International Conference on Computational Science, ICCS 2010

Individual-based simulation of sexual selection: a quantitative genetic approach

D. van Dijk[a,*], P.M.A. Sloot[a], J.C. Tay[b], M.C. Schut[c]

[a]*Computational Science, University of Amsterdam, Sciencepark 107, 1098 XG Amsterdam, The Netherlands*
[b]*Evolutionary and Complex Systems Program, School of Computer Engineering, Nanyang Technological University, Singapore 639798, Singapore*
[c]*Department of Artificial Intelligence, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands*

Abstract

Sexual selection has been mathematically modeled using quantitative genetics as well as population genetics. Two-locus simulation models have been used to study the evolution of male display and female preference. We present an individual-based simulation model of sexual selection in a quantitative genetic context. We show that under certain conditions Fisherian self-reinforcing sexual selection takes effect, predicted by Lande's analytic model of female choice. We also show that the dynamics involved in the co-evolution of male display and female preference is much more complex than mathematics would predict. We therefore argue that the study of sexual selection through individual-based simulation could give new and more realistic insight into a world dominated by overly simplified equations.

Keywords: sexual selection, individual-based modeling, agent-based modeling

1. Introduction

The male peacock is known for his extravagantly developed and brightly colored tail and has served as one of the text book examples of sexual selection. Peacocks don't have a direct use for their tails other than displaying them to other peacocks; the females in particular. Therefore they are considered secondary, and not primary, sexual organs. Even more striking is the fact that these tails almost certainly reduce survival of the peacock as they reduce maneuverability, power of flight and make the bird more conspicuous to predators. Also, growing the tail is at a cost. So why haven't they been eliminated by natural selection?

Darwin's answer was the theory of sexual selection. He argued that sexual selection "depends on the advantage which certain individuals have over other individuals of the same sex and species, in exclusive relation to reproduction." A structure produced by sexual selection in males exists not because of the struggle for existence, but because it gives the males that possess it an advantage over other males in the competition for mates. Darwin's idea is that the

*Corresponding author

Email addresses: d.vandijk@uva.nl (D. van Dijk), p.m.a.sloot@uva.nl (P.M.A. Sloot), joccing@ieee.org (J.C. Tay), schut@cs.vu.nl (M.C. Schut)

reduced survival of the peacocks with long, colorful tails is more than compensated by their increased 'advantage in reproduction'.

So why do females prefer males with long tails? And why would a disadvantageous trait evolve in the first place?

Since Darwin many Biologists have contributed to the field of sexual selection. Most importantly R. A. Fisher with his theory of run-away selection and Lande with a mathematical model of Fisher's ideas.

Theories in of sexual selection have been either verbal or mathematical. Verbal models have given an intuitive understanding of the processes involved, but often lack the ability to describe the complexity and dynamics of sexual selection in the real world. Mathematics on the other hand have proven useful in predicting evolutionary outcomes and are today the primary tool for biologists to describe and test their theories.

Nevertheless, setting up a mathematical model for an evolutionary system is a difficult process and even for the simplest of theoretical systems results in complicated equations. So far, only 2 locus systems have been accurately modeled and it is said that multi-locus, polygenic systems can not be modeled to the extent that they accurately predict evolutionary outcome. This presents a problem for biologists that seek methods to test their hypotheses and make predictions.

In this work a possible solution for this problem is put forward. Using the techniques of individual-based simulation, a model is presented that describes sexual selection and mate choice in a dynamic and realistic way.

2. Related work

At the base of sexual selection lies the idea that the success of an organism is not only measured by the number of offspring left behind, but by the quality or probable success of the offspring, also known as reproductive fitness. Theories that use this principle to explain sexual selection are the so called "Good Genes" and "Sexy Son" hypotheses, but often fail to explain why overly exaggerated traits can evolve that undoubtedly decrease fitness.

One of the first and certainly revolutionary theories to resolve this paradox was presented by R.A. Fisher in 1930. He hypothesized that these overly exaggerated traits were the outcome of an explosive positive feedback loop that could sometimes arise in the indirect genetic interaction between female preference and male display.

