1. Introduction

In nature, organisms display a broad range of variation in their morphology and behaviour. Darwin (1859) recognised that natural selection acts upon this phenotypic variation. This means that individuals that are better adapted to their environmental conditions have a higher probability to reproduce, passing their genes (and traits), on to the next generation. It is, however, naïve to think that all organisms are optimised for their respective environments. Various constraints may cause individuals to express phenotypes that are slightly maladapted and unable to reach their optimal state. In addition, some traits may be maladapted due to the relatively small effects that they have on the fitness of an individual. Looking at evolution through optimization is not a new concept and has been refined to make it more appropriate to unravelling evolutionary concepts (Maynard Smith, 1978b; Mitchell & Valone, 1990; Orzack & Sober, 2001; Stearns & Schmid-Hempel, 1987). One aim of studying evolution is thus to identify factors under selection that may influence specific traits.

In this thesis I focus mostly on two traits: dispersal and inbreeding. I aim to unravel which factors affect dispersal and inbreeding, how dispersal and inbreeding affect population structure and how this influences sampling schemes. I further empirically examine how dispersal and inbreeding affect the fitness in a model organism.

I use the introduction to describe the primary tool used, namely, individual-based simulation modelling, to investigate these traits and why it is ideal for optimisation studies. All the simulations are, however, grounded in reality and I introduce pollinating fig wasps as model organisms, which inspired most of these simulations. The brief description of their life history is also applicable to chapter 5, where molecular tools were used to investigate inbreeding and dispersal. Lastly, I provide an overview of the aims of each chapter.

Individual-based simulation modelling

"*In short, it shows what the world would look like, if it really did work the way in which we think it does*."

Peck (2004)

Individual-based simulation (also called individual-based modelling or agent-based modelling) is where organisms are created *in silico* and their interactions with each other and the simulated environment investigated. Individuals, as discrete units within a population or ecological systems, define these models. Each of these individuals is, in turn, defined by at least one parameter that varies between individuals, and is tracked during the history of interactions with the environment and other individuals. In the last four decades, individual-based simulation modelling has increased substantially (in part reflecting the advancements in computer processing power) and include models of ecological and evolutionary processes (DeAngelis & Mooij, 2005; Grimm, 1999; Grimm et al., 2006; Łomnicki, 1999; Mitchell, 1998; Peck, 2004; Prescott et al., 2007). These processes are complex and individual-based simulation is an extremely powerful tool with which biological systems may be mimicked and experimented on, to unravel questions in ecology and evolution (DeAngelis & Mooij, 2005; Grimm, 1999; Grimm et al., 2006; Kokko, 2007; Mitchell, 1998; Peck, 2004; Winsberg, 2003). Simulation modelling has a number of advantages over analytical models, especially when the system under investigation is complex. However, this method of investigation also has its disadvantages and limitations and is open to misuse as is the case for other research tools. Subsequently I will elaborate all these points, and the reasons why I used simulation modelling, starting with a description of how a basic simulation model works.

A simple simulation model

Individual-based simulations have a simple structure and it is immediately apparent that it is similar to natural processes driving evolution (figure 1.1). For brevity, I shall refer to "individual-based simulations" simply as "simulations". Each individual in

these simulations is characterised by a few traits. During the simulation, these traits are tracked for each individual organism throughout the life cycle of their digitised life history. Depending on the question addressed, offspring production of each individual is constrained by a number of factors including, the other individuals competing for mating and reproduction, simulated physical constraints on individuals, simulated environmental constraints, or a combination of some or all of these factors. After many generations the traits under investigation should be adapted to the enforced and emerging constraints.

The steps depicted in figure 1.1 are followed in all the models (bar the fitness calculation when fitness emerges as a consequence of different reproductive capabilities, chapter 3, or when loci under investigation are neutral, chapter 4):

1. A population of organisms is initialised and alleles are distributed to all individuals. Alleles can be assigned either randomly to the individuals, or all individuals can receive the same allele (necessary when local optima need to be identified, see stochasticity in simulation modelling).

2. If fitness values are implicitly stated (chapter 2), each individual's fitness is calculated.

3. Individuals are selected to mate. This is performed either at random (chapter 3, where fitness is an emerging property; chapter 4, where loci are neutral) or using a specific selection method when individuals have assigned fitness values (the selection method employed in this thesis is fitness proportional selection with "Roulette wheel" sampling, see stochasticity in simulation modelling and chapter 2).

4. Offspring are created and receive their alleles from their parents via normal Mendelian inheritance.

5. Each offspring is mutated at each locus with a predetermined probability (see stochasticity in simulation modelling).

6. Once the number of offspring reaches the number of individuals in the parental population the new population replaces the old. Steps 2 to 6 are iterated for a number of generations.
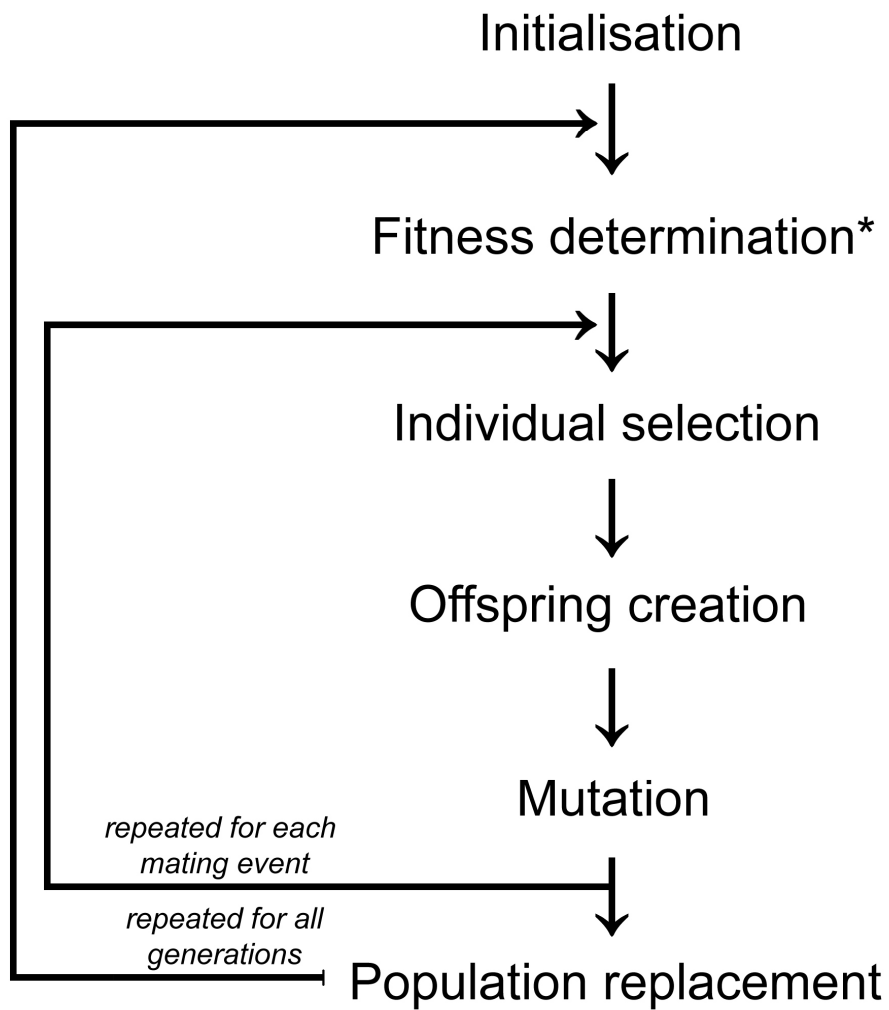
Initialisation

Fitness determination*

Individual selection

Offspring creation

Mutation

repeated for each mating event

repeated for all generations

Population replacement

Figure 1.1. Steps in a simple individual-based simulation model. *Explicit fitness determination is often excluded if it is an emerging property of the simulation.

The number of generations depends on how fast the traits of interest stabilise, and preliminary runs are performed to determine this. Data is collected at the end of the simulation or sometimes averaged over the latter part of the simulation (once stability is reached). Normally, simulations are repeated multiple times with slight adjustments to one parameter. In turn, each parameter set (i.e. each simulation with a change to any parameter) is repeated multiple times. Comparison between the adjustments and the trait of interest are investigated and analysed. If the trait is influenced by the adjusted parameter conclusions about this relationship, and its influence on the evolution of the system can be made.

Stochasticity in simulation modelling

A key feature of simulation modelling is its inherent demographic stochasticity and individual variation (DeAngelis & Mooij, 2005; Grimm, 1999). Including stochastic parameters and their effects are becoming increasingly important if we want to extend our knowledge of evolutionary theory (DeAngelis & Mooij, 2005; Housten & McNamara, 1985; Lenormand et al., 2009; Yoshimura & Shields, 1987). Individual-based simulations are therefore well suited to expand classical models, resulting in better prediction of population and ecosystem evolution (DeAngelis & Mooij, 2005; Judson, 1994; Peck, 2004). It is, however, important to note that the variation within the simulation models in this thesis is strictly constrained by the model parameters. This means that the trait being optimised has a finite number of solutions and that *de novo* development of a new trait to increase fitness and offspring production is impossible. These constraints are in agreement with Darwinian theory of evolution where phylogenetic inertia often limit better forms of adaptation, and optimisation is mostly from current individual variation (Darwin, 1859; Orzack & Sober, 2001).

Variation in natural populations is often lost due to drift or selection. During modelling, the aim is to create a realistic representation of these processes (drift and selection) while producing and maintaining enough variation in finite population sizes for a trait to be optimised (DeAngelis & Mooij, 2005). I will address two components, used by the models in this thesis, which maintain variation in a similar manner to natural systems, namely fitness proportion selection and mutation. This is by no means a complete account of the incorporation of stochastic variation in simulation

models, but is of importance because these two components force the maintenance of some variation (rather than emerging spontaneously from the model). It is therefore important that these components are an accurate reflection of natural systems.

In nature, fitter individuals are identified by their ability to produce more offspring. This does not mean that only the fittest individuals reproduce, but that they will, on average, produce proportionally more offspring. In simulation modelling there are various ways of selecting who should mate (Mitchell, 1998). It is, therefore, important to choose a selection method where sufficient variation is maintained to move away from local optima yet strong enough to optimise the traits under investigation (Mitchell, 1998). Fitness proportion selection with "roulette wheel" sampling (used in chapter 2) is well suited for simulations of biological systems because it enables fitter individuals to mate more often, while weaker individuals are not completely excluded. Briefly, each individual's fitness value is divided by the total fitness of the population (i.e. sum of all the individual's fitness values). This is the proportional fitness of each individual. Individuals are then assigned a slice of a "roulette wheel" according to their proportional fitness. To select an individual to mate, "the wheel is spun" (i.e. a random number is drawn). Whoever owns that number on the "roulette wheel" will mate. This is repeated for each mating event, leading to fitter individual being selected more often than weaker individuals (Mitchell, 1998).

In biological systems phenotypic variation is largely created by mutations. Unfortunately, the stochastic effects of mutations are often excluded from evolutionary models (Lenormand et al., 2009; Orr, 2005). In the simulations in this thesis, two types of mutational models are used; firstly jumping mutations (chapters 2 and 3, electronic appendix), where an allele can mutate to any of the possible alleles. Jumping mutations enable a population to search through large fitness landscapes, reducing the risk of getting stuck on a local optimum (Lenormand et al., 2009). It is however known that, when a population approaches the phenotypic optimum, mutations with large phenotypic effects will be more deleterious than mutations with small phenotypic effects. Therefore, if a trait is to be optimised completely, infinitesimally small mutations must occur (Fisher, 1930; Lenormand et al., 2009; Orr, 2005). For this reason a second mutational model is used, namely, stepwise mutations (chapters 2-4, see electronic appendix). When an allele mutates via stepwise mutation it changes to the next possible sequential value. In other words if

there are 11 possible alleles, determining dispersal probability, ranging from 0 to 1, an allele with the value 0.3 will change to 0.2 or 0.4 with equal probability.

In all the models the number of alleles can be changed (chapters 2-4, see electronic appendix), and if increased, the mutational steps become smaller (this unfortunately enlarges the parameter space, which requires more generations to explore). The mutations in simulation modelling are obviously not infinitely small, but the stepwise mutations model is a practical solution to enable the population to approach the optimum more efficiently than with random jumping mutations only. It should be noted that in all the simulations, where both mutational models were used, the probability of a stepwise mutation was four times that of a jumping mutation. Figure 1.2, from experiment 3.3 in chapter 3, is an example of where both mutational models in the simulation enable individuals to approach the optimal behaviour and encapsulate the mutational step-size decrease of Fisher's (1930) geometrical model of adaptation.

With increasing computing power, we are able to study evolutionary processes, including demographic stochasticity in more detail. Current commercial computers are often powerful enough to search through extremely large parameter spaces using large population sizes, thousands of alleles and many generations. The problem of local optima in studies of adaptation is therefore becoming less of a concern. We should however guard against exploiting current computing power and remember that simulations should reflect realistic conditions, as organisms in nature may get stuck on local optima. Depending on the biological question, comparison of multiple simulations, with realistic parameters may be more informative, including the dynamics of stochasticity in nature.
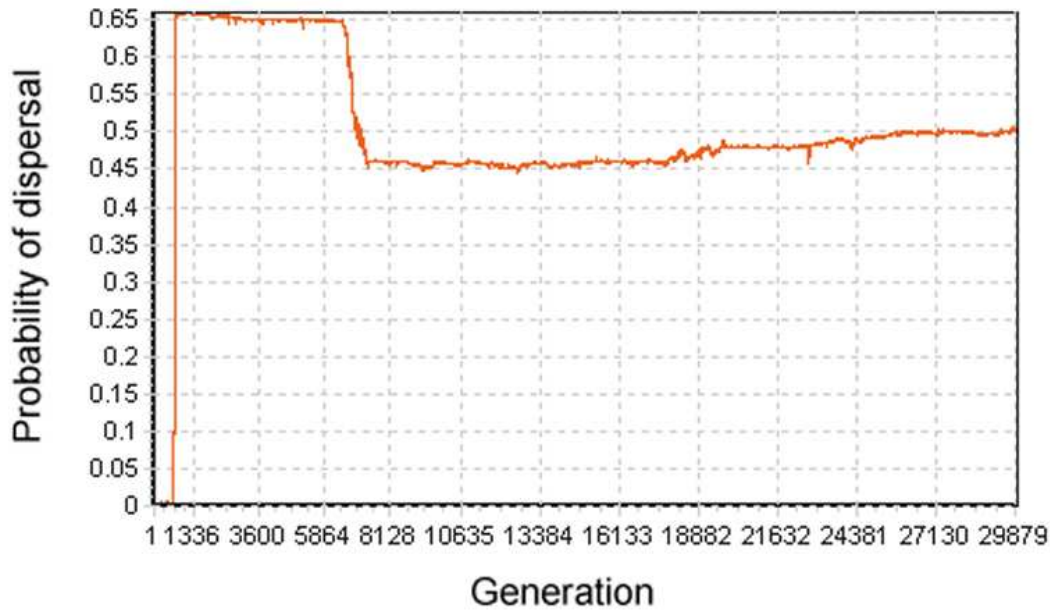
Figure 1.2. The optimal probability of dispersal under the simulated conditions (1 foundress mother dispersal cost of 0.5) is expected to be approximately 0.49 (see chapter 3). The population is initialised with a 0 probability of dispersal. Two large mutational steps can be seen early in the simulation but as the optimal phenotype is approached only small mutation steps are selected for.

A short note on inclusive fitness in simulation modelling

The concept of inclusive fitness, developed by Hamilton (1963, 1964), deals with the additional fitness advantage genes obtain by helping relatives of their bearers to mate. Using analytical models to track inclusive fitness is often difficult (Kokko, 2007). Inclusive fitness is, however, an emerging property of simulation modelling (Gros et al., 2008; Poethke et al., 2007), irrespective of the selection model employed. It is therefore not necessary for any formal statement in the model to deal with inclusive fitness. This property of simulation modelling is unfortunately easy to overlook.

