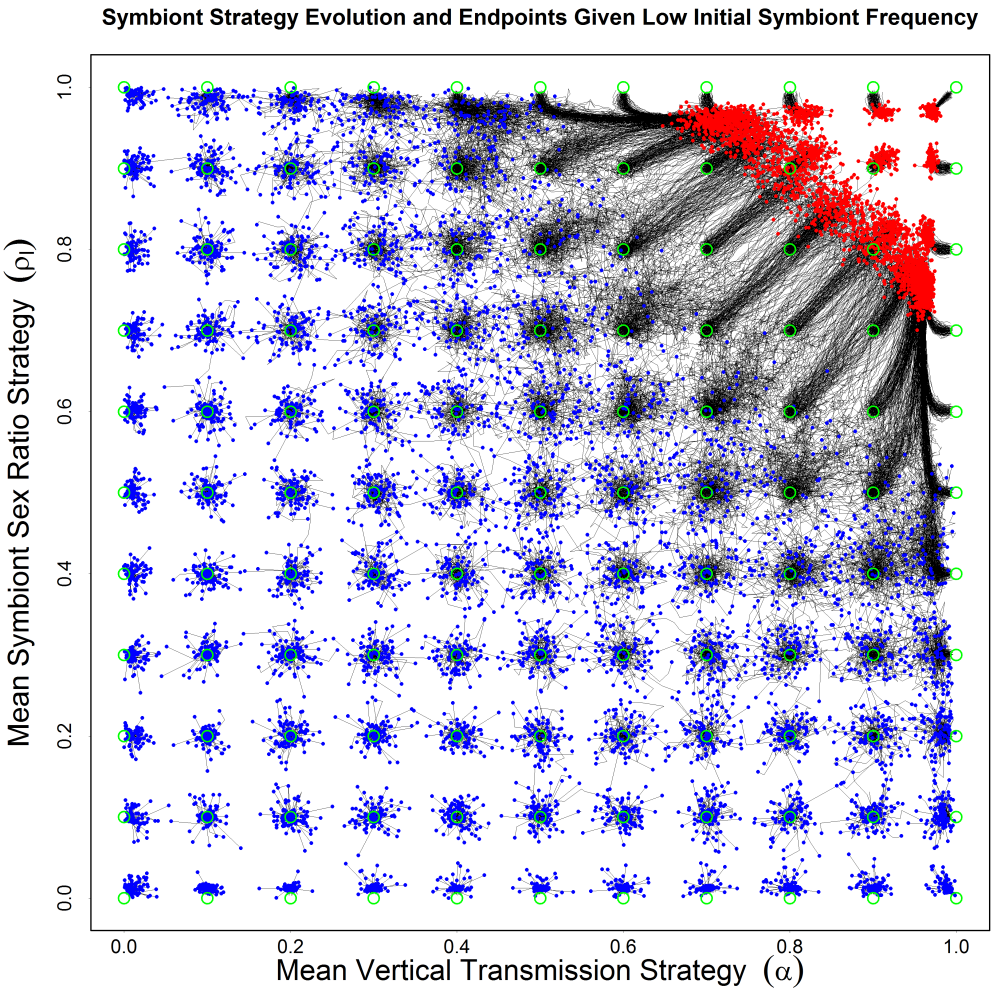

Figure 4.9: Trajectories of symbiont strategies, α and ρ_I, with 100 repeats for parameter values over the entire parameter space of ρ_I and α. Once again, we set $\rho_U = 0.5, b_I = 1.2, b_U = 1.1$, and $d = 1$. Red points indicate extinction of the host population, blue points indicate extinction of the symbiont, and green circles indicate initial conditions; no population reaches equilibrium, since either the symbiont dies out due to starting with too unexploitative strategies, or the host population dies out since the symbiont becomes too exploitative.

ulation crashes. Thus, the results of our stochastic and deterministic approaches corroborate with one another.

Therefore, we reach the conclusion that when symbiont strategies are allowed to evolve in an unstructured host population, they will either fail to invade the host population or they will exploit the host, distorting sex ratios to the extent that the host will go extinct. In the next section, we introduce population structure and dispersal to the model, and discuss further results.

4.5.3 Structured Population Results

As demonstrated by our simulation in an unstructured population, an invading symbiont will evolve to overexploit its host; this is an example of the tragedy of the commons. We hypothesise that introducing population structure and dispersal to our model will add a component of between-group selection which may favour less exploitative symbionts, or, conversely, increase within-group relatedness. The population is now organised into a 5x5 grid, holding 25 demes (or patches) with sub-population size 100. Every D_N generations, a proportion D_P of each deme is chosen and distributed uniformly at random amongst the eight neighbouring demes. We now adapt population regulation to act on each deme within the metapopulation, retaining the same method as before (see equation 4.8), but altering sub-population sizes. Specifically, members of the deme at position $\{i, j\}$, with sub-population size $N_{\{i,j\}}$, will have their fitness discounted by the fitness discounting function $R_{\{i,j\}}$:

$$R_{\{i,j\}} = \exp\{-2\log(r)\frac{N_{\{i,j\}} - 100}{100}\} \tag{4.9}$$

In words, this function states that for every 50 individuals over or below the neutral deme size of 100, every individual's fitness will be, respectively, divided or multiplied by the regulation parameter r. This shall be set to 1.5 for the first test, then subject to investigation.

We also now limit the strategy set for symbionts, to $\rho_I \in \{0.5, 0.75\}$ and $\alpha \in \{0.75, 1\}$. Thus, symbionts are either sex-ratio-disorting or not sex-ratio-distorting, and vertically transmit either perfectly or imperfectly. This is computationally more

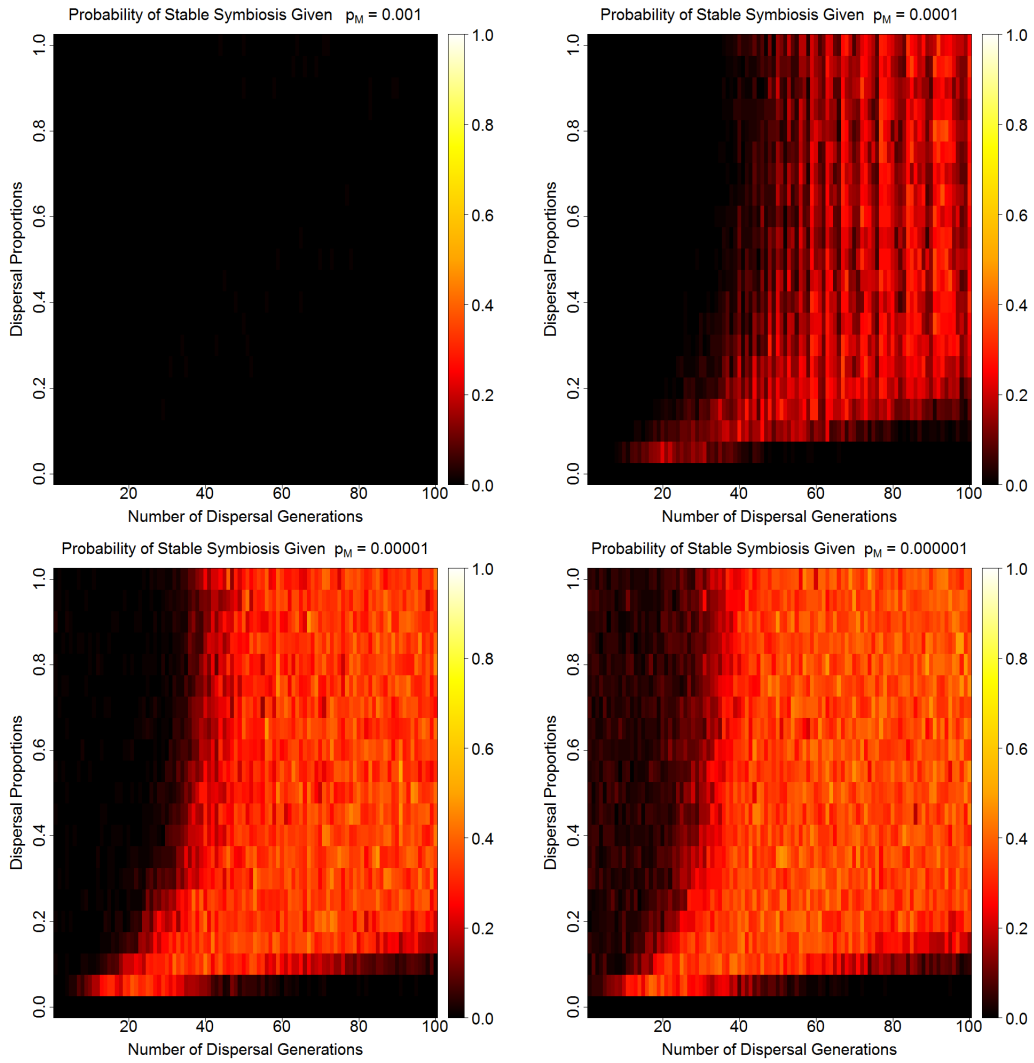

Figure 4.10: Probabilities of stable symbiont invasion given the number of generations between dispersal of a given proportion of each patch; the mutation rates considered in A, B, C, and D are, respectively, $10^{-3}, 10^{-4}, 10^{-5}$, and 10^{-6}. We also set $\rho_U = 0.5, b_I = 1.1, b_U = 1, d = 1$, and start with 1% infected hosts with symbionts equally likely to have ρ_I set to 0.5 or 0.75, and α set to 0.75 or 1. Equilibria involving stable frequencies of symbionts are reached with sufficiently low mutation, and sufficiently high numbers of generations between dispersal and dispersal proportions.

efficient, and simplifies our analysis, but will be relaxed in subsequent analyses. Mutations occur with rate p_M on each locus; when hosts are born that have vertically inherited a symbiont, the symbiont's α and ρ_I strategies each mutate with probability p_M..

We start by investigating the effects of the mutation rate p_M, the time between dispersal D_N, and the proportion of demes dispersed, D_P, choosing 100 repeats of each experiment given the choices of dispersal parameters, and evaluating the probability that the population survives to 1000 iterations with unexploitative symbionts present. Figure 4.10 shows these results. When p_M is too high, no choice of the dispersal parameters can allow the population to survive. By further investigation of specific experiments, we can see that every deme faces a problem in this scenario; mutation happens so often that symbionts that distort sex ratios and perfectly vertically transmit arise too often. These then proliferate, increasing in frequency to the extent that they bias the sex ratio so much that the deme dies out.

However, as the mutation rate decreases, exploitative symbionts become rarer; in populations where they are absent, it on average takes longer for them to arise. Non-exploitative symbionts, with $\rho_I = 0.5$ and $\alpha = 1$, are able to invade demes and persist because they provide a fitness benefit in the form of a higher birth rate. When dispersal occurs, they are given the opportunity to re-seed empty patches where exploitative symbionts have caused a host population to go extinct. However, this can only occur if D_N is sufficiently high; if it is not, then there may not be time for exploitative symbionts to increase in frequency and destabilise the host population before dispersal occurs, thus they are able to spread to nearby demes. This opportunity is smaller when D_P is low, which explains why the symbiont and host can coexist when D_N and D_P are both low.

Therefore, we choose $D_N = 50, D_P = 0.15$, and $p_M = 10^{-4.5}$ for the following experiments; the mutation parameter is chosen to be low enough that the symbiont can invade, yet high enough to test robustness of equilibria to mutations. Next, we consider how successful invasion of the symbiont is dependent on the infected birth rate, b_I, and the regulation parameter, r; these results are depicted in Figure 4.11. We also extend the set of possible strategies for symbionts to $\rho_I \in \{0.5, 0.75, 1\}$,

and $\alpha \in \{1/3, 2/3, 1\}$, in order to test that results are robust to even greater sex-distorting symbiont possibilities, and poorer vertical transmission possibilities.

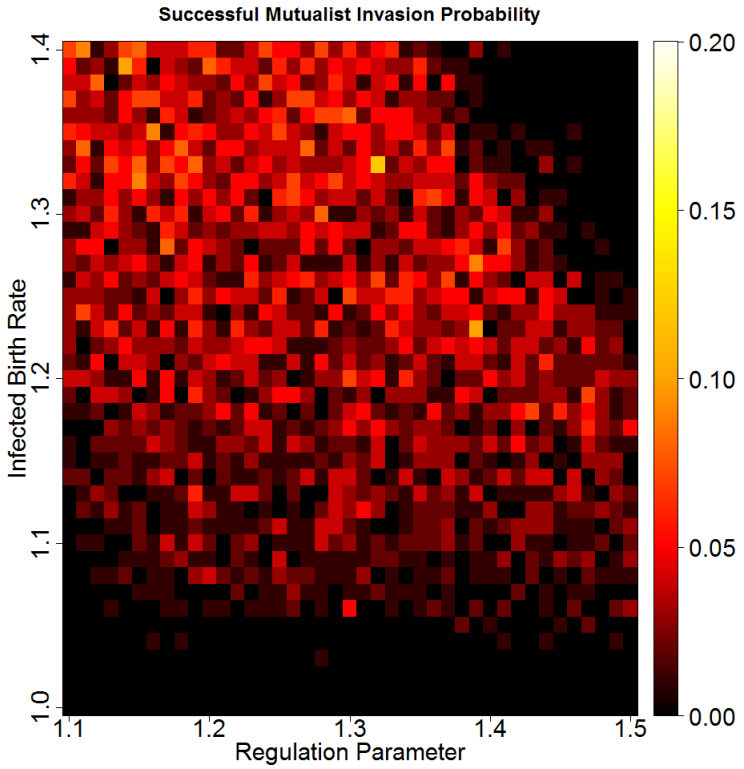

Figure 4.11: Probabilities of stable symbiont invasion given the regulation parameter, r, and the infected birth rate, b_I. Initially, symbionts are rare, vertically transmit infrequently, and rarely distort sex ratios. We also choose $b_U = 1, d = 1, p_M = 10^{-4.5}, D_N = 50$, and $D_P = 0.15$. Symbionts are more likely to invade with higher birth rate, so long as the regulation parameter r is chosen such that demes of sex-distorters cannot survive.