Since Fisher many verbal as well as mathematical models have been developed, most of them based on his original ideas of so called "Run-Away Sexual Selection" [1].

2.1. Population Genetics

Kirckpatrick [2] developed a simple analytic model of sexual selection, consisting of two loci, each with two alleles [1]. Kirkpatrick's model assumes that the males and females only contribute gametes (genetic information) to the next generation and that there is no direct relationship between a female's mating preference and her survivorship or fecundity[2]. The model also assumes that the preference and trait loci are not sex-linked, thus both sexes are carriers of both genes. Also, the two different genes reside on two different chromosomes, so there is no physical linkage between the two genes.

2.2. Quantitative Genetics

So far quantitative genetics has been most successful in understanding and describing the workings of sexual selection. Most models today are based on Lande's [3] work of modeling Fisherian sexual selection. Lande's study was the first to show "run-away" selection in a genetic model as Fisher [4] had hypothesized. It was shown that this process was driven genetic linkage. Genetic linkage, or linkage disequilibrium is a measure of correlation between two genetic traits that get associated with each other in certain evolutionary scenarios. With this model Lande showed that such an evolutionary system is unstable when females are sufficiently critical in their choice, or when viability selection (natural selection) is weak.

Using this framework, in the two decades after it's introduction, more than 20 different models of sexual selection have been produced. Some altering or adding variables more than others, but all based on Lande's interpretation of Fisher's ideas. For a review of these models see [5], [6] and [1].

---

[1]An allele is a version of the gene. Two alleles means that each locus (or gene) has two possble values.
[2]fertility

New mathematical models of sexual selection are still being produced and are still creating new insight. But one could ask if this ultimately will lead to a unified model that perfectly describes sexual selection in its many forms. Systems that are complex as the one studied here often lack the ability to be described by a single set of equations. The dynamics simply depend on too many variables that result in unpredictable behavior. In the past these complex systems have been successfully modeled using simulation instead of mathematics.

2.3. Simulation

In the field of sexual selection though simulation techniques have very rarely been used. Only a few studies exist where populations of sexually selecting individuals are simulated using an agent-based or any other method.
In [7] a computer simulation is presented of sexual selection on age-structured populations. It was studied what effect age-structures have on sexual selection and mate choice. Another simulation study of sexual selection was carried out in [8], but this time emphasis was laid on spatial structure, more specifically on density-dependent effects. Although agent-based simulations were carried out, evolution was not measured over multiple generations. Instead it was measured how mate competition and female choice were effected by different population densities in a single generation. Most successful in simulating sexual selection and female choice in an evolving population were Collins and Jefferson [9]. In their study a computer simulation was presented based on Kirkpatrick's two-locus model of female choice. Results showed that evolution occurred along the predicted lines of equilibrium and that under certain conditions alleles could invade the population and rapidly evolve to a point of equilibrium.
Collins and Jefferson's work has served as a basis for the models presented in our study.

3. Model: introduction

Although Collins and Jefferson were successful in testing Kirkpatrick's theory, their model lacked an appropriate realistic genetics model and was thus unable to test quantitative genetic theories of sexual selection.
In this section we present our model of sexual selection, based on Collins and Jefferson's work. We extend the two-locus model by simulating a quantitative genetic environment, inspired by Lande's analytical model of Fisherian sexual selection.

3.1. Model: hypotheses

Based on Lande's mathematical model of Fisherian sexual selection in a quantitative genetic context we formulate a number of hypotheses:

 i) Strong sexual selection will cause a strong genetic association between male display and female preference that can be measured through a high linkage disequilibrium.
 ii) Strong genetic association causes the Fisher process to take effect that will be measured through rapid growth or decline in the evolution of male display and female preference.
iii) Increasing genetic variability, through mutation, will decrease the strength of the Fisher process.
 iv) Increasing the strength of natural selection on the male trait will decrease the strength of the Fisher process.