To ensure that the inclusive fitness in the simulation model reflects the biological system, the same genetic setup must be used in the simulation. This is of utmost importance if the trait under investigation has an effect on the mating opportunities of related individuals, because the kin value of different family

8

members differs between these systems. The three genetic systems in this thesis are haploids, diploids and haplodiploids. A brief description of the relatedness within these genetic systems (without prior inbreeding) is as follows: Any individual is related to him/herself by 1 (applicable where selfing takes place, see chapter 2). The average relatedness of a haploid individual to any sibling or either parent is 0.5. It is the same for any diploid individual. In haplodiploids the average relatedness between brothers as well as mothers to daughters is 0.5. The relatedness of sisters to each other is 0.75. The relatedness of haploid males to their mothers is 1 (Crozier, 1970), and to their sisters 0.5, however they have a reduced reproductive value (Price, 1970). Intuitively we know that inbreeding will increase the relatedness of siblings. However, in haplodiploids the reproductive value of females increases more relative to their male counterparts in inbred populations (Hamilton, 1979; Taylor & Bulmer, 1980). In chapter 2 some of the effects of different inclusive benefits, due to different average relatedness, can be seen.

The point of this section, however, is the emerging property of inclusive fitness in simulation modelling. Therefore it needs to be stressed that, all other things being equal, individuals who have more offspring, nephews or nieces will have more of their genes in the next generation. If this is achieved by helping your relatives reproduce, the trait having this effect will be selected for without it being stated formally in the simulation model.

Advantages of simulation modelling

Using simulation modelling to study evolution has many advantages. I have already alluded to some of the advantages, such as, the allowance of greater variation (especially individual variation) than analytical models (DeAngelis & Mooij, 2005; Grimm, 1999) and the natural inclusion of inclusive fitness has also been dealt with. I will mention a number of other advantages of simulation modelling, all of which, to some degree motivated its use, in this thesis:

Individuals are discrete units in individual-based simulation. The system is thus investigated, using a 'bottom-up' approach where population characteristics emerge from the interactions of the individuals. This approach enables the investigator to

track the behaviour of individual organisms (Grimm, 1999), which could aid in unravelling the system dynamics.

When a researcher builds a simulation model, they usually tailor it to suit their research needs. In these models, we therefore have access to all the parameters, which we think are important in the system. Included are parameters that we cannot manipulate in natural systems due to logistic, physiological, budget or ethical constraints (Grimm, 1999; Peck, 2004; Winsberg, 2003). With all the parameters under our control, it should be clear that an experimental approach may help a great deal in understanding the system. In spite of this, modellers often fail to perform methodical experiments (DeAngelis & Mooij, 2005; Grimm, 1999). Additional advantages of the customisability and flexibility of these models are that different data input formats can be handled, specific data output formats can be created (DeAngelis & Mooij, 2005) and it becomes easy to integrate empirical data into the models (Grimm, 1999).

The realism of simulation models is another advantage. Analytical models are unmanageable when they become too complex (Judson, 1994; Łomnicki, 1999; Prescott et al., 2007), and cannot always capture all the fine scale life cycle detail necessary to explain processes in ecological systems (DeAngelis & Mooij, 2005; Łomnicki, 1999; Peck, 2004). Simulation modelling also automatically incorporates subtle interactions often left out from mathematical modelling. It is therefore possible to simulate the parameters from analytical models and compare results (Kokko, 2007), followed by adding more parameters to the simulation models. These additional parameters add more realism to the model (e.g. variation in behaviour and environmental conditions, realistic and finite population sizes, multiple fitness classes etc.), and may lead to the revision and extensions of current theories (Grimm, 1999, see chapters 2 and 3).

When realistic parameters are used, direct comparisons between the model system and the natural system can be made. As a result, simulation modelling can be used to guide planning of empirical experiments (Judson, 1994; Łomnicki, 1999; Peck, 2004; Winsberg, 2003). This was done in chapter 4, where realistic parameters were used to predict optimal sampling schemes in natural populations.

Disadvantages of simulation modelling

It is important to be realistic about the capabilities of any research tool. This section is therefore important to highlight potential pitfalls of individual-based simulation models and when their use is inappropriate. Strangely enough, many of the advantages simulation models have, also border on being reasons for not using it. Foremost is the detail that can be added to a simulation model. It is possible to create a simulation so realistic that it does, not only encapsulate the parameters and processes necessary to address the biological question, but to also have many additional factors which influence the system in a small or unimportant way (DeAngelis & Mooij, 2005; Kokko, 2007). This could result in a simulation that is as complex as the natural system from which no additional understanding of it may be gained (DeAngelis & Mooij, 2005; Kokko, 2007; Peck, 2004; Prescott et al., 2007), and is the reason that some biologists reject the use of simulation models (Judson, 1994; Mitchell, 1998).

The freedom to tailor simulation models also comes at a cost. It is often difficult to create and maintain simulation models, as the model structure is inherently more complex than that of analytical models (DeAngelis & Mooij, 2005; Grimm, 1999; Grimm et al., 2006; Peck, 2004). There is thus no standard way to create a simulation model, and this, together with the complexity, makes it difficult to document and communicate these models (Grimm et al., 2006; Judson, 1994). It is therefore clear that, if an analytical model can be used it should (Kokko, 2007; Peck, 2004), as they are simpler and clearer, and easy to communicate in the general language of mathematics (Grimm et al., 2006).

Assessment of simulation models

From the last two sections it is evident that, before individual-based simulation is used in a study, its possible contribution to unravelling the biological system under investigation should be assessed. Once completed, a useful simulation model would be poised between the last two sections where the benefits are maximised and the costs minimised. There are a number of ways to determine if a simulation model is

11

adequate, but before this is done it is essential to ensure that there are no functional problems (Grimm, 1999; Winsberg, 2003). It is usual to perform a barrage of trials, checking and rechecking the reaction when each parameter is changed or set to its extremes. This is probably the most tedious part of simulation modelling, and only after proper evaluation of the functioning, can experiments be designed from which understanding may be derived (Winsberg, 2003).

To assess the appropriateness of the model the following should be considered (summarised from: Grimm, 1999; Judson, 1994; Peck, 2004; Prescott et al., 2007; Winsberg, 2003): 1) Does the model represent the biological system adequately? The model must accurately capture the biological system and yet be simple enough to be used as a model from which insight into the system can be gained. 2) Do manipulations of the parameters and processes result in expected reactions? Simulations are created to mimic a biological system that has often been studied empirically and or analytically. The simulation model is therefore expected to respond in a similar fashion to changes in parameters, as the natural system or analytical models would. 3) Does the model make predictions? When multiple parameters are changed to encapsulate the complexities of the system, the simulation model should provide an estimate of how the system will react over a number of generations.

The power of simulation modelling does not lie in our ability to simulate natural or theoretical systems. Rather, by systematic manipulation and experimentation on these virtual representations, we attempt to gain insight in the modelled system and apply this knowledge to understanding how the world works (Grimm, 1999; Peck, 2004). Individual-based simulations are therefore very useful research tools when used in conjunction with empirical data. The biological system under investigation should be well studied and specific research questions defined. Once this is done it is important to plan the simulation in as much detail as possible before the actual modelling begins (this reduces the chance of requiring large changes once the model is developed). In addition to individual-based simulations, a "top-down" approach, using analytical models (often done before simulation modelling), improves knowledge of the system processes (Grimm, 1999). It is therefore necessary for biologists using various approaches to collaborate with each other to create a cohesive explanation of natural selection and evolution (Prescott et al., 2007).

Pollinating fig wasps

All of the simulation models in this thesis are grounded in reality, and were mostly inspired by the pollinating fig wasp system. Furthermore, the empirical data used in chapter 5 was obtained from the pollinating wasp species, *Platyscapa awekei*. A short description of their life history and interaction with their hosts is therefore justified.

The pollinating fig wasp and fig tree system is well studied in evolutionary biology and much is known about their life history and mating ecology (Greeff et al., 2003; Hamilton, 1979; Herre et al., 1997; Zammit & Schwarz, 2000). Consequently, the association between specific wasp and tree species is well documented (Cook & Rasplus, 2003; Cook & West, 2005; Corner, 1985; Janzen, 1979; Ramirez, 1970; Wiebes, 1979, but see Michaloud et al., 1996; Molbo et al., 2004; Rasplus, 1996 for exceptions). The role of the wasp in this mutualism is to transfer pollen from the generally inaccessible flowers of one tree to another receptive tree. In return, the wasps are provided with an environment to develop and mate in. During the development of a single crop of figs, pollinating wasps complete their whole life cycle. These events are similar in all pollinating wasp and fig species, and are depicted in figure 1.3 (the deviations from the general pollinating fig wasp life history observed in *P. awekei* are noted below). The development of figs are often categorised into 5 phases, A to E (Galil & Eisikowitch, 1968). During A-phase the fig syconium structure develops. The syconia can be seen as buds on the tree and are completely filled throughout most of this phase (Verkerke, 1989).

During-B phase the flowers inside the syconium becomes receptive to pollination. Volatile compounds released by the figs attract pollinating female wasps (Hossaert-McKey et al., 1994; Van Noort et al., 1989). The foundress females enter the syconium through a small opening, called an ostiole, losing their wings and antennae in the narrow passage (Wiebes, 1979). It is common for a single female or a small number of females to enter a syconium. Once inside, they lay their eggs in some of the flowers and simultaneously pollinate others (Herre, 1989; Kjellberg et al., 2001). A foundress mother is able to control the sex ratio of her
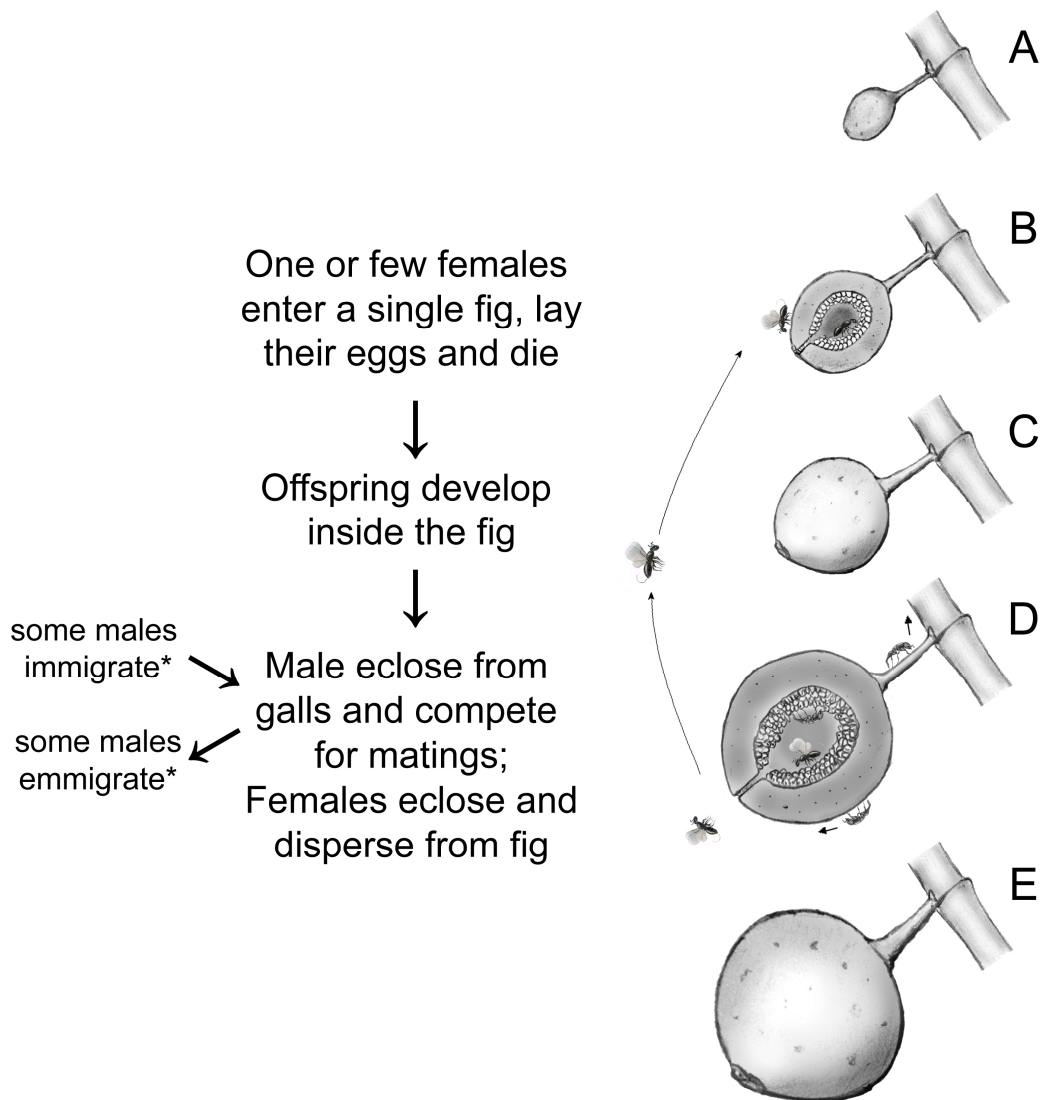
One or few females
enter a single fig, lay
their eggs and die

↓

Offspring develop
inside the fig

↓

some males
immigrate*

some males
emmigrate*

Male eclose from
galls and compete
for matings;
Females eclose and
disperse from fig

A

B

C

D

E

Figure 1.3. The life cycle of pollinating fig wasps and developmental stages of a fig. *P. awekei* are one of only a few species, of pollinating fig wasps, where the males are documented to disperse; see text for details.

offspring by selective fertilisation of her eggs. Fertilised eggs develop into diploid daughters while the unfertilised eggs develop into haploid males (Werren, 1987). Unable to exit from the syconium, a foundress female lays all her eggs in a single syconium and dies (but see Moore et al., 2006).

During C-phase the wasp larvae develop in individually galled flowers by feeding on the endosperm (Verkerke, 1989). At the same time the seeds develop in the flowers without wasp larvae.

D-phase begins when the males emerge from their galls. Pollinator males are wingless and are morphologically dissimilar to their female counterparts (Murray, 1990; Wiebes, 1979). After the males emerge they mate with the females that are still within their galls (Berg & Wiebes, 1992; Zammit & Schwarz, 2000). Since females are receptive only whilst still in their galls, males need to crawl between the galls to reach, and mate with the females (Galil, 1977; Hamilton, 1979; Herre et al., 1997). After mating, the females leave their galls and exit the syconium through a hole chewed by the males. Often the males complete their whole life cycle in their native fig. In a number of species, including *P. awekei*, males have been recorded to disperse from their native syconium and enter into other syconia (Greeff et al., 2003), where they will probably try to find additional mating opportunities. *P. awekei* is also one of a few species where the males engage in contest competition for mating opportunities (Greeff et al., 2003) and a strong association has been found between higher average sex ratios in fighting and dispersing species (Nelson & Greeff, 2009). Females leaving the syconium disperse to new receptive B-phase syconia and start the cycle anew.

Lastly, the syconium ripens, signalling E-phase. During this phase the syconium often swells, changes colour, and is consumed by a number of different animals, which disperse the seeds (Berg & Wiebes, 1992; Burrows & Burrows, 2003).

The life history of pollinating fig wasps have a number of features, which make the system ideal for testing evolutionary theories empirically, analytically and with simulation modelling. An important factor is the low number of foundress mothers per patch. The population is therefore extremely structured and most of the matings are with relatives. This leads to high levels of inbreeding in pollinating fig wasps, which is promoted by purging via the haploid males (Antolin, 1999; Herre et al., 1997; Werren, 1993). However, the observation of dispersing males in some species raises questions on optimal mating strategies and inbreeding depression, which is addressed in chapters 2 and 5. Another effect of the low foundress numbers is the high

levels of local mate competition, which is often reduced by female biased sex ratios or male dispersal (see below). These three factors are addressed in chapter 3. A last feature of pollinating fig wasp life history is the absence of overlapping generations, which make their system a pleasure to simulate.