Firstly, we note that a higher birth rate in general aids the stable presence of an unexploitative symbiont. This is hardly surprising; b_I essentially corresponds to a fitness benefit to infected individuals. However, there is still the risk of exploitative symbionts arising, and this is reflected by the fact that the symbiont is never present at equilibrium when the birth rate is less than 1.05; here, the symbiont dies out because it starts at such a low frequency; for it to proliferate within a given patch, it takes far more generations, and so the risk of an exploitative symbiont arising is greater. We can also note that the regulation parameter r must match b_I somehow; if both parameters are high, for example, then an exploitative symbiont can arise, invade a deme, and not kill the deme off because the combined benefit of an increased birth rate and fitness gains through the fitness discounting function are enough for the exploitative symbiont to remain, and infect other patches.

However, we have found many cases here where an unexploitative symbiont successfully invaded a population. Figure 4.12 provides one such example, taken from the wider analysis of Figure 4.11.

In Figure 4.12, we can see that the dynamics previously discussed are in evidence. Symbionts are initially rare, but the exploitative types usually invade and kill off patches where they are initially present. By the first round of dispersal, the mean α strategy is 1, and the mean ρ_I strategy is 0.5. Each successive round of dispersal is followed by an increase in the overall size of the population, until every patch has been re-seeded. Occasionally, exploitative symbionts arise, as shown by the mean symbiont ρ_I strategy occasionally increasing above 0.5, but each mutation does not successfully invade.

Thus, we have found a scenario in which unexploitative symbionts can invade a host population, and remain in an equilibrium stable to the invasion of sex-ratio-distorting symbionts.

4.5.4 Discussion

We have considered a simple system of maternally-transmitted symbionts in a sexual host population, in which the symbionts can manipulate sex ratios of the hosts, and vertically transmit to offspring of infected hosts. Our aim in modelling maternally-transmitted symbionts in such a way was to investigate a possible pathway for the evolution of mutualism, itself a necessary condition for an egalitarian major transition between species.

We started by considering a deterministic model, in which symbiont strategies cannot evolve. We found that excessive sex-ratio distortion by the symbiont leads to a host population crash, and that the coincidence of fitness interests between the host and symbiont increases as vertical transmission of the symbiont approaches 1.

Next, we introduced a stochastic model, which allowed symbiont strategies to evolve; we found that in the absence of population structure, symbionts evolve to be more and more exploitative until the host population goes extinct. However, when

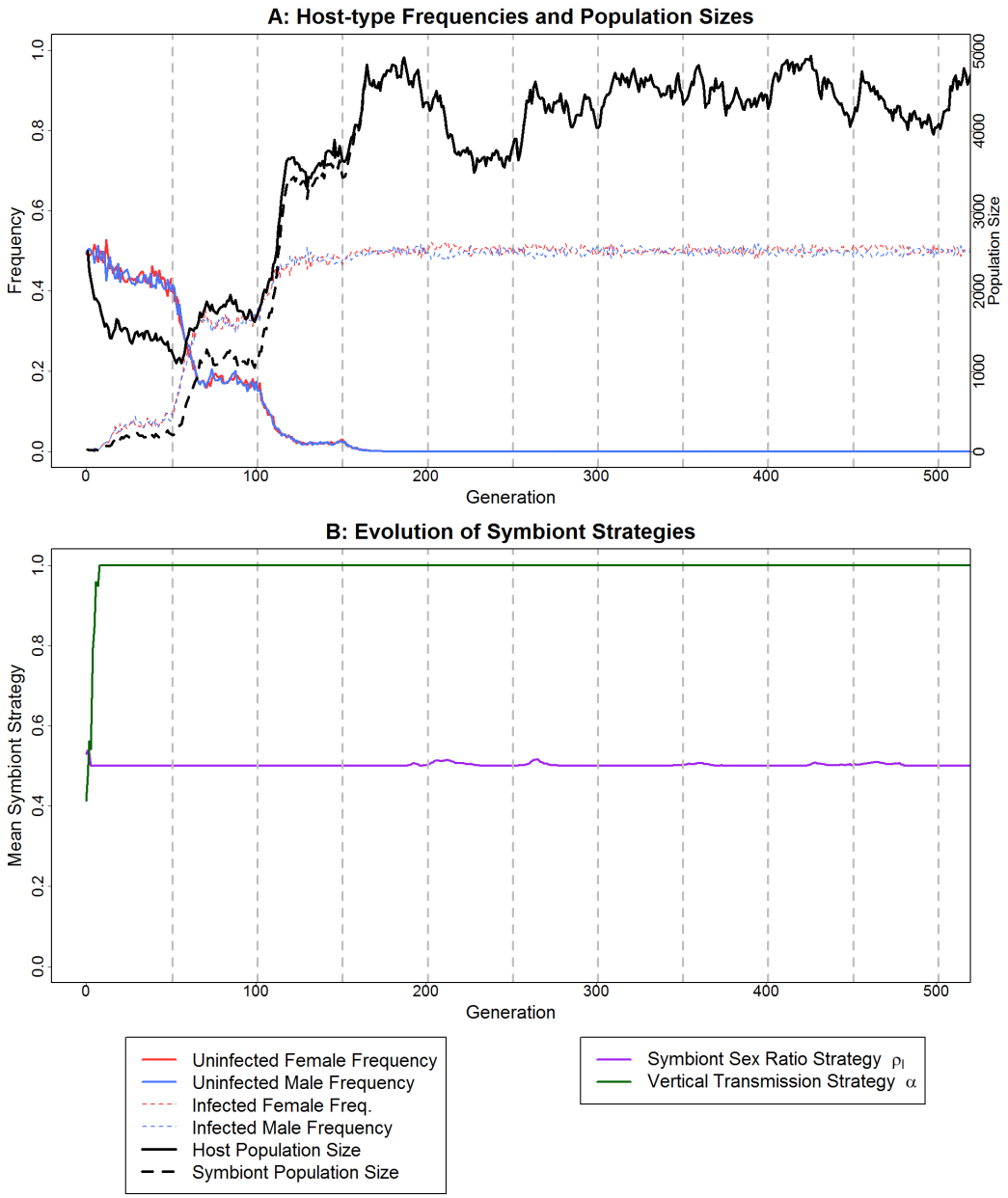

Figure 4.12: Evolution of host-type frequencies and population size (A), and symbiont strategies (B); results are displayed for 500 generations of a 1000-generation simulation, for reasons of legibility. Initially, symbionts are rare, vertically transmit infrequently, and rarely distort sex ratios. In this example, $r = 1.14, b_I = 1.36, b_U = 1, d = 1, p_M = 10^{-4.5}, D_N = 50$, and $D_P = 0.15$. There are 50 generations between dispersal (marked by vertical dotted lines) of proportion 0.15 of each population. However, they successfully reach fixation within the host population.

population structure and dispersal are introduced, between-group selection favours unexploitative symbionts, which under many circumstances can persist in a stable equilibrium.

This is an important finding; it suggests that with host population structure, a symbiont which provides a fitness benefit can successfully invade, since between-group selection can select against symbiont exploitation of the host. Thus, we have modelled a potential pathway to mutualism.

We now focus on comparisons between our model and those already existing in the literature, summarised in Chapter 2. Specifically, we discuss how our model adds to the current literature.

Like Yamamura (1993), we wanted to consider alignment of fitness interests between a host and symbiont. The context of our model is significantly different from that of Yamamura, who modelled mixed vertical and horizontal transmission, without sex-ratio distortion. Therefore, the host-symbiont relationship considered by Yamamura focuses on a less established host-parasite relationship; indeed, he argues that horizontal transmission is important to the initial spread of a parasite. The result that a coincidence of fitness interests between a host and symbiont can be achieved with high vertical transmission is similar to that of Yamamura, who found a threshold level of vertical transmission above which fitness interests between a host and parasite are aligned. However, our similar result comes about in a model which includes sex-ratio distortion, an additional means of symbiont exploitation of a host; thus, we have shown that the possibility of a coincidence in fitness interests between a host and symbiont is robust to the addition of even greater potential for host exploitation.

Another model which ours bears comparison to is that of Werren (1987). Werren focussed on the conflict between a maternally-transmitted sex-ratio gene, and a sex-ratio gene transmitted by both genders. This is analogous to the conflict over sex-ratio between host and symbiont in our model, as discussed in Section 2.3.4. Werren's analysis focuses on the evolution of compensatory sex-ratios; our model, which accounts for fitness costs associated with population-wide sex-ratio distortion,

differs in its results from both Werren's model, and that of Hatcher and Dunn (1995).

By including the possibility for symbiont strategies to mutate in our stochastic simulation, we have shown that a structured host population may survive in the presence of a potentially sex-distorting symbiont. This differs from results obtained by Hatcher et al. (2000), in which survival of the host population was solely dependent on patch turnover, with the symbiont invading and then perpetuating itself to other patches before killing off individuals in that patch; meanwhile, the host repopulates empty patches. Our model, by contrast, allows complete penetration of the symbiont in the host population and peaceful coexistence, with any sex-distorting symbionts that arise failing to perpetuate themselves either through randomness, or eventually killing off the population in their home patch. This is the primary result of our work, since it illustrates a potential pathway from initial parasitism based on sex-distortion to mutualism.

Our model is intended to involve a general form of sex-distortion, so should reflect the biological reality of each form of sex-distortion considered in Chapter 2 (cytoplasmic incompatibility, thelyokonous parthenogenesis, feminisation, and male-killing). For example, male-killing may be reflected in our model through parameter choices of high sex-ratio distortion, and fitness adjustments that account for the benefit to females of less kin competition, and the loss of offspring through male-killing itself. Cytoplasmic incompatibility may be an exception here, since it involves a mechanism which may more subtly affect host population dynamics. However, our analysis has focused only on the case in which symbionts provide a fitness benefit to a symbiont; as noted in Chapter 2, various strains of the bacterial endosymbionts *Spiroplasma*, *Wolbachia*, and *Cardinium* have all been known, or suspected, to confer such fitness benefits, for example through conferring parasite resistance to hosts (West, 2009; Walker et al., 2011; Xie et al., 2014).

Chapter 5

Conclusions and Future Directions

5.1 Conclusions

The original work we have presented in the modelling chapters in this thesis has focussed on major transitions between species; specifically, we have tackled the question of whether or not altruism can evolve between species (Chapter 3), and set out a potential pathway towards mutualism between a host species and a sex-distorting endosymbiont (Chapter 4).

In Chapter 3, we introduced a general model of between-species donation, and applied deterministic and stochastic approaches to analysing it. The deterministic approach involves an assumption of quasi-linkage equilibrium, which entails weak selection, while the stochastic approach relaxes this assumption and introduces stochasticity. Results found using the deterministic approach were confirmed through the stochastic approach, with some minor discrepancies discussed; largely, results were robust to the simultaneous introduction of strong selection and stochasticity.

We found that the donation behaviour was vulnerable to modifiers suppressing such behaviour arising on unassorted loci, presenting a similarity with greenbeard traits. In addition, we argued that the donation behaviour in fact amounts to within-species altruism, with donated fitness benefits travelling to members of the same species, using the other species as a vector for this altruism.

Within the context of the literature, as surveyed in Chapter 2, our conclusions

reflect those of Wyatt et. al., who contrasted the different conclusions provided by regression analyses of fitness in their model of between-species donation Wyatt et al. (2013). They showed that solely using same-species genetical predictors diagnosed the social behaviour as within-species altruism, whereas using mixed-species genetical predictors led to the social behaviour being identified as between-species altruism. A causal path from donated benefits to donation received by related individuals exists, suggesting that our model may fit with the first of these conclusions; however, we do not provide a rigorous test of this idea. We consider it arbitrary to terminate the path of received benefits before it reaches genetically related members of the same species.

We can also note the link between our model and that of Fletcher and Doebeli; while theirs is analysed as within-species without relatedness, it can be compared to the model (d) in Chapter 3, in which donation is determined by a single locus in each species. As shown by comparison with model (a) in Chapter 3, suppression modifiers could arise to render any potential between-species altruism transient.

In Chapter 4, we introduced a general model of maternally-transmitted sex-distorting endosymbionts, specifically focusing on the potential transition of an invading parasite towards mutualism with a host species. We initially found that the evolution of sex-distortion and vertical transmission led to a tragedy of the commons, with the symbiont driving the host population towards extinction. However, by introducing a patch-structure to the population, we found that a complete coincidence in fitness interests was possible, with non-sex-distorting symbionts at fixation at equilibrium, leading to increased fitness for the host population.