These hypotheses will show whether our model can successfully simulate Fisherian sexual selection and if Lande's analytic model is appropriate for describing the dynamics of our system.

3.2. Model: quantitative genetic

In order to simulate sexual selection with a quantitative genetics model we made a number of assumptions regarding genetics and other simulation variables. Next we will describe how we constructed and used this quantitative genetics model.

3.2.1. Gene encoding

Traits are determined by multiple (often many) genes that possibly reside on multiple chromosomes. Our genetics model assumes N_c chromosomes per individual, N_l loci per chromosome and N_a possible alleles per locus. On chromosomes lie the loci. These contain a version of the gene, a so called allele. In our model this allele is an integer value between 0 and N_a (including 0) and additively codes for a certain trait. In quantitative genetics, traits are coded by multiple genes. In our model half of the loci code for male display and half for female preference. Because we assume a diploid model, each gene is determined by two loci residing on two homologous chromosomes, thus the actual number of chromosomes is $2N_c$. Instead of assuming some sort of dominance relation between the two homologous alleles, we assume equal contribution by using the mean value of the two alleles.

3.2.2. Gene decoding

The phenotypic value of a trait (T_{phen}) is determined by the following equation:

$$T_{phen} = \frac{1}{N_c \cdot N_l \cdot (N_a - 1)} \sum_{c=1}^{N_c} \sum_{l=1}^{N_l} \tfrac{1}{2}(X_{clA} + X_{clB}) \tag{1}$$

Here X_{ijK} is the allele value at locus j on chromosome i. K indicates either the A or B version of the two homologous chromosomes.

3.2.3. Recombination

Collins and Jefferson's two-locus system used an independent recombination of the parental genomes for the children. To be able to guaranty this in a multi-locus model we introduce a cross-over procedure that recombines the parental genomes once each generation. This cross-over procedure is carried out on every chromosome. After randomly selecting a cross-over point the diploid chromosome is twisted on that point, so that loci after the cross-over point now reside on the other homologous chromosome.

3.2.4. Reproduction

Since in our model individuals can carry multiple chromosomes (because of diploidi this is actually always the case), reproduction and the actual passing on of the genes gets a bit more complicated. Like in real diploid organisms, in our model a child receives one homologous chromosome per parent of every chromosome pair. This results in a complete diploid set of chromosomes.

3.2.5. Mutation

We use mutation to add genetic variability to our population. The mutation probability describes the rate of mutation by determining the probability of each locus in each individual to mutate. When a locus mutates the allele value is increased or decreased (with equal probability) with one.

3.2.6. Viability selection

For the viability selection procedure we use a gaussian survival function $f_v(z)$ with mean μ and standard deviation ω that maximizes survival at optimum trait size. Where ω is the strength of survival, z the male trait size and μ the optimum trait size.

3.2.7. Mate selection

For the mate selection procedure too we resort to a gaussian function $f_s(z)$ with mean y and standard deviation v (inspired by [10] and [11]), that determines the probability of a male and female mating. v is the strength of female preference, z the male trait size and y the female preference size. The further away the size of a male trait is from the preference size of the female, the smaller the probability of them reproducing.

3.2.8. Generation

Each generation the following events take place: First natural selection (viability selection) kills a number of males based on their trait. This happens by drawing a uniform random number x_v for each male and removing the male from the population if $x_v > f_v(z)$ (where $0 <= x_v <= 1$ and z is the male trait value). Then all females select a male to mate with based on the mate selection procedure: Males are randomly drawn from the population with probability density function $f_s(z)$ (where z is the male trait value). Finally, recombination and mutation act on the population genotypes (as described above) to produce a new generation that replaces the old one.

In order to observe the dynamics of the model (and thus the dynamics of sexual selection) we introduce a number of measurements that will provide introspection into the most important processes of our simulations.

