Inferring female extra-pair mating behaviour from observed patterns of extra-pair paternity with a process-based model

Master of Science in Biology: Biodiversity, Evolution and Ecology

Daniel Einarsen Sellæg

Department of Biology
University of Bergen
October 2014

Contents

Acknowledgements 4
Summary 4
1. Introduction 5
1.1. General introduction 5
1.2. Mammalian versus avian mating systems 5
1.3. The explanations for EPP in birds 6
1.4. Detection of EPP via molecular methods 9
1.5 From observable EPP rates to the underlying EPC behaviour 10
1.6. Randomness in EPC behaviour and the effects of clutch size 10
1.7. Purpose of the study 12
2. Materials & methods 12
2.1. Model description 12
2.2. General Assumptions 13
2.3. Variation in population sample size 14
2.4. The use of parameter values from research articles 15
3. Results 18
3.1. When all females are promiscuous 18
3.2. Varying proportion of promiscuous females 22
3.3. Model motivated by empirical observations 27
4. Discussion 34
4.1. To observe the 'invisible' 34
4.2. Assumptions 35
4.3. Variation in female extra-pair mating behaviour 39
4.4. Comparison between species 40
4.5. Implications for field biologists 41
4.6. Conclusion 42
5. References 43
6. Appendix 48
6.1. Supplementary figures 48
6.2. MATLAB script 63

Acknowledgements

First and foremost, I would like to thank my supervisors Sigrunn Eliassen and Christian Jørgensen for their unparalleled knowledge and support. Their humour and happiness has been priceless throughout my work with them and the master thesis. Second, a big thanks to my nearest study mates Bjørn Snorre, Eirik and Sindre for their humour, knowledge, passion and support. Third, I would also like to thank all the professors, stipendiats and researchers at the research group Theoretical Ecology for creating a great atmosphere. And last but not least, a huge thanks to my parents for always supporting me.

Summary

There has been great focus on extra-pair paternity (EPP) in birds over the last decades. The knowledge gained has made researchers rethink certain notions about mating behaviour in birds. Initially it was believed that socially monogamous birds were also sexually monogamous. It is hard to observe EPC behaviour in nature and EPC may not necessarily lead to extra-pair fertilization. Therefore, molecular methods may help us estimate the proportion of females with EPC behaviour. It has been found through molecular methods that there is high variation in EPP levels between species and populations. In this study, we mainly investigate how well EPP levels can be used to make inferences about the underlying mating behaviour that causes EPP. We used a process-based model as a 'virtual laboratory' in order to simulate populations of socially monogamous and sexually polyandrous females. Our findings suggest that it is hard to make inferences about extra-pair copulation (EPC) behaviour (1) when the proportion of extra-pair young (EPY) in a population is small, (2) when the clutch size is small, and (3) when the sample size is limited. Researchers working in the field should be cautious when drawing conclusions about the proportion of females with EPC behaviour based on EPP levels.

1. Introduction

1.1. General introduction

The majority (about 93 %) of passerine subfamilies are socially monogamous (Lack, 1968; Griffith *et al*., 2002; Rosivall *et al*., 2009). Social monogamy among birds means that a female and a male form a pair-bond and stay together throughout their lives, during a breeding season or until one member of the pair dies. It was first assumed that socially monogamous birds were also genetically monogamous (Lack, 1968). However, females of many socially monogamous species engage in extra-pair copulations (EPCs), which are copulations with males other than the pair-bonded, social mate. Extra-pair fertilization, which might follow after EPC, occurred in at least 86 % of all surveyed socially monogamous passerine bird species (Griffith *et al*., 2002). Extra-pair paternity (EPP) is the observable outcome of extra-pair fertilization. EPP is most often quantified as the proportion of nests in a population with extra-pair young (EPY) or the proportion of EPY among all offspring in a population, and these quantities vary greatly both between populations and species. For example, the most promiscuous, socially monogamous bird species that is documented is the reed bunting (*Emberiza schoeniclus*), where a study found that the proportion of nests with EPY was 86 % and the proportion of EPY was 55 % (Dixon *et al*., 1994). The average EPP rates among socially monogamous bird species are 11.1 % of offspring and 18.7 % of broods (Griffith *et al*., 2002).

1.2. Mammalian versus avian mating systems

The proportion of social monogamy in mammals is an order of magnitude lower than in birds; only 9 % (up from 3 %; Kleiman, 1977) of surveyed species conform to this mating system (Lukas & Clutton-Brock, 2013). The prevalent mating system among mammals is that females breed solitarily (about 68 % of surveyed species), and the second-most prevalent mating system is that individuals live in social groups (23 % of surveyed species). Intolerance and competition may be high among females that breed solitarily, and males usually roam around and mate with several females which they do not pair-bond with. Social monogamy is often associated with genetic monogamy in mammals, and extra-dominant paternity rates range from 0 % to over 40 % in socially monogamous species, and from 0 % to over 80 % in socially polygynous and multiple male societies (Clutton-Brock & Isvaran, 2006). ‘Societies’ is defined as groups of individuals

where the members of such groups live in close proximity of each other. There is usually one dominant male in such a group. Females can mate with males other than the dominant male. Thus, the term 'extra-dominant paternity' is used in this case instead of 'extra-pair paternity'.

'Extra-pair' is a term commonly used to denote mating that happens outside of the social pair-bond. There are other terms to denote promiscuity that may suit species that conform to other non-monogamous mating systems. Polygynandry is defined as mating between at least two members of each sex. The superb fairy-wren (*Malurus cyaneus*) is a species where the female mate with males outside the breeding group (Dalziell & Cockburn, 2008). They live together in a breeding group where all males give paternal care to the female they have mated with. The term 'extra-group' mating may be used in this case.

Polygyny is when a male mates with more than one female, and can be used to describe the mating system socially and genetically. Polyandry is when a female mates with more than one male. The focus of this master thesis is on females of passerine species being socially monogamous while at the same time being sexually polyandrous.

1.3. The explanations for EPP in birds

Male benefits and costs

The male benefits of EPC are obvious. The male gets to fertilize more eggs, pass on his genes to the next generation and increase his reproductive success. The male does so apparently without any costs to himself.

Female benefits and costs

The benefits of EPC behaviour for a female might not be quite as obvious as the benefits for a male, because a female does not increase the number of offspring that she produces (Trivers, 1972; Forstmeier *et al*., 2014). There are costs to being promiscuous. After successful extra-pair fertilization, the within-pair (WP) male will have at least one offspring in his nest that is unrelated to him. The potential costs to the female are that the WP male can reduce his care for the young, or desert the female altogether, and thus provide no care for the brood (Brommer *et al*., 2010). Males of species which live longer are more likely to desert broods than males of

short-lived species when uncertainty in paternity arises (Mauck *et al*., 1999). Forstmeier *et al*. (2014) have reviewed the following costs and benefits to EPC behaviour. Potential costs may be the reception of *de novo* (new) deleterious mutations (Dean *et al*., 2010; Johnson & Gemmell, 2012), inbreeding depression if mated with a closely related extra-pair (EP) male (Szulkin *et al*., 2013), increased embryo mortality (Morrow *et al*., 2002), punishment by the WP male after EPC (Valera *et al*., 2003), increased harassment from males that have been denied by the female (Thornhill & Alcock, 1983), increased sibling competition (Briskie *et al*., 1994) and sexually transmitted diseases (Sheldon, 1993).

Generally speaking, EPC behaviour is expected to evolve when the benefits for the female outweigh the costs. The benefits of EPC behaviour are divided into direct and indirect (genetic) benefits. For example, direct benefits can be increased parental care because more males have a stake in the brood paternity (Nakamura, 1998b), greater access to breeding resources (Birkhead & Møller, 1992), direct protection from male harassment (Rowe *et al*., 1994), male transfer of food or nuptial food gifts that increase female fecundity (Wedell, 1997), substances that promote egg maturation and oviposition (Cordero, 1995), preventing other females from mating (Petrie *et al*., 1992) and avoidance of male punishment (Clutton-Brock & Parker, 1995). The following paragraphs summarize key hypotheses concerning both types of benefits.

Regarding direct benefits, the ‘fertility insurance hypothesis’ states that EPCs can function as insurance against WP male infertility (Jennions & Petrie, 2000). EPCs can also insure against oligospermy, i.e., low concentration of sperm. A relatively recent review argues that benefits are obtained only when the WP male is truly infertile (Hasson & Stone, 2009). Infertility is expected to be rare since there is strong selection against infertility (Jennions & Petrie, 2000). However, it is not so rare that it should be ignored. On average, about 15 % of eggs do not hatch (Ihle *et al*., 2012). However, hatching failure can also happen due to high embryo mortality.

One of the hypotheses concerning indirect benefits states that having EPC increases the genetic variation in a brood and that it can function as genetic bet-hedging (Jennions & Petrie, 2000). In a variable environment, some of the young may die but others may survive due to genetic superiority. An analogy to explain bet-hedging is that a female might not attempt to ‘put all of her

eggs in one basket' but to spread them around (increase genetic variation between each offspring), and thereby assuring that at least some of the offspring may survive.

The 'good genes hypothesis' states that females actively seek out males with so-called 'good genes' (Jennions & Petrie, 2000). Such genes may increase the attractiveness and viability of the female's offspring. It is possible that a female bonds with a genetically inferior male because the costs of locating and obtaining a superior male could be imposed by the environment, for example, in the form of higher predation risk (Jennions & Petrie, 1997). It could also be the case that if all females want to bond with a male, but that there is only one 'good' male, then all females cannot pair up with the 'best' male and must choose between lower quality males. This could apply uniformly to all females or vary depending on their ability to withstand environmental and social costs (Jennions & Petrie, 1997). This should select for EPCs with males that are more attractive than the WP male. It has been shown in older field studies that a female is more likely to perform an EPC with or switch to a male that is more attractive or dominant than her social male (e.g., Bollinger & Gavin, 1991; Houtman, 1992). A male with more elaborate secondary sexual traits is less likely to be cuckolded by his social female according to a meta-analysis (Møller & Ninni, 1998). Moreover, such males are more sought after by other females and are therefore more likely to cuckold their social female (Stutchbury, 1998). It has been found, however, that the difference between a WP male and an EP male is most likely that EP males are older and more experienced than WP males, independently of differences between them in other traits (Forstmeier *et al*., 2014).

Another hypothesis concerning indirect genetic benefits states that females may seek mating with males that have genes with higher compatibility than the genes of their WP males (Zeh & Zeh, 1996). This hypothesis differs from other genetic benefit hypotheses in that the fitness consequences of intra-genomic conflict depend on an interaction between maternal and paternal haplotypes and are thus non-additive (Zeh & Zeh, 1996). Some genes of the WP male may not be compatible with the genes of the female. Incompatibility often results in defective or inviable offspring. It is also thought that incompatibility leads to hatching failure even though sperm is present.

A more recent hypothesis states that by extra-pair mating, females set up incentives for males to work for the 'public good' in the neighbourhood (Eliassen & Jørgensen, 2014). From a male standpoint, the most beneficial thing to do when his offspring is spread around in many different nests following EPCs is to focus on the safety and productivity of the entire neighbourhood rather than concentrating resources towards their own social nest. Then, from a female perspective, the benefits of a cooperative neighbourhood may outweigh the risk of lost care from her social male (Eliassen & Jørgensen, 2014).

1.4. Detection of EPP via molecular methods

Molecular methods such as DNA fingerprinting have been used to identify paternity in more than 200 bird species (Cornwallis *et al*., 2010; Griffith *et al*., 2002). When identifying paternity, DNA microsatellites are often used as a source of genetic markers. These parts of the DNA are also known as simple tandem repeats (STR). They consist of repeating sequences of 1-13 base pairs where a sequence may be repeated up to 150 times (Lodish *et al*., 2008).

Microsatellite genotyping is prone to errors, and few studies have investigated where these errors come from and how to detect the errors (Hoffman & Amos, 2005). When amplifying DNA by using PCR amplification, a common problem is the stochastic failure of one allele to amplify. This will make it look as if heterozygous individuals only carry one allele, and this is known as 'allelic dropout' (Navidi *et al*., 1992; Walsh *et al*., 1992; Gerloff *et al*., 1995; Taberlet *et al*., 1996; Gagneux *et al*., 1997). Such individuals will be scored falsely as homozygotes, and the allele that failed to amplify is called a 'null' allele. For example, if the genotype of the offspring is A/null, then it will be scored as A/A and will be deemed incompatible with fathers that have genotypes B/null and C/null (scored B/B and C/C, respectively). However, the offspring's genotype may still be compatible with the genotypes of these males. 'Null' alleles can be detected as a significant deviation from Hardy-Weinberg equilibrium (Jones & Ardren, 2003). Another artefact source is 'misprinting', where products of PCR amplification can be misinterpreted as true alleles (Goossens *et al*., 1998; Bradley & Vigilant, 2002), electrophoresis artefacts (Fernando *et al*., 2001; Davison & Chiba, 2003), the wrongful scoring of allele banding patterns, data entry and other clerical errors (Hoffman & Amos, 2005).

Since there are many errors that could occur when assigning paternity, the actual EPP rates of most species could be lower or higher than each study suggests, but few studies have used error ranges with measured EPP rates.

1.5. From observable EPP rates to the underlying EPC behaviour

EPCs may occur hidden from view or during short periods of time. A study by Wetton & Parkin (1991) found that it was difficult to identify participants in communal displays of the house sparrow (*Passer domesticus*). EPCs were brief and hidden from view more often than copulations with a female's WP male. This is most likely the case for all birds in general, but especially for socially monogamous passerine species. Thus, it is easier to count the number of EPYs found in nests or nest boxes and use the proportion of nests with EPY to estimate a minimum proportion of promiscuous females. Oftentimes, this is the only option available for field biologists.

EPC may not necessarily lead to fertilization and subsequently EPY. There is only a certain probability that it does. EPCs that do not result in EPY leave no trace, and hence a female without EPY in her nest would be intuitively interpreted as being non-promiscuous. Sperm competition and cryptic female choice are two post-copulatory mechanisms that can change the distribution of within-pair paternity (WPP) and EPP from a probability based on the amount of sperm from each male.

It has been demonstrated by Dunn & Lifjeld (1994) that there is no linear relationship between EPCs and EPP across several species, and there is little or no evidence to this date that there is a relationship. However, it is widely assumed that there is a linear relationship between EPCs and actual rates of EPP "because extra-pair behaviour is generally studied only through molecular studies that attempt to sample behaviour through detached studies of paternity in the molecular laboratory" (pp. 275, Griffith, 2007).

1.6. Randomness in EPC behaviour and the effects of clutch size

To test whether the distribution of EPY over broods is random has been urged as a critical first step in studying mating dynamics (Westneat *et al*., 1990; Griffith *et al*., 2002). One can test for randomness by comparing the observed distribution of EPY across nests to a random (expected)

distribution. The expected distribution can be created by first calculating the average proportion of EPY over *n* broods, and second, to use this value as a probability estimate. The expected distribution is then generated by using one of three different processes. Most studies have used a binomial process, while a few have used a hypergeometric or a Poisson process (Brommer *et al*., 2007). These studies have concluded that the distribution of EPY is non-random. Significant deviations of the observed distribution from the generated expected distribution are seen as a dichotomy, meaning that more females than expected have either a lot of EPY or none (Brommer *et al*., 2007).

Clutch size varies greatly between all bird species. However, most passerine bird species have clutch sizes ranging from 4-10 eggs in a brood, while more than half of all bird species lay 2 or 3 eggs (Jetz *et al*., 2008). One would intuitively expect that species with larger clutch sizes have a larger number of EPY in their broods but that the percentage of EPY may not necessarily be larger than in broods of species with smaller clutch sizes. One would also expect that based on probability alone, there is a higher chance to detect EPY in a clutch of 10 eggs than in a clutch of 3 eggs even though the EPC behaviour of the females are exactly the same.