We introduced several previous models of symbiosis in the context of sex-distortion in Chapter 2. Our model successfully incorporates the feature arising in Yamamura's model, of a coincidence in fitness interests with high vertical transmission Yamamura (1993); at equilibrium, potential sex-ratio distorters are selected not to distort sex-ratio. We have significantly extended Werren's model of general sex-ratio distortion, focussing on the evolution of the symbiont rather than coevolution with the host Werren (1987). Like Hatcher et. al., who were analysing feminisation in an adapted version of Werren's model, we have utilised population structure Hatcher

et al. (2000); unlike Hatcher et. al., we have shown that by considering evolution of a symbiont's sex ratio strategy, a parasite can invade a host population, transitioning towards mutualism and reaching 100% prevalence.

5.2 Future Directions

As noted in Chapter 1, the process of a major transition can be split, broadly speaking, into three components; social group formation, maintenance, and transformation Bourke (2011). Current research tends to focus on social group formation (indeed, our two modelling chapters focus on the spread of social behaviours) or social group maintenance; less research has focussed on social group transformation Bourke (2011). We now briefly discuss several ideas relating to this area which are worthy of future research.

5.2.1 The Size-Complexity Hypothesis

The final stage in an evolutionary transition involves the emergence of a social group as an individual in its own right. One theory of how this occurs in the context of fraternal transitions is the 'size-complexity' hypothesis Bourke (2011). The size-complexity hypothesis states that there are feedback loops between group size and various characteristics known as 'complex' group features, which are associated with the completion of a major transition. External drivers may act to select for increased group size, launching a feedback loop, resulting in further increases to group size and complexity. Complex group features include those associated with reproductive division of labour, i.e. a lack of reproductive potential for certain elements (i.e. soma/worker) of a group (the labelling here applies, specifically, to the multicellularity and eusocial transitions, respectively), segregation of reproductive particles (i.e. germline/developing queens), and earlier divergence in the development of reproductive and non-reproductive particles. Greater numbers of particle types (i.e. cells/worker castes), and a greater number of particles themselves are also complex group features.

Many positive feedback effects have been established from increased collective size to the set of complex traits, both forwards and backwards Bourke (2011). For example, Bourke notes that increased group size increases the chances of deleteri-

ous mutations, so there will be selection for traits which suppress the potential for these mutations. A segregated, early-diverging germline is one such trait, since this decreases the number of cell divisions, thus decreasing the potential for mutation Bourke (2011).

While Bourke focusses on the increase in complexity of groups, Birch suggests that emphasis should also be given to an associated decrease in social complexity at the level of the particle Birch (2016). Birch suggests that in order for groups to cope with more complex tasks, particles specialise further; this reflects the idea of an increase in the number of particle types that Bourke notes is an aspect of complexity Bourke (2011). Indeed, Birch details an expanded feedback loop involving increased group size, loss of complexity at the lower level and increases in complexity at the higher level Birch (2016). Specifically, Birch invokes ideas about task structure, suggesting that, in the context of a eusocial society, a pool of reserve workers can be used to replace versatility of workers as a means of retaining robustness in task completion. Greater specialisation then allows for greater efficiency, the ability to complete more complex tasks, and a reduction in complexity at the lower level. Each of these links is argued for in terms of task structure.

Birch's verbal argument suggests that further investigation of each of the specific feedback loops involved in the social-complexity hypothesis remains an avenue for future research. Indeed, to take one example left open by Birch's approach, formal modelling of task structure in relation to a major transition could be useful; in general, mathematical models relating to social group transformation have been rare Birch (2012).

5.2.2 Fitness Decoupling

One aspect of the size-complexity hypothesis which has received attention through modelling work is that of the separation of the germ and soma, which Michod and Nedelcu (2003) have studied in *Volvox* (commonly known as volvocine).

Michod and Nedelcu argue that during the initial stages of a major transition involving volvocine, two fitness components, fecundity and viability, are 'coupled' at the level of the cell; in other words, they are such that cell fitness is maximised.

However, they then suggest that if the reproductive potential of some cells is suppressed, then those cells can focus more on group survival, through survival-related functions which arise from the nature of the group. Though not directly reproducing, these survival-focussed cells can contribute to reproduction indirectly, by improving quality of offspring. Similarly, this allows the germline cells to focus more on increasing reproductive output. Once germ/soma distinction is complete, then fitness has been 'decoupled'. That is, the two fitness components that make up the MLS2 (see Chapter 1 for more detailed discussion of this term) fitness of the group are quasi-independent of one another, so can, to some extent, be maximised independently. Elsewhere, Michod has proposed that at the end of a major transition, germ/soma specialisation must be complete to the extent that every cell is either totally specialised in fecundity or viability, and thus that every cell has zero fitness Michod (2006). Thus, fitness has been completely exported to the higher level.

Implicit in this argument is the idea that there is convexity in the viability/fecundity trade-off, which Michod demonstrates in volvocines Michod (2007). This means that the second derivative of fecundity as a function of viability is positive, and, necessarily, the same holds for viability as a function of fecundity; the first derivative of both functions must also be negative. These properties are present in figure 6.1, demonstrating a convex viability-fecundity trade-off graphically. In this case, maximising one of these components will yield a greater total contribution to viability and fecundity of the group than maximising both components simultaneously in a single cell (in the case of volvocines, it is noted that the product of the two components is an accurate reflection of fitness). Michod characterises this with the following equation Michod (2007):

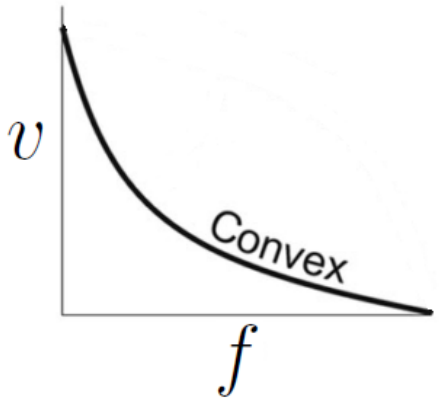

Figure 5.1: A convex fitness trade-off between viabililty and fecundity. Reproduced (with edits) from Michod (2007).

$$W = \bar{w} - \mathrm{Cov}(v, f) \tag{5.1}$$

Here, W is the collective fitness in the MLS2 sense, $\bar{w}$ is the mean cell fitness, v is the measure of viability in a single cell, and f is the measure of fecundity in a single cell. If collective members specialise to either contribute to viability or fecundity, then the covariance between v and f becomes negative and group fitness in the MLS2 sense increases.

Michod's work on germ-soma segregation has provided a mathematical condition for it to be beneficial, in the context of volvocine algae. This work appears to reflect Birch's idea of task specialisation, with the number of tasks in this case being two. Further work on other species may be necessary to assess the extent to which Michod's principles of fitness decoupling and the importance of convex fitness trade-offs can be extended.

5.2.3 An Alternative Decomposition of Collective and Particle-Level Selection

Shelton and Michod have investigated social group transformation with an alternative decomposition of selection; see Chapter 1 for discussion of the Price and contextual analysis approaches. Note that we retain the collective and particle terminology here Shelton and Michod (2014a). They introduce 'counterfactual' fitness to refer to collective fitness in the absence of collective-dependent effects, and use this to divide selection into two components, one dependent on collective membership and one independent of collective membership. They note that this decomposition is only appropriate when there is no within-collective selection on particles, as opposed to global selection on particles. Identifying global selection on particles is important in their discussion of a model of a simple cross-level by-product scenario, in which collective-level properties are simply aggregrate properties of particles.

Using their new decomposition of selection, Shelton and Michod compare a trait determining investment in fecundity or viability (a trade-off is assumed) across three models of dependence of individual fitness on emergent collective properties. Their analysis detects that the greater the dependence, the further equilibrium values of this trait deviate from a simple cross-level by-product scenario (i.e. no dependence

of individual fitness on emergent group properties). Using the model of volvocine algae, they demonstrate how this decomposition can be used to elaborate on how a group life-cycle can emerge; however, this model is not general. As a relatively new method of analysing selection in the context of social group transformation, the potential use of Shelton and Michod's decomposition in relation to contextual analysis and the Price equation should be subject to further investigation.

5.2.4 Monogamy hypothesis

Boosma hypothesised Boomsma (2009) that every eusocial lineage arose due to life-time monogamy of ancestors, which ensured that parents are equally related to siblings and offspring. This favours the evolution of altruism towards siblings rather than offspring under certain conditions; altruism towards siblings could take the form of somatic functions, and hence encourage a reproductive division of labour. An analogue of this condition for a fraternal major transition towards eusociality can be considered in terms of multicellularity; here, relatedness between siblings and offspring is equal in clonal societies; the condition of clonality has been shown to be important for the transition towards obligate multicellularity, as opposed to facultative multicellularity Fisher et al. (2013). Thus, a 'monogamy window', or a functional equivalent such as clonality among a group of cells, is a necessary (though not sufficient) condition for a fraternal ETI.

Boomsma notes that formal computational modelling may be required to investigate whether the monogamy hypothesis is oversimplistic, and could be extended to include low frequencies of double-mating or foundress association; thus, this remains an open topic within the major transitions view of evolution. However, it has been surveyed in the context of eusociality and found consistent with the data Hughes et al. (2008).

5.2.5 Organismality

Queller and Strassmann argue that 'organismality', an idea closely related to that of individuality (indeed, Queller and Strassmann refer to organismality in an attempt to avoid any implication of indivisibility), is characterised by low levels of within-organism conflict and high levels of within-organism cooperation Queller and

Strassmann (2009). While high cooperation and low conflict are linked, one is possible in the absence of the other; Queller and Strassmann give the example of human societies, which involve both high cooperation and conflict, while clones of non-social aphids show neither cooperation nor conflict Queller and Strassmann (2009). This definition subsumes all of the characteristics associated with the completion of the social group transformation phase of a major transition into two related features; it also suggests that organismality is defined by degree. This contrasts with the usual notions of individuality, which usually invoke specific assumptions of form or function, for example, spatial contiguity or germ-soma separation; intuitive counterexamples to these candidate conditions for individuality are often easily found Clarke (2010). As with the concept of the major transitions, Queller and Strassmann's framework is very general and not biased towards certain taxa West and Kiers (2009). This is another framework which could conceivably be useful in analysing future models of social group transformation.

Appendix A

Details of Further Models of Between-Species Donation

A.1 Introduction

In this section, we discuss details of further models of between-species donation, introduced in Chapter 3. The primary models we discuss in the main body of Chapter 3 are labelled (a) and (b), and refer to the basic model of between-species donation, with assortment covering only the donation locus in model (a), and the entire genome in model (b). Brief mention is made of four further models, which we now cover in greater detail. We start by providing a recap of models (c-f).

Model (c) involves genome-wide assortment, but, in contrast to models (a) and (b), individuals are uniformly at random assigned a single role in their lifetime, of potential donor or potential recipient. Individuals are either potential donors, with probability 0.5, or potential recipients, with probability 0.5. If they are potential donors, there are two possibilities; they either donate to the social partner of the other species, incurring a cost of c fitness if they are donators, or, if they are non-donators, they do not donate, in which case there is no fitness adjustment. If an individual is a potential recipient, they either receive b fitness if their social partner is a donator, or, if their partner is a non-donator, there is no fitness adjustment.

Model (d) is a single-locus model, in which individuals are donators if and only if they possess the donation allele; the suppression locus is entirely absent. Accordingly, assortment takes place only on the donation locus.

Model (e) considers a single population with genome-wide assortment; this is similar to model (b), with the exception that individuals are now paired up within the same species.

Model (f) involves genome-wide assortment, but donators no longer donate unconditionally; instead they only donate if their social partner is also a donator. This is similar to model (b), but the donation behaviour is now conditional on social partner phenotype.

A.2 Model Details

In Chapter 3, we use a deterministic approach; as described there, the details of assortment and donation are captured in the selection coefficients. Thus, to describe how the deterministic approach is implemented with regards to these four additional models, we run through construction of the selection coefficients; all general details relating to the four models are located in Chapter 3. Note that all notation used here follows from that used in Chapter 3.