3.2.9. Genotype and Phenotype variability

Like in Lande's analytic model of Fisherian sexual selection we look at mean male trait and mean female prefer-ence. We also define two measures of variance in our distributions of genotypes and phenotypes.
We define phenotypic variance (Var_{phen}) as well as additive genetic variance (per individual):

$$V_A = \frac{1}{N_C \cdot N_l} \sum_{c=1}^{n_c} \sum_{l=1}^{m_l} (X_{cl} - \bar{X}_{cl})^2 \qquad (2)$$

Here N_c is the number of chromosomes, N_l the number of loci per chromosome, X_{cl} the allele value at locus l on chromosome c and $\bar{X}$ the mean allele value.

3.2.10. Linkage Disequilibrium

In section 2.1 we introduced the concept of genetic linkage when it was used by Collins and Jefferson [9] to measure the degree of association between the genes of male trait and female preference. Genetic linkage has been shown to exist experimentally in species that exhibit strong sexually selective behavior [12, 13]. In quantitative genetics the LD is supposed to drive Fisherian self-reinforcing sexual selection and is defined as the covariance between two genetic traits. We use LD in the normalized form, where it is measured as the Pearson's correlation coefficient.

4. Results

Using our quantitative genetic simulation model we perform a number of experiments. Here we will describe the design and setup of our experiments and show the results.
To explore the behavior of our model we first run a number of single experiments. Later we perform a more thorough experimentation.
Table 1 shows the model variables and their values used in our simulations.

4.1. Explorative study

All simulations here were run on the LISA linux computing cluster at the SARA computing center [3] for 100,000 generations.

Our first simulation (experiment A1) tests the basic model with no natural selection and no sexual selection. To add genetic variance we set mutation to $1 \cdot 10^{-5}$. The genome consists of 4 chromosomes, 4 loci and 4 alleles. In figure 1A the result is shown. Next we add natural selection to the simulation by setting ω to 1.0 (experiment A2, figure 1 B). In the result of the simulation we clearly see that the mean male trait rapidly declines before it settles around a new value. Female preference appears unaffected by natural selection. Now that we have a basic evolving

---

[3]http://www.sara.nl

Fixed variables	
Population size	$N_p = 50,000$ (Females: $N_{pf} = 25,000$, Males: $N_{pm} = 25,000$)
Number of offspring	2; one male and one female
Independent variables	
Mutation	$\mu = \{0, 1 \cdot 10^{-5}\} = \{off, on\}$
Natural selection	$\omega = \{\inf, 1.0, 0.2\} = \{off, weak, strong\}$
Mating preference	$v = \{\inf, 0.2, 0.02\} = \{off, weak, strong\}$
Number of chromosomes	$N_c = \{1, 4, 16\}$
number of loci	$N_l = \{1, 4, 16\}$
number of alleles	$N_a = \{4\}$
Dependent variables	
Mean population phenotype	Mean male display (t), mean female preference (p)
Phenotypic variance	(male: V_{PM}, female: V_{PF})
Additive genetic variance	(male: V_{AM}, female: V_{AF})
Linkage disequilibrium	LD (as correlation coefficient)

Table 1: Model variables: Fixed variables are the ones that remain unchanged throughout all our experiments. The independent variables will be changed between experiments in order to investigate their effect on the dependent variables; the ones measured during and after simulations.

system, we want to see the effect of mate selection. In the next experiment (experiment A3) we turn off natural selection and set v to 0.2 (weak sexual selection). We observe a number of things. First we see that the additive genetic variance of the male display rapidly declines to a low value (data not shown). This happens because females choose their male mates on their display value and the population eventually will only carry attractive males. This process is called direct selection.

Next we observe that mean male display and mean female preference are synchronized and move together up and down. This can be seen in figure 2A. Here we also observe that as expected mean female preference moves first after which mean male display follows. In figure 2B we see the *LD* over time and conclude that weak sexual selection creates a positive (be it weak) *LD*, although we do not observe any Fisher process in the form of rapid growth or decline of the two traits. Next we increase the strength of female preference by setting v to 0.02 (Experiment A4, figure 1C). We immediately see a lot more movement in both traits and when we look at figure 2C we observe a large *LD* throughout the simulation. In our next experiment (experiment A5) we will add natural selection to our previous experiment by setting ω to strong (0.2). All other variables remain at our previous setting.