The average clutch size among species increases as one move from equator towards the poles (Jetz *et al*., 2008). Lack (1947; 1968) hypothesized that food abundance during the breeding season determines clutch size. High seasonality in the temperate region can cause high adult mortality and this will in turn lead to the evolution of high investment in reproduction and large clutch size. This is because the chance of surviving to the next breeding season is low. One could also argue that the amount of food available is due to low population density in the temperate region, which would increase the amount of resources per individual. This would allow birds in this region to nourish larger clutch sizes (Jetz *et al*., 2008). Thus, one would intuitively expect that the probability of detecting EPY in nests of tropical species, on average, is lower than in nests of species in the temperate zone.

1.7. Purpose of the study

The aims of the study are to (1) explore the relationship between observed patterns of EPP and the underlying female extra-pair mating behaviour, (2) to investigate the effects of varying clutch size and sample size on the distribution of EPY across and within nests, (3) to use parameter values of EPP from research articles to make inferences about the underlying EPC behaviour, and (4) to find out when one is more likely to conclude incorrectly about the proportion of promiscuous females based on EPP levels. More specifically, we want to find out what is the most likely proportion of females with EPC behaviour based on measured EPP levels, and study how clutch size, EPP level and sample size may affect our ability to make inferences about female EPC behaviour.

2. Materials & methods

We used a process-based numerical model to simulate EPC behaviour and the outcome of this behaviour among socially monogamous females of passerine bird species. The mating system under particular investigation is social monogamy with sexual promiscuity.

2.1. Model description

We consider a population of N females that are socially monogamous, but they may seek out EP males to mate with. The total number of females (N) in each population is set to 10 000. The reason why this value was chosen, is because there was no observable difference in the distributions when N was larger than 10 000. In this way, we can see how the limit values of the distributions behave when N approaches infinity. All females in a population are equal in quality and all have the same clutch size (c). We assume that there is one female per nest, and vice versa, so that N also represents the number of nests in a population. A proportion f of females is promiscuous, i.e., they mate with males other than their social male. The parameter p represents the proportion of all chicks in the population that are EPY. Thus, p/f is the promiscuous females' EPY proportion or the probability that any of her eggs will be fertilized by an extra-pair male. In the model simulations, we use pseudorandom numbers between 0 and 1 to determine first whether a female is promiscuous or not, and second whether each egg will be fertilized by an extra-pair male or not. If the pseudorandom value is smaller than or equal to f, the female is promiscuous and there is a probability p/f that each of her eggs will be fertilized by an extra-pair

male while the rest are fertilized by the social male. If the pseudorandom value is greater than *f*, the female is not promiscuous and all eggs will be fertilized by the social male. The ecological process can be illustrated in a branched diagram (Fig. 1).

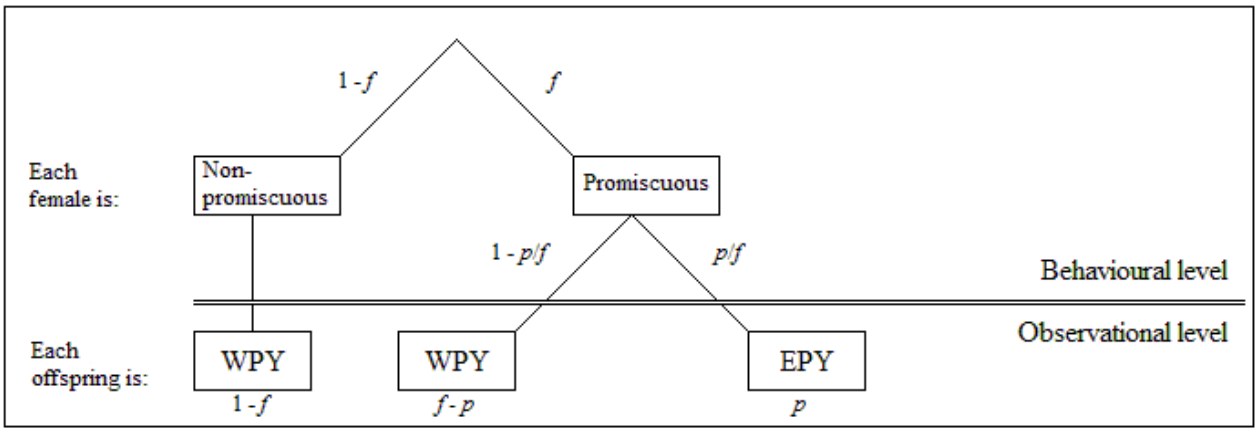


Figure 1. Conceptual presentation of the levels in the model. The model assumes two types of females. The proportion that is not promiscuous is denoted 1-*f*. The proportion that exhibits promiscuous behaviour is denoted *f*. This proportion copulates with extra-pair males, and the proportion of extra-pair young at the population level is denoted *p*. The paternity of the offspring is, unlike the female mating behaviour, easier to observe. The double line represents the divide between what is difficult to observe (upper part) and what is easy to observe (lower part). Promiscuous females can also be fertilized by the within-pair male, which results in WPY. If only WPY is observed in a nest, it is due to either that the female did not mate with an extra-pair male or that none of the eggs were fertilized by an extra-pair male.

We chose different proportions of EPY (*p*) as standard values to compare with each other throughout the analysis. These values were $p = 0.1$, $p = 0.4$ and $p = 0.8$. We did the same with clutch size. These values were $c = 4$ and $c = 10$.

2.2. General assumptions

There are simplifications to the model. This will most likely affect the predicted distributions. There is always going to be some sort of trade-off between simplicity and realism in a model (Hilborn & Mangel, 1997). A model that is too simple would exclude important factors, but a model that is too complex would not only take a long time to run, but would also be difficult to analyse and test.

All fertilizations are independent of each other

Fertilization success is not affected by previous fertilizations.

There are only two types of females

Females are either promiscuous or not. Those that are promiscuous will have a probability *p*/*f* that each egg is fertilized by an EP male, and this probability is constant for all promiscuous females in a population.

Constant clutch size within populations

All females within a population have the same clutch size. Between populations, however, the clutch size may vary. There are a maximum number of eggs in a nest, and this number has been chosen as a standard maximum value and it is 10 eggs. By using different clutch sizes between populations we can observe what effects different clutch sizes have on the distribution of EPY.

No intraspecific brood parasitism

We do not include egg dumping in the model. In other words, a female lays her eggs in her own nest.

No mortality among individuals

We do not include any differential mortality among females or differential mortality of WPY vs. EPY. All females and offspring are equally viable.

No infertility among individuals

No females or males are infertile.

2.3. Variation in population sample size

We wanted to investigate the effect of variation in sample size on the proportion of nests with EPY by dividing the whole population of N nests into N_{num} population samples with constant sample size. These population samples were not drawn randomly from the total number of nests (N), but N was divided up equally. The model can in this way be used as a 'virtual laboratory' to simulate realistic population sizes. The majority of empirical surveys has studied naturally small

population sizes (e.g., Charmantier & Blondel, 2003; Conrad *et al*., 2001). N_{pop} is the size of each population sample in terms of nests and N_{num} is the number of populations we study. N is still the total number of nests and can be written as:

$$N = N_{pop} \times N_{num} \quad (1)$$

We plotted everything that fell within two standard deviations as a shaded area around the mean. We compared two sample sizes (N_{pop} = 20; N_{pop} = 100).

2.4. The use of parameter values from research articles

The species and populations that were studied are presented in Table 1. We searched the databases of Oria, Google Scholar and Web of Science for data on measured EPP rates (*p* and proportion of nests with EPY) and other values that we could use as input parameter values in the model. Then we filtered out those species that we did not find many studies on or where the mating system of a particular species was not well described by the model. We could find independent EPP measurements from several populations of six species with a mating system that was well described by our model. These species were reed bunting (*Emberiza schoeniclus*), tree swallow (*Tachycineta bicolor*), pied flycatcher (*Ficedula hypoleuca*), collared flycatcher (*Ficedula albicollis*), great tit (*Parus major*), and blue tit (*Cyanistes caeruleus*). It must be noted that the pied flycatcher exhibits polygyny, which means that a male actively seeks out females and acquires at least one female to mate with (Stenmark *et al*., 1988). Only 3 populations each were chosen for both flycatcher species. The other species had 6 populations each. This is because not enough data were found for the flycatcher species. For details of the study sites, methodology of sample collection and parentage analysis, we refer to the original publications (Table 1). Some populations in Table 1 consist of combined data from many years because the sample sizes of each year were considered too small to be useful alone (N_{pop} < 10).

Table 1: Brief description of some of the data from the research articles that were chosen.

Population	Species*	Locality	Country	Study year(s)	References
1	CF	Niepolomice	Poland	2003-2006	Wilk *et al*., 2008
2	CF	Moravia	Czech Republic	2001-2002	Krist *et al*., 2005
3	CF	Gotland	Sweden	1994	Sheldon & Ellegren, 1999

Population	Species*	Locality	Country	Study year(s)	References
4	PF	Turku	Finland	2005-2006	Lehtonen *et al*., 2009
5	PF	Central Spain	Spain	2003	Moreno *et al*., 2010
6	PF	Central Spain	Spain	2010	Moreno *et al*., 2013
7	RB	Øvre Heimdalen	Norway	2001-2002	Kleven & Lifjeld, 2005
8	RB	Canton Zürich	Switzerland	2002-2005	Mayer & Pasinelli, 2013
9	RB	Gletterens	Switzerland	2004	Suter *et al*., 2009
10	RB	Gletterens	Switzerland	2005	Suter *et al*., 2009
11	RB	Gletterens	Switzerland	2006	Suter *et al*., 2009
12	RB	N/A	United Kingdom	N/A	Dixon *et al*., 1994
13	TS	Ontario	Canada	1990-1991	Dunn *et al*., 1994
14	TS	New Brunswick	Canada	1990-1995	Conrad *et al*., 2001
15	TS	Portland	Canada	1992-1993	Barber *et al*., 1996
16	TS	New Brunswick	Canada	1993	Conrad *et al*., 2001
17	TS	Prince George	Canada	2004	O'Brien & Dawson, 2007
18	TS	Wisconsin	USA	1997-1999	Whittingham *et al*., 2006
19	BT	Rouvière	France	2000	Charmantier & Blondel, 2003
20	BT	Rouvière	France	2001	Charmantier & Blondel, 2003
21	BT	Corsica	France	2000	Charmantier & Blondel, 2003
22	BT	Corsica	France	2001	Charmantier & Blondel, 2003
23	BT	Toledo	Spain	2010-2011	García-Navas *et al*., 2013
24	BT	Jomfruland	Norway	1994	Krokene *et al*., 1998
25	GT	Vlieland	Netherlands	1993-1994	Verboven & Mateman, 1997
26	GT	Wuppertal	Germany	1994	Strohbach *et al*., 1998
27	GT	Bahrdorf	Germany	1994	Lubjuhn *et al*., 1999
28	GT	Bahrdorf	Germany	1995	Lubjuhn *et al*., 1999
29	GT	Bahrdorf	Germany	1996	Lubjuhn *et al*., 1999
30	GT	Bahrdorf	Germany	1997	Lubjuhn *et al*., 1999

Note: Locality may refer to a specific location or a more general location.

* CF = collared flycatcher; PF = pied flycatcher; RB = reed bunting; TS = tree swallow; BT = blue tit; GT = great tit.

Table 2: Measured values from research articles that were chosen.

Population	Species	Proportion of EPY (p)	Proportion of nests with EPY	Mean clutch size (c_{pop})	Sample size (N_{pop})
1	CF	0.15	0.34	6.1	78
2	CF	0.24	0.51	6.1	27
3	CF	0.15	0.32	5.8	79
4	PF	0.04	0.13	4.4	191
5	PF	0.07	0.22	6.0	58
6	PF	0.13	0.28	4.5	59
7	RB	0.29	0.54	4.6	72
8	RB	0.37	0.56	3.6	181
9	RB	0.33	0.55	3.8	38
10	RB	0.45	0.71	3.7	56
11	RB	0.36	0.60	3.7	49
12	RB	0.54	0.86	3.7	58
13	TS	0.46	0.71	5.2	39
14	TS	0.51	0.74	4.9	106
15	TS	0.68	0.84	4.4	25
16	TS	0.55	0.84	5.1	13

Population	Species	Proportion of EPY (p)	Proportion of nests with EPY	Mean clutch size (c_{pop})	Sample size (N_{pop})
17	TS	0.35	0.85	5.4	40
18	TS	0.48	0.78	5.1	46
19	BT	0.12	0.40	9.0	25
20	BT	0.16	0.52	8.0	25
21	BT	0.21	0.68	5.9	25
22	BT	0.29	0.68	5.6	25
23	BT	0.11	0.46	6.5	26
24	BT	0.10	0.38	10.0	18
25	GT	0.03	0.08	6.2	82
26	GT	0.05	0.33	9.0	39
27	GT	0.08	0.32	7.1	65
28	GT	0.06	0.27	8.6	36
29	GT	0.07	0.33	6.8	75
30	GT	0.08	0.44	8.6	52

Note: The mean clutch size (c_{pop}) is calculated by dividing the number of offspring by the number of nests. The EPP rates were found either directly from the research articles or by dividing the number of EPY found by the total number of young and by dividing the number of nests with EPY found by the total number of nests.

The sample sizes varied greatly among the studies that we found (Table 2). We only included studies that reported EPP levels as the proportion of EPY (p) and the proportion of nests with EPY, the population's mean clutch size (c_{pop}) and the sample size (N_{pop}). The model was run with these parameters to create a scatterplot of points that represent simulated populations with emergent proportions of EPY and emergent proportions of nests with EPY. Each simulated population has an f-value assigned to it. This f-value represents the expected proportion of females with EPC behaviour in a population. The frequency of each f-value can be plotted. In order to do this, we need an error range for both the measured proportion of EPY and the measured proportion of nests with EPY to determine whether or not to count each f-value assigned to populations that fall within the range. The error ranges are represented as squares in all figures. The error ranges are based on the following two equations:

$$\Delta x = 2k_{\mathrm{x}} \frac{1}{N_{\mathrm{pop}}-1} \quad (2)$$

$$\Delta y = 2k_{\mathrm{y}} \frac{1}{cN_{\mathrm{pop}}-1} \quad (3)$$

Δx is the error range in x-axis direction in the scatterplot, while Δy is the error range in y-axis direction. N_{pop} is the sample size and c is the population's mean clutch size (c_{pop}) rounded off to

the nearest whole number. The parameters k_x and k_y can be used to alter the length of the error ranges in *x*-axis and *y*-axis direction, respectively. As standard values, $k_x = 3$ and $k_y = 4.5$ was chosen. When we run the model with parameter values from research articles listed in Table 2, the model is going to create many simulated populations. These populations have an emergent proportion of nests with EPY and an emergent proportion of EPY. Each time these values fall within the error ranges described above, an *f*-value will be counted. A frequency distribution of all counted *f*-values can then be plotted in a bar diagram. The frequency distribution of *f* can be interpreted as a probability density function. We based the error ranges on the assumption that the measured values in the research articles could be wrong due to error sources that may exist when identifying paternity (see section 1.4). We used MATLAB for simulations, analysis, and graphical presentation.

3. Results

In section 3.1, the proportion of females (*f*) with EPC behaviour is set to 1.0, which means that all females have mated with EP males. In section 3.2, we consider the situation when we change the proportion of females (*f*) that have EPC behaviour. In section 3.3, the values for the proportion of EPY in the population (*p*), the proportion of nests with EPY, the clutch size (*c*) and the sample size of each population (N_{pop}) are taken from research articles. The purpose of this is to use the model to predict a frequency distribution of potential *f*-values found within each simulated population.

