

Identification of the factors that lead to dispersal and inbreeding

by

Ronald Michael Nelson

Submitted in partial fulfilment of the requirements for the degree PhD Genetics

In the Faculty of Natural and Agricultural Science

University of Pretoria

Pretoria

August 2009

Summary

Individual-based simulation modelling is an excellent method for testing hypotheses, while including realistic and stochastic population parameters. This thesis considers the evolution of dispersal or inbreeding through individual-based simulation modelling.

The occurrence of exclusive inbreeding and exclusive outbreeding is found in a number of organisms and are referred to as mixed mating. Mixed mating is suggested to be in response to low levels of inbreeding depression as well as simultaneous inbreeding- and outbreeding depression while intermediately related mating partners are not available. The results of this thesis show that stable mixed mating strategies evolve in the presence of both inbreeding and outbreeding depression, as well as, under conditions where low levels of inbreeding depression are present. Also, inclusive fitness allows higher levels of inbreeding in genetic systems where the mating partners are more related to each other.

Dispersal evidently evolves in response to inbreeding depression. A number of other factors, such as local mate competition and the cost of dispersal also influence the rate of dispersal. In addition to these factors, it is shown in this thesis that male dispersal evolves when there is variation in patch sex ratios. Simulation data also supports parent offspring conflict models, as males have reduced dispersal rates when they, rather than their parents, determine the dispersal rate.

Population structure is affected by dispersal rates. Using individual-based simulation modelling and various sampling strategies, reveals that few molecular markers, for a few individuals, are sufficient to accurately detect population subdivision, especially when the sub-populations are large. It is, however, indicated that planning prior to sampling are important for proper assessment of population structure.

Lastly, molecular data from the pollinating fig wasp *Platyscapa awekei* reveals that this species suffers from low levels of inbreeding depression. However, when this data are simulated, stable mixed mating did not evolve although it is observed in *P. awekei*. Sex ratio variation, high local mate competitions and male only broods are therefore suggested to drive male dispersal. It is consequently advantageous to use various techniques to unravel the evolution of a trait and gain insight into the system.

To: My Grandfather

Acknowledgements

I thank my supervisor, Jaco Greeff, for his support during my research, allowing me the to freedom to choose my topics of research, and subsequently to wander into the realm of individual-based simulation modelling. Many hours were spent discussing various aspects of evolution, biology and statistics and I learned a great deal from him and it has been a privilege to work together.

I thank my family for all the love and encouragement. To my parents, many thanks for the support, both financial and otherwise, provided over the years. To my sibling, Cindy, thanks for all the visits in the lab and often participating in various discussions on evolution and related matters.

I am very grateful to everyone who read, commented and pointed to errors in the manuscripts: Chapter 2 – Jaco Greeff, Carina Schlebusch, Marié Warren and two anonymous reviewers; Chapter 3 – Jaco Greeff, Carina Schlebusch and Marié Warren; Chapter 4 – Jaco Greeff, Carina Schlebusch, Marié Warren and Pam de Waal and Chapter 5 – Jaco Greeff, Carina Schlebusch and Marié Warren. A special thanks to Marié Warren, who patiently corrected my English throughout the thesis.

To the Department of Genetics at the University of Pretoria and my fellow lab members, past and present Annie, Aret, Christoff, David, Duncan, Heike, Joscha, Marié, Mel, Motoaki, Okkie, Pam, Soné, Vinet, and Zee, many thanks for making the work environment a fun place to be and your companionship during this study.

I am grateful to the National Botanical Gardens of South Africa that allowed me to collect figs and fig wasps on their premises. I also extend my appreciation to the University of Pretoria and the National Research Foundation for their financial support during this study.

Lastly, I owe a massive debt of gratitude to my wife, Carina. I am thankful for her constant support of my research while doing her own PhD. Her love, encouragement and understanding made life a pleasure during this study.

Preface

Each chapter in this thesis, except for the introduction and conclusion, is written as a journal article and is either submitted or in preparation for submission and as such includes its own introduction, and discussion section. In chapters 2 and 3 the results and discussion sections are combined respectively as a discussion naturally follows a description of the generated data.

Jaco Greeff, my supervisor, is included as co-author for the submitted manuscripts for chapters 2, 3 and 4. The reasons for this are twofold. First, the models were developed with suggestions from him, although the underlying ideas as well as the simulation models themselves are my own. Secondly, all of the financing and facilities required to complete these studies were provided through funding secured by him.