In figure 1D and 2D we see mean traits and *LD* respectively. We observe a decrease in Fisher process as movement is less dynamic and *LD*s are smaller, although still significantly high.

4.2. Large-scale experimentation

Now that we have an understanding of the workings of our model we perform a more rigorous experimentation. Using the parallel computing power of Lisa [4] we perform a number of experiments. Each experiment is run 100 times for 10,000 generations with the fixed variable settings described in the experimental design.

Table 2 shows the parameter settings used for the independent variables (column 2-6) and experiment results (column 7-16), where variables were measured at the 10,000th generation.

---

[4]http://www.sara.nl/userinfo/lisa/description/
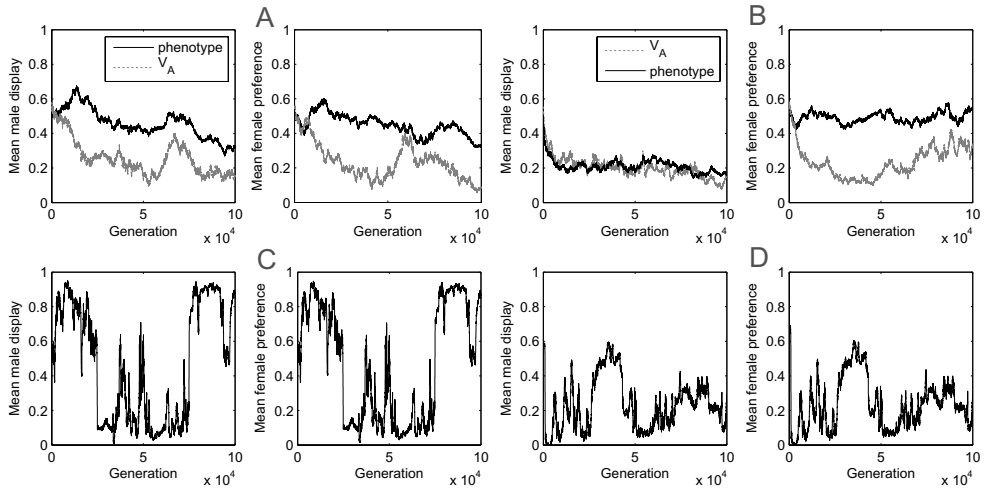680 nodes of 2 Intel Xeon 3.4 GHz cores and 4GByte ram

Figure 1: Four simulations of quantitative genetic sexual selection. Mean male display and mean female preference shown over time. A and B also show the additive genetic variance (V_A) over time. A: experiment A1 (*no mutation, no natural selection, no mate selection*). B: experiment A2 (*mutation, no natural selection, no mate selection*). C: experiment A4 (*mutation, no natural selection, strong mate selection*). D: experiment A5 (*mutation, strong natural selection, strong mate selection*).
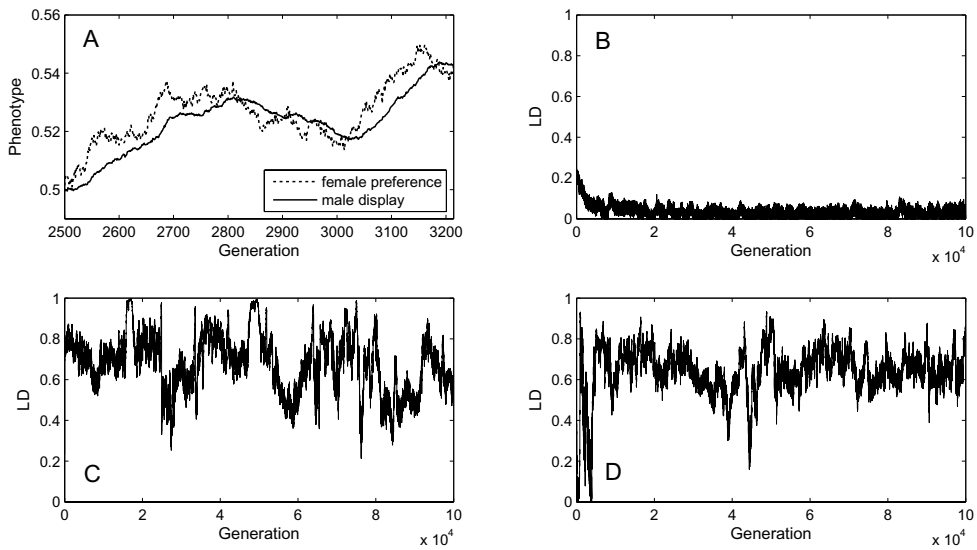