A.2.1 Model (c): The Single-Role Model

For model (c), assortment parameters will be the same as those for model (b); in other words, they are described by equations (3.3,3.4,3.8). This is because, similarly, there is genome-wide assortment across two populations. However, since the donation behaviour now occurs on average half the time it would have occurred in model (b), fitness equations are now as follows:

$$W_i(0,0) = w_0 + \frac{P_i((0,0),(1,1))b}{2} \tag{A.1}$$

$$W_i(0,1) = w_0 + \frac{P_i((0,1),(1,1))b}{2} \tag{A.2}$$

$$W_i(1,0) = w_0 + \frac{P_i((1,0),(1,1))b}{2} \tag{A.3}$$

$$W_i(1,1) = w_0 + \frac{P_i((1,1),(1,1))b - c}{2} \tag{A.4}$$

Consequently, selection coefficients are calculated as follows:

$$a_{1,i} = \frac{\tilde{b}(P_i((1,0),(1,1)) - P_i((0,0),(1,1)))}{2+\tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} + p_{2,i}a_{\{1,2\},i} \tag{A.5}$$

$$a_{2,i} = \frac{\tilde{b}(P_i((0,1),(1,1)) - P_i((0,0),(1,1)))}{2+\tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} + p_{1,i}a_{\{1,2\},i} \tag{A.6}$$

$$a_{\{1,2\},i} = \frac{\tilde{b}\big(P_i((1,1),(1,1)) + P_i((0,0),(1,1))\big)}{2+\tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} + \frac{\tilde{b}\big(-P_i((1,0),(1,1)) - P_i((0,1),(1,1))\big) - \tilde{c}}{2+\tilde{b}f_{i'}(1,1) - \tilde{c}f_i(1,1)} \tag{A.7}$$

The expressions describing the maximum values of $\tilde{b}$ and $\tilde{c}$ required to satisfy the first requirement of QLE turn out to be similarly scaled; our conditions are $\tilde{b} \ll 1$ and $\tilde{c} \ll (2-\tilde{b})/2$. Since $a \leq \max(\frac{2\tilde{b}-\tilde{c}}{2-\tilde{c}}, \frac{\tilde{b}+\tilde{c}}{2-\tilde{c}})$ and $D_i \leq 1/4$, we have that:

$$D_i < \frac{4a}{\max(\frac{2\tilde{b}-\tilde{c}}{2-\tilde{c}}, \frac{\tilde{b}+\tilde{c}}{2-\tilde{c}})}$$

Thus, the second condition for QLE is satisfied. As before, we can find a condition for donation or non-suppression to be selected for which depends on some function of α and genotype frequencies being larger than c/b, but it is not tractable. The only thing that we can conclude from it is that the c/b ratio is crucial once again.

A.2.2 Model (d): The Single-Locus Model

In this case, we can simply use the equations (3.3,3.4,3.8) for our assortment equations; the notation used to carry out genome-wide assortment also applies when there is only one locus. However, the set of genotype here is simply $G = \{0,1\}$. This does change exactly how we specify equations for fitnesses of individuals:

$$W_i(0) = w_0 + P_i(0,1)b \tag{A.8}$$

$$W_i(1) = w_0 + P_i(1,1)b - c \tag{A.9}$$

As there is only one locus in this model, there is no linkage disequilibrium. Thus,

the assumptions of QLE do not need to be made; the same approach based on the Kirkpatrick et al. (2002) method may be followed. Since there is a single locus, we must adapt notation; here, a_i relates to the selection coefficient on the donation locus in the ith population, and p_i refers to the donation allele frequency in the ith population. Using equations (3.9,A.8,A.9), we find the single selection coefficient for each population i, relating to the only locus present, as follows:

$$a_i = \frac{\tilde{b}(P_i(1,1) - P_i(0,1)) - \tilde{c}}{1 + p_{i'}\tilde{b} - p_i\tilde{c}} \tag{A.10}$$

Firstly, we can see that if $p_i = p'_i$, then $a_i > 0$ is equivalent to $\alpha > c/b$. In other words, if the two allele frequencies are equal, then the donation allele, and, consequently, donators, will receive positive selection if α is sufficiently high. Since $0 < c < b$, we know that this is always satisfied when $\alpha = 1$, i.e. when there is maximum association. When $p_i \neq p'_i$, we can once again find a condition for donators to receive positive selection which is dependent on a boundary c/b; as before, the other side of the expression is intractable and does not lead to any further conclusions.

A.2.3 Model (e): The Single-Species Model

Since this model entails a single population, the frequencies of genotypes will be the same in the two interacting groups of donors and recipients. Thus, we can essentially take the equations (3.3,3.4,3.8), but remove the i subscripts where they occur, since we do not have to account for two populations. As a result, we obtain the following equations describing associations in model (e):

$$\beta(\alpha, u, v) = \alpha I_{\{u=v\}} f(u) \tag{A.11}$$

$$g(u) = (1-\alpha) f(u) \tag{A.12}$$

$$P(u, v) = \alpha I_{\{u=v\}} + (1-\alpha) f(v) \tag{A.13}$$

Expressions for average fitnesses remain similar to those for model (b), though once again we remove subscripts:

$$W(0,0) = w_0 + P((0,0),(1,1))b \tag{A.14}$$

$$W(0,1) = w_0 + P((0,1),(1,1))b \tag{A.15}$$

$$W(1,0) = w_0 + P((1,0),(1,1))b \tag{A.16}$$

$$W(1,1) = w_0 + P((1,1),(1,1))b - c \tag{A.17}$$

We can then calculate the selection coefficients using equations (3.9,A.11-17); note that we denote the frequency of the donation allele as p_1, and the frequency of the non-suppression allele as p_2, since there is only one population:

$$a_1 = p_2 a_{\{1,2\}} \tag{A.18}$$

$$a_2 = p_1 a_{\{1,2\}} \tag{A.19}$$

$$a_{\{1,2\}} = \frac{\alpha\tilde{b} - \tilde{c}}{1 + (\tilde{b} - \tilde{c})f(1,1)} \tag{A.20}$$

To satisfy the first requirement of QLE, we must have that $a \ll 1$. We can see that $a = \max|a_{\{1,2\}}|$, since p_1 and p_2 are between 0 and 1. We can also see that a is at a maximum when $f(1,1) = 0$. Thus, we must have that $\alpha\tilde{b} - \tilde{c} \ll 1$ and $\tilde{c} - \alpha\tilde{b} \ll 1$. These expressions can be reduced to the conditions $\tilde{b} \ll 1$ and $c < b$. Since $a \leq (\alpha\tilde{b} - \tilde{c})$, we can deduce that $D \leq 4a/(\alpha\tilde{b} - \tilde{c})$, thus satisfying the second requirement for QLE.

We can analytically find a condition for the evolution of donation by using equations (3.9,A.18,A.20). Using, once again, z_2 to denote the value of the allele on the suppression/non-suppression locus, we find the following condition for positive selection on the donation allele:

$$\begin{aligned} & \bar{W}a_{1,i} + a_{\{1,2\},i}(z_2 - p_2) > 0 \\ \iff \quad & \bar{W}a_{\{1,2\},i}z_2 > 0 \end{aligned}$$

For an individual with $z_2 = 0$, the allele on the donation locus does not change the individual's fitness, but when $z_2 = 1$, we find the condition $\alpha > c/b$ for the

donation allele to be beneficial. We can similarly deduce the same condition (using equations (3.9,A.19,A.20)) for the evolution of non-suppression, since the two loci have the same role in the system. Thus, the condition for donators to receive positive selection in the absence of stochasticity is $\alpha > c/b$. Note that this is equivalent to the same condition for positive selection of non-suppression found for model (a).

A.2.4 Model (f): The Partner Rejection Model

This model involves genome-wide assortment, thus we start with the same equations relating to assortment as model (b): equations (3.3,3.4,3.8). However, the introduction of partner rejections changes the fitness equations. Now, no matter what genotype non-donators are partnered with, they will not receive any benefit from the donation behaviour; nor will they pay any cost, by virtue of them not being donators. Donators, on the other hand, will only donate when paired with other donators; thus, when paired with another donator they gain b fitness from receiving donated benefits, and lose c fitness from donating themselves.

$$W_i(0,0) = w_0 \tag{A.21}$$

$$W_i(0,1) = w_0 \tag{A.22}$$

$$W_i(1,0) = w_0 \tag{A.23}$$

$$W_i(1,1) = w_0 + P_i((1,1),(1,1))(b-c) \tag{A.24}$$

Therefore, the selection coefficients are as follows:

$$a_{1,i} = p_{2,i} a_{\{1,2\},i} \tag{A.25}$$

$$a_{2,i} = p_{1,i} a_{\{1,2\},i} \tag{A.26}$$

$$a_{\{1,2\},i} = \frac{(\tilde{b}-\tilde{c})P_i((1,1),(1,1))}{1+(\tilde{b}-\tilde{c})f_i(1,1)P_i((1,1),(1,1))} \tag{A.27}$$

Thus, $a \leq (\tilde{b}-\tilde{c})$, and so for $a \ll 1$, we must have $\tilde{b} \ll 1+\tilde{c}$. Recall that we also have $c < b$, so $\tilde{b} \ll 1$ is sufficient to satisfy the first condition of QLE. Since $D_i \leq 1/4$, we satisfy the second condition of QLE, since:

$$D_i \leq \frac{4a}{\tilde{b}-\tilde{c}}$$

Using equations (3.9,A.25,A.27), a condition for donation to receive positive selection in population i is:

$$\begin{aligned} & a_{\{1,2\},i} z_2 > 0 \\ \iff \quad & z_2(\tilde{b}-\tilde{c}) P_i((1,1),(1,1)) > 0 \end{aligned}$$

This is always positive so long as $z_2 = 1$ and the frequency of donators in the other population is non-zero (which is always the case in the deterministic approach unless they are initially set to 0). Note that $\tilde{b} - \tilde{c} > 0$ necessarily since $b > c$. The same condition applies for non-suppression with z_1 replacing z_2. Thus, the donation and non-suppression alleles will always be equal or beneficial to individuals in a population, so will always be selected for so long as the frequency of donators in the other population is non-zero.

A.3 Results

We now display numerical examples of each model; these are designed to be compared to similar figures in Chapter 3. We display results for both the deterministic and stochastic approaches within each subsection. The stochastic approaches to each model require no further set-up than that given for models (a) and (b) in Chapter 3; however, note that we use smaller values of N and n_{SUS} than those seen in the main body of the thesis. This is because of the long computation times entailed by higher N and n_{SUS}. The primary motivation for the stochastic approach is to check the robustness of the results in Chapter 3 to the simultaneous introduction of stochasticity and strong selection.

A.3.1 Model (c): The Single-Role Model

Figure A.1 shows a numerical example of this model, using the same initial conditions and parameters as Figure 3.3 in the main body of the thesis. The clear result

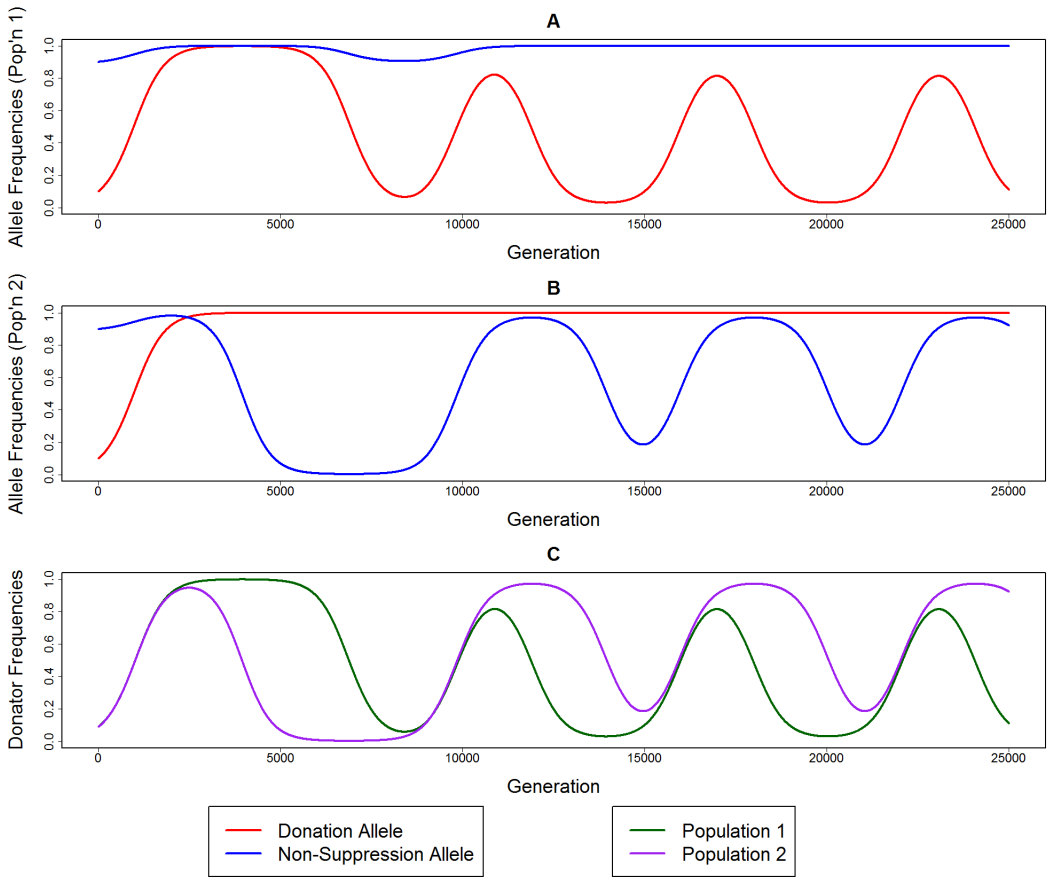

Figure A.1: Evolution of allele frequencies (A/B), and donator frequencies (C) (using the deterministic approach with model (c), in which assortment is genome-wide and individuals have single lifetimes roles); initial frequencies are $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1$, and $\alpha = 1$. Individuals being assigned a single role has no effect on the evolutionary outcomes (see Fig. 3.3) beyond halving the selection strength.

here is that selection strength is halved.