3.1. When all females are promiscuous

We first consider the situation where we assume that all females have EPC behaviour ($f = 1.0$). A comparison between different EPP levels is made (Fig. 2). The probability that a nest contains at least one EPY increases as the clutch size increases (Fig. 2). This is because more opportunities are granted for the eggs to be fertilized by an EP male. This may cause one to overlook the large proportion of females that are promiscuous, as none of the eggs in the nest gets fertilized by and EP male despite the female being promiscuous (grey areas in Fig. 2). If one were to use the proportion of nests with EPY as an indicator of the proportion of females that are promiscuous, the proportion often gets underestimated. If the clutch size is low and the EPY proportion (*p*) is low, then extra-pair mating behaviour would most likely remain undetected in most of the nests

(Fig. 2a). Almost 40 % of the nests have no EPY although the female mated with an extra-pair male, even when the EPY proportion (*p*) was 0.1 and the clutch size was 10 (Fig. 2a). A relatively high clutch size would be needed (*c* = 6 to 10) for the proportion of nests with undetected extra-pair mating behaviour to be less than 5 % if the EPY proportion (*p*) is 0.4 (Fig. 2b). When *p* = 0.8, as in some extreme cases, all nests have EPY when the clutch size is 3 or larger (Fig. 2c). Knowing that among socially monogamous passerine species EPP rates averages 11.1 % of offspring and 18.7 % of nests (Griffith *et al*., 2002) and clutch size being in the range from 2 to 10 eggs on average, some species will fall within the grey area in Fig. 2a which suggests a substantial proportion of undetected EPC behaviour.

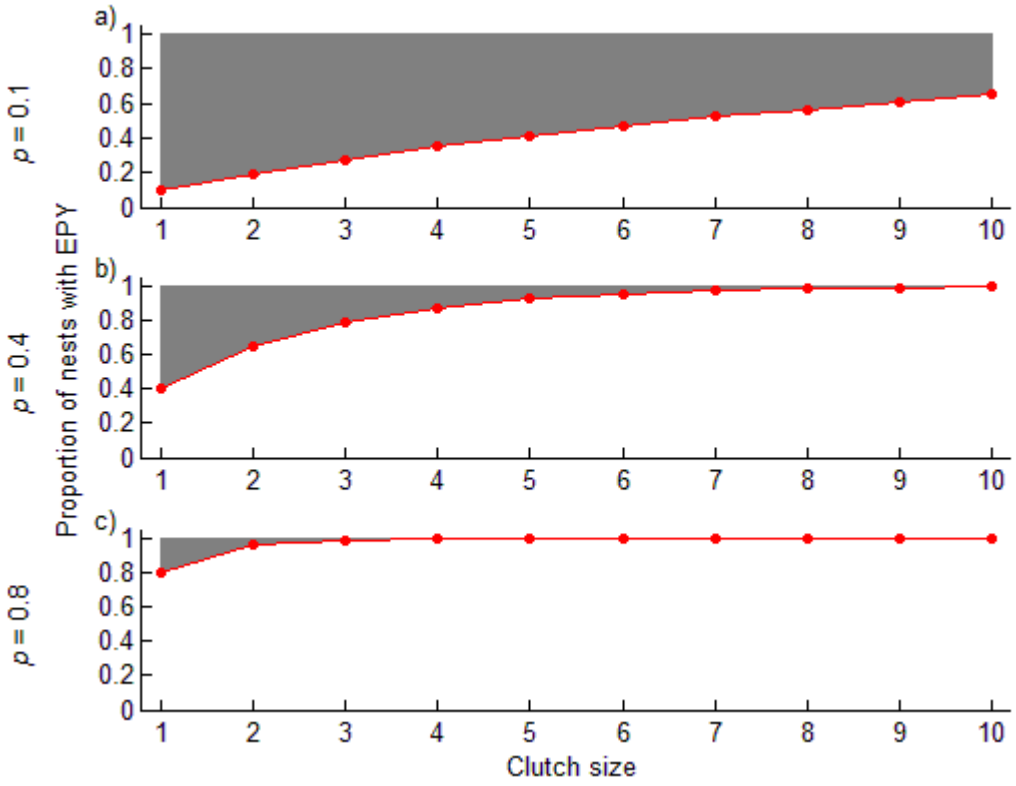


Figure 2. Proportion of nests with EPY versus clutch size for three different levels of EPP (*p*). The grey area shows the amount of EPC behaviour that is not detected because a female can have EPC behaviour but it does not manifest itself as extra-pair young in the nest. As the proportion of EPY (*p*) increases, the amount of undetected EPC behaviour decreases. Parameters: *f* = 1.0; *N* = 10 000.

When *p* is small and the clutch size is small, there are not many EPYs in the nests, and most of the nests have no EPY or only one EPY (Fig. 3). When *p* increases, there is a shift in the distribution towards higher frequencies of larger numbers of EPY in each nest, while there are few or almost no nests that have zero EPY. The effect is more pronounced when the clutch size is larger. When the clutch size is small, we can see that the number of EPY occurring at the highest frequency (modal value) is approximately *p* multiplied by *c*, which is the expected mean of a binomial distribution. When *c* = 10, we can see that this is exactly the case.

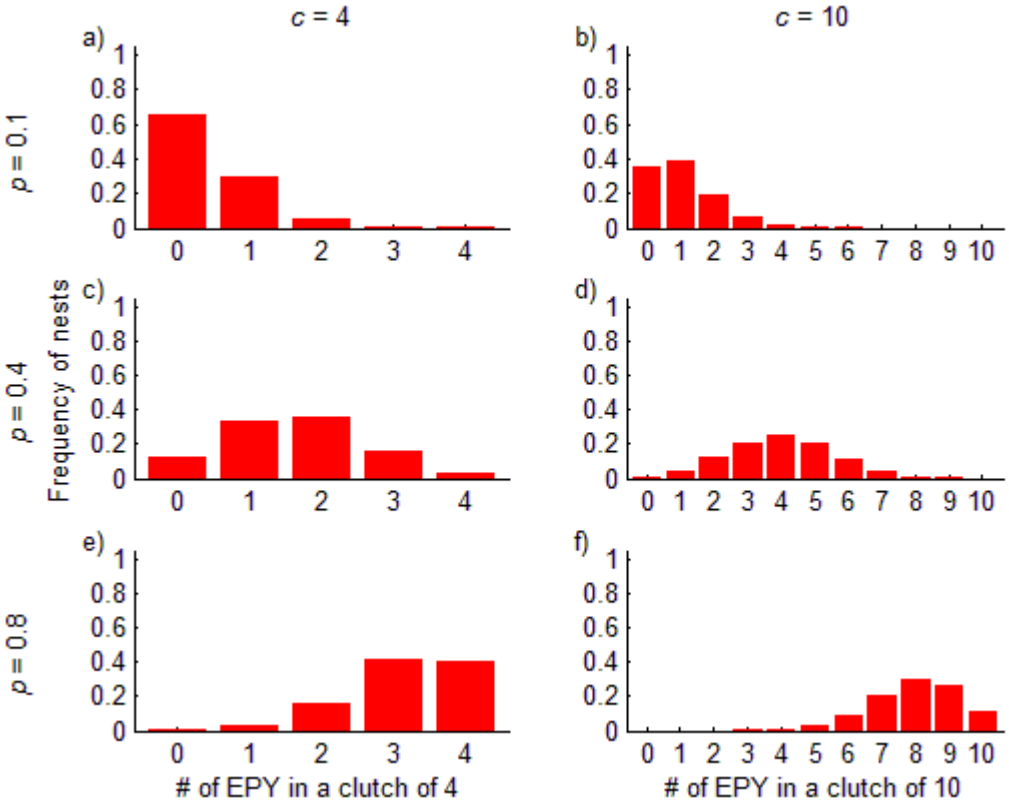


Figure 3. Frequency of nests that contain a given number of EPY when all females have EPC behaviour. As *p* increases, there is a shift in the distribution towards a larger average numbers of EPY in the clutch. Parameters: *f* = 1.0; *N* = 10 000.

After using the model to see general features and how everything behaves when N is large, we consider the variance introduced by limited sample size. One general pattern is that the smaller the sample size is, the more variation there is in the distribution (Fig. 4). A sample size of 100 has lower variance than a sample size of 20. However, the constraints of sample size disappear when p is large, and in particular when c is large.

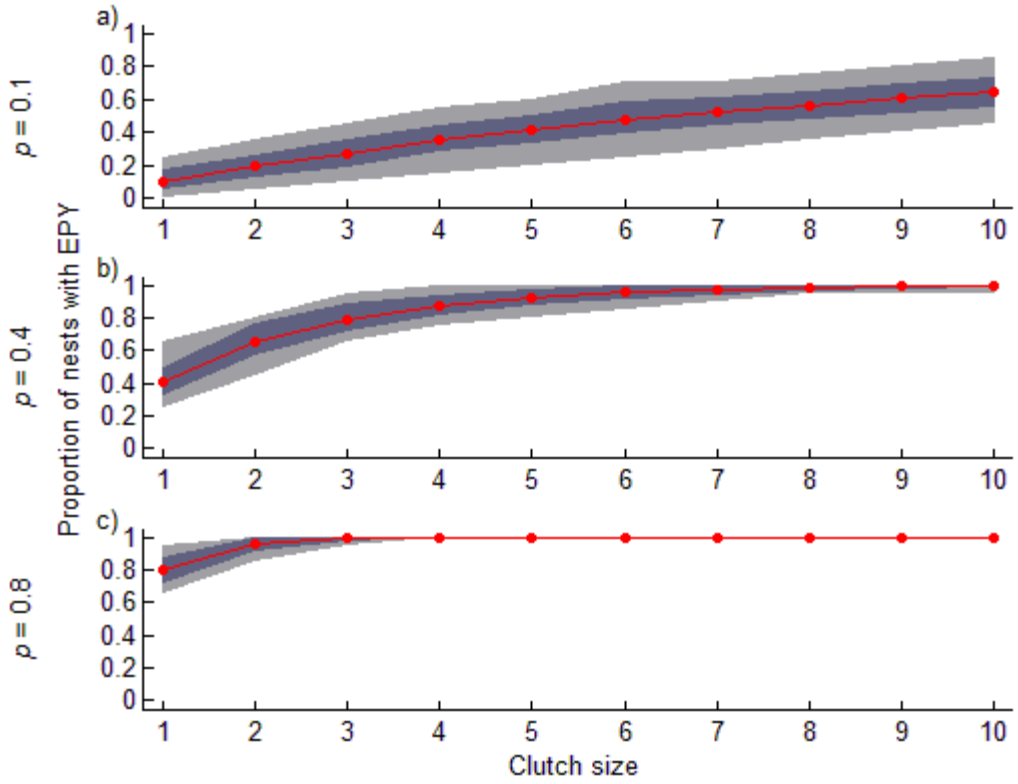


Figure 4. The effect of limited sample size on proportion of nests with EPY when all females have EPC behaviour. Three different p-values have been chosen for comparison. Variance increases as sample population size decreases. The dark grey area represents the variation (95th percentile) when the sample population size is 100, and the light grey area when the sample population size is 20. Parameters: f = 1.0; N = 10 000.

3.2. Varying proportion of promiscuous females

We now introduce the possibility that there are two types of females in the population, a proportion of promiscuous females (*f*) having EPC while the rest only copulates with their social male. When the *f*-value decreases, the proportion of nests with EPY becomes smaller (Fig. 5a). When $f = 0.4$ and $p = 0.4$, we can see that the proportion of nests with EPY remains constant for all clutch sizes (Fig. 5b). This is because the EPY proportion (p/f) of each promiscuous female is 1.0, which means that all females that are promiscuous have only EPY in their nests regardless of clutch size. Five *f*-values have been chosen for the plots. However, for the last two plots, there are some *f*-values that have not been plotted because the EPY proportion (p/f) of promiscuous females is larger than 1. When *p* and *f* are the same, the amount of undetected EPC behaviour is 0, and all promiscuous females have only EPY. However, when $p = 0.1$, $f = 0.2$, and the clutch size is small, there are some small amounts of undetected EPC behaviour. The larger the *f*-value becomes, the smaller the *p*-value becomes, and the smaller the clutch size becomes, the more undetected EPC behaviour there is.

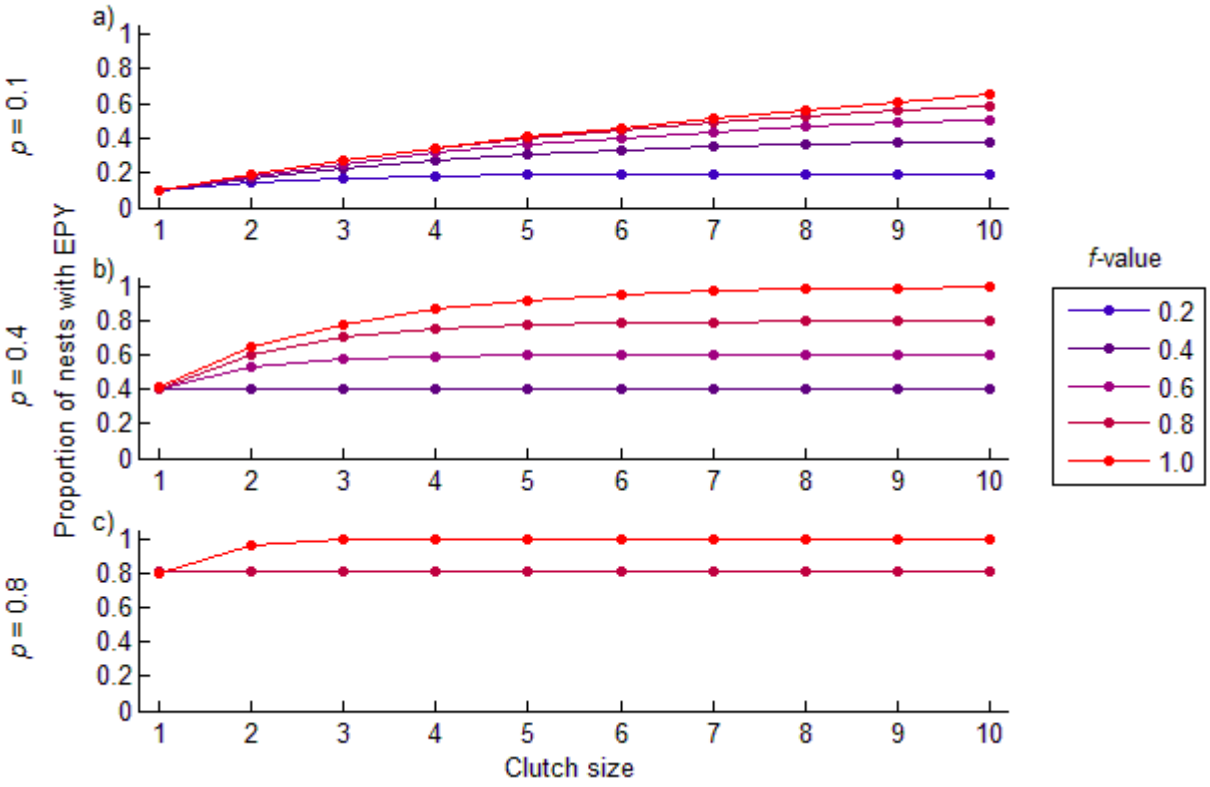


Figure 5. Proportion of nests with EPY versus clutch size. There are some *f*-values that have not been plotted, and this is because $p/f > 1$. When the *p*-value is small, we need a larger clutch size in order to reach a given proportion of nests with EPY. Parameters: $N = 10\,000$; $c = 10$.

The proportion of females that have EPC behaviour also influences the distribution of EPY across nests (Fig. 6). Even though there is a change in the clutch size, the shape of the distribution is almost the same for both clutch sizes. When p = 0.1 and c = 4, we can see that a lot of nests have no EPY in them while fewer nests have at least one EPY in them (Fig. 6a). As the f-value decreases and the p-value increases, a dichotomy in the distribution starts to manifest itself (Fig. 6c and 6d). By dichotomy (bimodal probability distribution) it is meant that there is a large proportion that does not have EPY at all, and the rest is having a lot of EPY or only EPY. This is valid for most clutch sizes. When the p-value becomes large enough, the graph for the two lowest f-values disappears. This is because the EPY proportion per female (p/f) is larger than 1.