Aims of this thesis

The aims of this thesis are to extend theories on the evolution of inbreeding and dispersal. Four specific questions are addressed, three using individual-based simulation modelling (chapters 2-4) and one with empirical data obtained from pollinating fig wasps (chapter 5). Following, is a short introduction to each question (due to their diverse nature each chapter is preceded with a complete account of the relevant background from the literature).

1. When should mixed mating evolve?

In nature we see a rich variety of mating systems. At the extremes are exclusive inbreeding and exclusive outbreeding, and in-between is a continuum of various levels of both (Thornhill, 1993). In this thesis, I investigate the conditions necessary for the evolution of stable mixed mating, defined as: a strategy where individual sometimes mate with close individuals (or self in the case of hermaphrodites) while at other times they mate with completely unrelated individuals. In other words, a strategy where individuals employ either of the extreme mating options on a regular basis. Mixed mating strategies are found in a number of organisms where individuals do not have access to intermediate related mating partners (Godfray & Cook, 1997; Goodwillie et al., 2005; Greeff et al., 2009; Hardy, 1994; Jain, 1976).

The default choice of all organisms should be inbreeding, due to the kin advantage gained by mating with relatives (Bateson, 1983; Bengtsson, 1978; Fisher, 1941; Kokko & Ots, 2006; Parker, 1979; Pusey & Wolf, 1996; Waller, 1993; Waser et al., 1986; Wolf, 2000; see also section: 'A short note on inclusive fitness in simulation modelling'). This decision is however offset by the possible fitness losses due to inbreeding depression (Bengtsson, 1978; Fisher, 1941; Pusey & Wolf, 1996; Waser et

al., 1986). Conversely, outbreeding depression would shift the strategy towards inbreeding (Charlesworth & Charlesworth, 1987; Knight, 1799; Mather, 1955; Williams, 1975). In some circumstances both inbreeding and outbreeding depression may be present simultaneously and matings with distantly related individuals may be optimal (Bateson, 1983; Price & Waser, 1979). As mentioned, in some organisms, this is not an option.

When investigating mating systems it is therefore necessary to explore the fitness level of offspring resulting from inbreeding or outbreeding. A number of models have laid the foundation for predicting, when inbreeding, outbreeding or mixed mating should occur, from the ensuing offspring fitness levels (Campbell, 1986; Charlesworth & Charlesworth, 1987; Damgaard et al., 1992; Feldman & Christianson, 1984; Holsinger, 1988; Holsinger et al., 1984; Lande & Schemske, 1985; Latta & Ritland, 1993; Lloyd, 1979; Maynard Smith, 1977; Maynard Smith, 1978a; Taylor & Getz, 1994; Uyenoyama & Waller, 1991a; Uyenoyama & Waller, 1991b; Waser et al., 1986). In chapter 2, a simulation approach is used to extend current analytical models. This is done by exploring: optimal mating strategies ranging from inbreeding depression to outbreeding depression (including where intermediate inbreeding has the highest fitness); stochastic demographic variables; multiple fitness values for serially inbred individuals; sibmating diploids and haplodiploids, in addition to, selfing diploids.

2. Does the sex ratio affect dispersal?

In 1871, Darwin noted the equal proportions of males and females in many species. He proposed the first theory of the adaptive nature of sex ratios (Darwin, 1871), but being unable to convince himself, he retracted it (Darwin, 1874). Fisher (1930) showed that natural selection would promote the under-represented sex, which would lead to equal sex ratios. The theory of sex ratios was, however, vastly extended by Hamilton (1967), who included non-equal reproductive values of males and females and structured populations to his models. He showed that, in a number of organisms, only one or a few females contribute offspring to discrete patches in a population. If mating took place in these patches, without prior dispersal, competition for mating opportunities was mostly between relatives. This phenomenon is termed Local Mate

Competition (LMC), and females often bias their sex ratios to limit its occurrence (for examples in fig wasps see: Frank, 1985; Herre, 1985; Herre et al., 1997).

Simultaneously, theories on when dispersal should evolve were also advanced. Once again, LMC was recognised early on as an important factor, in this case, driving the evolution of dispersal (Hamilton & May, 1977; Van Valen, 1971; see for reviews: Clobert et al., 2001; Ronce, 2007). Two other factors, eminent in the evolution of dispersal and usually included in dispersal models, are inbreeding depression and the cost of dispersal (Bengtsson, 1978; Clobert et al., 2001; Frank, 1986; Gandon, 1999; Hamilton & May, 1977; Motro, 1983; Motro, 1991; Perrin & Mazalov, 2000; Ronce, 2007; Taylor, 1988). It is apparent that, while low foundress numbers (i.e. high LMC) and the occurrence of inbreeding depression will increase the rate of dispersal, high cost of dispersal will oppose the evolution thereof.

Theories of dispersal and sex ratio evolution often overlap and a number of models have investigated the co-evolution of these factors (Hamilton, 1967; Perrin & Mazalov, 2000; Taylor, 1994; Taylor & Bulmer, 1980; Wild & Taylor, 2004). While these models show that the sex ratio is affected by the rate of dispersal, dispersal is often not affected by the sex ratio. In chapter 3, the effect of sex ratios on dispersal is explored. This is done with simulation modelling where: binomially distributed sex ratios are produced; males decide to disperse in response to their native patch sex ratio; different foundress mothers per patch are allocated. All this is done in the absence of inbreeding depression.

3. How many samples are necessary to accurately detect population structure?

Original ideas for measuring genetic distance between populations were developed by Wright (1931). With the advent of molecular genetics a number of practical methods were derived to estimate population differentiation. Collectively called summary statistics, they differ from Wright's Fst and include: Gst, where heterozygosity is used to estimate population subdivision (Nei, 1987); θ, where variance components are used to determine population structure (Weir & Cockerham, 1984); and Rst, where rare alleles are used, taking a stepwise mutation model of microsatellites into account,

to determine if there is population structure (Slatkin, 1995). (It should be noted that the method described by Weir & Cockerham (1984) is commonly used in analysis of population subdivision. Their estimate will be referred to as the default Fst from here onwards).

New generation programs have changed the way in which population structure is determined (Neigel, 2002; Pearse & Crandall, 2004). These programs, of which STRUCTURE is currently the most popular, uses likelihood methods to infer population subdivision (Falush et al., 2003; Kaeuffer et al., 2007; Neigel, 2002; Pearse & Crandall, 2004; Pritchard et al., 2000). In spite of this, summary statistics are still routinely used in combination with newer techniques (Pearse & Crandall, 2004; Neigel, 2002; Balloux & Lugon-Moulin, 2002). The power of all these methods is, however, still dependent on the number of loci and individuals sampled.

The aim of chapter 4 is to provide simple guidelines when planning a sampling scheme. The program used was originally developed to simulate a large population that split, which were then used to investigate how drift, stepwise mutation and low migration rates between the two populations increase their genetic distance. The program was later amended by adding an extensive sampling scheme generator. This allowed multiple different samplings (i.e. different sample sizes and or different number of loci) from the simulated population. The output generated comprises of summary statistics, which can be compared and analysed to determine the minimum sampling requirements. Alternatively, the allele values of the sampled individual can be obtained, that can be analysed independently using current genetic programs. In chapter 4 a number of sampling schemes are investigated and general sampling guidelines are presented.

4. Are mixed mating and male dispersal in Platyscapa awekei, due to intermediate inbred individuals having the highest fitness?

Pollinating fig wasps are famous for being extremely inbred (Herre et al., 1997; Molbo et al., 2002; Molbo et al., 2004), but a few including *P. awekei*, have been noted to display mixed mating strategies (Greeff, 2002; Greeff et al., 2009; Greeff et al., 2003; Herre et al., 1997; Jansen van Vuuren et al., 2006; Molbo et al., 2002;

Molbo et al., 2004; Moore et al., 2006; Zavodna et al., 2005). A number of studies suggest possible reasons for the dispersal phenotype in *P. awekei* and these include less female biased sex ratios (Nelson & Greeff, 2009) and reduced LMC between brothers (Moore et al., 2006; Nelson & Greeff, 2009), a consequence of fighting morphology (Greeff et al., 2003; Nelson & Greeff, 2009) and simultaneous inbreeding and outbreeding depression (Greeff et al., 2009). The last study however had ambiguous results and it is not clear if *P. awekei* suffers from outbreeding depression only, or a combination of both inbreeding and outbreeding depression (Greeff et al., 2009).

In chapter 5, I examine the fitness of various levels of inbred *P. awekei* females. This is accomplished by comparing the number of offspring (as a proxy for fitness) to the number of homozygous microsatellite loci for each mother. The results are used in the simulation model "SibMate" (chapter 2) and the predicted mating system of *P. awekei* is compared and discussed.

2. Evolutionary stable mixed mating in a variety of genetic systems

Abstract

In nature, individuals often have to choose between mating with close relatives and unrelated individuals. In some species however, a mixture of close inbreeding and outbreeding co-exists and is referred to as mixed mating. A number of theoretical models can explain the existence of mixed mating. We simulate the evolution of mating preferences for three genetic systems, diploid selfing, diploid sibmating and haplodiploid sibmating. Mating preferences are determined by a single locus, while the fitness of an individual depends on the level of inbreeding. Fitness combinations that allow stable mixed mating strategies to evolve include low levels of inbreeding depression and almost all scenarios where intermediate inbreeding is optimal. We find that stable mixed mating readily evolves when a minimum of at least three fitness levels are specified (in contrast to two fitness levels: one for inbreeding and one for outbreeding). We also find that stable mixed mating can evolve for selfing diploids, sibmating diploids and sibmating haplodiploids. The high relatedness of the selfing individuals lead to lower levels of inbreeding followed by sibmating diploids and lastly sibmating haplodiploids. Comparing our results with empirical data give an indication of when mixed mating can be expected.

Introduction

When individuals choose a mate they often have to decide between a close relative or an unrelated individual. In plants and simultaneous hermaphrodites, selfing is an additional option. Various mating strategies could therefore evolve, ranging from exclusive inbreeding to exclusive outbreeding, and a myriad of mating strategies are indeed found in nature (see: Keller & Waller, 2002; Thornhill, 1993; for reviews). The optimal mating strategy is dependent on genetic and environmental pressures acting on an individual, as well as, the dynamics and breeding strategy of the population (Keller & Waller, 2002).

The fitness of an individual with a specific mating strategy is affected by inclusive fitness benefits, outbreeding depression and inbreeding depression. Inclusive fitness benefits are gained from extra mating opportunities to related individuals and favours the evolution of inbreeding (Bateson, 1983; Bengtsson, 1978; Fisher, 1941; Kokko & Ots, 2006; Parker, 1979; Pusey & Wolf, 1996; Waller, 1993; Waser et al., 1986; Wolf, 2000). Outbreeding can be detrimental due to the break-up of co-adapted gene complexes or the loss of adaptation to local environments (Bengtsson, 1978; Fisher, 1941; Pusey & Wolf, 1996; Waser et al., 1986). Inbreeding depression and its causes, such as overdominance, are well documented (Charlesworth & Charlesworth, 1987; Knight, 1799; Mather, 1955; Williams, 1975). If inbreeding depression is severe it will lead to inbreeding avoidance (see Kokko & Ots, 2006). Conversely, outbreeding depression should lead to exclusive inbreeding (Lynch, 1991).

It has been suggested that if inbreeding and outbreeding depression are present simultaneously, these two factors will oppose each other and an intermediate level of inbreeding could be the optimal mating strategy (Bateson, 1983; Price & Waser, 1979). In these circumstances, matings with close kin are avoided but so are matings with completely unrelated individuals (Bateson, 1983). However, individuals often only have full sibs or unrelated individuals available and not individuals of the optimal intermediate genetic distance. Here, a mixed mating strategy, consisting of close inbreeding or complete outbreeding of individuals, may be a "frustrated" way to solve the evolutionary puzzle (Greeff et al., 2009). Under these conditions, a population should have both inbreeding and outbreeding as mating strategies, each occurring at a stable frequency.

A number of models did not find equilibria where mixed mating was stable. Rather, they showed that if inbreeding depression became too severe the equilibrium switched from complete inbreeding to complete outbreeding. For instance, it is well known that self compatible diploids species should switch from outbreeding to inbreeding if the fitness of inbred individuals are more than half that of the outbred individuals, and *vice versa* (Charlesworth & Charlesworth, 1987; Feldman & Christianson, 1984; Holsinger et al., 1984; Lande & Schemske, 1985; Lloyd, 1979; Maynard Smith, 1977; Maynard Smith, 1978a; Waser et al., 1986). Taylor & Getz, (1994) explored sibmating in diploid and haplodiploid individuals with two fitness classes, namely inbred and outbred. They found no fitness combination where mixed mating was stable for sibmating diploids. For sibmating haplodiploids they found that stable mixed mating may occur under very specific conditions. Stable mixed mating occur when inbreeding reduces the fitness of the offspring between 30% and 33%, and when the population starts as mostly outbreeding individuals (Taylor & Getz, 1994).

Maynard Smith (1977) explored models where serial inbreeding led to a continuous reduction in fitness. He showed that mixed mating systems could evolve as pure inbreeding or outbreeding populations were not resistant against opposite strategies invading. In support, Latta & Ritland (1993) showed, using recursive equations for multiple loci, that a decline in fitness of serial inbred individuals led to stable mixed mating. They found that mixed mating evolved at higher frequencies when controlled by increasing numbers of genes. In addition, they showed that mixed mating remained stable in the face of purging if inbreeding depression was sufficiently small (Latta & Ritland, 1994). In a similar study Damgaard et al. (1992) also found stable mixed mating for selfing individuals at different levels of inbreeding depression. In these studies (as in ours) a fixed relationship between fitness and the level of inbreeding was assumed. This allows for easy comparison to experimental data, but ignores the potential coevolution between fitness and inbreeding. Genotypes for inbreeding and genotypes that tolerate inbreeding becomes linked and may determine the direction in which the mating system evolves (Campbell, 1986; Holsinger, 1988; Uyenoyama & Waller, 1991a; Uyenoyama & Waller, 1991b). However, empirical data should include these dynamic effects and may then easily be combined with the fixed relationship models, such as we will develop below.

By considering serial classes of inbreeding Campbell (1986) also found that mixed mating could be maintained, even in systems where inbreeding depression coevolves with the mating preference loci. Campbell's models, however, only explored serial inbreeding for selfing diploids and it would be difficult to adjust for sibmating diploids or haplodiploids, as their fitness needs to take into account the inbred status of individuals two generations prior.

In this study we investigate the optimal mating strategy that emerges in three genetic systems (selfing diploids, sibmating diploids and sibmating haplodiploids) with different fitness levels for each class of serially inbred individuals. These genetic systems have different inclusive fitness benefits to inbreeding: A female is related to herself by 1 and to her diploid brother by 0.5. In haplodiploids, a male related to this sister by 0.5, but this value needs to be discounted by the reduced reproductive value of haplodiploid males (Price, 1970), giving a kin value of 0.25. We should thus expect that inbreeding should evolve less readily in sibmating haplodiploids than in sibmating diploids than in selfing diploids.

We used individual-based simulation modelling to determine the optimal mating strategy for individuals of three genetic systems over a range of fitness values. Note that fitness values were not allowed to co-evolve with the mating strategy. Fitness values were chosen to reflect inbreeding depression, outbreeding depression, inbreeding depression with purging, and situations where intermediate inbreeding has the highest fitness.