In this thesis I extend a number of existing ideas with the use of individual-based simulation modelling within the Delphi environment, compiled for the Windows operating system. As such, these programs and source code are available in the electronic appendix submitted with this thesis. Although these programs are merely the tools used to test ideas, the following can be mentioned regarding the code itself: the program discussed in chapter 4 was created first, followed by the program for chapter 2 and lastly the program for chapter 3. Starting with no formal education in programming and no experience with the Delphi environment the program code may be somewhat informal. Therefore, the code for the last program is more efficient than the first. Additionally, in the general introduction, I deal with individual-based simulation modelling as a tool to study evolutionary processes to avoid repetition in the chapters where it was employed.

Lastly, to keep all the theory and models grounded in reality, much of the research in this thesis refer to pollinating fig wasp biology. While the simulation models are not exclusive for the pollinating fig wasp system they are applicable to them. Furthermore, chapter 5 deals with empirical data from this system and for these reasons I introduce the pollinating fig wasp system in the in general introduction.

Table of Contents

Summary	II
Acknowledgements	IV
Preface	V
Table of Contents	VI
List of Tables and Figures	X
1. INTRODUCTION	1
Individual-based simulation modelling	2
A simple simulation model	2
Stochasticity in simulation modelling	5
A short note on inclusive fitness in simulation modelling	8
Advantages of simulation modelling	9
Disadvantages of simulation modelling	11
Assessment of simulation models	11
Pollinating fig wasps	13
Aims of this thesis	16
1. When should mixed mating evolve?	16
2. Does the sex ratio affect dispersal?	17
3. How many samples are necessary to accurately detect population structure?	18
4. Are mixed mating and male dispersal in <i>Platyscapa awekei</i> , due to intermediate inbred individuals having the highest fitness?	19
2. EVOLUTIONARY STABLE MIXED MATING IN A VARIETY OF GENETIC SYSTEMS	21
Abstract	21
Introduction	22

Model description	24
State variables	25
Simulation process and scheduling	25
Design concepts	27
Initialisation and input	27
Sub-models	29
Experiments	29
Experiment 2.1: Optimal mating strategies with three inbreeding classes	30
Experiment 2.2: The importance of initial conditions on the final equilibria	30
Experiment 2.3: Optimal mating strategies with 21 inbreeding classes	31
Program	31
Results and Discussion	33
3. SEX RATIO DEPENDANT DISPERSAL WHEN SEX RATIOS VARY BETWEEN PATCHES	42
Abstract	42
Introduction	43
Model description	44
State variables and scales	44
Simulation process and scheduling (population life history)	47
Design Concepts	47
Initialization Input	48
Sub models	48
Experiments	51
Experiment 3.1: Binomial sex ratios	51
Experiment 3.2: Exact sex ratios	51
Experiment 3.3: Different cost of dispersal	52
Statistics	52

Results and Discussion	53
4. ADEQUATE SAMPLE SIZES FOR ACCURATE DETECTION OF POPULATION SUBDIVISION: A SIMULATION BASED EXPLORATION OF SUMMARY STATISTICS	61
Abstract	61
Introduction	62
Materials and methods	64
Model description	64
General parameters	65
Experiments	65
Experiment 4.1: Simulation of different F_{st} values	65
Experiment 4.2: Simulation of different population sizes	66
Experiment 4.3: Simulation of a single population for estimates of optimal sampling	66
Results	68
Experiment 4.1: Effect of the F_{st} values on the standard deviation of the F_{st}	68
Experiment 4.2: Effect of population size on the standard deviation of the F_{st}	71
Experiment 4.3: Effect of sample size vs. number of loci on the accuracy of the F_{st}	71
Discussion	74
Effect of population size on accuracy	74
Effect of the F_{st} on the accuracy	75
Effect of sample size and loci number on accuracy	76
Guidelines on sampling	77
5. INBREEDING DEPRESSION DOES NOT PROMOTE MIXED MATING AND DISPERSAL IN A MALE POLLINATING FIG WASP, <i>PLATYSCAPA AWEKEI</i>	79
Abstract	79

Introduction	80
Materials and Methods	82
Sample collection	82
Genotype reconstruction	84
Genotyping	86
Statistics	89
Results	89
Discussion	91
6. CONCLUSIONS	100
Mixed mating	100
Dispersal	101
Sampling	102
Platyscapa awekei	103
7. REFERENCES	105
APPENDIX A	117
ELECTRONIC APPENIDIX	(inside back cover)