Figure A.2 shows a corresponding example utilising the stochastic approach. There are superficial differences between Figure A.2 and both Figures A.1 and 3.8. However, the primary finding that donator frequencies fluctuate remains evident. The difference between allele frequencies in Figure 3.8 and those displayed here are pronounced; rather than donation behaviour being under the control of one locus in each species, changes in the frequencies of alleles on both loci have an effect on the frequencies of donators in population 1 in Figure A.2. One possible reason for this is the smaller population size than that used in Figure 3.8, resulting in greater stochasticity. The selection strength being halved may also contribute to stochasticity having a greater effect. However, the primary result of this figure is that donator frequencies fluctuate, following a similar pattern to that evident in Figures 3.3, 3.8, and A.1. Thus, these results are robust to the introduction of

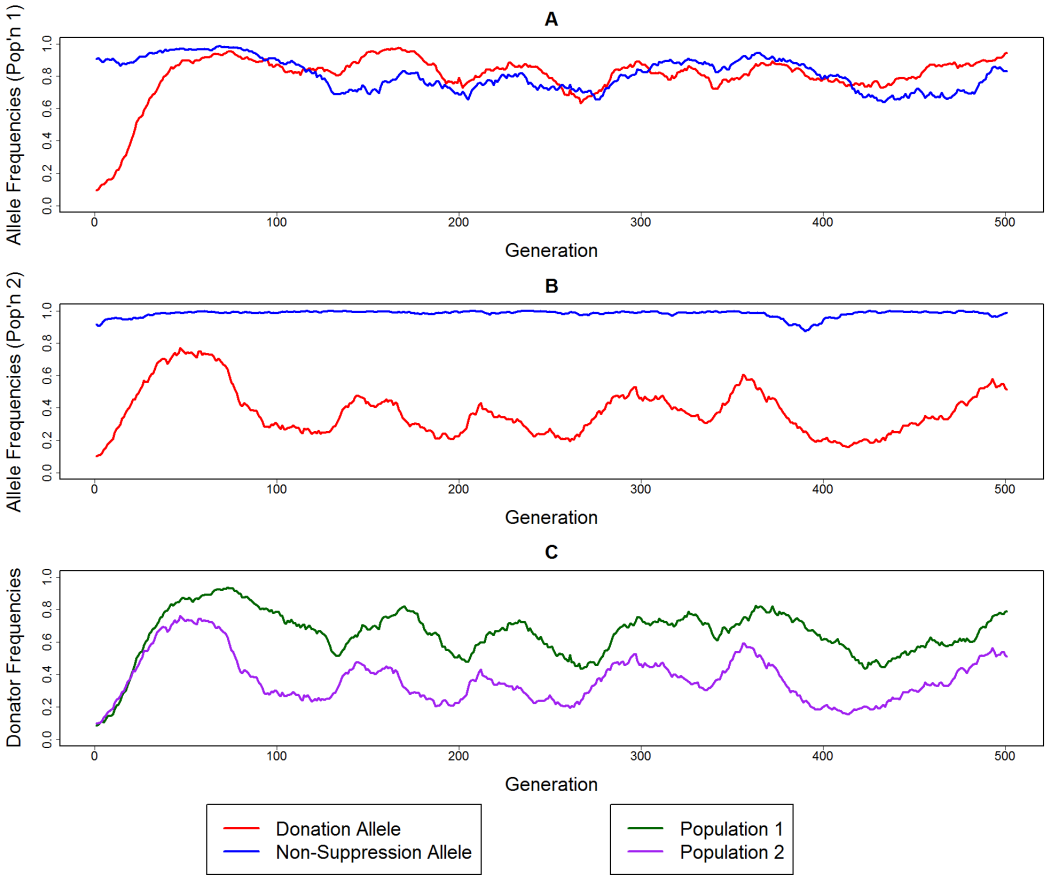

Figure A.2: Evolution of allele frequencies (A/B), and donator frequencies (C) (using the stochastic approach with model (c), in which assortment is genome-wide and individuals have single lifetimes roles); initial frequencies are $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 1, c = 0.5, w_0 = 1$, and $\alpha = 1$. In addition, $p_m = 0.001, N = 1000, n_{SUS} = 200$. This figure bears comparison to Fig.s A.1 and 3.8; we can see that the selection strength appears halved, and stochasticity has not significantly altered the outcome from that of A.1.

stochasticity, strong selection, and single lifetime roles, which in fact simply the adjust the strength of stochasticity and selection further.

A.3.2 Model (d): The Single-Locus Model

Figure A.3 shows three examples of changes in donator frequencies in the single-locus model. Note that donators here are simply bearers of the donation allele. The mathematical analysis of this model showed that the donation allele receives positive selection when donator frequencies are the same between populations and $\alpha > c/b$; these examples illustrate that donator frequencies may converge to one another when initial donator frequencies are equal, then receive solely positive selection. Thus, the condition that $\alpha > c/b$ is required for donators to increase in frequency may apply more than initially thought. This reflects the results of model (b), where genome-wide assortment in the two-locus model led to intermediate levels

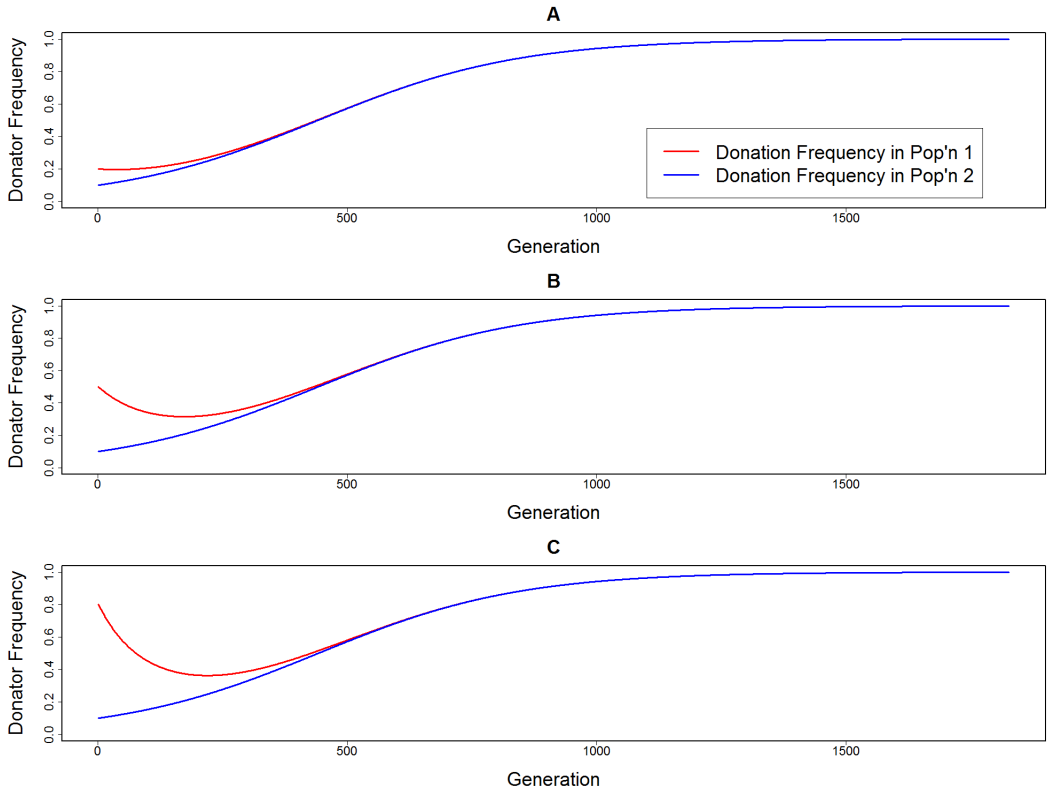

Figure A.3: Evolution of donation allele frequencies (using the deterministic approach with model (d), in which donation is governed by a single locus); initial frequencies are $p_1 = 0.2$, 0.5, and 0.8 in A, B and C respectively, $p_2 = 0.1$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1$, and $\alpha = 1$. Donation reaches fixation when $\alpha > c/b$, even when initial donator frequencies differ between populations.

of donation when $\alpha > c/b$; both of these results resemble Hamilton's rule.

Figure A.4 shows the same example as Figure A.3, with the introduction of stochasticity; notably, results are different between the two models, with cyclic dynamics much more akin to those found in the two-locus case being present. It may be illustrative here to consider equation (3.29), and the insight that fitnesses of donation and non-donation need only be equal within species for there to be no selection on either trait; thus, there may be equilibria, like those in the two-locus model (b), which entail different frequencies of donation between species. The equilibria shown in Figure A.3 appear to correspond to the type 2 equilibrium shown in Figure 3.5, which entail equal donation between species, whereas the cyclic dynamics that appear in Figure A.4 may reflect stochasticity interrupting the type 1 equilibrium shown in Figure 3.5.

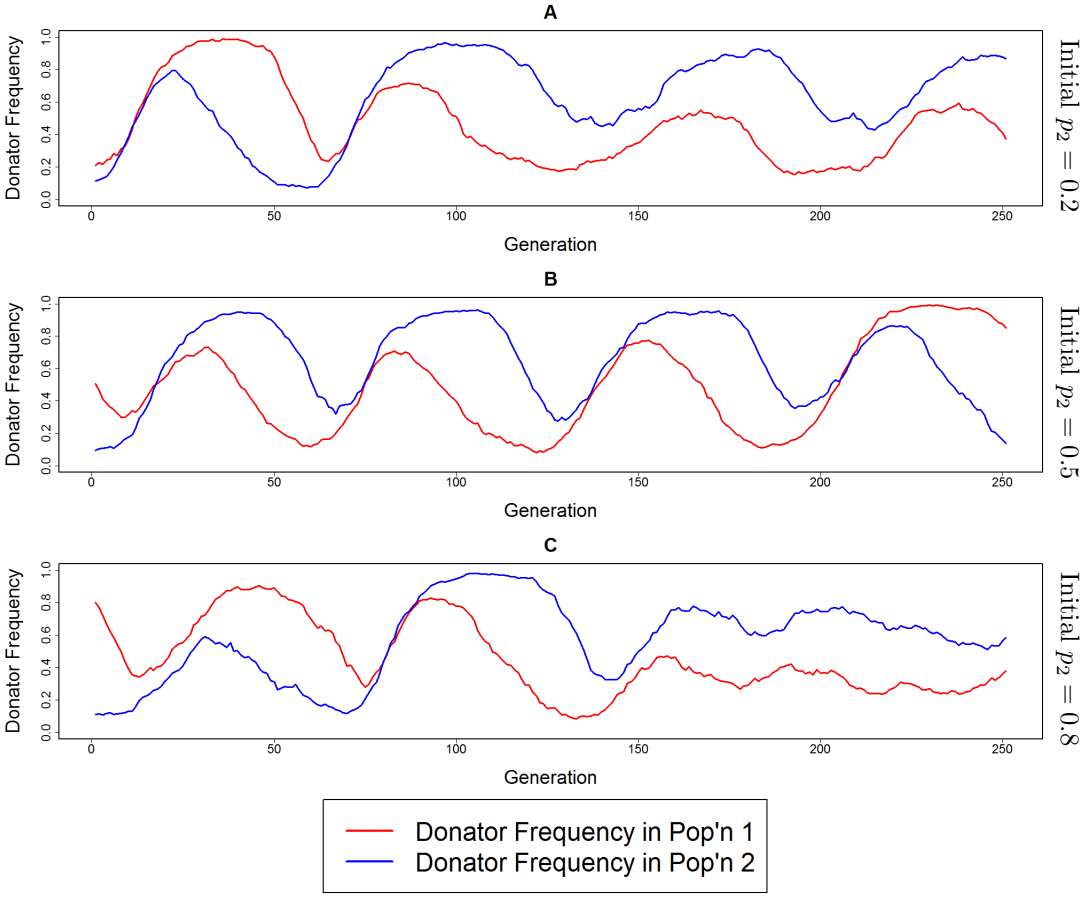

Figure A.4: Evolution of donation allele frequencies (using the stochastic approach with model (d), in which donation is governed by a single locus); initial frequencies are $p_1 = 0.2$, 0.5, and 0.8 in A, B and C respectively, $p_2 = 0.1$, and parameters are $x = 0.05, b = 1, c = 0.5, w_0 = 1$, and $\alpha = 1$. In addition, $p_m = 0.001, N = 1000, n_{SUS} = 200$. Donation is present in both species when $\alpha > c/b$, however stochasticity acts to destabilise the equilibria found in the deterministic case, which entailed equal frequencies of donators between species.

A.3.3 Model (e): The Single-Species Model

Figure A.5 shows an example of the single-species model. Mathematical analysis predicted that donators would reach fixation if $\alpha > c/b$; this is shown to be the case here. Similarly, Figure A.6 introduces stochasticity and strong selection, and depicts very similar results.