Model description

The purpose of this model is to understand when mixed mating will evolve in the context of different fitness specifications for individuals. To test the effects of inclusive fitness we consider three genetic systems that differ in kin benefits when inbreeding occurs: selfing diploids, sibmating diploids and sibmating haplodiploids. The fitness values used were different for a range of successive inbred individuals rather than having only one fitness value for inbred individuals which is irrespective of the number of times inbred.

State variables

The model consists of two structural levels: individuals and populations. Each individual has the following state variables: allele-A, allele-B, ancestry and fitness. The allele-variables are memory allocations for two alleles at a single locus and are only defined as A and B to distinguish between the two assigned values. Each allele allocated to an individual has a value from 0 to 1 that defines the probability that the individual will sibmate (for diploids the phenotype is the average of the two alleles, i.e. additive). This phenotype is expressed in the females (i.e. the females decide if they want to sibmate or outbreed). In cases where the individuals are haplodiploid, the B-allele for the males is not defined. The ancestry variable records the breeding history and it is reset to 0 when an individual is mated with an unrelated individual. Each successive sibmating, or selfing, increases the ancestry variable by 1 and is inherited by the offspring. The fitness of each individual is dependant on the ancestry variable and is thus a function of the organism's degree of inbreeding. This is a key component in this study as it allows individuals that have different degrees of inbreeding to have different fitness values. In each simulation the fitness level for each inbred class is explicitly defined in the experimental procedure (see below).

When the population is sibmating it is composed of two arrays of individuals (male and female). When the population is selfing the arrays collapse into one. The population is furthermore characterised by mutation rate (for the locus determining selfing or sibmating), number of generations, ploidy of individuals, number of evenly spaced alleles (e.g. for 11 alleles their values would be: 0; 0.1; 0.2; …; 1) and fitness of the inbreeding classes.

Simulation process and scheduling

The model proceeds in generation time steps (figure 2.1). In each generation step the events are processed as follows: Fitness is assigned for all individuals in the population, mating commences and the newly created population replaces the old population (i.e. no overlapping generations). The following sub-steps are followed during mating until the new population is equal in size to the old population: selection
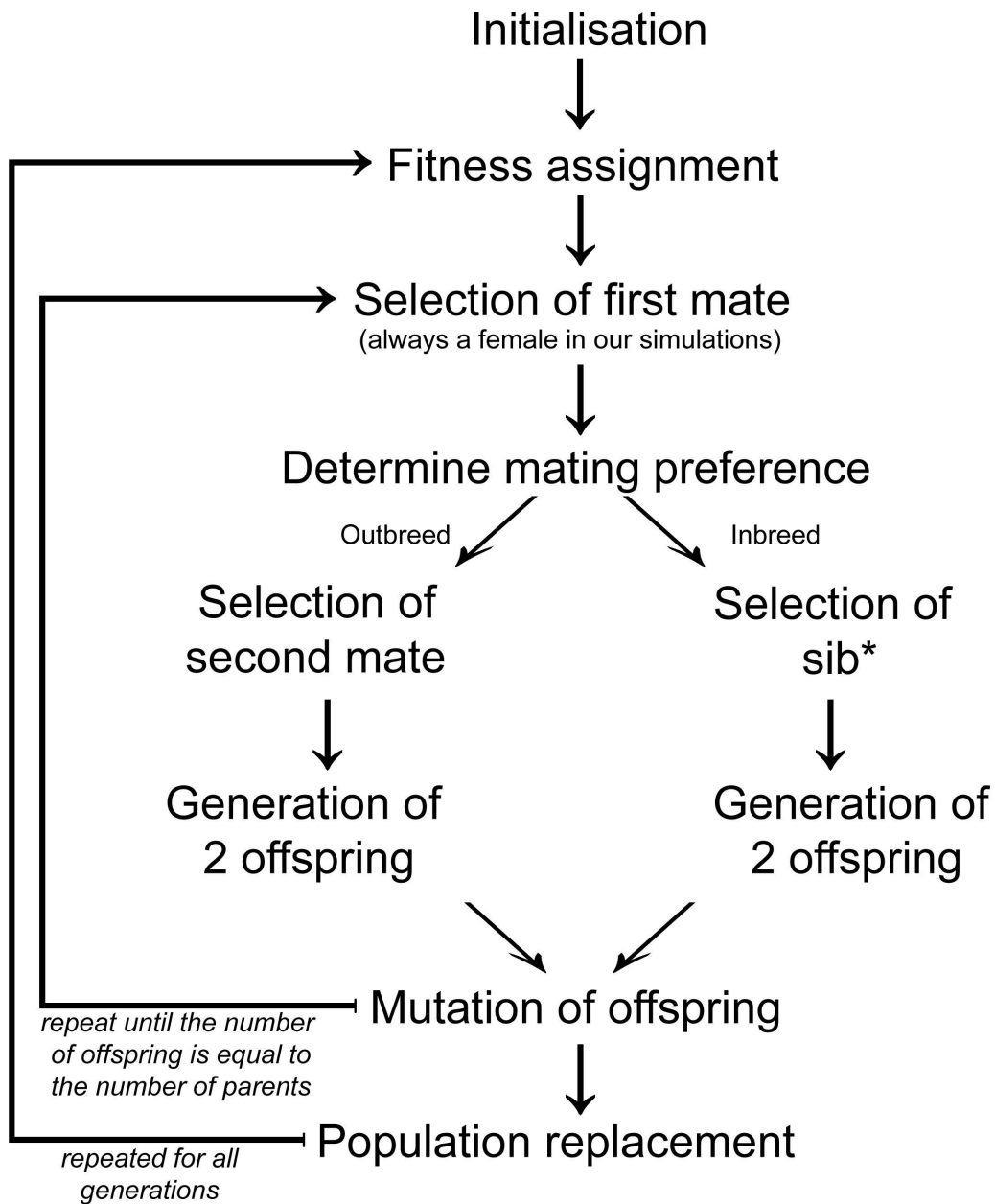
Initialisation

↓

Fitness assignment

↓

Selection of first mate
(always a female in our simulations)

↓

Determine mating preference

Outbreed ↙   ↘ Inbreed

Selection of
second mate

Selection of
sib*

↓   ↓

Generation of
2 offspring

Generation of
2 offspring

↘   ↙

Mutation of offspring

*repeat until the number
of offspring is equal to
the number of parents*

↓

Population replacement

*repeated for all
generations*

Figure 2.1. Process sequence of model. *If the population is selfing rather than sibmating the first mate will also be the second mate when inbreeding.

of an individual to mate, selection of a mate, creating two offspring from parental genotypes and mutation of the offspring.

Design concepts

Optimal mating strategies emerge from the population dynamics but the population dynamics are entirely characterized by rules specifying an individual's behaviour. As mentioned, the fitness of each inbreeding class is specified explicitly and does not change during the simulation. The fitness of individuals more inbred than the last defined class have the same fitness as the individuals in last defined class. A fitness proportion scale with "roulette wheel" sampling is used to obtain the mating individual (Mitchell, 1998). In brief, this means that the fitness of each individual is weighted, relative to the total fitness of the population, and assigned a proportion of the total fitness (i.e. fitter individuals have more numbers on the roulette wheel assigned to them). A random value is drawn between 0 and the total fitness of the population and the individual whose assigned proportion includes this value is selected to mate. Note that the fitness of haploid males was unaffected by their ancestry as all their loci are hemizygous and homozygosity is unaffected by their inbreeding history.

Data gathered for analysis included the probability of inbreeding (α) in the population (calculated as the average inbreeding of the last 500 generations) as specified by the alleles of the individuals.

It is important to note that inclusive fitness does not have to be introduced explicitly since kin-advantage will emerge by default in any individual-based simulation (Gros et al., 2008; Poethke et al., 2007).

Initialisation and input

The population input values are given in table 2.1. The initial inbreeding frequency of the population was established by assigning two alleles to each individual at the first generation. These values were from a uniform distribution ranging from 0 to 1 (in this chapter we used values with a 0.1 interval or 0.01 interval, therefore providing 11 and

Table 2.1. State variables and default input values.

Sate variable	Experiment 2.1	Experiment 2.2	Experiment 2.3
	Initial input values		
Population parameters			
Population size	10000	10000	10000
# Generations	5000	5000	10000
Mutation rate	0.001	0.001	0.0001
# Possible alleles	11	11	101
Individual parameters			
Allele-A	from uniform distribution	0, 1 and uniform distribution	from uniform distribution
Allele-B	from uniform distribution	0, 1 and uniform distribution	from uniform distribution
Ancestry	0	0	0
Fitness	derived from ancestry	derived from ancestry	derived from ancestry
Fitness parameters			
W_0 (Outbred)	1	1	1
W_1 (Once inbred)	0.01 to 2 (step size 0.01)	0.2 to 2 (step size 0.2)	*b*, see figure 2.2
W_2 to W_∞	0.01 to 2 (step size 0.02)	0.2 to 2 (step size 0.2)	*b*, see figure 2.2

101 possible allele values respectively, see table 2.1). In some experiments the initial inbreeding frequency was specified and here all the alleles in the population had the same value. The individual's ancestry was set to 0 (i.e. initial population is outbred). The females' genotype was used to determine if she would sibmate or not.

Sub-models

Mutations: The probability that an allele was selected to mutate was 0.0001 per individual per generation. The mutational model used was a combination of stepwise and jumping mutations. Once an allele was selected to undergo a mutation, there was a 20% chance for it to become any of the possible alleles. Alternatively, the allele mutated one step to either of the adjacent alleles with equal probability.

Mating: When two individuals mated, each transferred one allele to the new offspring, if the offspring was diploid. The haploid offspring received an allele only from the diploid parent. In both cases the allele that each parent donated was chosen at random, i.e. normal Mendelian inheritance. Each mating pair always produced one male and one female offspring.

The use of probability distributions during mutation, mate selection and the choice of females, causes noise in the model so that each run will reach a stable point along a different trajectory and may not reach the same equilibrium. Therefore the inbreeding optima have a rugged appearance across the parameter space.

Experiments

We first investigate the possibility of a stable mixed mating strategy evolving with three, rather than two inbred classes (outbred, once inbred, twice or more times inbred). This was done for all three genetic systems by simulation of many different fitness combinations (high resolution, 20000 parameter sets per genetic system). Second, we investigate the effect of the mating strategy of the initial population on the evolution of subsequent strategies. This was done for all three genetic systems simulations (low resolution, 100 parameter sets per genetic system) of the three fitness classes with specified initial mating strategies. Last, by simulating a population

with 21 fitness classes for all three genetic systems, we show that although stable mixed mating strategy requires at least 3 fitness classes, more classes will easily lead to stable mixed mating. These simulations also allow us to compare our simulations to empirical estimates of the inbreeding coefficient and the mating system.

Experiment 2.1: Optimal mating strategies with three inbreeding classes

All three genetic systems (selfing diploids, sibmating diploids and sibmating haplodiploids) had the following fitness parameters during simulation. The fitness of the outbred individuals (W_0) was fixed at 1. Simulations were performed with stepwise increases of the fitness level of either the once-inbred individuals (W_1) or twice-inbred individuals (W_2). We explored the following fitness ranges $0.01 \leq W_1 \leq 2$ and $0.01 \leq W_2 \leq 2$. The resolution for W_1 was 0.01 for each successive simulation and 0.02 for W_2 (larger increases were used due to computational time constraints). In the selfing diploid and the sibmating haplodiploid the following fitness combinations were not simulated: $W_1 = 0.01$ to 2 and $W_2 = 1$ to 2 only where $W_1 < W_2$ (these areas were however simulated at lower resolution in experiment 2.2, see results and discussion). All the individuals that had a higher inbreeding level than twice-inbred individuals had the same fitness as the twice-inbred individuals (i.e. W_3 to $W_\infty = W_2$).

Experiment 2.2: The importance of initial conditions on the final equilibria

We simulated three inbreeding classes with different initial inbreeding frequencies for all three genetic systems. We fixed the fitness of outbred individuals (W_0) at 1. The same fitness ranges for once- (W_1) and twice- (W_2) inbred individuals were explored as in experiment 2.1 ($0.01 \leq W_1 \leq 2$ and $0.01 \leq W_2 \leq 2$) but the stepwise increase for both W_1 and W_2 were 0.2. For all three genetic systems all the simulations were repeated twice. Once with the initial inbreeding frequency of the population set to 0 and once set to 1.

Experiment 2.3: Optimal mating strategies with 21 inbreeding classes

We assigned 21 fitness classes (W_0 to W_{20}) for all three genetic systems. The fitness for W_0 was fixed at 1.00 for all the simulations. The fitness of each inbred class was calculated as: $W_i = e^{-bfi}$, where *-b* is the slope of fitness decline on the inbreeding coefficient (*f*) (Charlesworth & Charlesworth, 1987). We then performed 8 simulations, each with a greater level of inbreeding depression, i.e. a larger value of *b* (figure 2.2). The cost of inbreeding for each genetic system ranged from $b = 0.1$ to $b = 2$ (figure 2.2, we explored slopes outside this range but the population fixed a single strategy at these points, see results and discussion). Since the fitness is calculated from the inbreeding coefficient, *f,* and not the number of times an individual is inbred, *f* needs to be calculated for each class of inbredness: for sibmating individuals *f* was calculated as: $f_i = f_{i-1}/2 + f_{i-2}/2 + 1/4$ and for selfing individuals as: $f_i = (1 + f_{i-1})/2$, where $f_{-1} = f_0 = 0$ from Lynch & Walsh (1998). Note that the inbreeding coefficient for offspring of full sibmating in haplodiploids is equal to that of diploids (Wright, 1969). The inbreeding coefficient was almost 1 after 20 serial sibmating events (figure 2.2a) and reached the same level after only 8 serial selfing events (figure 2.2b). Each simulation was repeated four times.