Figure 2: A: Male trait following the female preference. B,C,D: Linkage disequilibrium measured over time. A,B: experiment A3 (*mutation, no natural selection, weak mate selection*). C: experiment A4 (*mutation, no natural selection, strong mate selection*). D: experiment A5 (*mutation, strong natural selection, strong mate selection*).

exp.id	μ	ω	v	N_c	N_l	t	σ_t	p	σ_p	$V_{p,t}$	$V_{p,p}$	$V_{a,t}$	$V_{a,p}$	LD
B1	0	inf	inf	4	4	0.50	0.10	0.50	0.09	0.02	0.02	0.35	0.33	-0.00
B2	$1 \cdot 10^{-5}$	inf	inf	4	4	0.49	0.08	0.48	0.049	0.03	0.03	0.43	0.45	-0.00
B3	$1 \cdot 10^{-5}$	1.0	inf	4	4	0.22	0.03	0.51	0.07	0.01	0.02	0.23	0.37	0.00
B4	$1 \cdot 10^{-5}$	0.2	inf	4	4	0.09	0.02	0.49	0.07	0.00	0.02	0.12	0.38	0.00
B5	$1 \cdot 10^{-5}$	inf	0.2	4	4	0.52	0.08	0.52	0.08	0.00	0.02	0.02	0.34	0.06
B6	$1 \cdot 10^{-5}$	inf	0.02	4	4	0.55	0.32	0.55	0.31	0.10	0.10	0.33	0.34	0.72
B7	$1 \cdot 10^{-5}$	1.0	0.02	4	4	0.47	0.29	0.47	0.28	0.13	0.14	0.40	0.41	0.74
B8	$1 \cdot 10^{-5}$	0.2	0.02	4	4	0.16	0.11	0.16	0.11	0.04	0.04	0.25	0.27	0.66
B9	$1 \cdot 10^{-5}$	1.0	0.2	1	16	0.50	0.12	0.50	0.13	0.00	0.02	0.02	0.35	0.06
B10	$1 \cdot 10^{-5}$	1.0	0.2	16	1	0.49	0.12	0.49	0.12	0.00	0.02	0.01	0.31	0.05
B11	$1 \cdot 10^{-4}$	inf	0.02	4	4	0.48	0.20	0.48	0.20	0.14	0.14	0.56	0.57	0.78

Table 2: μ is the mutation rate, ω survival, v strength of female preference, N_c number of chromosomes and N_l the number of loci per chromosome. t is mean male trait, σ_t is the standard deviation of the male trait, p is mean female preference, σ_p the standard deviation of female preference, $V_{p,t}$ is the mean phenotypic variance of the male trait, $V_{p,p}$ the mean phenotypic variance of female preference, $V_{a,t}$ the mean additive genetic variance of the male trait, $V_{a,p}$ the mean additive genetic variance of female preference, LD the linkage disequilibrium measured as Pearson's correlation coefficient.

5. Conclusions

From the results of our explorative study and large-scale simulation we conclude that:

1) Strong mate selection causes strong genetic linkage.
2) Genetic linkage (a high *LD*) coincides with run-away Fisherian sexual selection.
3) Increased mutation decreases the strength of the Fisher process.
4) Strong natural selection decreases the strength of the Fisher process.