List of Tables and Figures

Chapter 1	Pg.
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.	4
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.	8
Figure 1.3. The life cycle of pollinating fig wasps and developmental stages of a fig. * <i>P. awekei</i> are one of only a few species, of pollinating fig wasps, where the males are documented to disperse; see text for details.	14
Chapter 2	
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.	26
Table 2.1. State variables and default input values.	28
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 ($W = e^{-bf}$).	32

- Figure 2.3. Contour plots of the probability of inbreeding (α) for a) Sibmating 34
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.
- Figure 2.4. Comparison of optimal α (indicated by contours) when different initial 38
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.
- Figure 2.5. Probability of inbreeding (α), for different cost of inbreeding (b). For 40
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.

Chapter 3

- Figure 3.1. Life history of individuals (process sequence of model). 46
- Table 3.1. Population and patch parameters for each simulation. *Average 49
number of males per patch when binomial variation was introduced.
Each parameter set was simulated 10 times.
- Table 3.2. ANOVA table for model from experiment 3.1. 54
- Table 3.3. ANOVA table for model from experiment 3.2. 54
- Table 3.4. ANOVA table for models from experiment 3.3. 54

- 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. 55
- 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). 56
- 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). 57

Chapter 4

- Table 4.1. Parameters used for each simulated population in experiment 4.2. The F_{st} value for all the simulations at the separation of the single population into two sub populations was 0. Final F_{st} was calculated using all individuals and loci. 67
- Figure 4.1. Standard deviation of F_{st} over a number of generations (calculated using repeated sampling of different sample sizes, from 10 to 100). It can be seen that the variation over generations was very small except at generation 5000, where the populations were separated (also figure 4.2 for actual F_{st} values). 69
- Figure 4.2. The increase in summary statistics when a population was separated into two subpopulations. The summary statistics were calculated using all the loci from all the individuals at different generations and a clear increase in the population structure can be seen. 70

Figure 4.3. Decrease in standard deviation of the F_{st} as the number of individuals sampled increases for an increasing number of loci sampled for different sized populations (subpopulation size indicated in the legend from 1000 to 10000 with the F_{st} of the population when all individuals and loci are used). Larger populations have a lower standard deviation for all combinations of loci and individuals sampled. 72

Figure 4.4. Contour plot of the standard deviation of the F_{st} . The standard deviation decreases when either the number of loci genotyped or the number of individuals sampled increases. 73

Chapter 5

Table 5.1. Population source data with averages and standard deviation (SD) for the clutch size and number of underdeveloped individuals as well as the percentage of virgin females (i.e. male only broods) and average homozygosity per tree. 83

Figure 5.1 a) Range of traps, constructed from eppendorf tubes, used on different sized fruits. b) Traps fitted on figs of Tree 2. Scale = 1 cm. 85

Table 5.2. Primer sequences, label colour, size and number of alleles of the 6 microsatellites loci used to determine foundress genotypes. Reaction pools indicate the combinations of primers amplified together, with their respective annealing temperatures. 87

Table 5.3. PCR conditions, and reagents. The denaturing, annealing and elongation steps were repeated 30 times. *quantity for each primer in the pool 3. 88

Table 5.4. Parameters used for simulations. The maximum and minimum fitness values are from the 95% confidence intervals of *model 1*. Each parameter set was simulated 5 times. The number of homozygous loci for each successive sibmating was determined from Greeff et al. (2009). 90

Table 5.5. ANOVA table and model estimates for *model 5.1* (Adjusted $R^2 = 0.188$) 92

- Figure 5.2. Relationship between clutch size and number of homozygous loci for 93
each tree (*model 5.1*). Tree 1: dash-dot line, Tree 2: Long-dash line,
Tree 3: short-dash line, Tree 4: dotted line, Tree 5: solid line.
- Table 5.6. ANOVA table and estimates for the linear model: Under-developed ~ 94
tree (Adjusted $R^2 = 0.045$).
- Table 5.7. ANOVA table and estimates for the generalised linear model: 95
Proportion under-developed per clutch ~ tree (data transformed with
logit transformation)
- Table 5.8. Probability of sibmating from simulated parameter estimates (see table 95
5.4).
- Figure 5.3. Different trees used in study: a) Tree 1; b) Tree 2; c) Tree 3; d) Tree 4 97
and e) Tree 5. Scale = 1 meter, measured from marker attached to each
tree.