Program

The code for the simulation program was written in Delphi Professional version 7.0. Although the code used in all the simulations is similar, minor alterations were made to enable repeated simulations with different fitness combinations or initial inbreeding strategies. The application "SibMate" (standard version for single simulations only and no automatic repeats) is available in the electronic appendix. It may be used to simulate populations with up to 21 breeding and fitness classes, which can be easily compared to empirical data.
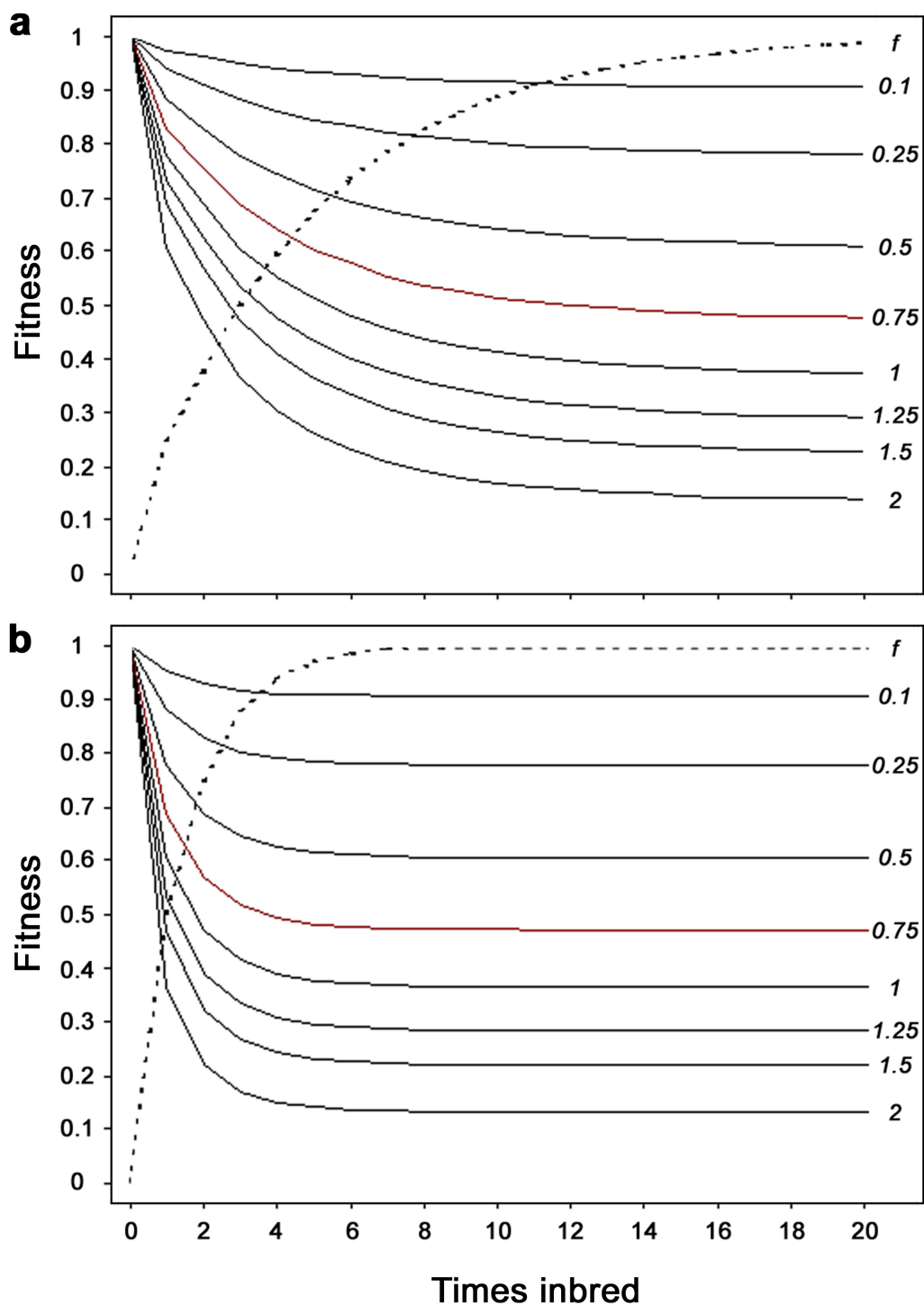
Figure 2.2. The increase in the inbreeding coefficient, f, during 20 sequential inbreeding events (dotted lines) and the corresponding decrease in fitness (solid lines) for a) both sibmating systems and b) the selfing system. Values in italics indicate the cost of inbreeding, b, that relates the inbreeding coefficient to a fitness value ($W = e^{-bf}$).

Results and Discussion

Our results showed that mixed mating strategies can be stable for a range of fitness values, that the degree of inbreeding increases with the inclusive fitness benefits of inbreeding, that mixed mating is expected for realistic values of inbreeding depression and that the starting conditions have a minor influence on the final equilibria that are reached.

From experiment 2.1 we produced the plots in figure 2.3. The contours present the probability of inbreeding (α), which was calculated for each simulation from the average value of the alleles in the whole population of the last 500 generations (i.e. generation 4500 to 5000, while stability was reached after approximately 1000 generations). Each of the simulations had different fitness values for the once- (W_1) and twice- (W_2) inbred individuals as indicated on the y- and x-axes respectively, with a total of 10 000 simulations for each genetic system. Note that figure 2.3d indicates the general fitness trends within each parameter space as bar charts and may be used as a key for the other plots in figure 2.3.

Mixed mating was stable over a large area of the parameter space: this included low levels of inbreeding depression, as well as, scenarios where intermediate inbreeding had the highest fitness. This result is in contrast to models with only two fitness values, one for inbreeding and one for outbreeding, where mixed mating is not stable.

The three genetic systems are indicated in figure 2.3a-c. By comparing these graphs it is clear that inbreeding was more frequent in selfing than sibmating diploids, and that it was more common in diploid than haplodiploid sibmating individuals. This is as we expected and is because the kin benefits are the highest for the selfing diploids and the lowest for the sibmating haplodiploids.

A number of published models explored only two classes of inbredness, i.e. exclusive inbreeding W_{inbred} or exclusive outbreeding W_0 for selfing diploids (Charlesworth & Charlesworth, 1987; Feldman & Christianson, 1984; Lloyd, 1979; Maynard Smith, 1977; Maynard Smith, 1978a; Taylor & Getz, 1994) and for sibmating diploids and haplodiploids (Taylor & Getz, 1994). The diagonal lines in the figure 2.3 correspond to these models as the fitness for inbreeding once, twice or more are the same (i.e. $W_{inbred} = W_1 = W_2$ to W_∞). We find that selfing diploids switch from
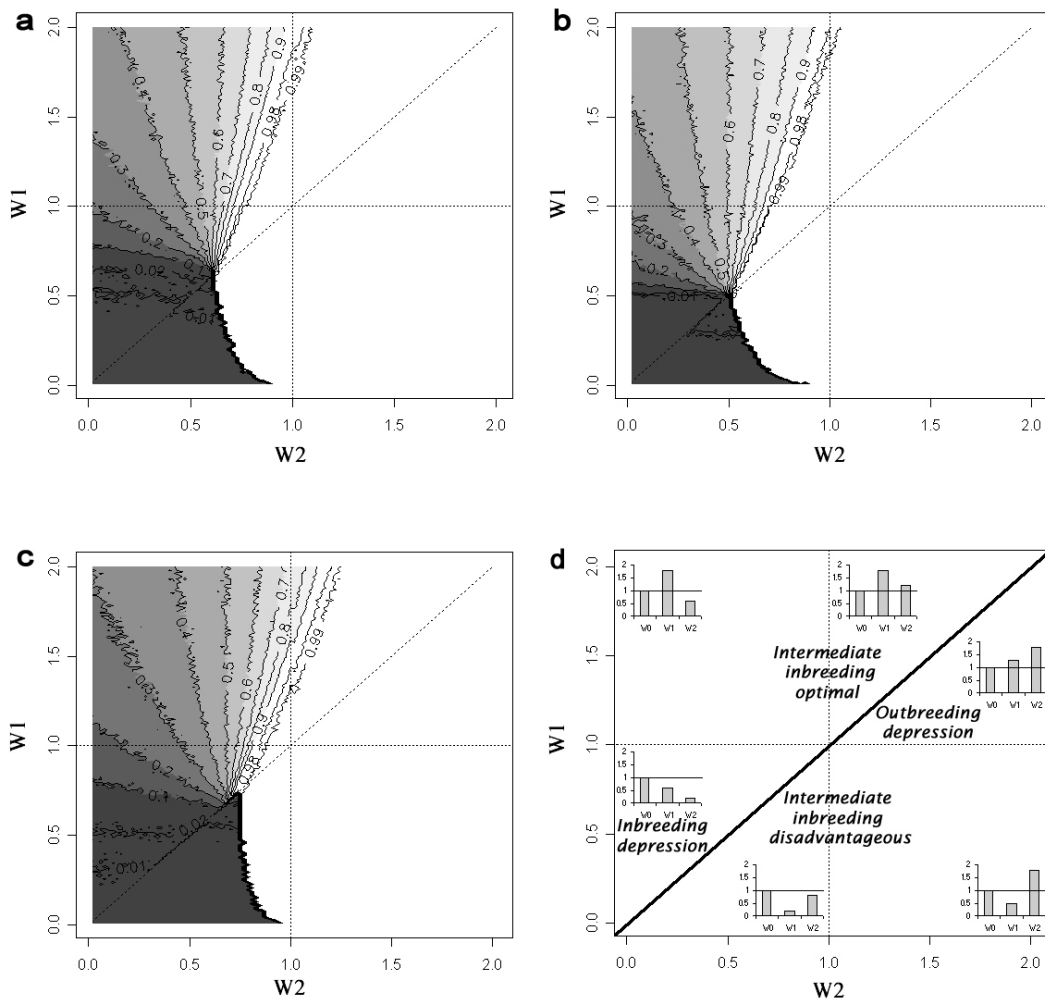
Figure 2.3. Contour plots of the probability of inbreeding (α) for a) Sibmating diploids b) selfing diploids and c) sibmating haplodiploids from simulations of different fitness values for once- (W_1) or twice- (W_2) inbred individuals. d) Key for plots a-c with bar charts in each area representing the general fitness trends of that parameters space. The diagonal line in each plot indicates where there is only one fitness level for all the inbreeding classes and one for the outbreeding class. White areas represents exclusive inbreeding ($\alpha = 1$) while the darkest area represents exclusive outbreeding ($\alpha = 0$). Shades of grey show different levels of mixed mating.

selfing diploids switch from exclusive inbreeding to exclusive outbreeding at $W_{inbred} =$ 0.51 (figure 2.3b). This is similar to the aforementioned studies where the switch is found at an inbreeding depression of 50% for selfing diploids.

Our results for sibmating also agreed with previous models where only two classes of inbredness were explored (Taylor & Getz, 1994) and we found the switch from exclusive inbreeding to outbreeding at levels of inbreeding depression close to that of Taylor & Getz (1994) ($W_{inbred} = 0.61$ for sibmating diploids and at $W_{inbred} \approx$ 0.70 for the sibmating haplodiploids, along the diagonals of figures 2.3a and 2.3c respectively). The higher switch points for the selfing diploids than the sibmating haplodiploids than the sibmating diploids were due to the different inclusive fitness advantages as mentioned above.

In the midpoint of each plot in figure 2.3 the fitness of all the inbreeding classes was the same ($W_0 = W_1 = W_2 = 1$). Even though fitness values were equal the kin advantage favoured inbreeding for all three genetic systems.

The area in the lower left quadrant above the diagonal is where different levels of inbreeding depression were simulated and outbreeding evolved (darkest areas in figure 2.3a-c), especially at more severe inbreeding depression. However, in all three genetic systems there were stable mixed mating strategies when inbreeding depression was reduced (shaded areas, figure 2.3a-c). The reason for mixed mating in this area is a high fitness value of the once-inbred individuals (W_1), regardless of a low value for twice-inbred individuals (W_2). However, as twice inbred individual's fitness increased, inbreeding as a strategy could be maintained and exclusive inbreeding evolved.

The contour lines for α at 0.01 and 0.02 were drawn to give an indication of the gradient of the switch from outbreeding to mixed mating. The selfing diploid system (figure 2.3b) had a sudden switch from outbreeding to mixed mating, and had the smallest area of exclusive outbreeding. Both the sibmating diploid and haplodiploid systems (figures 2.3a and 2.3c) had slightly less steep gradients. The sibmating models had larger areas of exclusive outbreeding with the haplodiploid having the largest (again the inclusive fitness differences between the genetic systems were the cause for the different results).

Two areas were simulated where intermediate inbreeding was optimal (i.e. once-inbred individuals were the fittest; W_1 higher than W_0 and W_2). First, the top left quadrant: here, in addition to the once-inbred individuals being the fittest, outbred

individuals had a higher fitness than the repeated inbred individuals ($W_1 > W_0 > W_2$). A large part of the parameter space in this area led to the evolution of stable mixed mating (indicated by shades of grey, figure 2.3a-c). This result confirmed that stable mixed mating would evolve if intermediate inbreeding were optimal, given that intermediately related individuals are not available or recognisable. Again the high value of the once-inbred individuals (W_1) favoured inbreeding, but a low value of the twice-inbred individuals (W_2) prevented it from being fixed and mixed mating evolved as a stable strategy. However, where the twice-inbred individuals (W_2) had high fitness, inbreeding was maintained in the population and mixed mating became more biased towards inbreeding. The different kin advantages again caused selfing populations to fix inbreeding at lower levels of W_2, followed by sibmating diploids and then sibmating haplodiploids.

The second area where once-inbred individuals had the highest fitness was the top right quadrant above the diagonal. Here, however, the twice-inbred individuals were fitter than outbred individuals ($W_1 > W_2 > W_0$, see figure 2.3d) and inbreeding was the most prevalent strategy that evolved (figure 2.3a-c). In these circumstances the reduction in fitness associated with outbreeding in order to proceed to the highest fitness of the once-inbred class was too great to allow stable mixed mating. This is a special case of outbreeding depression within an area where intermediate inbreeding is specified as optimal, but never reached. The contour lines for α at 0.98 and 0.99 were drawn to give an indication of the swiftness of the switch from exclusive outbreeding to mixed mating. The switch in all three genetic systems was rapid and relatively similar.

Outbreeding depression was simulated in the parameter space at the top right quadrant below the diagonal (see figure 2.3d). For sibmating diploids we found, as expected, that only inbreeding evolved as a stable strategy (the same results are found for sibmating haplodiploids and selfing diploids at lower resolution, figure 2.4, and inferred for figure 2.3).

Fitness values in the lower right quadrant are unlikely to be found in nature. Here, once inbred individuals had the lowest fitness while repeatedly inbred individuals had the highest fitness ($W_2 > W_0 > W_1$, see figure 2.3d). Therefore, only the sibmating diploid system was simulated to include the lower right quadrant in experiment 2.1. We expected all three genetic systems to fix exclusive inbreeding as

the only stable strategy and verified it with low-resolution simulations of this fitness area for the two remaining genetic systems (experiment 2.2).

Purging of recessive deleterious mutations is found in many natural situations (Crnokrak & Barrett, 2002), and this may lead to exclusive inbreeding. In the lower left quadrant below the diagonal, various levels of purging were simulated (i.e. once inbred individuals had the lowest fitness while outbred individuals had the highest W_0 > W_2 > W_1, see figure 2.3d) and we see there was a single switch from outbreeding to inbreeding without any stable mixed mating (figure 2.3).

In natural populations we also expect that there should be some degree of purging and a loss of inbreeding depression (Lloyd, 1979), particularly if there are multiple serial inbreeding events. Purging is slow if the cost of inbreeding is low and inbreeding depression may be maintained (Charlesworth et al., 1990; Lande & Schemske, 1985; Latta & Ritland, 1994). Even lethal mutations, which will lead to a large decrease in fitness of the inbred individuals, may be maintained in moderately selfing populations (Latta & Ritland, 1994).

The haplodiploid genetic system is very effective at purging deleterious alleles as these alleles are exposed to selection in the haploid males. This will mean that haplodiploid taxa will tend to be have less inbreeding depression (Antolin, 1999; Bruckner, 1978; Werren, 1993) and will be situated more to the right and above diploid populations in figure 2.3. This will agree with the observation that haplodiploids in natural populations often inbreed (Hamilton, 1967; Werren, 1993; but see chapter 5).