A.3.4 Model (f): The Partner Rejection Model

Figure A.7 shows an example of the dynamics of the partner rejection model. Since the donation behaviour is now conditional on both being a donator and having a social partner who is a donator, and since $b - c > 0$, the mathematical analysis showed that the donation and non-suppression alleles will always receive positive selection. These results are borne out by the numerical example; indeed, even if α

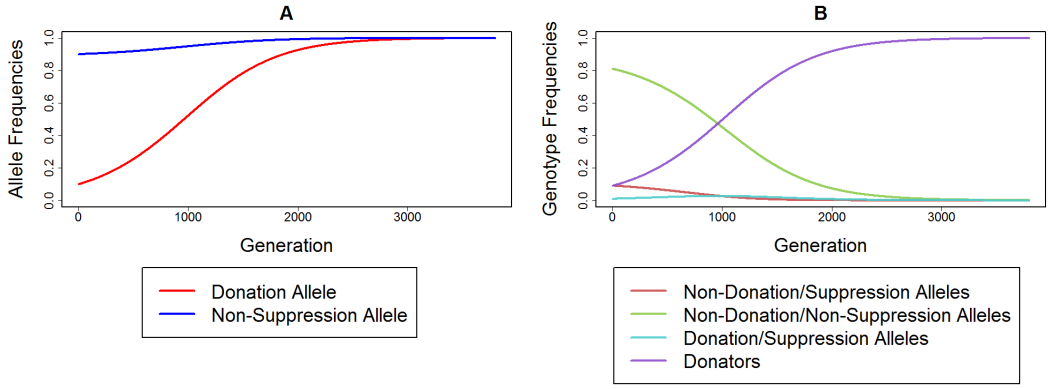

Figure A.5: Evolution of allele frequencies (A) and genotype frequencies (B) (using the deterministic approach with model (d), in which genome-wide assortment occurs within a single species); initial frequencies are $p_1 = 0.1$ and $p_2 = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1$, and $\alpha = 1$. Donators reach fixation since $\alpha > c/b$.

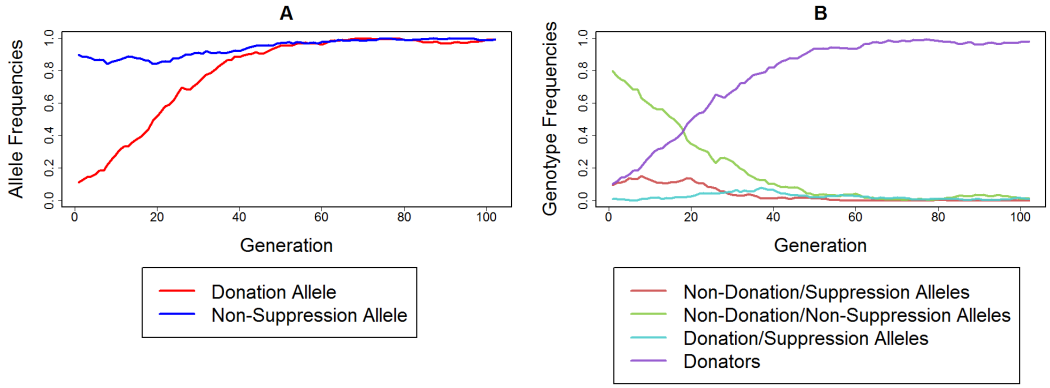

Figure A.6: Evolution of allele frequencies (A) and genotype frequencies (B) (using the deterministic approach with model (d), in which genome-wide assortment occurs within a single species); initial frequencies are $p_1 = 0.1$ and $p_2 = 0.9$, and parameters are $x = 0.05, b = 1, c = 0.5, w_0 = 1$, and $\alpha = 1$. In addition, $p_m = 0.001, N = 1000, n_{SUS} = 200$. These results are very similar to those shown if figure A.5.

is set to 0, both the donation and non-suppression alleles will reach fixation. Once again, the example shown by Figure A.8 of the stochastic approach to analysing this model gives the same results as the deterministic approach.

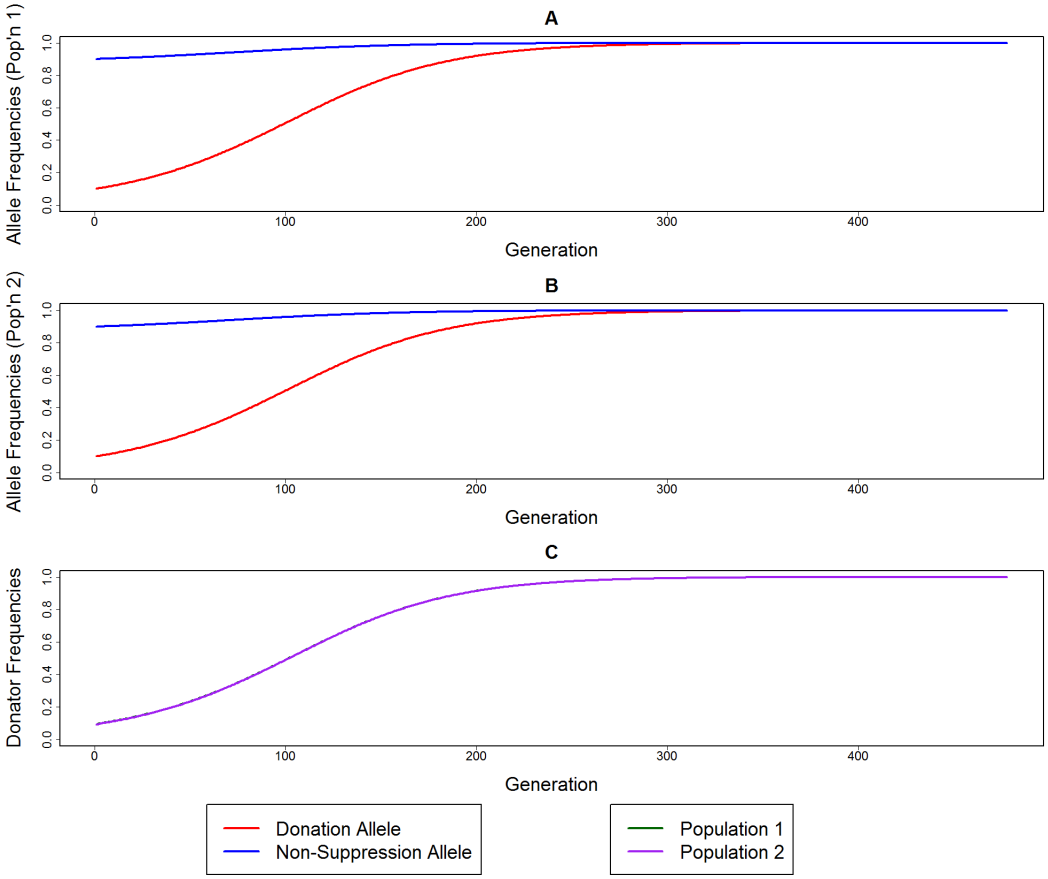

Figure A.7: Evolution of allele frequencies (A) and genotype frequencies (B) (using the deterministic approach with model (d), in which genome-wide assortment occurs within a single species); initial frequencies are$p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 0.01, c = 0.005, w_0 = 1$, and $\alpha = 1$. Note firstly that donator frequencies in the two populations are close enough to equal throughout the process that donator frequencies in population 1 are hidden behind those of population 2 in A.7C.

A.4 Conclusions

We briefly summarise the conclusions of each of the models (c-f); for a full discussion, see Section 3.6.

Model (c) is a variation of model (b); it involves genome-wide assortment, but individuals are either potential donors or potential recipients, with roles chosen uniformly at random. The outcome of this alteration is that the selection strength is halved; otherwise, results remain identical to those found for model (b). The point of including this model is to emphasise the simple point that the important quantity

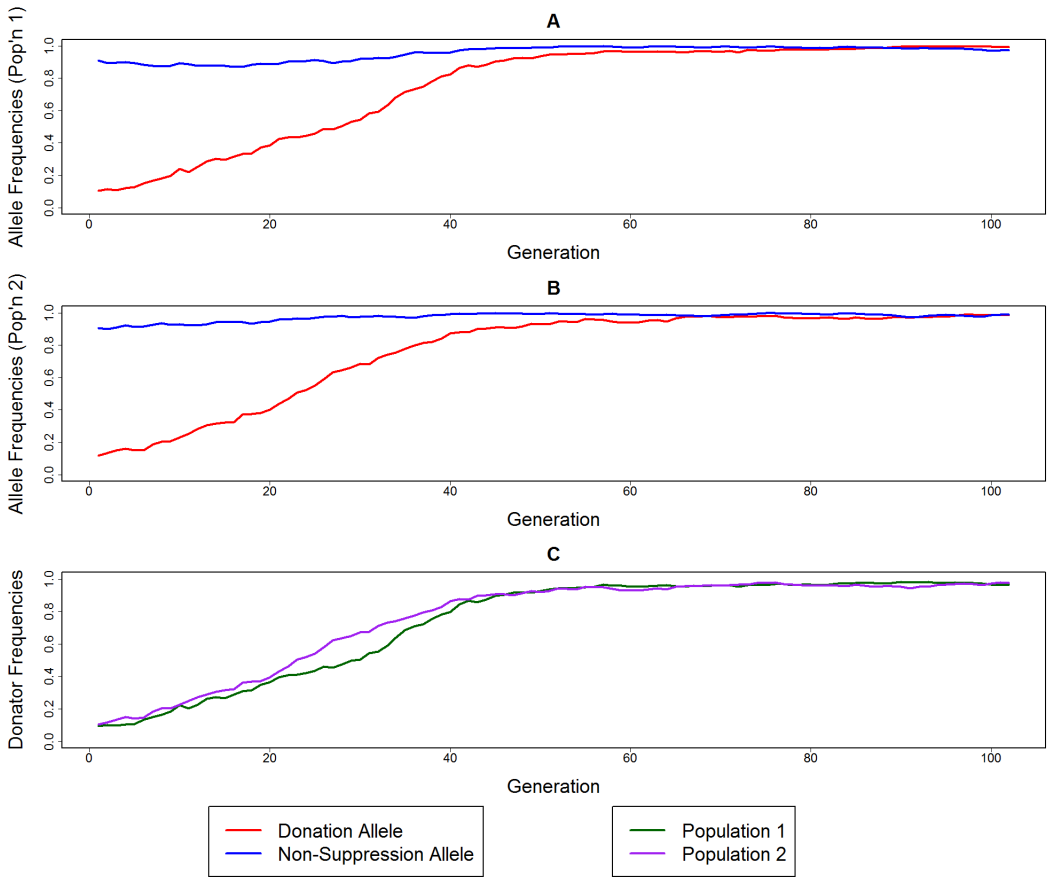

Figure A.8: Evolution of allele frequencies (A) and genotype frequencies (B) (using the deterministic approach with model (d), in which genome-wide assortment occurs within a single species); initial frequencies are $p_{1,1} = 0.101, p_{1,2} = 0.1, p_{2,1} = 0.901, p_{2,2} = 0.9$, and parameters are $x = 0.05, b = 1, c = 0.5, w_0 = 1$, and $\alpha = 1$. In addition, $p_m = 0.001, N = 1000, n_{SUS} = 200$.

for selection on traits is mean lifetime fitness.

The single-locus model (d) relates to both models (a) and (b). A progression can be formulated: model (d) involves donation determined by a single locus, with assortment based on that locus, while model (a) considers what changes when a new locus arises to suppress the donation behaviour which is not subject to assortment. Finally, model (b) considers what happens if this suppression locus becomes subject to assortment. Model (d) shows that in the simple single-locus case, something like Hamilton's rule is in operation: if $\alpha > c/b$, then donation reaches either fixation, or, in the present of stochasticity, intermediate frequencies, if allele frequencies are the same between the two populations.

Model (e) is another variation of model (b); this considers donation within a single species, when it is determined by donation and suppression loci; both loci are subject to assortment. This is essentially equivalent to considering model (b) with

equal allele and genotype frequencies between the two populations. Once again, donators reach fixation if $\alpha > c/b$.

Finally, model (f) consider rejection of interactions. This is essentially a facultative greenbeard scenario (Gardner and West, 2010), since donators adjust their behaviour in response to the phenotype of social partners. Here, the donation and non-suppression alleles always receive positive selection given the assumption of the model.

Bibliography

Aanen, D.K. and Bisseling, T. (2014). The birth of cooperation. *Science*, 345(6192):29–30.

Abbot, P., Abe, J., Alcock, J., Alizon, S., Alpedrinha, J.A.C., Andersson, M., Andre, J.B., van Baalen, M., Balloux, F., Balshine, S., et al. (2011). Inclusive fitness theory and eusociality. *Nature*, 471(7339):E1–4.

Aktipis, C.A., Boddy, A.M., Jansen, G., Hibner, U., Hochberg, M.E., Maley, C.C., and Wilkinson, G.S. (2015). Cancer across the tree of life: cooperation and cheating in multicellularity. *Philos. Trans. R. Soc. B Biol. Sci.*, 370(1673):1–21.

Arakaki, N., Miyoshi, T., and Noda, H. (2001). *Wolbachia*-mediated parthenogenesis in the predatory thrips *Franklinothrips vespiformis* (Thysanoptera: Insecta). *Proc. R. Soc. B*, 268(1471):1011–1016.

Axelrod, R. and Hamilton, W.D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.

Baker, J.E. (1987). Reducing bias and inefficiency in the selection algorithm. In *Proc. Second Int. Conf. Genet. Algorithms*, pages 14–21.

Biernaskie, J.M., West, S.A., and Gardner, A. (2011). Are greenbeards intragenomic outlaws? *Evolution*, 65(10):2729–2742.

Birch, J. (2012). Social revolution. *Biol. Philos.*, 27(4):571–581.

Birch, J. (2014). Has Grafen formalized Darwin? *Biol. Philos.*, 29(2):175–180.