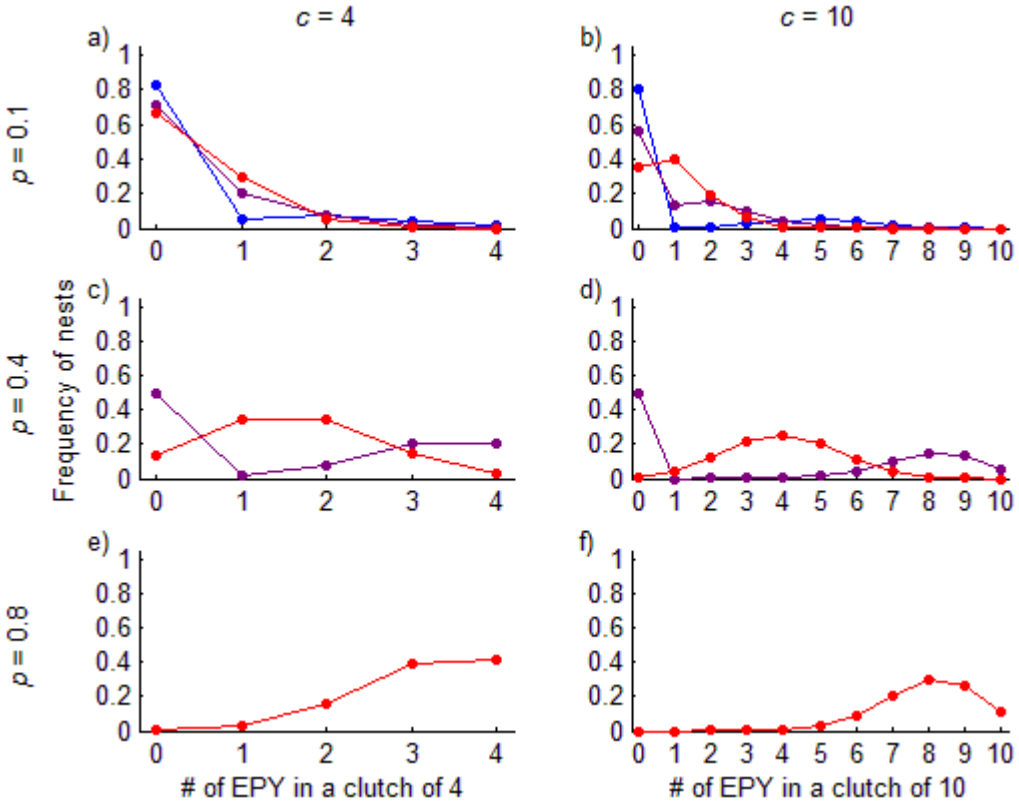


Figure 6. Frequency of nests that contain a given number of EPY when the proportion of promiscuous females (*f*) varies. a) When the clutch size is small, there is little difference in the shape of the distribution and the frequency values. b) As clutch size increases, the difference in the frequency of nests between f-values becomes larger. c) As the p-value increases, a dichotomy forms. d) This dichotomy is more pronounced and the shape is smoother when clutch size increases. e) and f) When the p-value and f-value are large, most nests have many EPY. Blue line is f = 0.2, purple line is f = 0.5 and red line is f = 1.0. Other parameters: N = 10 000.

A population can have a certain proportion of nests with EPY, but the proportion of females (*f*) that have EPC behaviour can be larger (Fig. 7). When the level of EPP is small ($p = 0.05$) and the clutch size is small ($c = 4$), we can see that the proportion of females (*f*) with EPC behaviour varies greatly but that the proportion of nests with EPY stays almost the same (Fig. 7a). This is not the case when the clutch size is larger ($c = 10$) and the level of EPP is 0.05 (Fig. 7b). This is due to the fact that as the clutch size increases, there are more opportunities for at least one egg in a nest to be fertilized by an EP male.

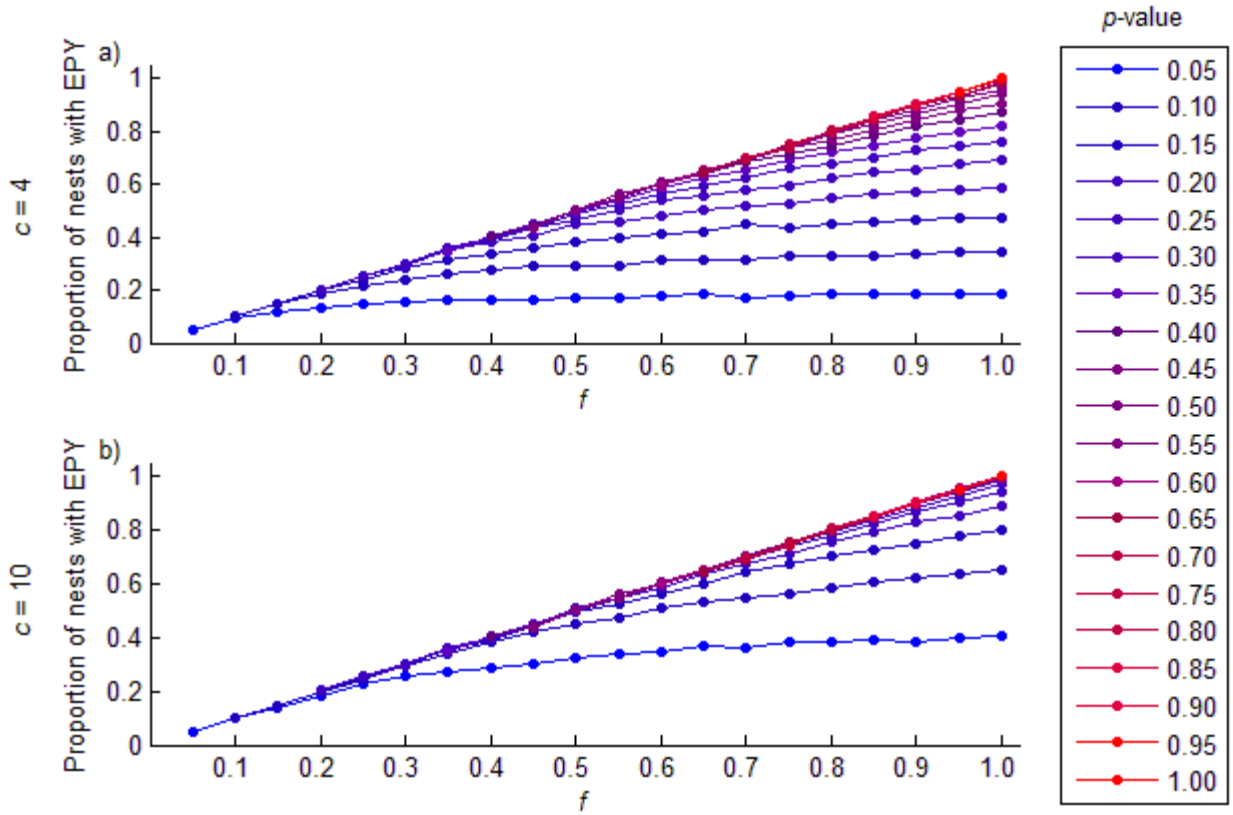


Figure 7. Proportion of nests with EPY versus proportion of females (*f*) with EPC behaviour. Each line represents a different *p*-value. a) When the clutch size is small and the *p*-value is small, the proportion that has EPC behaviour can be different even when the proportion of nests with EPY is the same or almost the same. Parameters: $N = 10\,000$.

As clutch size increases while p = 0.1 and f = 1.0, the proportion of nests with EPY increases (Fig. 8). This is because more opportunities are granted for at least one egg to be fertilized by an EP male. This figure can be used by an empiricist to look up an *f*-value given the values for the proportion of EPY (*p*) in the population, the rounded-off average clutch size (*c*) and the proportion of nests with EPY. When the *f*-value is small and when the *p*-value increases, there is almost no change in how large the proportion of nests with EPY is. This means that the same females that have EPY get more EPY in their broods. The only difference between Fig. 7 and 8 is that on the x-axes of Fig. 7 the proportion of females with EPC behaviour is used, while on the x-axes of Fig. 8 the proportion of EPY (*p*) is used. The figures represent two different perspectives in that the *f*-value is known in Fig. 7 but not in Fig. 8.

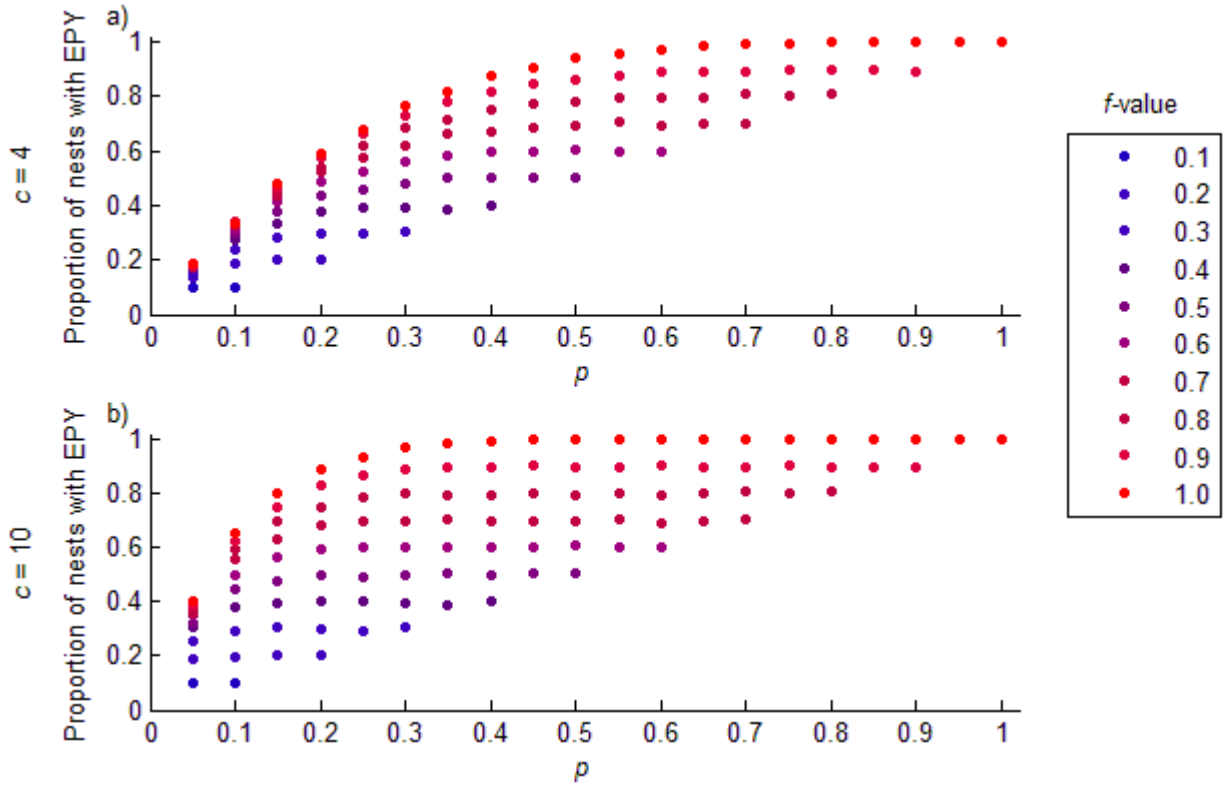


Figure 8. Proportion of nests with EPY versus proportion of EPY (*p*). a) The distribution of potential *f*-values when the clutch size is 4, and (b) when the clutch size is 10. Each point represents a population. Parameters: N = 10 000.

3.3. Model motivated by empirical observations

In this section we use parameter values taken directly from research articles, except for the first two figures. The points in the middle of the squares represent EPP levels from 3 imagined measurements, and all other coloured points in the scatterplot represent simulated populations (Fig. 9a). Three squares with different colours (Fig. 9a) have corresponding frequency distributions of *f*-values (Fig. 9b). The sample size chosen here is N_{pop} = 500, which is a much larger value than what is found in most research articles. When the sample size is large, the distributions of the *f*-values are relatively smooth (Fig. 9b).

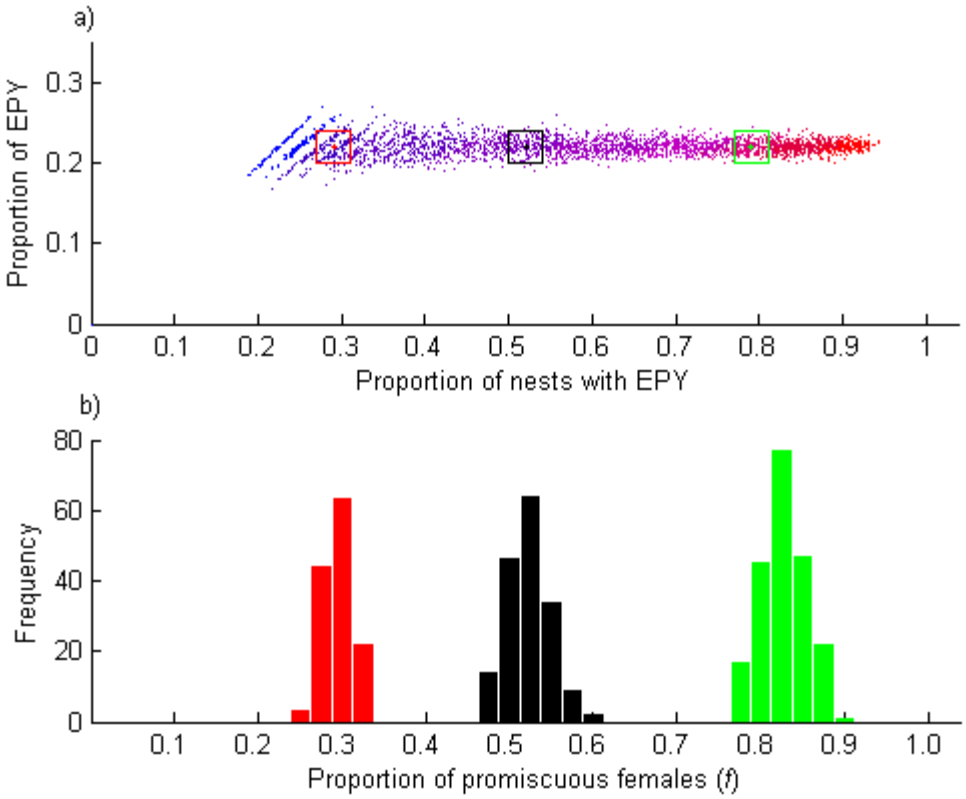


Figure 9. Frequency distribution of *f*. a) Proportion of EPY versus the proportion of nests with EPY. The size of each square is the same and is determined by the point in the middle of each square and the corresponding Δ*x* and Δ*y* of each square. The different colours of the points in the scatterplot correspond to different *f*-values where blue represents the smallest *f*-value and red represents the largest *f*-value. b) The distributions of each *f*-value that is found within the corresponding squares. The values (EPP levels) used to create the points within each square in Fig. 9a have been chosen for practical reasons. Parameters: N_{pop} = 500; c = 10; p = 0.22; proportion of nests with EPY (from left to right) = 0.29, 0.52 and 0.79.

The increase in sample size has two effects on the frequency distribution of f (Fig. 10). First, the frequency distribution becomes smoother as a result of less stochastic variation between population samples. Second, the size of the square becomes smaller because the error margin of Δx and Δy becomes smaller. This causes fewer f-values to be counted, and those f-values that get counted, do so less often. We have drawn a green square together with the black square in an attempt to separate the effects (Fig. 10b). The sample sizes chosen here are $N_{\text{pop}} = 20$ and $N_{\text{pop}} = 100$. When the sample size is small, the f-values get distributed over a wider range (Fig. 10c). Due to the high variation, it becomes harder to say anything for sure about the underlying EPC behaviour. When the sample size is larger, the f-values get distributed over a narrower range (Fig. 10d) and it becomes easier to make inferences about the underlying EPC behaviour.