By comparing figures 2.4a, c and e to figures 2.4b, d, and f respectively, it is clear that there is a small part of the parameter space where the final equilibrium was dependent on the initial conditions. This is where W_1 was very low but the difference between W_0 and W_2 was not so big that an individual who are able to attain the optimum fitness will have an extreme advantage over the rest of the individuals in the population. This area was largest for selfing diploids, followed by sibmating diploids and then sibmating haplodiploids. Theory predicts that the initial mating strategies of the population could affect the evolution of subsequent mating strategies (Campbell, 1986; Holsinger, 1988; Lande & Schemske, 1985; Uyenoyama, 1986; Waller, 1993). In these studies it was due to coevolution and linkage of alleles determining fitness and alleles determining mating strategy.
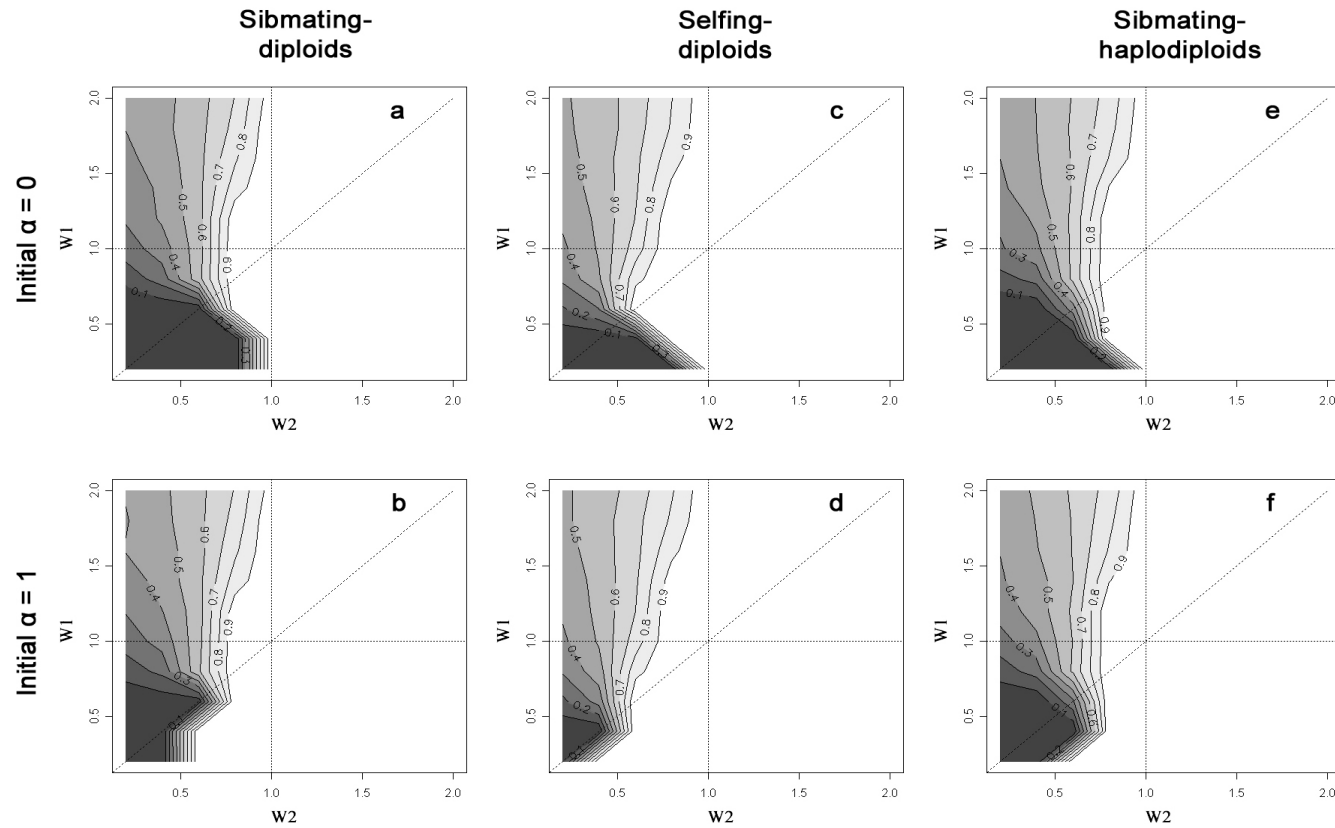
Figure 2.4. Comparison of optimal α (indicated by contours) when different initial mating strategies are specified. The initial inbreeding frequency of the population is 0 in plots a, c & e while it is 1 in plots b, d & f. Similar levels of mixed mating can be seen when comparing the same genetic system: sibmating diploids (a & b), selfing diploids (c & d) and sibmating haplodiploids (e & f). For all three genetic systems the final mating strategy is similar regardless of the starting α, except in the lower left quadrant below the diagonal.

Latta and Ritland (Latta & Ritland, 1993) showed that mixed mating can be stable for selfing diploids when there is a monotonical decline in fitness for serial inbreeding. We followed a similar strategy to Damgaard et al. (1992) to show mixed mating may be stable for selfing diploids. We simulated 21 breeding classes each with reduced fitness relative to the previous class (figure 2.5). We found stable mixed mating for sibmating diploids and sibmating haplodiploids at similar costs of inbreeding to the selfing individuals (figure 2.5). Specifically, selfing individuals switched from exclusive outbreeding to mixed mating when b became smaller than 1.5 (figure 2.5), and this was not surprising that here the fitness of the inbred classes was slightly more than 50% of the outbred class ($W_1 = 0.53$ at $b = 1.25$, see also figure 2.2b).

The higher kin advantage that selfing diploids obtained from inbreeding as compared to sibmating individuals, led to the expectation that they will have higher levels of inbreeding. This apparent predisposition to inbreeding was, however, cancelled by the rapid increase of the inbreeding coefficient of selfing compared to sibmating (see figure 2.2). Comparing the sibmating genetic systems, haplodiploid individuals once more had a lower level of mixed mating than diploid individuals (figure 2.5) due to their lower inclusive benefits of inbreeding. All three genetic systems reached exclusive outbreeding at $b = 2$ and exclusive inbreeding at $b = 0.1$.

Populations in nature show a wide range of inbreeding costs. In a meta-analysis, Crnokrak & Roff (1999) found that plants have an average b of 0.552 (SE ± 0.106). At this low inbreeding cost we would expect more than 85% of all matings to be selfing (figure 2.5), suggesting that the average plant will inbreed on a regular basis. Crnokrack and Roff found average b for homeotherms to be 0.818 (SE ± 0.472) and 0.661 (SE ± 0.121) for poikilotherms, suggesting mixed mating to be common in animals (although a wide range of inbreeding costs is present in homeotherms). The cost of inbreeding in mammals is, however, much higher, with an average b of 1.98 and with some species reaching values as high as 14 (Greeff & Bennett, 2000) and 15.16 (Ralls et al., 1998). We found that inbreeding was prevented at $b = 2$ and expected, therefore, that although sibmating could occur in mammals it would be infrequent. Much lower inbreeding depression was found in insects with $b = 0.29$ for diploids and $b = -0.014$ for haplodiploids which would lead to chronic inbreeding as is often found in haplodiploids (Hamilton, 1967; Werren, 1993).
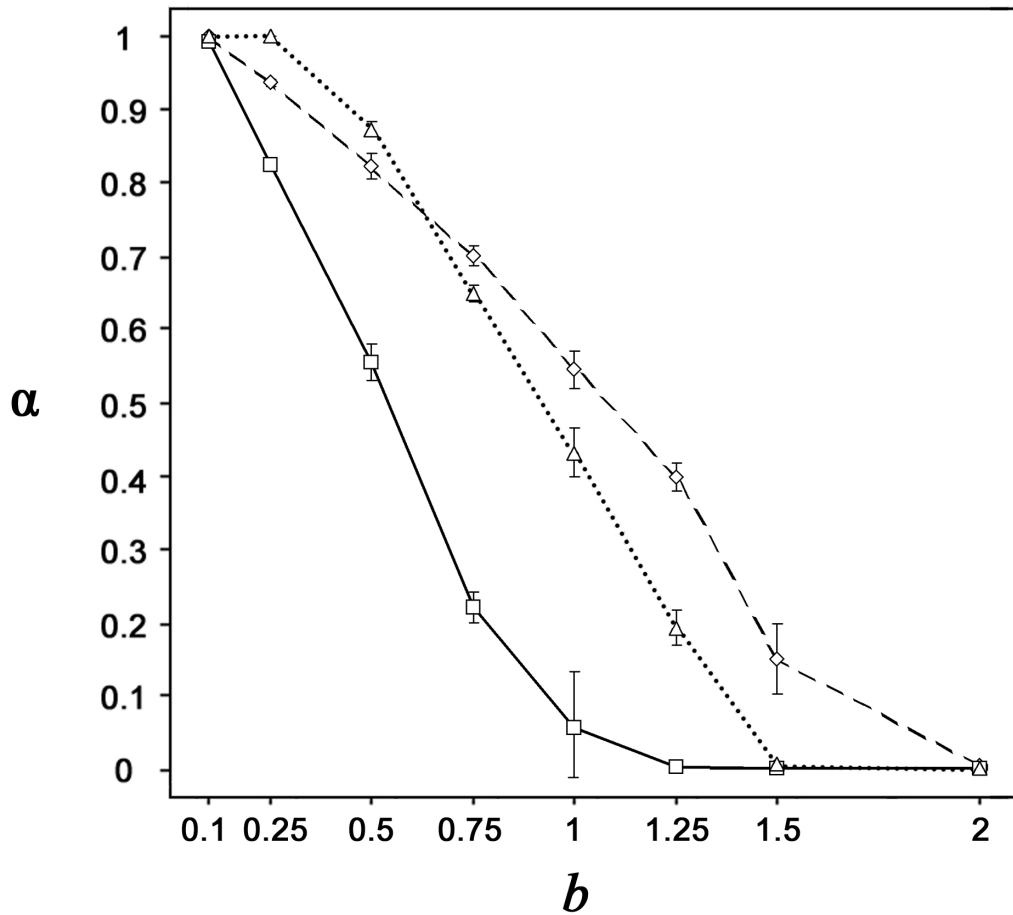
Figure 2.5. Probability of inbreeding (α), for different cost of inbreeding (b). For selfing diploids (triangles, dotted line), sibmating diploids (diamonds, dashed line), and sibmating haplodiploids (squares, solid line). Error bars indicate standard deviation of the four repeats.

In conclusion, we showed that stable mixed mating readily evolves when intermediate inbreeding is optimal for a number of mating systems. We also find, in agreement with previous models, that mixed mating may be stable even with low levels of inbreeding depression but show that a minimum of three fitness classes are required. The benefits of inbreeding increases as the females' relatedness to her mate increases. By comparing observed values of inbreeding depression in nature to our results it seems that some inbreeding should occur in a fair number of species, but that many species also lay close to the switch between outbreeding and a mixed mating system.

3. Sex ratio dependant dispersal when sex ratios vary between patches

Abstract

Theory predicts that sex ratio is affected by dispersal. In turn, individuals will disperse when the cost of dispersal is low or their relatedness is high. However, dispersal has been shown to be unaffected by sex ratio. This is most likely due to the assumption of identical sex ratios between patches in the whole population. In natural populations the sex ratios produced at different sites by different mothers are expected to be binomially distributed. We show, using individual-based simulations, realistic population sizes and variation in sex ratios between patches, that the dispersal of males increases when the sex ratio in their local environment increases. In addition, our results corroborate analytical models that show that relatedness and the cost of dispersal are important factors driving the evolution of dispersal. Lastly, we highlight the decrease in dispersal when offspring, rather than parents, determine when to disperse.

Introduction

Several models consider the coevolution between sex ratios and dispersal (Hamilton, 1967; Perrin & Mazalov, 2000; Taylor, 1994; Taylor & Bulmer, 1980; Wild & Taylor, 2004). Interestingly, these models do not predict any feedback effect of sex ratio on dispersal. This means that while sex ratio evolves in response to kin conflict and dispersal, dispersal evolves in response to kin conflict only and not in response to sex ratio. This is most likely because these models assume a homogenous population where all patches have the same sex ratio. In reality however, sex ratios often vary between patches. Analytical models dealing with dispersal also assume all females will be mated, as well as, very large clutch sizes so that sex ratio may be treated as a continuous trait (Frank, 1986; Gandon, 1999; Gandon & Michalakis, 2001; Leturque & Rousett, 2003; Motro, 1982; Ronce, 2007; Taylor, 1988; Taylor, 1994). These three simplifying assumptions have good heuristic benefits but they may oversimplify the problem. To address these problems, we develop a simulation model that allows males to disperse conditionally, depending on their local sex ratio, with realistic clutch sizes, and with the additional requirement that females must mate to produce offspring.

Analytical models clearly explain how dispersal is promoted by high relatedness of individuals competing for a resource, while being hindered by high costs of dispersal (Clobert et al., 2001; Frank, 1986; Gandon, 1999; Hamilton & May, 1977; Motro, 1983; Motro, 1991; Taylor, 1988). Here, individuals may increase their inclusive fitness by leaving their native patches, thereby providing their siblings with additional mating opportunities when the cost of dispersal is low. As the cost of dispersal increases, the advantage to remain in the native patch and to mate locally increases, due to the lower chance of dispersing successfully and finding a mate. In these models the mothers often determine the rate of dispersal.

However, when dispersal is determined by the offspring, optimal dispersal rates are lower, indicating parent-offspring conflict (Gandon, 1999; Hamilton & May, 1977; Motro, 1983; Taylor, 1988). This makes intuitive sense as parents gain more if enough offspring remain in the native patch to mate each other, while the remainder disperse (thereby reducing local mate competition and increasing the chance of finding additional mating opportunities). On the other hand, offspring would like to

remain in the native patch and not pay the cost of dispersal if mating opportunities are locally available.

Inbreeding depression and a number of other environmental factors (such as variable environmental conditions, predation and parasitism), have also been proposed to affect dispersal (Bengtsson, 1978; Clobert et al., 2001; Perrin & Mazalov, 2000). Reduced dispersal could give rise to structured populations, which in turn will affect the sex ratio of the local patches (Hamilton, 1967). A number of studies have examined the effect of assortative mating on dispersal directly or its influence on sex biased dispersal (Greeff, 1995; Gros et al., 2008; Leturque & Rousett, 2003; Perrin & Mazalov, 2000; Wild et al., 2004; Wild & Taylor, 2004).

To model complex natural processes, such as dispersal, phenotypic variation of a trait is often excluded (Mitchell & Valone, 1990). However, proper understanding of the mechanisms influencing dispersal requires refinement of these models by including realistic life history parameters with normal variation (Gros et al., 2008; Judson, 1994; see also: Lenormand et al., 2009; Ronce et al., 2001). In this chapter we use individual-based simulation modelling to explore how sex ratio variation may affect the evolution of dispersal. We also corroborate current models of dispersal where the cost of dispersal varies. We simulated offspring-determined dispersal in a haploid population where: males determine their dispersal and disperse before mating; all the females disperse from their natal patches after mating; and there are a finite number of males and females in the population and in each patch.

Model description

State variables and scales

The model comprised three structural levels: individual, patch and population. Each individual is characterised by the following state variables: identity number, sex, mating status (mated or unmated), patch number (identity of the patch where the individual reside), dispersal genes (an array of genes each responding to a different range of sex ratios) and male sperm (genes obtained by the females after mating). It is important to note that each dispersal gene has a value from 0 to 1 indicating the

probability of dispersal. However, each dispersal gene responds to a different range of sex ratios, and therefore only one gene at a time determines if an individual disperses from a patch. For example: if the first gene of a male individual has a value of 0.1 and responds to a sex ratio from 0 to 0.33 the male will have a 10% chance of dispersing. If the sex ratio is not within this range another gene, with its own probability of dispersal, will affect the individual's dispersal. This enables us to obtain optimal dispersal rates for a number of different sex ratios. In addition, the dispersal genes determine only male dispersal (all the females disperse after mating, irrespective of their dispersal genes). We assume that males have unlimited matings, while females can mate only once, and the mating status variable is therefore only utilised by female individuals.

Each patch has the following variables: patch identity number, number of foundresses, foundress identity numbers, patch size (number of individuals within patch), number of males, offspring identity numbers, expected sex ratio, observed sex ratio and patch viability. Sex ratio is defined as: (males)/(males + females). Only offspring (i.e. the current generation) are considered as members of the patch, and used in the determination of the observed sex ratio or patch viability. The expected sex ratio is set before the start of each simulation. The observed sex ratio is determined after the offspring are created but before male dispersal takes place (figure 3.1). Two modes of sex ratio production were simulated: In experiment 3.1, the sex of each offspring was assigned with a certain probability (i.e. the expected sex ratio). The sex ratio's produced at each patch were binomially distributed around the expected sex ratio. Consequently, the observed sex ratio for each patch was calculated as the proportion of males in that patch. In experiment 3.2, all the foundress females produced the expected sex ratio and the observed and expected sex ratio were therefore the same. A patch is considered as viable when at least one male and one female are present and is evaluated after offspring creation, male migration and each mating event.