Our methodology of extending collins and Jefferson's model to polygenic simulation showed that quantitative genetic sexual selection can be studied successfully using individual-based simulation instead of mathematics. Although the model simulates a complex system and is therefore based on many variables, we succeeded to focus on the essential parts and showed that the complex dynamics can be understood intuitively.

Exploration of our model through experimentation showed that sexual selection and especially the Fisher process can be simulated. Conditions for Fisherian sexual selection, hypothesized by Lande, proved to be accurate for our simulation model. Firstly, strong sexual selection causes genetic association that in turn causes rapid change of traits, a result of the Fisher process. Secondly, an increase in genetic variance, caused by mutation, or an increase in natural selection cause a decrease in the strength of the Fisher process.

More rigorous experimentation, where we repeated each experiment 100 times, confirmed that our model behaves like the theory put forward by Lande. Interestingly, we have shown that the process of sexual selection in fact is much more dynamic than a set of equations would predict. Lande assumes many variables, like additive genetic variance and genetic association (*LD*) to be constant throughout generations. Using such assumptions evolutionary equilibria are either walk-towards or run-away. In a more complex and also realistic situation, like our model simulates, genetic variability and genetic association are dynamic variables. Consequently, an evolutionary process can constantly change from walk-towards to run-away, explaining our observation of short spurts of rapid increase or decay.

However, we still make a number of assumptions to simplify our model. The influence of the fixed population size and non-overlapping generations is unknown and should be investigated in future studies. Also, adding a spatial dimension to our model might have significant influence as it has already been shown that density-dependent effects play an important role in sexual selection [8, 1].

Simplifying a dynamical system with mathematics might often be useful and create understanding, but will also greatly abstract theoretical understanding from the real world. Especially in sciences like evolutionary biology one

must never be fully detached from the physical world and the problems we are trying to solve.
With our quantitative genetic simulation model we introduce an alternative and possibly more realistic technique for studying polygenic sexual selection and female choice.

Acknowledgements

References

[1] S. Cotton, J. Small, A. Pomiankowski, Sexual selection and condition-dependent mate preferences, Current Biology 16 (17) (2006) 755–765.
[2] M. Kirkpatrick, Sexual Selection and the Evolution of Female Choice, Evolution 36 (1) (1982) 1–12.
[3] R. Lande, Models of Speciation by Sexual Selection on Polygenic Traits, Proceedings of the National Academy of Sciences of the United States of America 78 (6) (1981) 3721–3725.
[4] R. Fisher, The evolution of sexual preference, Eugenics Review 7 (1915) 184–192.
[5] L. Mead, S. Arnold, Quantitative genetic models of sexual selection, Trends in Ecology & Evolution 19 (5) (2004) 264–271.
[6] H. Kokko, M. Jennions, R. Brooks, Unifying and testing models of sexual selection, Annu. Rev. Ecol. Evol. Syst 37 (2006) 43–66.
[7] S. Martins, T. Penna, Computer Simulation of Sexual Selection on Age-Structured Populations, Int. J. Modern Physics C 9 (1998) 491–496.
[8] H. Kokko, D. Rankin, Lonely hearts or sex in the city? Density-dependent effects in mating systems, Philosophical Transactions: Biological Sciences 361 (1466) (2006) 319–334.
[9] R. Collins, D. Jefferson, The Evolution of Sexual Selection and Female Choice, Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life.
[10] M. Andersson, Sexual Selection, Princeton University Press, 1994.
[11] J. Maynard Smith, Evolution and the Theory of Games, Vol. 64, 1976.
[12] S. Cotton, D. Rogers, J. Small, A. Pomiankowski, K. Fowler, Variation in preference for a male ornament is positively associated with female eyespan in the stalk-eyed fly diasemopsis meigenii, Proceedings of the Royal Society B 273 (1591) (2006) 1287.
[13] D. Rogers, M. Denniff, T. Chapman, K. Fowler, A. Pomiankowski, Male sexual ornament size is positively associated with reproductive morphology and enhanced fertility in the stalk-eyed fly teleopsis dalmanni, BMC Evolutionary Biology 8 (1) (2008) 236.