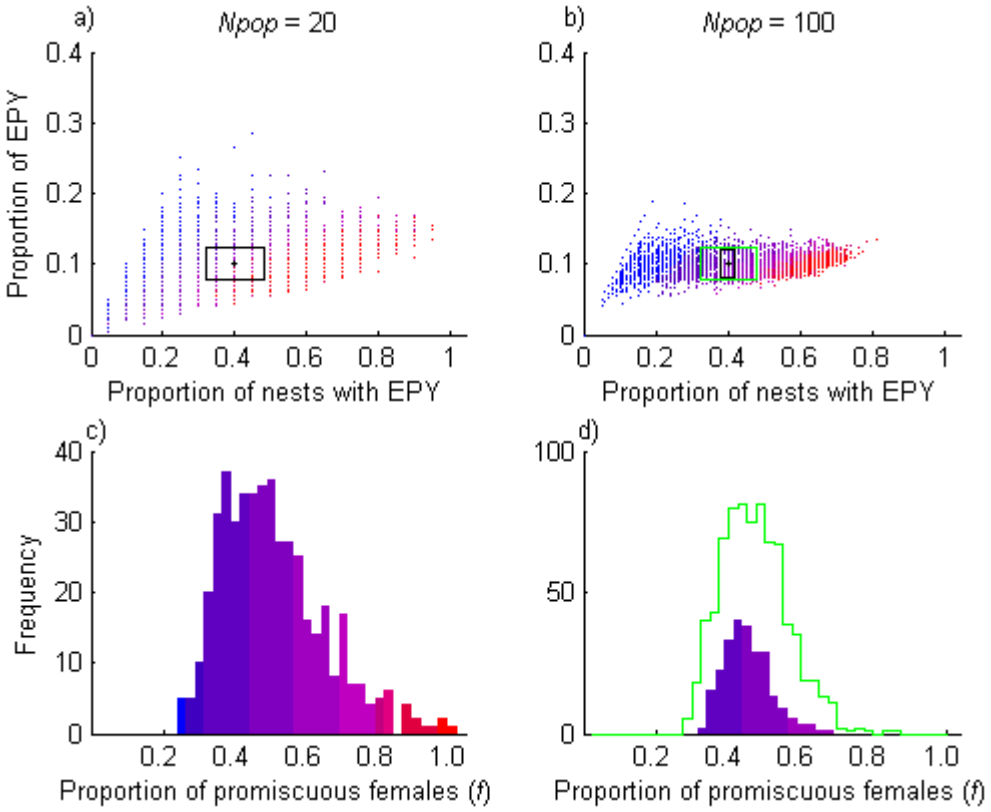


Figure 10. Comparison between two different sample sizes. a) Distribution of simulated population samples when each population sample (points in scatterplot) has size N_{pop} = 20. b) Distribution of simulated population samples when each population sample has size N_{pop} = 100. c) The corresponding frequency distribution of *f* has a lot of noise, and this tells us that it is not easy to predict an *f*-value when the population sample size is small (given all other parameter values). d) The frequency distribution of *f* contains less noise when the sample size is larger. The green line shows the distribution of the *f*-values found within the green box in fig. 10b. The bars show the distribution of the *f*-values found within the smaller black square. The size of the green square in fig. 10b is equal to the size of the black square in fig. 10a. Parameters: p = 0.1; c = 10; proportion of nests with EPY = 0.4, N_{num} = 100.

Next, we take a look at the simulation results after using parameter values taken from a research article (Fig. 11). As an example, the species shown here is reed bunting (*Emberiza schoeniclus*). The frequency distributions of the other populations found in Table 2 are listed in section 6.1 in the Appendix.

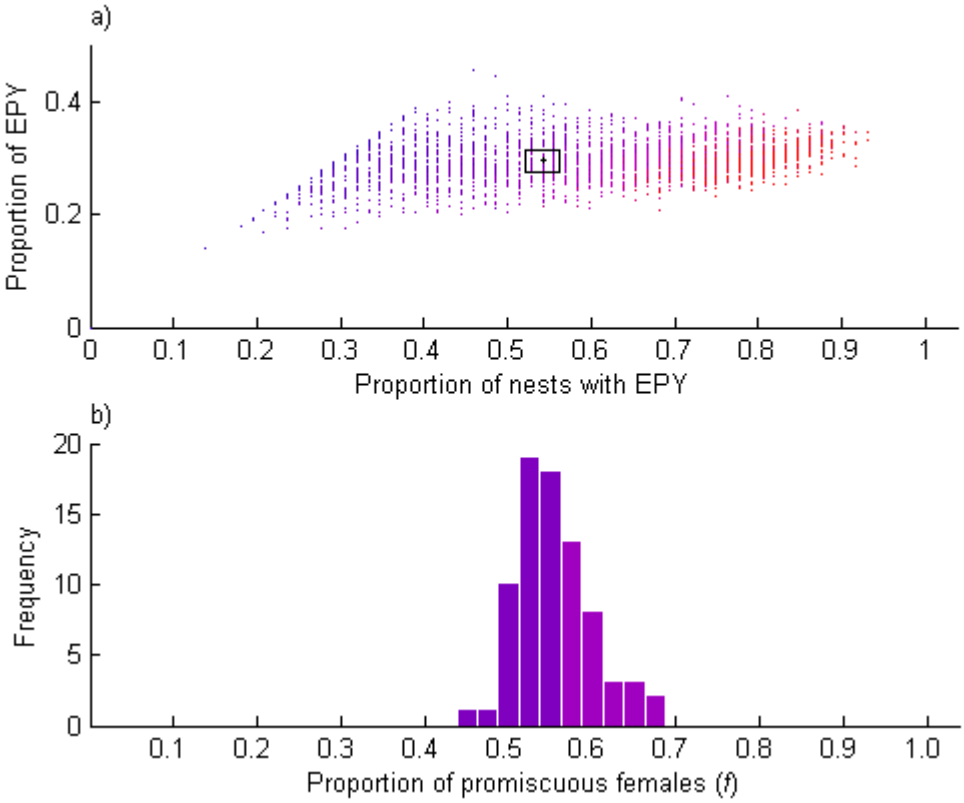


Figure 11. Frequency distribution of *f*. a) the distribution of each simulated population. b) the distribution of all *f*-values that was counted when certain EPP levels fell within the square in Fig. 11a. Parameters: p = 0.295; c = 5; proportion of nests with EPY = 0.542; N_{pop} = 72; N_{num} = 100.

There is a tendency for the *f*-value ranges to become large when the proportion of EPY (*p*) is small (Fig. 12). We have investigated six different species. The species investigated here is reed bunting (*Emberiza schoeniclus*), collared flycatcher (*Ficedula albicollis*), pied flycatcher (*Ficedula hypoleuca*), tree swallow (*Tachycineta bicolor*), great tit (*Parus major*) and blue tit (*Cyanistes caeruleus*). The predicted range of *f*-values is plotted for the corresponding *p*-values taken from the research articles. The points represent the modal *f*-values, i.e., the values that occurred with the highest frequency in the simulations. The reed bunting has medium-sized proportions of EPY (*p*), and the *f*-value ranges are fairly short (Fig. 12a). The collared flycatcher has more variation in the length of it *f*-value ranges (Fig. 12b). The pied flycatcher has small values of *p*, but the *f*-value ranges are fairly long, which suggests that there is a possibility that the proportion of females with EPC behaviour is larger than what the EPP levels suggest (Fig. 12c). The tree swallow has the largest variation in *p*-values, but the *f*-value ranges are relatively short and include large *f*-values (Fig. 12d). The great tit and blue tit have relatively small *p*-values, but the range of *f*-values includes many large *f*-values (Fig. 12e and 12f). This suggests that the proportion of females with EPC behaviour can be much larger than what the proportion of EPY suggests. Some species do not have $f = 1.0$ included in their *f*-value range (Fig. 12a and 12b), which suggests that not all females in a population have EPC behaviour.

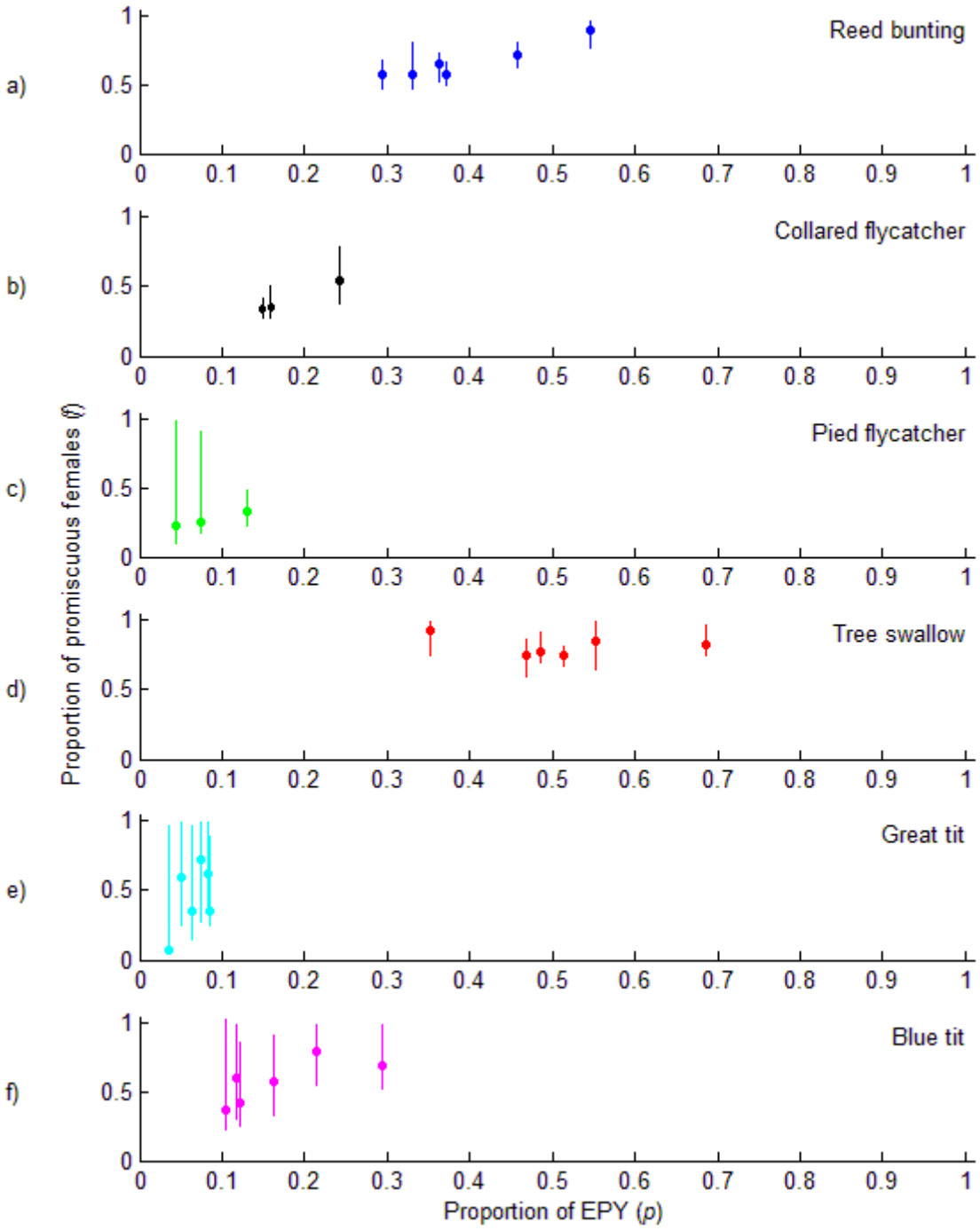


Figure 12. Proportion of promiscuous females versus the proportion of EPY in modelled populations using parameter values from research articles. Each line represents a simulation using values from a real population. The points represent the modal *f*-values found in each simulation. The lines represent the range of the *f*-values that was counted when the values from the research paper fell within a certain error range.

The range of *f*-values become large when the proportion of EPY (*p*) is small (Fig. 13). In the tree swallow, the nest boxes were located in close proximity to each other (ref). The great tit and the pied flycatcher have similar breeding ecologies. There are in total 3 species where $f = 1.0$ is included in the *f*-value ranges. The reed bunting and the collared flycatcher are the only 2 species that do not have $f = 1.0$ included in their *f*-value ranges. The model predicts that it becomes hard to make inferences about the size of the proportion of females (*f*) with EPC behaviour when the proportion of EPY (*p*) is small.

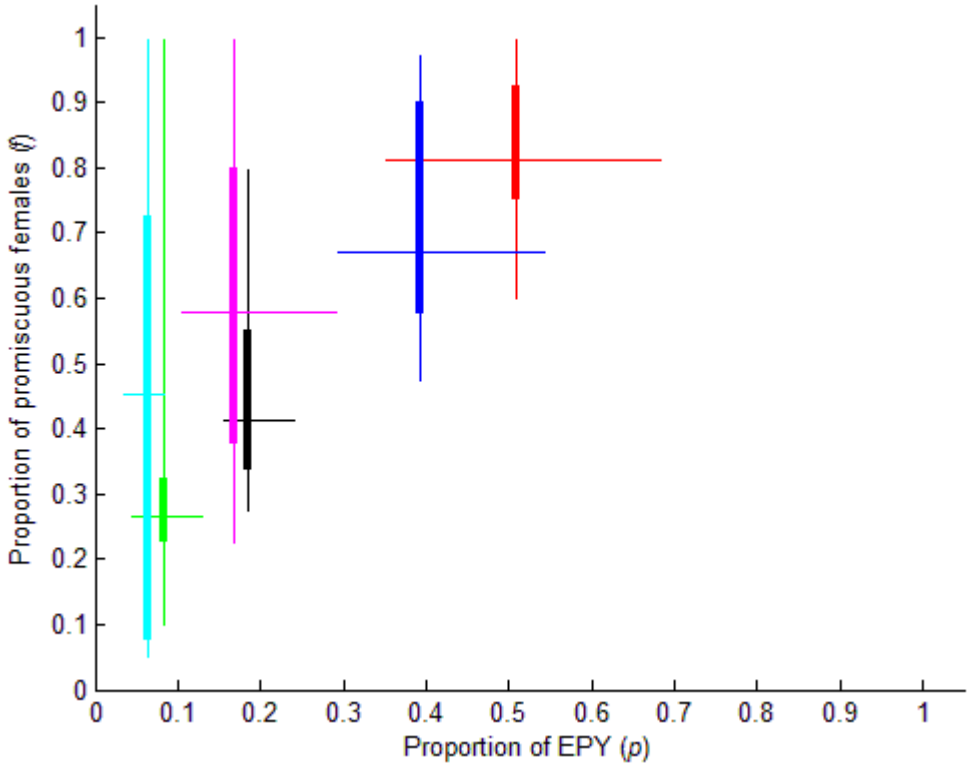


Figure 13: Predicted *f*-value ranges versus *p*-values from research articles. The different colours represent different species, and within each species there are several populations. The ends of each line show the minimum and maximum values for both *f* and *p*-values. Cyan line is great tit, green line is pied flycatcher, magenta line is blue tit, black line is collared flycatcher, blue line is reed bunting and red line is tree swallow. Where the lines intersect is the average modal *f*-value and average proportion of EPY (*p*) of each species. The thick lines represent the range of the modal *f*-values.

4. Discussion

4.1. To observe the 'invisible'

The proportion of nests with EPY does not need to be representative for the proportion of females with EPC behaviour (*f*). Using a mathematical model, we have shown that the proportion of females with EPC behaviour is larger than the measured proportion of nests with EPY for a wide range of common parameter combinations. There is a chance that a large proportion of females in a population sample can have EPC behaviour while the proportion of nests with EPY is small (Fig. 12 and 13). However, in most cases where the population sample size (N_{pop}) is large enough to be representative, we can see that the larger the proportion of females (*f*) with EPC behaviour becomes, the less often these proportions get counted, i.e., larger *f*-values have lower frequencies than smaller *f*-values (as in Fig. 10c and 10d). Sample size is a factor that may have a large impact on the difference between a measured proportion of nests with EPY and a proportion of promiscuous females. It is easier to conclude incorrectly about what the most probable proportion of females (*f*) with EPC behaviour is when the population sample size is small. When the population sample size is smaller than 10 nests, it becomes virtually impossible to create a meaningful distribution of proportions of females (*f*) with EPC behaviour.

When we combine this with a small clutch size, it becomes even harder to conclude correctly. When the clutch size is 1 or 2 eggs, there will be more stochastic variation between each nest. This in turn affects the variation between each population sample. When the proportion of EPY (*p*) in a population sample is small, there is a tendency for the model to count many large *f*-values. A long tail in the distribution will usually manifest itself.