The population is composed of an array of individuals and an array of patches. The population is furthermore characterised by the following variables: population size (i.e. the size of the individual array), number of patches (i.e. the size of the patch array), clutch size, number of generations, number of genes, number of alleles,
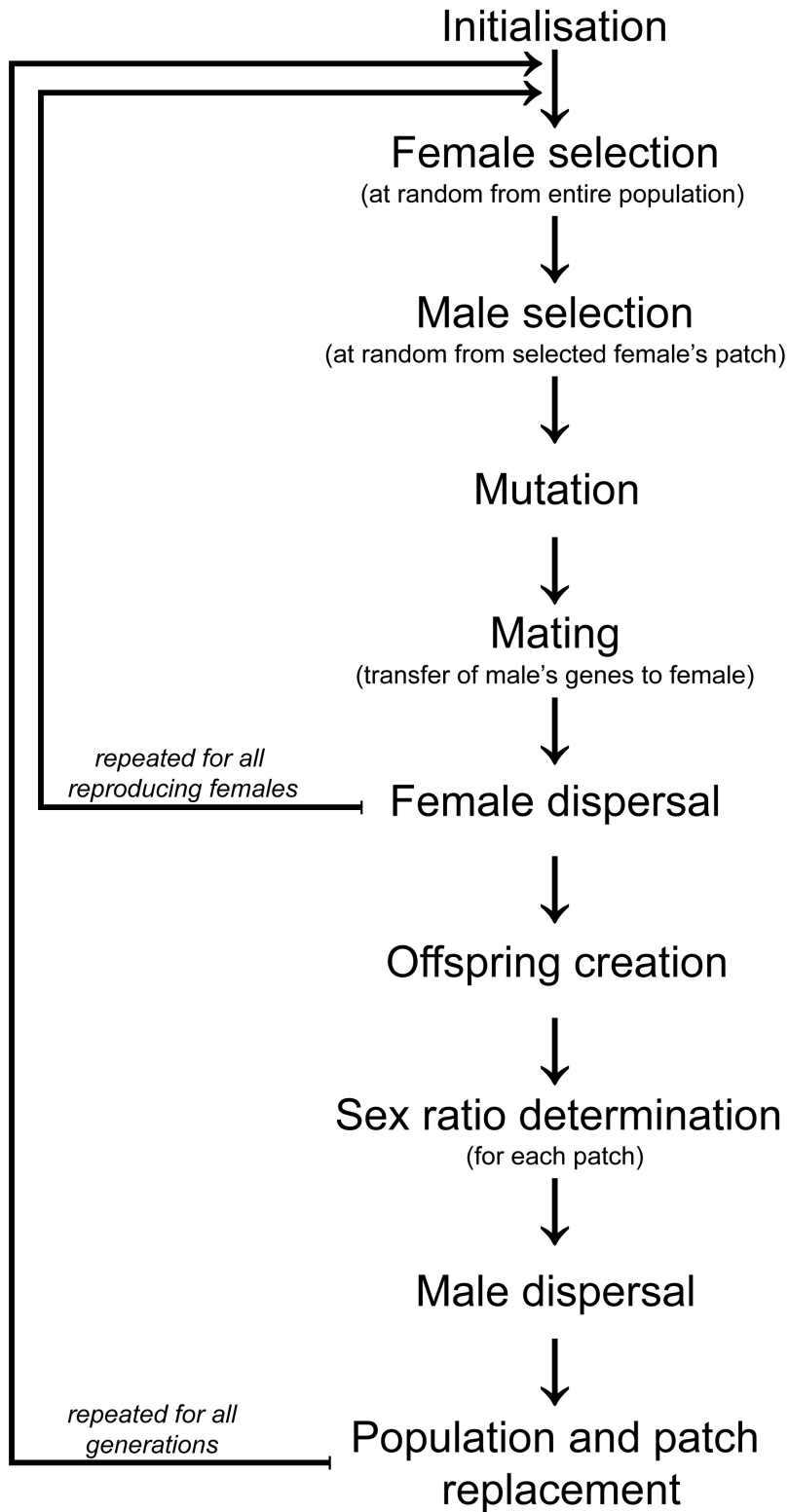
Initialisation

Female selection
(at random from entire population)

Male selection
(at random from selected female's patch)

Mutation

Mating
(transfer of male's genes to female)

repeated for all reproducing females

Female dispersal

Offspring creation

Sex ratio determination
(for each patch)

Male dispersal

repeated for all generations

Population and patch replacement

Figure 3.1. Life history of individuals (process sequence of model).

mutation rate and cost of dispersal. From these variables the fraction of reproducing females are calculated as: population size/clutch size. Patch viability is determined at every generation and the simulation continues only if viable patches remain.

Simulation process and scheduling (population life history)

The model process sequence is depicted in figure 3.1. The simulation proceeds in discreet generational time steps (i.e. no overlapping generations). In each generation step the events are processed as follows: Females are selected at random from the population to mate. Each female mates once with a male from her local patch (chosen at random from that patch). After mating, the mated females disperse to colonize new patches where offspring are created. Females produce exactly the same clutch size. The offspring sex ratio of each patch is determined. Male offspring have a single chance to disperse from their natal patch in response to their natal patch's sex ratio. Males that disperse are assigned a new patch at random. Finally, the offspring in the new patches replaces the old population and their patches and the whole process starts anew. The process is repeated a number of generations (see table 3.1). Mutation is a sub-step that proceeds mating (see below).

Design Concepts

Optimal male dispersal strategies emerge from the population dynamics and their stochastic nature since there is no direct fitness assessment of an individual (i.e. there is no fitness value assigned to the individuals, rather, their fitness becomes apparent through the amount of offspring they have). At each gene, an allele defines the probability that a male disperse. There are 101 possible alleles, each of which is a value from 0 to 1 at regular intervals. As mentioned, each dispersal gene only responds to a specific sex ratio range (see response intervals below). All genes are unlinked and new alleles arise in the population through mutation only. Therefore, each simulation needs to run for a number of generations to ensure that the mutational space is adequately searched and equilibrium reached. Ultimately, males should have a dispersal strategy in response to their local sex ratio that will maximise their

inclusive fitness. It should be noted that kin selection do not have to be introduced explicitly as kin-advantage will emerge by default in any individual-based simulation (Gros et al., 2008; Poethke et al., 2007). The probability of dispersal, mutation and the offspring sex ratio are all drawn from binomial distributions. Data gathered for analysis included the dispersal probability for each gene (averaged for all the individuals in the last generation). The complete genotype for each of the individuals in the last generation is also recorded to ensure each gene reached a single optimum (i.e. at each gene only one allele was found throughout the population). In addition, a graph displaying the average value for each gene at every generation is also saved from which stability of the strategies can be evaluated. All gene values stabilised long before data was recorded from the simulation.

Initialization Input

The simulation input values are given in table 3.1. The following variables were kept constant for all the simulations: The initial probability of dispersal (0), number of possible alleles per simulation (101), mutation rate (0.0005). During all the simulations the number of reproducing females (1000) was kept constant, irrespective of the population size or number of foundress per patch. Each set of variables in table 3.1 was simulated 10 times.

Sub models

Mutation rate: The mutational model was a combination of stepwise and jumping mutations. Once an allele was selected to undergo a mutation there was an 80% chance of the allele to step to either of the adjacent alleles and a 20% chance for it to step to any of the other 101 possible alleles.

Dispersal cost: Each male that dispersed had a probability of being excluded from the mating population given by the dispersal cost. All the settling costs and transit costs are subsumed in the single dispersal cost parameter.

Table 3.1. Population and patch parameters for each simulation. *Average number of males per patch when binomial variation was introduced. Each parameter set was simulated 10 times.

Experiments	Population parameters						Patch parameters			
	Population size	Patches	Clutch size	Generations	Genes	Cost	Foundresses	Males/patch	Offspring/patch	Sex ratio
Experiment 3.1	100000	1000	100	100000	5	0.2	1	25*	100	0.250
	100000	500	100	100000	5	0.2	2	50*	200	0.250
	100000	333	100	100000	5	0.2	3	75*	300	0.250
Experiment 3.2	30000	1000	30	30000	1	0.2	1	5	30	0.167
	15000	1000	15	30000	1	0.2	1	5	15	0.333
	10000	1000	10	30000	1	0.2	1	5	10	0.500
	8000	1000	8	30000	1	0.2	1	5	8	0.625
	7000	1000	7	30000	1	0.2	1	5	7	0.714
	60000	1000	60	30000	1	0.2	1	10	60	0.167
	30000	1000	30	30000	1	0.2	1	10	30	0.333
	20000	1000	20	30000	1	0.2	1	10	20	0.500
	15000	1000	15	30000	1	0.2	1	10	15	0.667
	12000	1000	12	30000	1	0.2	1	10	12	0.833
	90000	1000	90	30000	1	0.2	1	15	90	0.167
	45000	1000	45	30000	1	0.2	1	15	45	0.333
	30000	1000	30	30000	1	0.2	1	15	30	0.500
	23000	1000	23	30000	1	0.2	1	15	23	0.652
	18000	1000	18	30000	1	0.2	1	15	18	0.833

Table 3.1. *continued*

	30000	1000	30	30000	1	0	1	15	30	0.5
	30000	1000	30	30000	1	0.1	1	15	30	0.5
	30000	1000	30	30000	1	0.2	1	15	30	0.5
	30000	1000	30	30000	1	0.3	1	15	30	0.5
	30000	1000	30	30000	1	0.4	1	15	30	0.5
	30000	1000	30	30000	1	0.5	1	15	30	0.5
	30000	1000	30	30000	1	0.6	1	15	30	0.5
	30000	1000	30	30000	1	0.7	1	15	30	0.5
	30000	1000	30	30000	1	0.8	1	15	30	0.5
Experiment 3.3	30000	1000	30	30000	1	0.9	1	15	30	0.5
	30000	1000	30	30000	1	1	1	15	30	0.5
	30000	500	30	30000	1	0	2	15	30	0.5
	30000	500	30	30000	1	0.1	2	15	30	0.5
	30000	500	30	30000	1	0.2	2	15	30	0.5
	30000	500	30	30000	1	0.3	2	15	30	0.5
	30000	500	30	30000	1	0.4	2	15	30	0.5
	30000	500	30	30000	1	0.5	2	15	30	0.5
	30000	500	30	30000	1	0.8	2	15	30	0.5

Mating: During mating only one of the parents, chosen at random, transferred an allele to the new offspring. This donating parent was selected separately for each of the genes enabling normal Mendelian inheritance for unlinked genes.

Response intervals: For simulations where individuals had more than one gene (experiment 3.1) a response interval for each gene was calculated. These intervals were chosen so that the area under the binomial expectation roughly divided into equal parts (as many parts as there were genes). As a result each gene will be expressed with equal frequency and will be under equal selection strength. A specific gene will thus be "switched on" if the observed sex ratio is within the interval. The gene specifies the probability of dispersal for the individual. (Note that the intervals were not assigned from the median of the distribution but started from 0. This skewed the intervals but ensured that the area of each interval were exactly the same size).

Experiments

Experiment 3.1: Binomial sex ratios

Variation in the sex ratio between the patches was caused by enabling the foundress females to produce binomial sex ratios with an average of 0.25 males. Male dispersal could therefore evolve to respond to a range of sex ratios. Five genes, each responding to a different sex ratio range, were optimised per simulation (see above for response interval). To change the number of foundress females per patch the number of reproducing females was kept constant while the number of available patches was reduced. To test the effect of kin selection, simulations were done for 1, 2 and 3 foundress females (see table 3.1 for full parameter set).

Experiment 3.2: Exact sex ratios

To show that dispersal is not affected by the sex ratio when all the females in the population produce the same sex ratio, exact sex ratios were simulated. In each simulation all the foundress females produced the same sex ratio (the expected sex

ratio) and no variation between patches was produced. We therefore had only one dispersal gene that was optimised per simulation. Sex ratios were however different between simulations (table 3.1). Three sets of simulations were performed where the sex ratio was different but the males per patch were kept constant (5, 10 and 15 males respectively for sex ratio range, table 3.1). The sex ratio within each set was therefore changed by keeping the number of sons constant and increasing the clutch size with extra daughters (see table 3.1 for full parameter set).

Experiment 3.3: Different cost of dispersal

To test the effect of dispersal cost when dispersal is under offspring control, we simulated different costs of dispersal (table 3.1). In each simulation all the foundress females produced the exact sex ratio (0.5), and therefore only a single dispersal gene was optimised per simulation. To test the effect of kin selection at different dispersal costs, simulations were performed for 1 and 2 foundress females (see table 3.1 for full parameter set).

Statistics

All statistics on the generated data were performed in R version 2.4.1. For each simulation we recorded the average allele value (i.e. the probability of dispersal) from all the individuals at the final generation for each of the sex ratio ranges. Most genes had only one, or a few similar alleles (e.g. 0.53 and 0.54), fixated throughout the population. The average would therefore reflect the common dispersal strategy.

In experiment 3.1, dispersal was modelled with the following linear model: dispersal $\sim$ foundress:sexratio^2 + foundress:sexratio + sexratio^2 + foundress + sexratio; where foundress was a factor depicting the number of foundresses in each patch. All the higher order terms were tested for significance (see table 3.2). For each of the 5 genes we used the weighted midpoint of the gene range, with weighting equal to the frequency of each class, (from the binomial distribution) as the explanatory variable.

The following linear model was used to model the data from experiment 3.2: dispersal $\sim$ sexratio^2 + sexratio + males^2 + males; where males indicated the number of

males in the patch (table 3.1). The model was reduced to the minimum adequate model by removing non-significant terms, starting with the higher order terms (table 3.3).

The following models were fitted to the data from experiment 3.3: dispersal ~ $cost^2$ + cost; for the 1 foundress simulations and: dispersal ~ $cost^3$ + $cost^2$ + cost; for the 2 foundress simulations (see table 3.4). All the data was untransformed.

Results and Discussion

When variation between patches was present, our results indicated that males will disperse more from patches where there are more males (figure 3.2, table 3.2). The data was best explained by the following model, where each row represents the estimates for 1, 2 and 3 foundress simulations respectively:

$$\text{dispersal} = \begin{bmatrix} -10.014 \\ 18.597 \\ 51.120 \end{bmatrix} \text{sexratio}^2 + \begin{bmatrix} 7.049 \\ -7.852 \\ -25.288 \end{bmatrix} \text{sexratio} + \begin{bmatrix} -0.406 \\ 0.363 \\ 2.451 \end{bmatrix}$$

It was also apparent that dispersal increased when the relatedness of the males in a patch was higher, as males from single foundress patches dispersed more readily than males from two foundress patches, followed by males from three foundress patches (figure 3.2, see also figure 3.4), as has been found in previous studies (Clobert et al., 2001; Frank, 1986; Hamilton & May, 1977; Motro, 1982; Taylor, 1988). In other words, when there are fewer founding mothers per patch the competing males have a higher average relatedness and will disperse more to reduce local mate competition.

When the sex ratio variation between the patches was removed (experiment 3.2), we found that sex ratio had no effect on the evolution of dispersal (figure 3.3, table 3.3, dispersal = -0.002males^2 + 0.045males + 0.373). In these circumstances emigration from the natal patch on account of sex ratio does not pay because, 1) the

Table 3.2. ANOVA table for model from experiment 3.1.

	df	F	P
Adjusted R^2 = 0.946			
Retained terms			
# foundress:sexratio^2	2	8.056	<0.001
# foundress:sexratio	2	7.707	<0.001

Table 3.3. ANOVA table for model from experiment 3.2.

	df	F	P
Adjusted R^2 = 0.579			
Retained terms			
# males^2	1	33.513	<0.001
# males	1	57.930	<0.001
Deleted terms			
sexratio^2	1	1.510	0.2211
sexratio	1	1.550	0.2151

Table 3.4. ANOVA table for models from experiment 3.3.