Birch, J. (2016). *The Philosophy of Social Evolution (Manuscript in preparation)*. Oxford University Press.

Birch, J. and Marshall, J.A.R. (2014). Queller's separation condition explained and defended. *Am. Nat.*, 184(4):531–540.

Birch, J. and Okasha, S. (2014). Kin selection and its critics. *Bioscience*, 65(1):22–32.

Boomsma, J.J. (2009). Lifetime monogamy and the evolution of eusociality. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 364(1533):3191–3207.

Bouchon, D., Rigaud, T., and Juchault, P. (1998). Evidence for widespread *Wolbachia* infection in isopod crustaceans: molecular identification and host feminization. *Proc. R. Soc. B*, 265(1401):1081–1090.

Bourke, A.F.G. (2011). *Principles of Social Evolution*. Oxford University Press.

Bourke, A.F.G. (2014a). Hamilton's rule and the causes of social evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 369(1642):20130362.

Bourke, A.F.G. (2014b). The genes-eye view, major transitions and the formal darwinism project. *Biol. Philos.*, 29(2):241–248.

Bourke, A.F.G. (2015). Sex investment ratios in eusocial Hymenoptera support inclusive fitness theory. *J. Evol. Biol.*, 28(11):2106–2111.

Bronstein, J.L. (1994). Conditional outcomes in mutualistic interactions. *Trends Ecol. Evol.*, 9(6):214–217.

Buss, L. (1987). *The Evolution of Individuality*. Princeton University Press.

Charnov, E.L. (1979). Simultaneous hermaphroditism and sexual selection. *Proc. Natl. Acad. Sci. USA*, 76(5):2480–2484.

Clarke, E. (2010). The problem of biological individuality. *Biol. Theory*, 5(4):312–325.

Crawford, C. and Krebs, D., editors (2013). *Handbook of evolutionary psychology: Ideas, issues, and applications.* Psychology Press.

Damuth, J. and Heisler, I.L. (1988). Alternative formulations of multilevel selection. *Biol. Philos.*, 3(4):407–430.

Darwin, C. (1859). *On the Origin of Species*. John Murray.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press.

De Massy, B., Baudat, F., and Nicolas, A. (1994). Initiation of recombination in *Saccharomyces cerevisiae* haploid meiosis. *Proc. Natl. Acad. Sci. USA*, 91(25):11929–11933.

Dhooge, A., Govaerts, W., and Kuznetsov, Y.A. (2003). MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. *ACM Trans. Math. Softw.*, 29(2):141–164.

Duplouy, A., Iturbe-Ormaetxe, I., Beatson, S.A., Szubert, J.M., Brownlie, J.C., McMeniman, C.J., McGraw, E.A., Hurst, G.D.D., Charlat, S., O'Neill, S.L., and Woolfit, M. (2013). Draft genome sequence of the male-killing *Wolbachia* strain *w*Bol1 reveals recent horizontal gene transfers from diverse sources. *Bmc Genomics*, 14(1):20.

Ewald, P.W. (1987). Transmission modes and evolution of the parasitism-mutualism continuum. *Ann. N. Y. Acad. Sci.*, 503:295–306.

Famah Sourassou, N., Hanna, R., Breeuwer, J.A.J., Negloh, K., de Moraes, G.J., and Sabelis, M.W. (2014). The endosymbionts *Wolbachia* and *Cardinium* and their effects in three populations of the predatory mite *Neoseiulus paspalivorus*. *Exp. Appl. Acarol.*, 64(2):207–221.

Ferdy, J. and Godelle, B. (2005). Diversification of transmission modes and the evolution of mutualism. *Am. Nat.*, 166(5):613–627.

Fine, P.E. (1975). Vectors and vertical transmission: an epidemiologic perspective. *Ann. N. Y. Acad. Sci.*, 266:173–194.

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

Fisher, R.M., Cornwallis, C.K., and West, S.A. (2013). Group formation, relatedness, and the evolution of multicellularity. *Curr. Biol.*, 23(12):1120–1125.

Fletcher, J.A. and Doebeli, M. (2009). A simple and general explanation for the evolution of altruism. *Proc. Biol. Sci.*, 276(1654):13–19.

Fletcher, J.A. and Zwick, M. (2006). Unifying the theories of inclusive fitness and reciprocal altruism. *Am. Nat.*, 168(2):252–262.

Fogel, L.J., Owens, A.J., and Walsh, M.J. (1966). *Artificial Intelligence Through Simulated Evolution*. John Wiley and Sons.

Foster, K.R. (2005). Hamiltonian medicine: Why the social lives of pathogens matter. *Science*, 308(5726):1269–1270.

Foster, K.R. (2009). A defense of sociobiology. *Cold Spring Harb. Symp. Quant. Biol.*, 74:403–418.

Foster, K.R. and Wenseleers, T. (2006). A general model for the evolution of mutualisms. *J. Evol. Biol.*, 19(4):1283–1293.

Frank, S.A. (1987). Individual and population sex allocation patterns. *Theoretical population biology*, 74:47–74.

Frank, S.A. (1990). Sex allocation theory for birds and mammals. *Annual Review of Ecology and Systematics*, 21(1):13–55.

Frank, S.A. (1994). Genetics of mutualism: the evolution of altruism between species. *J. Theor. Biol.*, 170(4):393–400.

Friedman, D. (1998). On economic applications of evolutionary game theory. *J. Evol. Econ.*, 8(1):15–43.

Fry, A.J. and Rand, D.M. (2002). *Wolbachia* interactions that determine *Drosophila melanogaster* survival. *Evolution*, 56(10):1976–1981.

Gardner, A. (2009). Adaptation as organism design. *Biol. Lett.*, 5(6):861–4.

Gardner, A. (2014). Life, the universe and everything. *Biol. Philos.*, 29(2):207–215.

Gardner, A. and West, S.A. (2010). Greenbeards. *Evolution*, 64(1):25–38.

Gardner, A., West, S.A., and Wild, G. (2011). The genetical theory of kin selection. *J. Evol. Biol.*, 24(5):1020–1043.

Giorgini, U.M., Bernardo, U., Monti, M.M., Nappo, A.G., and Gebiola, M. (2010). *Rickettsia* symbionts cause parthenogenetic reproduction in the parasitoid wasp *Pnigalio soemius* (Hymenoptera: Eulophidae). *Appl. Environ. Microbiol.*, 76(8):2589–2599.

Goodnight, C., Schwartz, J., and Stevens, L. (1992). Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. *Am. Nat.*, 140(5):743–761.

Gotoh, T., Noda, H., and Ito, S. (2007). *Cardinium* symbionts cause cytoplasmic incompatibility in spider mites. *Heredity (Edinb).*, 98(1):13–20.

Grafen, A. (1985). A geometric view of relatedness. *Oxford Surv. Evol. Biol.*, (2):28–89.

Grafen, A. (1991). Modelling in behavioural ecology. In J.R. Krebs and N.B. Davies, editors, *Behav. Ecol. An Evol. Approach*, pages 5–31. Blackwell Scientific.

Grafen, A. (1999). Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proc. R. Soc. B Biol. Sci.*, 266(1421):799–803.

Grafen, A. (2002). A first formal link between the Price equation and an optimization program. *J. Theor. Biol.*, 217(1):75–91.

Grafen, A. (2006). Optimization of inclusive fitness. *J. Theor. Biol.*, 238(3):541–563.

Grafen, A. (2014a). The formal darwinism project in outline. *Biol. Philos.*, 29(2):155–74.

Grafen, A. (2014b). The formal darwinism project in outline: response to commentaries. *Biol. Philos.*, 29(2):281–292.

Griffin, A.S., West, S.A., and Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature*, 430(7003):1024–1027.

Hagimori, T., Abe, Y., Date, S., and Miura, K. (2006). The first finding of a *Rickettsia* bacterium associated with parthenogenesis induction among insects. *Curr. Microbiol.*, 52(2):97–101.

Hamilton, W.D. (1963). The evolution of altruistic behavior. *Am. Nat.*, 97(896):354–356.

Hamilton, W.D. (1964a). The genetical evolution of social behaviour. I. *J. Theor. Biol.*, 7(1):1–16.

Hamilton, W.D. (1964b). The genetical evolution of social behaviour. II. *J. Theor. Biol.*, 7(1):17–52.

Hamilton, W.D. (1967). Extraordinary sex ratios. *Science*, 156(April):477–488.

Hamilton, W.D. (1970). Selfish and Spiteful Behaviour in an Evolutionary Model. *Nature*, 228(5277):1218–1220.

Hamilton, W.D. (1972). Altruism and related phenomena, mainly in social insects. *Annu. Rev. Ecol. Syst.*, (3):193–232.

Hamilton, W.D. (1975). Innate social aptitudes of man: an approach from evolutionary genetics. In R. Fox, editor, *Biosocial Anthrpology*, pages 115–132. John Wiley and Sons.

Hamilton, W.D. (1996). Sorority Avenue. In M. Ridley, editor, *Narrow Roads Gene L. Collect. Pap. WD Hamilt. Vol 1 - Evol. Soc. Behav.*, pages 255–266. Oxford University Press.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.

Hartmann, A., Wantia, J., Torres, J.A., and Heinze, J. (2003). Worker policing without genetic conflicts in a clonal ant. *Proc. Natl. Acad. Sci.*, 100(22):12836–12840.

Hatcher, M.J. and Dunn, A.M. (1995). Evolutionary consequences of cytoplasmically inherited feminizing factors. *Philos. Trans. R. Soc. B Biol. Sci.*, 348(1326):445–456.

Hatcher, M.J., Taneyhill, D.E., Dunn, A.M., and Tofts, C. (1999). Population dynamics under parasitic sex ratio distortion. *Theor. Popul. Biol.*, 56(1):11–28.

Hatcher, M.J., Dunn, A.M., and Tofts, C. (2000). Co-existence of hosts and sex ratio distorters in structured populations. *Evol. Ecol. Res.*, 2(2):185–205.

Hatchwell, B.J., Gullett, P.R., and Adams, M.J. (2014). Helping in cooperatively breeding long-tailed tits: a test of helping in cooperatively breeding long- tailed tits. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 369(1642):20130565.

Heisler, I. and Damuth, J. (1987). A method for analyzing selection in hierarchically structured populations. *Am. Nat.*, 130(4):582–602.

Hoffmann, A.A., Clancy, D., and Duncan, J. (1996). Naturally-occurring *Wolbachia* infection in *Drosophila simulans* that does not cause cytoplasmic incompatibility. *Heredity*, 76:1–8.

Hornett, E.A., Moran, B., Reynolds, L.A., Charlat, S., Tazzyman, S., Wedell, N., Jiggins, C.D., and Hurst, G.D.D. (2014). The evolution of sex ratio distorter suppression affects a 25 cM genomic region in the butterfly *Hypolimnas bolina*. *PLoS Genetics*, 10(12):e1004822.

Hughes, W.O.H., Oldroyd, B.P., Beekman, M., and Ratnieks, F.L.W. (2008). Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science*, 320(5880):1213–6.

Huigens, M.E., Luck, R.F., Klaassen, R.H., Maas, M.F., Timmermans, M.J., and Stouthamer, R. (2000). Infectious parthenogenesis. *Nature*, 405(6783):178–179.

Huneman, P. (2014). Formal Darwinism as a tool for understanding the status of organisms in evolutionary biology. *Biol. Philos.*, 29(2):271–279.

Hurst, G.D.D. and Jiggins, F.M. (2000). Male-killing bacteria in insects: Mechanisms, incidence, and implications. *Emerg. Infect. Dis.*, 6(4):329–336.

Hurst, G.D.D. and Majerus, M.E.N. (1993). Why do maternally inherited microorganisms kill males? *Heredity (Edinb).*, 71(1):81–95.

Hurst, G.D.D. and Werren, J.H. (2001). The role of selfish genetic elements in eukaryotic evolution. *Nat. Rev. Genet.*, 2(8):597–606.

Hurst, G.D.D., Graf von der Schulenburg, J.H., Majerus, T.M., Bertrand, D., Zakharov, I.A., Baungaard, J., Völkl, W., Stouthamer, R., and Majerus, M.E. (1999a). Invasion of one insect species, *Adalia bipunctata*, by two different male-killing bacteria. *Insect Mol. Biol.*, 8(1):133–139.

Hurst, G.D.D., Jiggins, F.M., Graf von der Schulenburg, J.H., Bertrand, D., West, S.A., Goriacheva, I.I., Zakharov, I.A., Werren, J.H., Stouthamer, R., and Majerus, M.E.N. (1999b). Male-killing *Wolbachia* in two species of insect. *Proc. R. Soc. B Biol. Sci.*, 266(1420):735–740.

Hurst, G.D.D., Johnson, A., Schulenburg, J., and Fuyama, Y. (2000). Male killing *Wolbachia* in *Drosophila*: a temperature sensitive trait with a threshold density. *Genetics*, 156(2):699–709.

Hurst, L.D. (1991). The incidences and evolution of cytoplasmic male killers. *Proc. R. Soc. B Biol. Sci.*, 244(1310):91–99.

Jaenike, J. (2007). Fighting back against male-killers. *Trends Ecol. Evol.*, 22(4):167–169.

Jaenike, J. (2012). Population genetics of beneficial heritable symbionts. *Trends Ecol. Evol.*, 27(4):226–232.