Passerine bird species which breed in urban areas or breed where nest boxes are in close proximity of each other are probably easier to observe than arboreal species or species that do not utilize nest boxes. However, the observation of EPC behaviour in the house sparrow (*Passer domesticus*) was not so easy in a study by Wetton & Parkin (1991). This study showed that communal displays of the house sparrow are rather conspicuous. However, identifying the participants in these displays was extremely difficult due to the number of participants and the briefness of the encounters. The average EPY proportion over the 5 years the study was

conducted, however, was found to be approximately 13.6 %. The average proportion of nests with EPY over the same 5 years was found to be approximately 26.1 %. However, the proportion of females with EPC behaviour was highly underestimated, due to the difficulties of observing EPC behaviour.

Mathematical modelling would suit a difficult situation like the one described above. By using the model, we have shown that most *f*-values (proportions of females with EPC behaviour) of simulated populations are larger than the observed proportion of nests with EPY. However, in most cases, there is a relatively low probability for the *f*-values to be extremely large. Nonetheless, in some cases the *f*-value ranges include $f = 1.0$ and this value has been counted many times relative to the modal *f*-value, which means that there is a relatively high probability for the *f*-values to be extremely large (e.g., see Appendix figures 17, 30, 34 and 36).

Fossøy *et al*. (2006) conducted a field experiment in order to find the true proportion of females with EPC behaviour. In this experiment, WP males were fitted with rubber tubes around their cloaca in order to hinder WP males from fertilizing the eggs of their mates. Since WP males fitted with this device were not able to fertilize the eggs, those eggs that were found in these nests had to be either infertile (due to the female being non-promiscuous) or fertilized by at least one EP male. They found that 87 % of the experimental pairs compared to only 36 % of unmanipulated pairs produce EPY. This difference suggests that females having no EPY may still be sexually promiscuous, which supports our findings.

4.2. Assumptions

A model is not meant to be too complicated but also not too simple. There are certain things that have been excluded from the model.

No mortality

We assume that there is no mortality among females, males and chicks. Females could die before they copulate with an extra-pair male or before they copulate in general. Including mortality biased towards females with EPC behaviour would cause the emergent EPP levels and the emergent proportion of females with EPC behaviour to become smaller. Factors that affect

mortality may be predation and disease. One would intuitively expect that females seeking out EPCs are more exposed to predation since they must move around a lot more in the environment than what they would otherwise do if they did not seek out EPC.

No difference in quality

We assume that all of the offspring that a female can give birth to are of the same quality. Low quality offspring could have a smaller probability of survival compared to high quality offspring. Some of the eggs containing EPY or WPY may not hatch, and this could affect the observed EPP levels in a population to which we compare our model. Most studies usually take DNA samples from live young.

Constant clutch size

We assume that the clutch size is constant within a population. Having a variable clutch size would definitely affect the distributions. The average clutch size in a population could vary on a temporal and/or spatial scale due to variation in resources. If random variation in clutch size between each nest was included, the distributions would most likely vary between each simulation run. Variation in clutch size within a population would also cause the distribution of EPY to vary between each nest.

No decline in the rate of promiscuity as egg-laying continues

When a female starts laying eggs, there is no decline in the frequency of EPC behaviour as the number of eggs in the nest becomes larger. Nonetheless, Magrath *et al*. (2009) found that, on average, most of the EPY were found in the first half of the clutch, meaning that the first few eggs a female laid got fertilized by EP males. Females usually lay one egg per day and start incubating the eggs once all the eggs are laid. One could argue that a female must lie on her eggs when she has started to lay eggs. Therefore, the probability that a female leaves her nest in search for EPC should be small. However, in our model, there are no restrictions on the female to seek out EPCs. Breeding synchrony can also play a role. Breeding synchrony is defined as the level of synchronization of female fertility in a population. This means that if more females in a population are fertile at the same time, then the breeding synchrony is higher. Stutchbury & Morton (1995) argued that EPC behaviour should be more common when females nest

synchronously. They found a strong positive correlation between the degree of synchrony and the rate of extra-pair fertilizations. When there are higher EPC opportunities for extra-pair males, they are expected to cluster in response to these opportunities. In turn, this would provide females with greater opportunities to seek out genetically high quality (extra-pair) males. Thus, if the breeding synchrony is low, the females that are already breeding should encounter extra-pair males at a low rate.

No post-copulatory selection

In our model we assume that there is equal sperm competition and no female cryptic choice. This means that the probability of fertilizing an egg is equal for all males. However, the EPY proportion (*p*/*f*) of each promiscuous female states that over the long run there will be a proportion of all the eggs that are fertilized by at least one extra-pair male and the rest by the within-pair male. How would fertilizations that are dependent on each other affect the distributions of EPY? If there was bias toward females to mate with more EP males after mating with EP males, the emergent proportion of EPY in a population would become larger, and vice versa if females mated with more WP males after mating to such males. Smaller clutch sizes would most likely cause huge variation in the distributions of EPY within and between simulation runs.

Last male sperm precedence is when the last male a female mates with will have an advantage over the preceding males the female mated with (Birkhead *et al*., 1999). We assume in our model that the last male a female mate with before laying an egg is not going to fertilize the egg more readily than the preceding males the female mated with. It has been mentioned that "last male sperm precedence in birds occurs in the laboratory under conditions that are biologically improbable among individuals in the wild" (Birkhead *et al*., 1999). It may therefore be realistic when the model assumes this does not happen. Moreover, we do not know how many times a female mate, with whom or in what order before she lays an egg, simply because each time a random value appear in step two of our model (see "methods" section), it is a fertilization that occurs and not a copulation.

The EPC behaviour is driven by the female

We assume that the EPC behaviour is driven by the female. Both the male and the female can seek out EPCs (Stewart *et al*., 2010), but the female is the party which rejects or accepts the male (Smith, 1988). In this way, it is the female's strategy that determines the outcome of a meeting between the two sexes. One could argue that copulations that are extra-pair are forced extra-pair copulations. This means that a male copulates with a female by force and that EPC behaviour may not actually be a voluntary reproductive strategy used by females. It was traditionally believed that EPCs were a reproductive strategy used by males to potentially augment their own reproductive success (Walsh *et al*., 2006). But is it likely that the female do not seek out EPCs voluntarily and that all EPCs are forced? Many authors have reviewed the potential benefits of EPCs to females (e.g., Jennions & Petrie, 2000; Griffith *et al*., 2002). Since there are many benefits of EPCs to females, one would expect them to participate in EPCs voluntarily in order to gain these benefits.

There are only two variants of female mating behaviour

In our model, the possible mating behaviours a female can possess are either to have EPC behaviour or not. However, those that have EPC behaviour can be promiscuous to a certain extent in the sense that the number of EPY varies between nests. The model has no in-built mechanism or trigger (other than the probability *p*/*f*) which says that a promiscuous female should consistently stop being promiscuous after mating with either an extra-pair or within-pair male, or after a certain time. In the real world, it could be the case that some females must mate many times with an EP male to produce at least one EPY while others may not need to.

Brommer *et al*. (2007) outlined a model that accounts for random variation in the number of EPC across broods in addition to random variation in the number of EPY within broods. They concluded that the observed distribution of EPY becomes less likely to deviate from random when the traditional null model gets extended to also allow for random variation on the level of the brood in the number of EPC. In contrast, other studies that have not incorporated random variation in the number of EPC across broods have concluded the opposite (e.g., Sheldon & Ellegren, 1999; Bouwman *et al*., 2006). The main difference between our study and the study by Brommer *et al*. (2007) is that we do not account for the number of EPCs each promiscuous

female has while the study by Brommer *et al*. (2007) does. In our model, the probability that a promiscuous female has EPY (*p*/*f*) does not depend on the number of EPCs.

There is no difference in fertility between females

We assume that all females are fertile. There is no female that suddenly stops being fertile or is not fertile to begin with. However, infertility rates are often remarkably high in animal populations (Morrow *et al*., 2002). One could, for example, include in the model that females with EPC behaviour are less fertile than non-promiscuous females or vice versa. If a fraction of promiscuous females were infertile, they would not produce EPY or WPY. This would most likely change the proportions of nests with EPY and proportions of EPY.

The probability that eggs hatch is 1 and equal for all eggs

We assume that there are no eggs that do not hatch. All eggs will be fertilized one way or the other. However, eggs usually do not hatch because they do not get fertilized or because of high embryo mortality (Birkhead *et al*., 2008; Forstmeier & Ellegren, 2010). On average, about 15 % of eggs do not hatch (Ihle *et al*., 2012). If the egg is not fertilized, one should not be able to assess the paternity of the egg. In the model, the numbers (0 or 1) that are assigned represent either WPY or EPY, respectively. Thus, the numbers represent what has happened after hatching. The model would probably become too complicated if we included that some eggs do not hatch.

4.3. Variation in female extra-pair mating behaviour

The reason why females vary in their mating behaviour across populations of the same species and between species has been a subject of much debate (Arnold & Owens, 2002). The variation between species can be caused by the phylogenetic differences between them (Arnold & Owens, 2002; Griffith *et al*., 2002), and the variation between populations of the same species can be caused by immediate ecological factors during a breeding season, although the relationship is not always consistent (Griffith *et al*., 2002). However, the ecological factors can also affect the variation between two or more species during a breeding season. The ecological factors during a breeding season can be the breeding density, breeding synchrony, the quality of the males, resource availability, predation or maybe even abiotic factors. In the model, the variation between females is random. Our results suggest that there is a difference in the predicted *f*-value ranges

between species (Fig. 13). There are also differences among populations of the same species (Fig. 12). Some species have large differences in predicted *f*-value ranges.

4.4. Comparison between species

The model predicts that it is hard to say anything about the proportion of females that are promiscuous (*f*) when the EPY proportion in the population (*p*) is small. This means that the range of *f*-values in the frequency distribution becomes large when *p* is small. However, the frequency distribution of *f* should be viewed as a probability distribution. The larger *f*-values are only counted a few times so the probability that a large proportion of females are promiscuous is relatively small. The most promiscuous species we have included in our comparison is the reed bunting. It has a rather large range of *p*-values, but the *p*-values are still much larger than the *p*-values of most of the other species included. Only the tree swallow have similar or larger *p*-values than the reed bunting. 3 species have ranges of *f*-values that include the largest *f*-value that a population can have, which is $f = 1.0$. Only the collared flycatcher has *f*-values that do not reach $f = 1.0$. Great tit and blue tit both have ranges of *f*-values that include $f = 1.0$ or is close to this value many times. The model therefore predicts that it is difficult to be completely sure that the EPP levels measured in a population can precisely represent the proportion of promiscuous females at all times and/or in every location.

There are some species that are similar to each other in both *p*-values and *f*-value ranges. These species are pied flycatcher, great tit and blue tit. These species have similar life histories and breeding ecology. They can breed in nest boxes in garden areas or in forests that are mixed, deciduous or evergreen (e.g., Lubjuhn *et al*., 1999; Charmantier & Blondel, 2003; Lehtonen *et al*., 2009; García-Navas *et al*., 2013). Their breeding densities may vary slightly in time and space. They differ slightly when it comes to clutch size. The blue tit tends to have larger average clutch sizes than the pied flycatcher and the great tit (Table 2). EPC behaviour might be determined by ecological factors rather than the pursuit of ‘good genes’ of EP males as was once thought.

4.5. Implications for field biologists

In this section we will discuss implications for field biologists. Field biologists often encounter time constraints and have lots of things to consider when conducting field experiments or surveys. For example, it is not necessarily easy to choose the correct sample size. Most studies conducted have used sample sizes smaller than 75 nests (e.g., Conrad *et al*., 2001; Hill *et al*., 2010; Suter *et al*., 2009; O'Brien & Dawson, 2007). It would be easy to say that sample sizes ranging from 500 nests to 10 000 nests would be the best because it works well in our model, but using these sample sizes would be rather unrealistic in the real world and would probably expend too much time and resources in a field survey. The model suggests that a sample size between 50 and 150 nests would be useful. However, one needs to take into account that most breeding populations within confined breeding habitats are naturally small, i.e., in the range of 10-25 nests (e.g., Barber *et al*., 1996; Conrad *et al*., 2001; Charmantier & Blondel, 2003). Thus, the use of sample sizes larger than 150 nests is in some cases not feasible. Anything smaller than 50 nests could work, but when we reach sample sizes of lower than 10 nests we must start to question if sample sizes like this can be useful enough to draw conclusions about EPC behaviour. There will be a lot of stochastic variation when samples sizes are very small and one could question whether or not the dynamics that produce the outcome are representative of what happens in nature.

Sample size is not going to be the only thing that will affect the distribution. Clutch size is also going to affect the distribution in combination with sample size. When the clutch size is small, there is going to be a lot of noise in the frequency distribution of f and even worse when the sample size is small. The model predictions show that species with larger clutch sizes could give a better approximation of the population's f-value than species that have smaller clutch sizes. However, sample size is not everything to consider. The model simulates many populations with a given sample size and creates a frequency distribution of f. Hence, another thing to consider is how many times a field survey should be conducted with the same sample size each time. Another problem is how to draw conclusions about EPC behaviour each year based on EPP levels that vary between years.

The error range

We have used an error range in both axis directions for all plots concerning the populations from the research articles (see Appendix section 6.1). This error range is meant to represent error sources that may exist when identifying paternity and these error sources could cause the measured values to differ from what they actually are. Some populations can have larger or smaller proportions of nests with EPY and proportions of EPY. The error ranges that have been chosen are ultimately dependent on sample size and clutch size. But the parameters k_x *and* k_y (see "methods" section) can be used to alter the size of the error ranges. An increase in the value of the parameters would cause the error ranges to increase and thereby include more *f*-values and their frequencies in the frequency distribution of *f* (Fig. 10).

4.6. Conclusion

The model predicts that it is not easy to make inferences about EPC behaviour when the proportion of EPY is small. There is a relatively high probability that all females in a population can have EPC behaviour even though the proportion of EPY is small. We have also shown that clutch size can influence the probability of detecting at least one EPY in a nest. As clutch size increases, the probability of finding at least one EPY increases. There is a chance that the proportion of females that have EPC behaviour can be larger than the proportion of nests with EPY. This means that empiricists must be careful in their conclusions about EPC behaviour when basing their conclusions on EPP levels. However, each female that have EPC behaviour may still vary in the number of EPCs they perform in order to produce one EPY. Females in real populations can have the potential to be promiscuous, but they can be strategic and refrain from EPCs or accept EPCs for a variety of reasons.

5. References

ARNOLD, K. E. & OWENS, I. P. F. 2002. Extra-pair paternity and egg dumping in birds: life history, parental care and the risk of retaliation. *Proceedings: Biological Sciences,* 269**,** 1263-1269.

BARBER, C. A., ROBERTSON, R. J. & BOAG, P. T. 1996. The high frequency of extra-pair paternity in tree swallows is not an artifact of nestboxes. *Behavioral Ecology and Sociobiology,* 38**,** 425-430.

BENNETT, P. M. & OWENS, I. P. F. 2002. *Evolutionary Ecology of Birds: Life Histories, Mating Systems and Extinction,* Oxford, Oxford University Press.

BIRKHEAD, T. R., HALL, J., SCHUT, E. & HEMMINGS, N. 2008. Unhatched eggs: methods for discriminating between infertility and early embryo mortality. *Ibis,* 150**,** 508-517.

BIRKHEAD, T. R., MARTINEZ, J. G., BURKE, T. & FROMAN, D. P. 1999. Sperm mobility determines the outcome of sperm competition in the domestic fowl. *Proceedings: Biological Sciences,* 266**,** 1759-1764.

BIRKHEAD, T. R. & MØLLER, A. P. 1992. *Sperm Competition in Birds: evolutionary causes and consequences,* New York, Academic Press.

BLAKEY, J. K. 1994. Genetic evidence for extra-pair fertilizations in a monogamous passerine, the great tit *Parus major*. *Ibis,* 136**,** 457-462.

BOLLINGER, E. K. & GAVIN, T. A. 1991. Patterns of extra-pair fertilizations in bobolinks. *Behavioral Ecology and Sociobiology,* 29**,** 1-7.