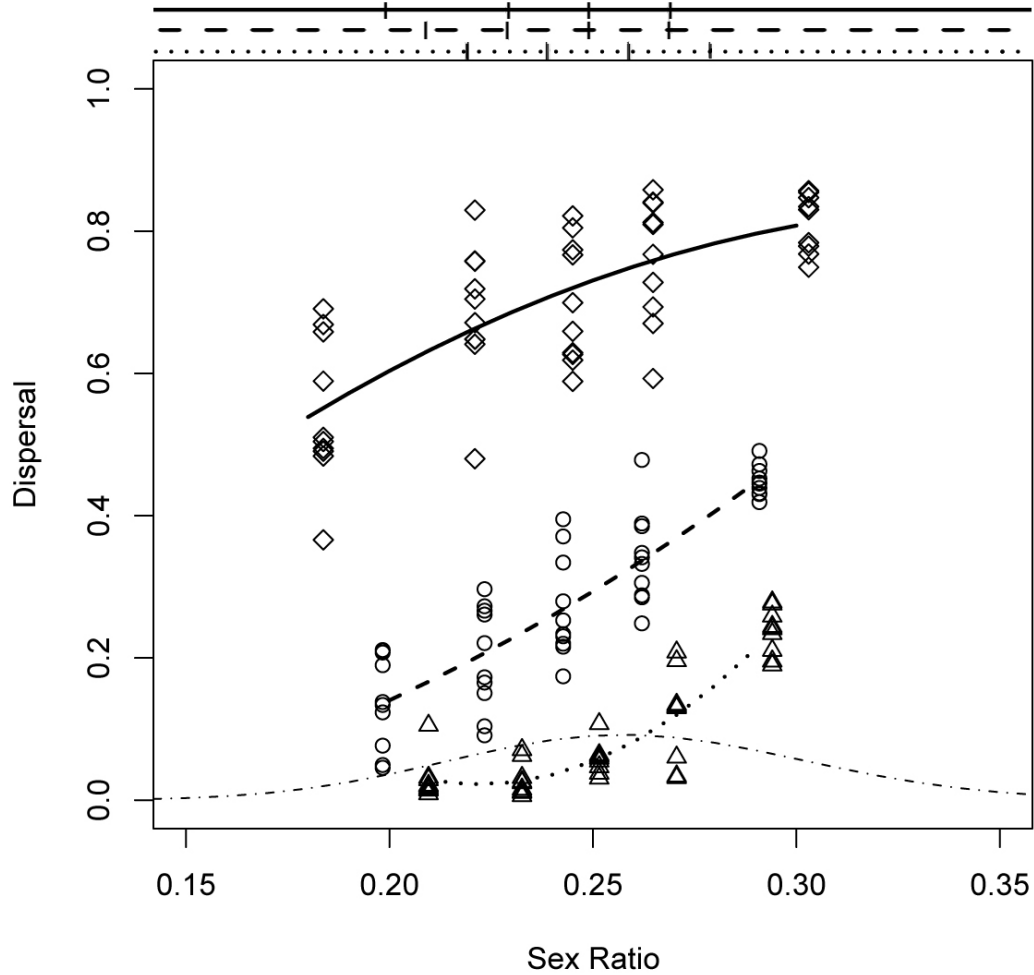|	df	F	P
1 Foundress Model: *Adjusted R^2 = 0.861*			
Retained terms			
cost^2	1	120.743	<0.001
cost	1	17.994	<0.001
2 Foundress Model: *Adjusted R^2 = 0.975*			
Retained terms			
cost^3	1	50.278	<0.001
cost^2	1	127.822	<0.001
cost	1	516.601	<0.001

Figure 3.2. Modelled relationships between dispersal and sex ratio with binomial variance (from experiment 3.1, see table 3.2). Diamonds and solid lines: 1-foundress simulations; circles and dashed lines: 2-foundress simulations; triangles and dotted lines: 3-foundress simulations. Dash-dot line indicates the binomial distributed sex ratio with an average of 0.25 males. The response intervals for each of the simulations are indicated above the graph.
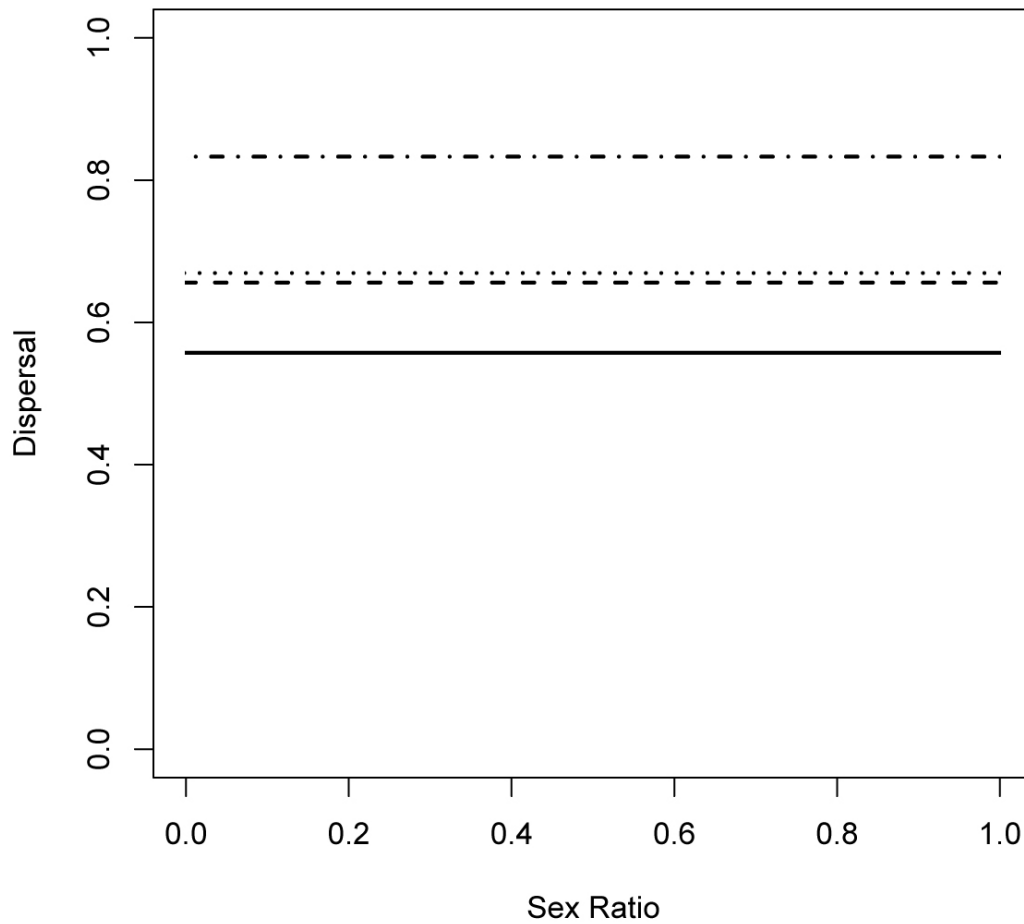
Figure 3.3. Modelled relationships between dispersal and sex ratio with no variation between patches (from experiment 3.2, see table 3.3). Solid line: 5 males per patch; dashed line: 10 males per patch; dotted line: 15 males per patch. Dash-dot line indicates the predicted level of dispersal (0.833) at for one-foundress patches at dispersal cost of 0.2 (see text).
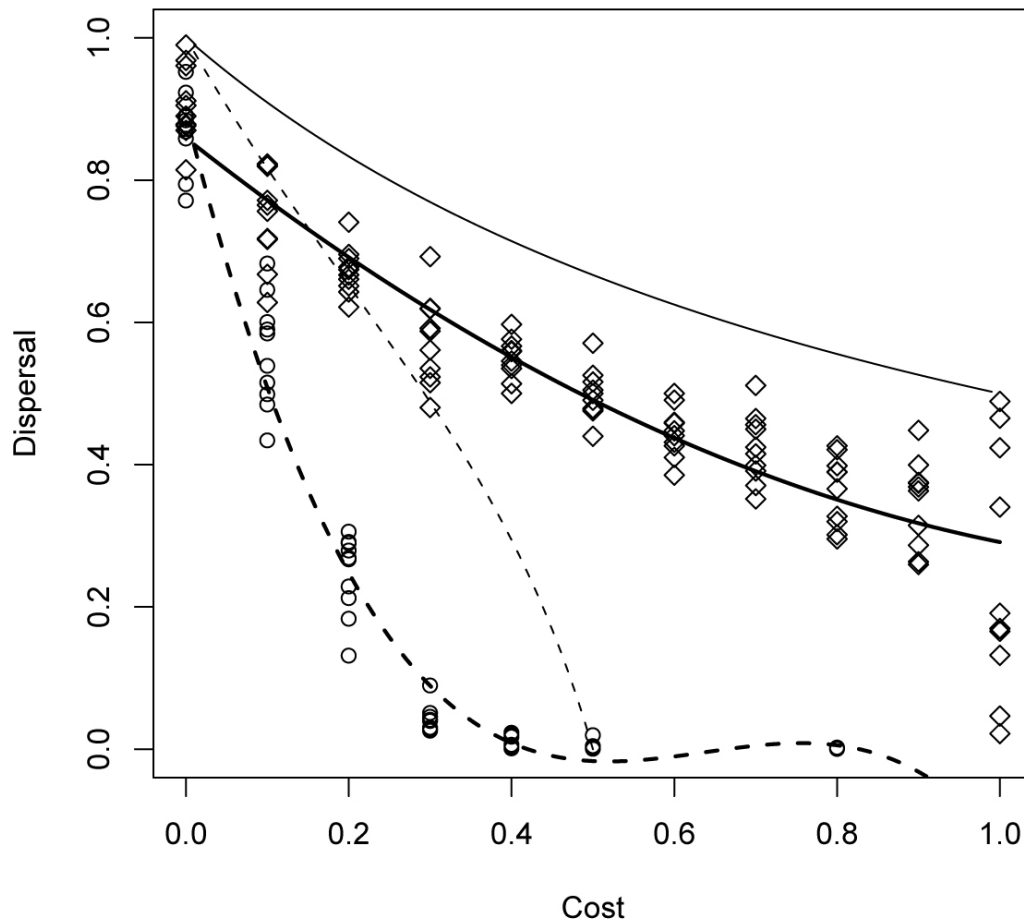
Figure 3.4. Modelled relationship between dispersal rate and cost of dispersal (from experiment 3.3, see table 3.4). Diamonds and solid lines: 1 foundress; circles and dashed lines: 2 foundresses. Bold lines indicate our results, thin lines from predicted dispersal if under maternal control (see text).

amount of competitors are exactly the same, 2) the amount of mating opportunities are exactly the same and 3) there is a cost to dispersal. However, here we found that the more males present per patch (irrespective of the sex ratio), the higher the dispersal rates (figure 3.3, table 3.3).

Most models dealing with dispersal assume infinite clutch sizes and exact population-wide sex ratios (Frank, 1986; Gandon & Michalakis, 2001; Leturque & Rousett, 2003; Motro, 1982; Ronce, 2007; Taylor, 1988; Taylor, 1994) to make these models tractable (Greeff, 1998; Mitchell & Valone, 1990; Ronce, 2007). It is however known that stochastic demographic conditions may affect the evolution of many traits under selection (Lenormand et al., 2009), including dispersal strategies (Gros et al., 2008; Ronce, 2007) and even that facultative dispersal may reduce the fitness impact due to inaccurate sex ratios (Greeff & Compton, 2002). When we included variation in our simulation, we found that males responded differently in populations where there was variation between patches compared to populations with homogenous patch sex ratios (compare figures 3.2 and 3.3). In figure 3.2 we observe that sex ratio had a significant, positive relationship, with increased dispersal, but was significantly different for each of the foundress treatments (table 3.2). In this experiment we had foundress mothers (irrespective of the amount of foundress mothers per patch) produce an average sex ratio of 0.25 males but with a binomial distribution as indicated by the dash-dot line in figure 3.2. Males had 5 dispersal genes, each responding to, and optimised for, one sex ratio range (indicated above the graph in figure 3.2).

Our results are supported by field observations. Moore et al. (2006) showed that pollinating males had higher dispersal rates when fewer females were available in a patch. Similarly, Lawrence (1987) observed a significant increase in dispersal flights of male milkweed beetles as the sex ratio became less female biased.

High levels of dispersal are observed in all our results in response to high levels of local mate competition (Perrin & Mazalov, 2000). The model from Wild & Taylor (2004) predicts the amount of dispersal for all our simulations in experiment 3.2 to be 0.833 (where male dispersal = $(C-1/N)/(C^2-1/N)$, with the cost ($C = 0.2$), the number of foundress ($N = 1$), and mothers determining the dispersal rate of their sons). There are two possible explanations for the lower dispersal rates relative to the predicted rate. First, as the number of males per patch decreased the chance that all the males disperse from that patch increased. When this happens the females from that patch

would either be unmated and die or be mated by a complete stranger thereby negating the possible advantage obtained from sibmating. For example, if the average dispersal rate is 0.833, the chance that all males will disperse away from a patch with 5 males is 47%. In a patch with 10 males it is 16% and in a patch with 15 males it is only 6.5%. It is clear that this risk is much larger in patches with realistic numbers of males, and lower dispersal rates in these patches are expected.

A second reason for lower dispersal rates than predicted is that mothers are often in control of the offspring dispersal rates in these models (Clobert et al., 2001; Frank, 1986; Gandon, 1999; Hamilton & May, 1977; Motro, 1983; Motro, 1991; Taylor, 1988). This is sometimes justified when the mothers control the morphology and thus the dispersal tactic of her sons (Pienaar & Greeff, 2003). In our simulations the males 'decided' to disperse or not. From experiment 3.3 it is clear that females would want their sons to be more altruistic towards each other and disperse more than they are willing to (figures 3.4, table 3.4, one foundress model: dispersal $= 0.337\text{cost}^2$ $- 0.905\text{cost} + 0.859$; two foundress model: dispersal $= -3.869\text{cost}^3 + 7.397\text{cost}^2 -$ $4.551\text{cost} + 0.893$). This parent-offspring conflict has been noted before by several authors (Gandon, 1999; Hamilton & May, 1977; Motro, 1983; Taylor, 1988).

Our results from experiment 3.1 and 3.3 also confirm that higher relatedness increased the willingness to disperse from a natal patch (figures 3.2 and 3.4). From experiment 3.1 we observed that when the males were less related (i.e. more foundress mothers) they had a larger reaction to changes in the sex ratio (note also the steeper slopes of the models). In spite of this, less related males had a lower dispersal probability in general (figure 3.2). From experiment 3.3 we found that males from 2-foundress patches had much lower dispersal rates than those of 1-foundress patches. As the cost of dispersal increased less related males rejected dispersal as a strategy completely (figure 3.4). This was in agreement with theory, mentioned previously, as related males do not only disperse to increase their own fitness directly but also their inclusive fitness from the extra matings their brothers may have by not competing with them (Frank, 1986; Hamilton & May, 1977; Motro, 1982; Taylor, 1988). Note however, that if the cost of dispersal is 1, that males from 1-foundress patches vary highly in their dispersal rates (figure 3.4, cost = 1). At this high cost, no males from neighbouring patches can invade and only brothers compete for mating opportunities. There is thus no selection on the trait and drift is the only evolutionary force in

operation. The only motivation for reduced dispersal is the risk of all males dispersing from the patch and that none of the sisters are mated.

In conclusion, most models show that dispersal is mainly affected by the cost of dispersal and the relatedness of the competing individuals (Clobert et al., 2001; Frank, 1986; Hamilton & May, 1977; Motro, 1982; Taylor, 1988). Additionally we know that inbreeding depression increases the rate of dispersal (Bengtsson, 1978; Clobert et al., 2001; May, 1979; Motro, 1991; Perrin & Goudet, 2001; Perrin & Mazalov, 2000). However, in our simulations the cost of dispersal is kept constant, and there are no negative fitness effects due to inbreeding. Lastly, sex ratios in natural populations are often binomially distributed (Hardy, 1992), and we highlight the feedback of this variation on the evolution of dispersal. We show that if males are able to estimate the sex ratio in their own patch that they will disperse if the sex ratio is high, but only if there is sex ratio variation between the patches.