Jaenike, J., Unckless, R., Cockburn, S.N., Boelio, L.M., and Perlman, S.J. (2010). Adaptation via symbiosis: Recent spread of a *Drosophila* defensive symbiont. *Science*, 329(5988):212–215.

Jeon, K.W. (2011). Endosymbionts. *eLS*, pages 1–11.

Keller, L. and Ross, K. (1998). Selfish genes: a green beard in the red fire ant. *Nature*, 394(6693):573–575.

Kellner, K. and Heinze, J. (2011). Absence of nepotism in genetically heterogeneous colonies of a clonal ant. *Ethology*, 117(6):556–564.

Kimura, M. (1965). Attainment of quasi linkage equilibrium when gene frequencies are changing by natural selection. (575).

Kirkpatrick, M., Johnson, T., and Barton, N. (2002). General models of multilocus evolution. *Genetics*, 161(4):1727–1750.

Lewontin, R.C. (1964). The interaction of selection and linkage. I. general considerations; heterotic models. *Genetics*, 49(1):49–67.

Lewontin, R. (1970). The units of selection. *Annu. Rev. Ecol. Syst.*, pages 1–18.

Lipsitch, M., Nowak, M.A., Ebert, D., and May, R.M. (1995). The population dynamics of vertically and horizontally transmitted parasites. *Proc. R. Soc. B Biol. Sci.*, 260(1359):321–327.

Margulis, L. (1981). *Symbiosis in Cell Evolution: Life and its environment on the early earth*. W. H. Freeman & Co.

Marin, B., Nowack, E.C.M., and Melkonian, M. (2005). A plastid in the making: Evidence for a second primary endosymbiosis. *Protist*, 156(4):425–432.

Marshall, J.A.R. (2011). Group selection and kin selection: formally equivalent approaches. *Trends Ecol. Evol.*, 26(7):325–32.

Marshall, J.A.R. (2015). *Social Evolution and Inclusive Fitness Theory: An Introduction*. Princeton University Press.

Martin, O.Y., Puniamoorthy, N., Gubler, A., Wimmer, C., Germann, C., and Bernasconi, M.V. (2013). Infections with the microbe *Cardinium* in the Dolichopodidae and other Empidoidea. *J. Insect Sci.*, 13:47.

Maynard Smith, J. and Price, G.R. (1973). The logic of animal conflict. *Nature*, 246(5427):15–18.

Maynard Smith, J. (1982). *Evolution and The Theory of Games*. Cambridge University Press.

Maynard Smith, J. and Szathmáry, E. (1995). *The Major Transitions in Evolution*. Oxford University Press.

McShea, D.W. and Changizi, M.A. (2003). Three puzzles in hierarchical evolution. *Integr. Comp. Biol.*, 43(1):74–81.

Michod, R.E. (2006). On the transfer of fitness from the cell to the multicellular organism. *Biol. Philos.*, 20(5):967–987.

Michod, R.E. (2007). Evolution of individuality during the transition from unicellular to multicellular life. *Proc. Natl. Acad. Sci. U. S. A.*, 104 Suppl:8613–8618.

Michod, R.E. and Hamilton, W.D. (1980). Coefficients of relatedness in sociobiology. *Nature*, 288(5792):694–697.

Michod, R.E. and Nedelcu, A.M. (2003). On the reorganization of fitness during evolutionary transitions in individuality. *Integr. Comp. Biol.*, 43(1):64–73.

Narita, S., Kageyama, D., Nomura, M., and Fukatsu, T. (2007). Unexpected mechanism of symbiont-induced reversal of insect sex: Feminizing *Wolbachia* continuously acts on the butterfly *Eurema hecabe* during larval development. *Appl. Environ. Microbiol.*, 73(13):4332–4341.

Nowak, M.A., Tarnita, C.E., and Wilson, E.O. (2010). The evolution of eusociality. *Nature*, 466(7310):1057–1062.

Okasha, S. (2004). Multilevel selection and the partitioning of covariance: a comparison of three approaches. *Evolution*, 58(3):486–494.

Okasha, S. (2006). *Evolution and the Levels of Selection.* Clarendon Press.

Okasha, S. (2015). The relation between kin and multilevel selection: an approach using causal graphs. *Br. J. Philos. Sci.*, 67(2):435–470.

Okasha, S. and Paternotte, C. (2014a). Adaptation, fitness and the selection-optimality links. *Biol. Philos.*, 29(2):225–232.

Okasha, S. and Paternotte, C. (2014b). The formal Darwinism project: editors' introduction. *Biol. Philos.*, 29(2):153–154.

O'Neill, S.L., Giordano, R., Colbert, A.M., Karr, T.L., and Robertson, H.M. (1992). 16S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects. *Proc. Natl. Acad. Sci. U. S. A.*, 89(7):2699–2702.

Orlove, M.J. and Wood, C.L. (1978). Coefficients of relationship and coefficients of relatedness in kin selection: a covariance form for the RHO formula. *J. Theor. Biol.*, 73(4):679–686.

Pepper, J.W., Sprouffske, K., and Maley, C.C. (2007). Animal cell differentiation patterns suppress somatic evolution. *PLoS Comput. Biol.*, 3(12):e250.

Price, G.R. (1970). Selection and Covariance. *Nature*, 227(5257):520–521.

Price, G.R. (1972). Extension of Covariance Selection Mathematics. *Ann. Hum. Genet.*, 35(4):485–490.

Queller, D.C. (1985). Kinship, reciprocity and synergism in the evolution of social behaviour. *Nature*, 318(6044):366–367.

Queller, D.C. (1992). A general model for kin selection. *Evolution*, 46:376–380.

Queller, D.C. (1997). Cooperators since life began. *Q. Rev. Biol.*, 72(2):184–188.

Queller, D.C. (2000). Relatedness and the fraternal major transitions. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 355(1403):1647–55.

Queller, D.C. (2011). Expanded social fitness and Hamilton's rule for kin, kith, and kind. *Proc. Natl. Acad. Sci. U. S. A.*, 108 Suppl:10792–10799.

Queller, D.C. and Strassmann, J.E. (2009). Beyond society: the evolution of organismality. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 364(1533):3143–3155.

Queller, D.C., Ponte, E., Bozzaro, S., and Strassmann, J.E. (2003). Single-gene greenbeard effects in the social amoeba *Dictyostelium discoideum*. *Science*, 299(5603):105–106.

Quickfall, C.G. (2016). Models of the Major Evolutionary Transitions: Code and Selected Workspaces. Retrieved from osf.io/xtcmy.

R Core Team (2016). R: A language and environment for statistical computing. (v2.15.1).

Rabeling, C. and Kronauer, D.J. (2011). Thelytokous parthenogenesis in eusocial Hymenoptera. *Annu. Rev. Entomol.*, 58(1):273–292.

Rice, S.H. (2004). *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates.

Rigaud, T. and Juchault, P. (1995). Success and failure of horizontal transfers of feminizing *Wolbachia* endosymbionts in woodlice. *J. Evol. Biol.*, 8:249–255.

Roubik, D.W. (1982). The ecological impact of nectar-robbing bees and pollinating hummingbirds on a tropical shrub. *Ecology*, 63(2):354–360.

Roughgarden, J. (1975). Evolution of marine symbiosis - a simple cost-benefit model. *Ecology*, 56(5):1201–1208.

Rousset, F., Bouchon, D., Pintureau, B., Juchault, P., and Solignac, M. (1992). *Wolbachia* endosymbionts responsible for various alterations of sexuality in arthropods. *Proc. R. Soc. B Biol. Sci.*, 250(1328):91–98.

Shelton, D.E. and Michod, R.E. (2014a). Group selection and group adaptation during a major evolutionary transition: Insights from the evolution of multicellularity in the volvocine algae. *Biol. Theory*, 9(4):452–469.

Shelton, D.E. and Michod, R.E. (2014b). Levels of selection and the formal Darwinism project. *Biol. Philos.*, 29(2):217–224.

Smukalla, S., Caldara, M., Pochet, N., Beauvais, A., Guadagnini, S., Yan, C., Vinces, M.D., Jansen, A., Prevost, M.C., Latgé, J.P., et al. (2008). FLO1 is

a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell*, 135(4):726–37.

Soetaert, K., Petzoldt, T., and Setzer, R.W. (2010). Package deSolve: Solving initial value differential equations in R. *J. Stat. Softw.*, 33(9):1–25.

Taylor, P.D. (1981). Intra-sex and inter-sex sibling interactions as sex-ratio determinants. *Nature*, 291(5810):64–66.

Taylor, P.D. (1993). Female-biased sex ratios under local mate competition: An experimental confirmation. *Evol. Ecol.*, 7(3):306–308.

The Mathworks Inc. (2010). MATLAB R2015b (v7.10.0).

Trivers, R.L. (1971). The Evolution of Reciprocal Altruism. *Q. Rev. Biol.*, 46(1):35–57.

Trivers, R.L. and Hare, H. (1976). Haploidploidy and the evolution of the social insect. *Science*, 191(4224):249–263.

Trivers, R.L. and Willard, D.E. (1973). Natural Selection of Parental Ability to Vary the Sex Ratio of Offspring. *Science*, 179(4068):90–92.

Von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Wade, M. (1985). Soft selection, hard selection, kin selection, and group selection. *Am. Nat.*, 125(1):61–73.

Walker, T., Johnson, P.H., Moreira, L.A., Iturbe-Ormaetxe, I., Frentiu, F.D., McMeniman, C.J., Leong, Y.S., Dong, Y., Axford, J., Kriesner, P., et al. (2011). The wMel *Wolbachia* strain blocks dengue and invades caged *Aedes aegypti* populations. *Nature*, 476(7361):450–453.

Watanabe, M., Kageyama, D., and Miura, K. (2013). Transfer of a parthenogenesis-inducing *Wolbachia* endosymbiont derived from *Trichogramma dendrolimi* into *Trichogramma evanescens*. *J. Invertebr. Pathol.*, 112(1):83–87.

Weeks, A.R. and Breeuwer, J.A. (2001). *Wolbachia*-induced parthenogenesis in a genus of phytophagous mites. *Proc. Biol. Sci.*, 268(1482):2245–2251.

Wenseleers, T. and Van Oystaeyen, A. (2011). Unusual modes of reproduction in social insects: Shedding light on the evolutionary paradox of sex. *BioEssays*, 33(12):927–937.

Werren, J.H. (1987). The coevolution of autosomal and cytoplasmic sex-ratio factors. *J. Theor. Biol.*, 124(3):317–334.

Werren, J.H. (1997). Biology of *Wolbachia*. *Annu. Rev. Entomol.*, 42(124):587–609.

Werren, J.H., Hurst, G.D.D., Zhang, W., Breeuwer, J.A.J., Stouthamer, R., and Majerus, M.E.N. (1994). Rickettsial relative associated with male killing in the ladybird beetle (Adalia bipunctata). *J. Bacteriol.*, 176(2):388–394.

Werren, J.H., Baldo, L., and Clark, M.E. (2008). *Wolbachia*: master manipulators of invertebrate biology. *Nat. Rev Microbiol*, 6(10):741–751.

West, S.A. (2009). Sex Allocation. *Princet. Univ. Press*.

West, S.A. and Kiers, E.T. (2009). Evolution: what is an organism? *Curr. Biol.*, 19(23):R1080–2.

West, S.A., Griffin, A.S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.*, 20(2):415–432.

West, S.A., Fisher, R.M., Gardner, A., and Kiers, E.T. (2015). Major evolutionary transitions in individuality. *Proc. Natl. Acad. Sci. U. S. A.*, 112(33):1–8.

Williams, G. and Nesse, R. (1991). The dawn of Darwinian medicine. *Q. Rev. Biol.*, 66(1):1–22.

Wright, S. (1922). Coefficients of inbreeding and relationship. *Am. Nat.*, 56(645):330–338.

Wyatt, G.A.K., West, S.A., and Gardner, A. (2013). Can natural selection favour altruism between species? *J. Evol. Biol.*, 26(9):1854–1865.

Xie, J., Butler, S., Sanchez, G., and Mateos, M. (2014). Male killing *Spiroplasma* protects *Drosophila melanogaster* against two parasitoid wasps. *Heredity*, 112(4):399–408.

Yamamura, N. (1993). Vertical transmission and evolution of mutualism from parasitism. *Theor. Popul. Biol.*, 44(1):95–109.

Yen, J.H. and Barr, A.R. (1973). The etiological agent of cytoplasmic incompatibility in *Culex pipiens*. *J. Invertebr. Pathol.*, 22(2):242–250.

Zeng, K. and Charlesworth, B. (2011). The joint effects of background selection and genetic recombination on local gene genealogies. *Genetics*, 189(1):251–266.

Zchori-Fein, E., Gottlieb, Y., Kelly, S.E., Brown, J.K., Wilson, J.M., Karr, T.L., and Hunter, M.S. (2001). A newly discovered bacterium associated with parthenogenesis and a change in host selection behavior in parasitoid wasps. *Proc. Natl. Acad. Sci. USA*, 98(22):12555–12560.

Zchori-Fein, E. and Perlman, S.J. (2004). Distribution of the bacterial symbiont *Cardinium* in arthropods. *Mol. Ecol.*, 13(7):2009–2016.