BOUWMAN, K. M., BURKE, T. & KOMDEUR, J. A. N. 2006. How female reed buntings benefit from extra-pair mating behaviour: testing hypotheses through patterns of paternity in sequential broods. *Molecular Ecology,* 15**,** 2589-2600.

BRADLEY, B. J. & VIGILANT, L. 2002. False alleles derived from microbial DNA pose a potential source of error in microsatellite genotyping of DNA from faeces. *Molecular Ecology Notes,* 2**,** 602-605.

BRISKIE, J. V., NAUGLER, C. T. & LEECH, S. M. 1994. Begging intensity of nestling birds varies with sibling relatedness. *Proceedings: Biological Sciences,* 258**,** 73-78.

BROMMER, J. E., ALHO, J. S., BIARD, C., CHAPMAN, J. R., CHARMANTIER, A., DREISS, A., HARTLEY, I. R., HJERNQUIST, M. B., KEMPENAERS, B., KOMDEUR, J., LAAKSONEN, T., LEHTONEN, P. K., LUBJUHN, T., PATRICK, S. C., ROSIVALL, B., TINBERGEN, J. M., VELDE, M., OERS, K., WILK, T. & WINKEL, W. 2010. Passerine extrapair mating dynamics: a bayesian modeling approach comparing four species. *The American Naturalist,* 176**,** 178-187.

BROMMER, J. E., KORSTEN, P., BOUWMAN, K. M., BERG, M. L. & KOMDEUR, J. 2007. Is extrapair mating random? On the probability distribution of extrapair young in avian broods. *Behavioral Ecology,* 18**,** 895-904.

CHARMANTIER, A. & BLONDEL, J. 2003. A contrast in extra-pair paternity levels on mainland and island populations of mediterranean blue tits. *Ethology,* 109**,** 351-363.

CLUTTON-BROCK, T. H. & ISVARAN, K. 2006. Paternity loss in contrasting mammalian societies. *Biology Letters,* 2**,** 513-516.

CLUTTON-BROCK, T. H. & PARKER, G. A. 1995. Sexual coercion in animal societies. *Animal Behaviour,* 49**,** 1345-1365.

CONRAD, K. F., JOHNSTON, P. V., CROSSMAN, C., KEMPENAERS, B., ROBERTSON, R. J., WHEELWRIGHT, N. T. & BOAG, P. T. 2001. High levels of extra-pair paternity in an isolated, low-density, island population of tree swallows (*Tachycineta bicolor*). *Molecular Ecology,* 10**,** 1301-1308.

CORDERO, C. 1995. Ejaculate substances that affect female insect reproductive physiology and behavior: honest or arbitrary traits? *Journal of Theoretical Biology,* 174**,** 453-461.

CORNWALLIS, C. K., WEST, S. A., DAVIS, K. E. & GRIFFIN, A. S. 2010. Promiscuity and the evolutionary transition to complex societies. *Nature,* 466**,** 969-972.

DALZIELL, A. H. & COCKBURN, A. 2008. Dawn song in superb fairy-wrens: a bird that seeks extrapair copulations during the dawn chorus. *Animal Behaviour,* 75**,** 489-500.

DAVISON, A. & CHIBA, S. 2003. Laboratory temperature variation is a previously unrecognized source of genotyping error during capillary electrophoresis. *Molecular Ecology Notes,* 3**,** 321-323.

DEAN, R., CORNWALLIS, C. K., LØVLIE, H., WORLEY, K., RICHARDSON, D. S. & PIZZARI, T. 2010. Male reproductive senescence causes potential for sexual conflict over mating. *Current Biology,* 20**,** 1192-1196.

DIXON, A., ROSS, D., O'MALLEY, S. L. C. & BURKE, T. 1994. Paternal investment inversely related to degree of extra-pair paternity in the reed bunting. *Nature,* 371**,** 698-700.

DUNN, P. O. & LIFJELD, J. T. 1994. Can extra-pair copulations be used to predict extra-pair paternity in birds? *Animal Behaviour,* 47**,** 983-985.

DUNN, P. O., ROBERTSON, R. J., MICHAUD-FREEMAN, D. & BOAG, P. T. 1994. Extra-pair paternity in tree swallows: why do females mate with more than one male? *Behavioral Ecology and Sociobiology,* 35**,** 273-281.

ELIASSEN, S. & JØRGENSEN, C. 2014. Extra-pair mating and evolution of cooperative neighbourhoods. *PLoS ONE,* 9**,** e99878.

FEDY, B. C. & STUTCHBURY, B. J. M. 2005. Territory defence in tropical birds: are females as aggressive as males? *Behavioral Ecology and Sociobiology,* 58**,** 414-422.

FERNANDO, P., EVANS, B. J., MORALES, J. C. & MELNICK, D. J. 2001. Electrophoresis artefacts - a previously unrecognized cause of error in microsatellite analysis. *Molecular Ecology Notes,* 1**,** 325-328.

FORSTMEIER, W. & ELLEGREN, H. 2010. Trisomy and triploidy are sources of embryo mortality in the zebra finch. *Proceedings: Biological Sciences,* 277**,** 2655-2660.

FORSTMEIER, W., NAKAGAWA, S., GRIFFITH, S. C. & KEMPENAERS, B. 2014. Female extra-pair mating: adaptation or genetic constraint? *Trends in Ecology & Evolution,* 29**,** 456-464.

FOSSØY, F., JOHNSEN, A. & LIFJELD, J. T. 2006. Evidence of obligate female promiscuity in a socially monogamous passerine. *Behavioral Ecology and Sociobiology,* 60**,** 255-259.

GAGNEUX, P., BOESCH, C. & WOODRUFF, D. S. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology,* 6**,** 861-868.

GARCÍA-NAVAS, V., ORTEGO, J., FERRER, E. S. & SANZ, J. J. 2013. Feathers, suspicions, and infidelities: an experimental study on parental care and certainty of paternity in the blue tit. *Biological Journal of the Linnean Society,* 109**,** 552-561.

GERLOFF, U., SCHLÖTTERER, C., RASSMANN, K., RAMBOLD, I., HOHMANN, G., FRUTH, B. & TAUTZ, D. 1995. Amplification of hypervariable simple sequence repeats (microsatellites) from excremental DNA of wild living bonobos (*Pan paniscus*). *Molecular Ecology,* 4**,** 515-518.

GIBBONS, W. J. & ANDREWS, K. M. 2004. PIT tagging: simple technology at its best. *BioScience,* 54**,** 447-454.

GOOSSENS, B., WAITS, L. P. & TABERLET, P. 1998. Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology,* 7**,** 1237-1241.

GRIFFITH, S. C. 2007. The evolution of infidelity in socially monogamous passerines: neglected components of direct and indirect selection. *The American Naturalist,* 169**,** 274-281.

GRIFFITH, S. C., OWENS, I. P. F. & THUMAN, K. A. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology,* 11**,** 2195-2212.

HASSON, O. & STONE, L. 2009. Male infertility, female fertility and extrapair copulations. *Biological Reviews,* 84**,** 225-244.

HILBORN, R. A. Y. & MANGEL, M. 1997. *The Ecological Detective: confronting models with data (MPB-28)*, Princeton University Press.

HILL, C. E., GJERDRUM, C. & ELPHICK, C. S. 2010. Extreme levels of multiple mating characterize the mating system of the saltmarsh sparrow (*Ammodramus caudacutus*). *The Auk,* 127**,** 300-307.

HOFFMAN, J. I. & AMOS, W. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology,* 14**,** 599-612.

HOUTMAN, A. M. 1992. Female zebra finches choose extra-pair copulations with genetically attractive males. *Proceedings: Biological Sciences,* 249**,** 3-6.

IHLE, M., KEMPENAERS, B. & FORSTMEIER, W. 2013. Does hatching failure breed infidelity? *Behavioral Ecology,* 24**,** 119-127.

JENNIONS, M. D. & PETRIE, M. 1997. Variation in mate choice and mating preferences: a review of causes and consequences. *Biological Reviews,* 72**,** 283-327.

JENNIONS, M. D. & PETRIE, M. 2000. Why do females mate multiply? A review of the genetic benefits. *Biological Reviews,* 75**,** 21-64.

JETZ, W., SEKERCIOGLU, C. H. & BÖHNING-GAESE, K. 2008. The worldwide variation in avian clutch size across species and space. *PLoS Biol,* 6**,** e303.

JOHNSON, S. L. & GEMMELL, N. J. 2012. Are old males still good males and can females tell the difference? *BioEssays,* 34**,** 609-619.

JONES, A. G. & ARDREN, W. R. 2003. Methods of parentage analysis in natural populations. *Molecular Ecology,* 12**,** 2511-2523.

KLEIMAN, D. G. 1977. Monogamy in mammals. *The Quarterly Review of Biology,* 52**,** 39-69.

KLEVEN, O. & LIFJELD, J. T. 2005. No evidence for increased offspring heterozygosity from extrapair mating in the reed bunting (*Emberiza schoeniclus*). *Behavioral Ecology,* 16**,** 561-565.

KRIST, M., NÁDVORNÍK, P., UVÍROVÁ, L. & BUREŠ, S. 2005. Paternity covaries with laying and hatching order in the collared flycatcher *Ficedula albicollis*. *Behavioral Ecology and Sociobiology,* 59**,** 6-11.

KROKENE, C., RIGSTAD, K., DALE, M. & LIFJELD, J. T. 1998. The function of extrapair paternity in blue tits and great tits: good genes or fertility insurance? *Behavioral Ecology,* 9**,** 649-656.

LACK, D. 1947. The significance of clutch size, parts 1 and 2. *Ibis,* 89**,** 302-352.

LACK, D. 1968. *Ecological Adaptations for Breeding in Birds,* London, Methuen Ltd.

LEHTONEN, P. K., PRIMMER, C. R. & LAAKSONEN, T. 2009. Different traits affect gain of extrapair paternity and loss of paternity in the pied flycatcher, *Ficedula hypoleuca*. *Animal Behaviour,* 77**,** 1103-1110.

LODISH, H., BERK, A., KAISER, C. A., KRIEGER, M., SCOTT, M. P., BRETSCHER, A., PLOEGH, H. & MATSUDAIRA, P. 2008. *Molecular Cell Biology,* New York, Freeman.

LUBJUHN, T., STROHBACH, S., BRÜN, J., GERKEN, T. & EPPLEN, J. T. 1999. Extra-pair paternity in great tits (*Parus major*): a long term study. *Behaviour,* 136**,** 1157-1172.

LUKAS, D. & CLUTTON-BROCK, T. H. 2013. The evolution of social monogamy in mammals. *Science,* 341**,** 526-530.

MAGRATH, M. J. L., VEDDER, O., VAN DER VELDE, M. & KOMDEUR, J. 2009. Maternal effects contribute to the superior performance of extra-pair offspring. *Current Biology,* 19**,** 792-797.

MAUCK, R. A., MARSCHALL, E. A. & PARKER, P. G. 1999. Adult survival and imperfect assessment of parentage: effects on male parenting decisions. *The American Naturalist,* 154**,** 99-109.

MAYER, C. & PASINELLI, G. 2013. New support for an old hypothesis: density affects extra-pair paternity. *Ecology and Evolution,* 3**,** 694-705.

MORENO, J., MARTÍNEZ, J.-G., MORALES, J., LOBATO, E., MERINO, S., TOMÁS, G., VÁSQUEZ, R. A., MÖSTL, E. & OSORNO, J. L. 2010. Paternity loss in relation to male age, territorial behaviour and stress in the pied flycatcher. *Ethology,* 116**,** 76-84.

MORENO, J., MARTÍNEZ, J. G., GONZÁLEZ-BRAOJOS, S., RUIZ-DE-CASTAÑEDA, R., CANTARERO, A. & SÁNCHEZ-TÓJAR, A. 2013. Extra-pair matings, context-dependence and offspring quality: a brood manipulation experiment in pied flycatchers. *Behaviour,* 150**,** 359-380.

MORROW, E. H., ARNQVIST, G. & PITCHER, T. E. 2002. The evolution of infertility: does hatching rate in birds coevolve with female polyandry? *Journal of Evolutionary Biology,* 15**,** 702-709.

MØLLER, A. P. & NINNI, P. 1998. Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. *Behavioral Ecology and Sociobiology,* 43**,** 345-358.

NAKAMURA, M. 1998a. Multiple mating and cooperative breeding in polygynandrous alpine accentors. I. Competition among females. *Animal Behaviour,* 55**,** 259-275.

NAKAMURA, M. 1998b. Multiple mating and cooperative breeding in polygynandrous alpine accentors. II. Male mating tactics. *Animal Behaviour,* 55**,** 277-289.

NAVIDI, W., ARNHEIM, N. & WATERMAN, M. S. 1992. A multiple-tubes approach for accurate genotyping of very small DNA samples by using PCR: statistical considerations. *The American Journal of Human Genetics,* 50**,** 347-359.

O'BRIEN, E. L. & DAWSON, R. D. 2007. Context-dependent genetic benefits of extra-pair mate choice in a socially monogamous passerine. *Behavioral Ecology and Sociobiology,* 61**,** 775-782.

PETRIE, M., HALL, M., HALLIDAY, T., BUDGEY, H. & PIERPOINT, C. 1992. Multiple mating in a lekking bird: why do peahens mate with more than one male and with the same male more than once? *Behavioral Ecology and Sociobiology,* 31**,** 349-358.

PRENTICE, E. F., FLAGG, T. A. & MCCUTCHEON, C. S. 1990. Feasibility of using implantable passive integrated transponder (PIT) tags in salmonids. *American Fisheries Society Symposium,* 7**,** 317-322.

ROSIVALL, B., SZÖLLŐSI, E., HASSELQUIST, D. & TÖRÖK, J. 2009. Effects of extrapair paternity and sex on nestling growth and condition in the collared flycatcher, *Ficedula albicollis*. *Animal Behaviour,* 77**,** 611-617.

ROUSSEL, J. M., HARO, A. & CUNJAK, R. A. 2000. Field test of a new method for tracking small fishes in shallow rivers using passive integrated transponder (PIT) technology. *Canadian Journal of Fisheries and Aquatic Sciences,* 57**,** 1326-1329.

ROWE, L., ARNQVIST, G., SIH, A. & KRUPA, J. J. 1994. Sexual conflict and the evolutionary ecology of mating patterns: water striders as a model system. *Trends in Ecology & Evolution,* 9**,** 289-293.

SHELDON, B. C. 1993. Sexually transmitted disease in birds: occurrence and evolutionary significance. *Philosophical Transactions: Biological Sciences,* 339**,** 491-497.

SHELDON, B. C. & ELLEGREN, H. 1999. Sexual selection resulting from extrapair paternity in collared flycatchers. *Animal Behaviour,* 57**,** 285-298.

SMITH, S. M. 1988. Extra-pair copulations in black-capped chickadees: the role of the female. *Behaviour,* 107**,** 15-23.

STENMARK, G., SLAGSVOLD, T. & LIFJELD, J. T. 1988. Polygyny in the pied flycatcher, *Ficedula hypoleuca*: a test of the deception hypothesis. *Animal Behaviour,* 36**,** 1646-1657.

STEWART, S. L. M., WESTNEAT, D. F. & RITCHISON, G. 2010. Extra-pair paternity in eastern bluebirds: effects of manipulated density and natural patterns of breeding synchrony. *Behavioral Ecology and Sociobiology,* 64**,** 463-473.

STROHBACH, S., CURIO, E., BATHEN, A., EPPLEN, J. & LUBJUHN, T. 1998. Extrapair paternity in the great tit (*Parus major*): a test of the “good genes” hypothesis. *Behavioral Ecology,* 9**,** 388-396.

STUTCHBURY, B. J. M. 1998. Extra-pair mating effort of male hooded warblers, *Wilsonia citrina*. *Animal Behaviour,* 55**,** 553-561.

STUTCHBURY, B. J. M. & MORTON, E. S. 1995. The effect of breeding synchrony on extra-pair mating systems in songbirds. *Behaviour,* 132**,** 675-690.

SUTER, S. M., BIELAŃSKA, J., RÖTHLIN-SPILLMANN, S., STRAMBINI, L. & MEYER, D. R. 2009. The cost of infidelity to female reed buntings. *Behavioral Ecology,* 20**,** 601-608.

SZULKIN, M., STOPHER, K. V., PEMBERTON, J. M. & REID, J. M. 2013. Inbreeding avoidance, tolerance, or preference in animals? *Trends in Ecology & Evolution,* 28**,** 205-211.

TABERLET, P., GRIFFIN, S., GOOSSENS, B., QUESTIAU, S., MANCEAU, V., ESCARAVAGE, N., WAITS, L. P. & BOUVET, J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research,* 24**,** 3189-3194.

THORNHILL, R. & ALCOCK, J. 1983. *The Evolution of Insect Mating Systems,* Cambridge, Harvard University Press.

TRIVERS, R. 1972. Parental Investment and Sexual Selection. *In:* CAMPBELL, B. (ed.) *Sexual Selection and the Descent of Man.* Chicago: Aldine Publishing Company.

VALERA, F., HOI, H. & KRIŠTÍN, A. 2003. Male shrikes punish unfaithful females. *Behavioral Ecology,* 14**,** 403-408.

VERBOVEN, N. & MATEMAN, A. C. 1997. Low frequency of extra-pair fertilizations in the great tit *Parus major* revealed by DNA fingerprinting. *Journal of Avian Biology,* 28**,** 231-239.

WALSH, C. J., WILHELM, S. I., CAMERON-MACMILLAN, M. L. & STOREY, A. E. 2006. Extra-pair copulations in common murres I: a mate attraction strategy? *Behaviour,* 143**,** 1241-1262.

WALSH, P. S., ERLICH, H. A. & HIGUCHI, R. 1992. Preferential PCR amplification of alleles: mechanisms and solutions. *Genome Research,* 1**,** 241-250.

WEDELL, N. 1997. Ejaculate size in bushcrickets: the importance of being large. *Journal of Evolutionary Biology,* 10**,** 315-325.

WESTNEAT, D. F., SHERMAN, P. W. & MORTON, M. L. 1990. The ecology and evolution of extra-pair copulations in birds. *Current Ornithology,* 7**,** 331-369.

WETTON, J. H. & PARKIN, D. T. 1991. An association between fertility and cuckoldry in the house sparrow, *Passer domesticus*. *Proceedings: Biological Sciences,* 245**,** 227-233.

WETTON, J. H., PARKIN, D. T. & CARTER, R. E. 1992. The use of genetic markers for parentage analysis in *Passer domesticus* (house sparrows). *Heredity,* 69**,** 243-254.

WHITTINGHAM, L. A., DUNN, P. O. & STAPLETON, M. K. 2006. Repeatability of extra-pair mating in tree swallows. *Molecular Ecology,* 15**,** 841-849.

WILK, T., CICHOŃ, M. & WOLFF, K. 2008. Lack of evidence for improved immune response of extra-pair nestlings in collared flycatcher *Ficedula albicollis*. *Journal of Avian Biology,* 39**,** 546-552.

ZEH, J. A. & ZEH, D. W. 1996. The evolution of polyandry I: intragenomic conflict and genetic incompatibility. *Proceedings: Biological Sciences,* 263**,** 1711-1717.

6. Appendix

6.1. Supplementary figures

The following figures are based on the data from Table 2. The order of the figures is the same as the order of the populations in Table 1 and 2. All parameter values for each population are found in Table 2.

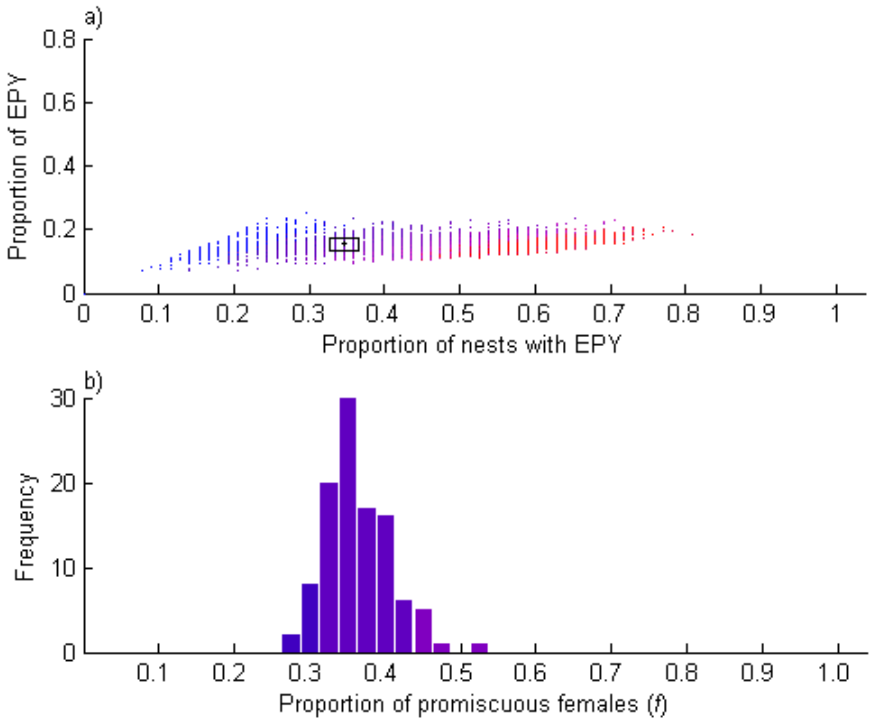


Figure 14: Population 1 from Wilk *et al.*, 2008.

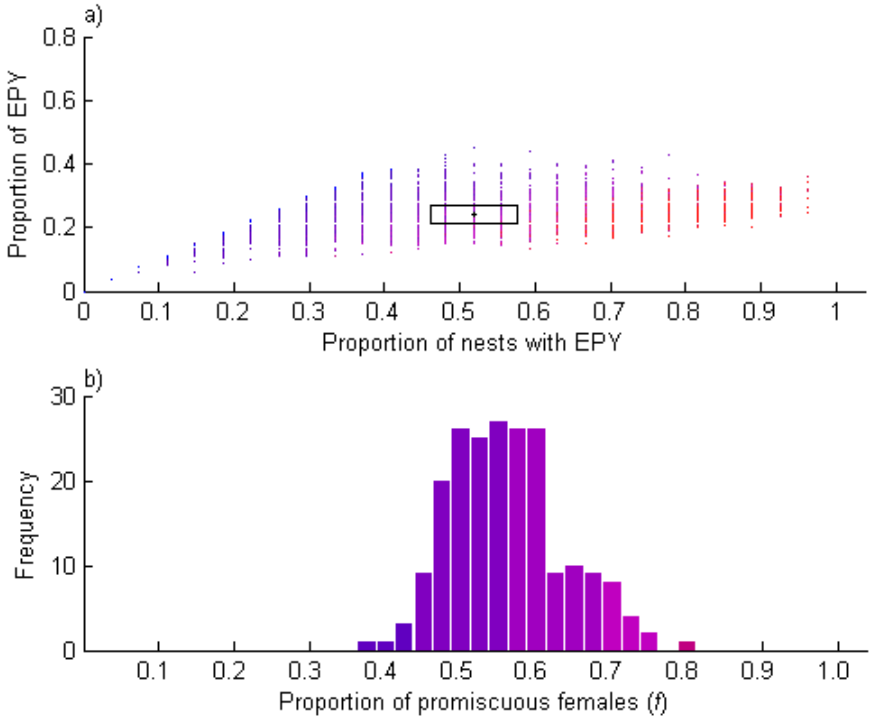


Figure 15: Population 2 from Krist *et al.*, 2005.

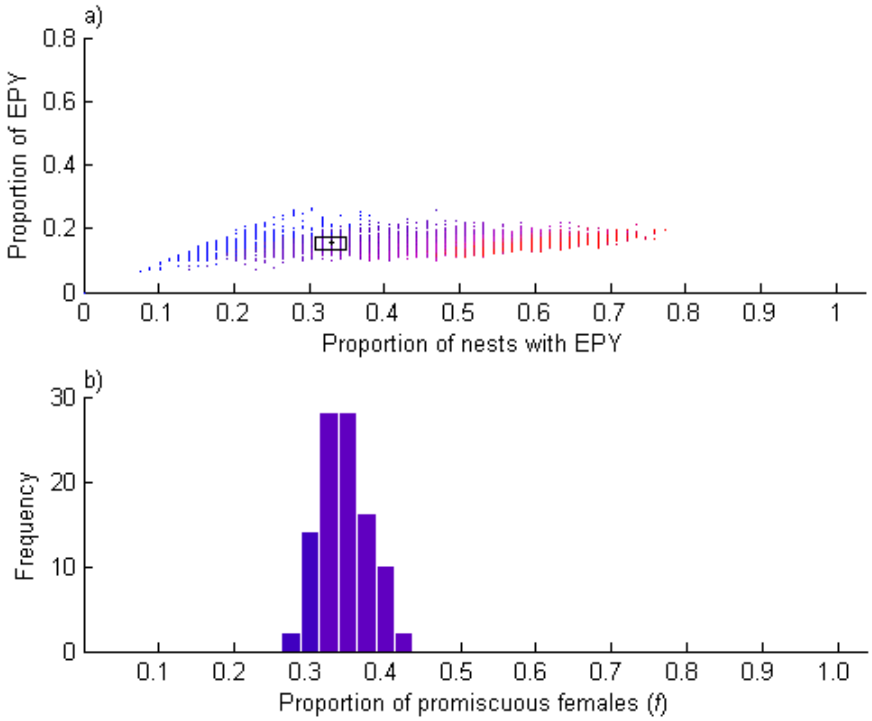


Figure 16: Population 3 from Sheldon & Ellegren, 1999.

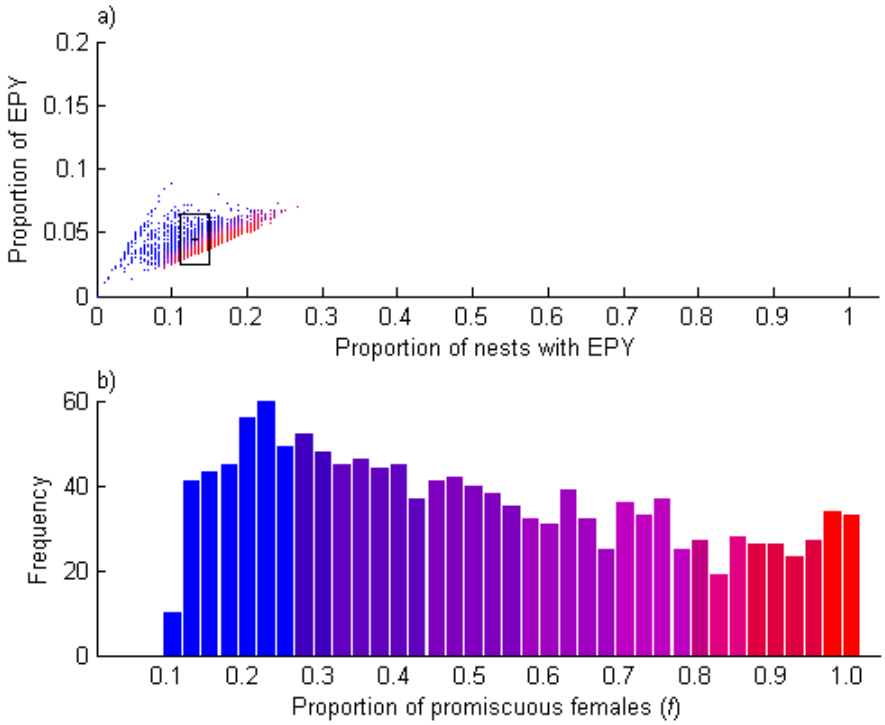


Figure 17: Population 4 from Lehtonen *et al.*, 2009.

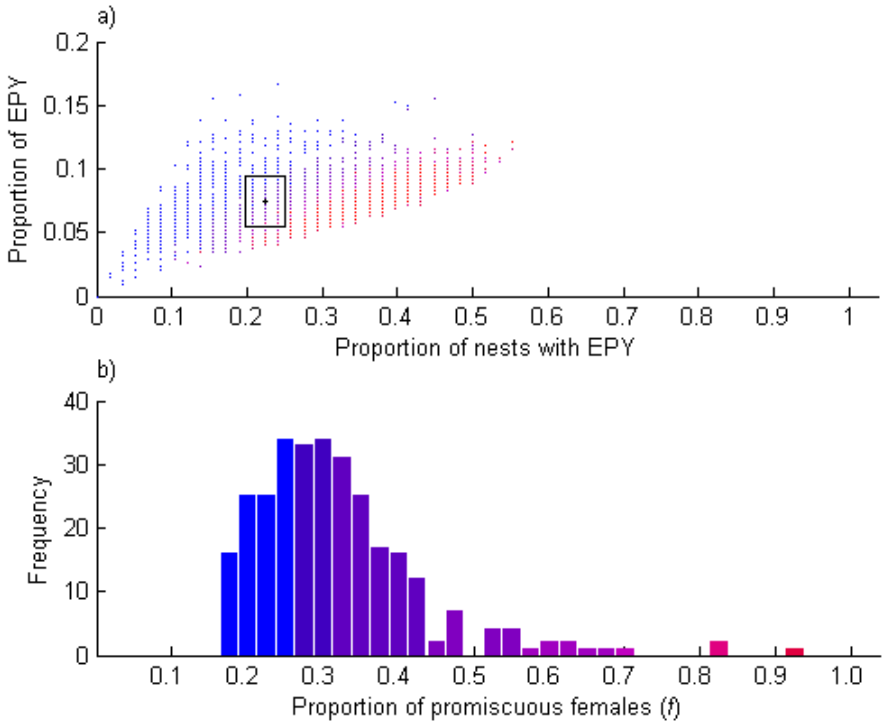


Figure 18: Population 5 from Moreno *et al.*, 2010.

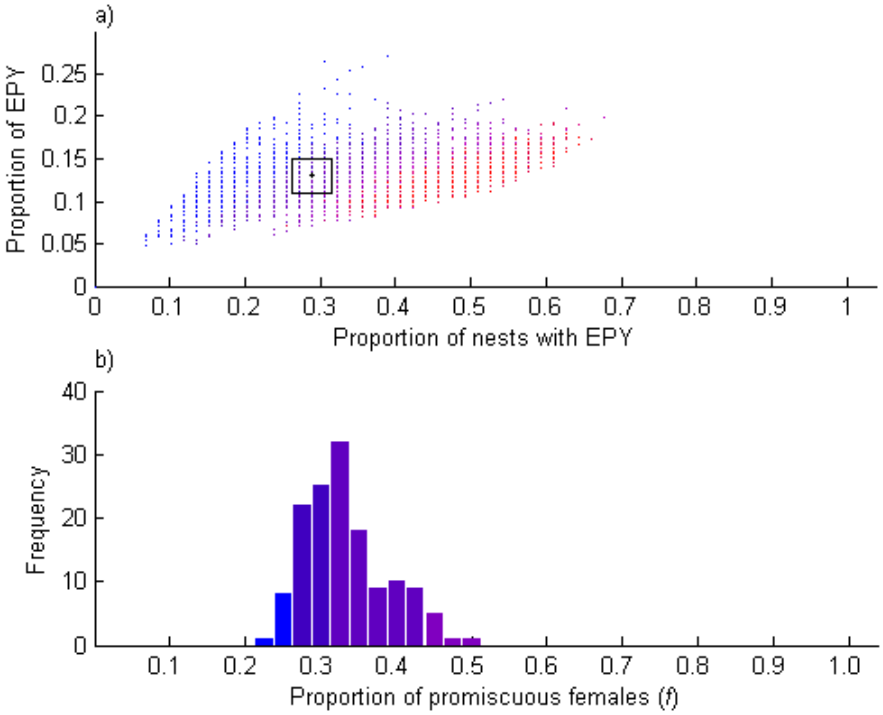


Figure 19: Population 6 from Moreno *et al.*, 2013.

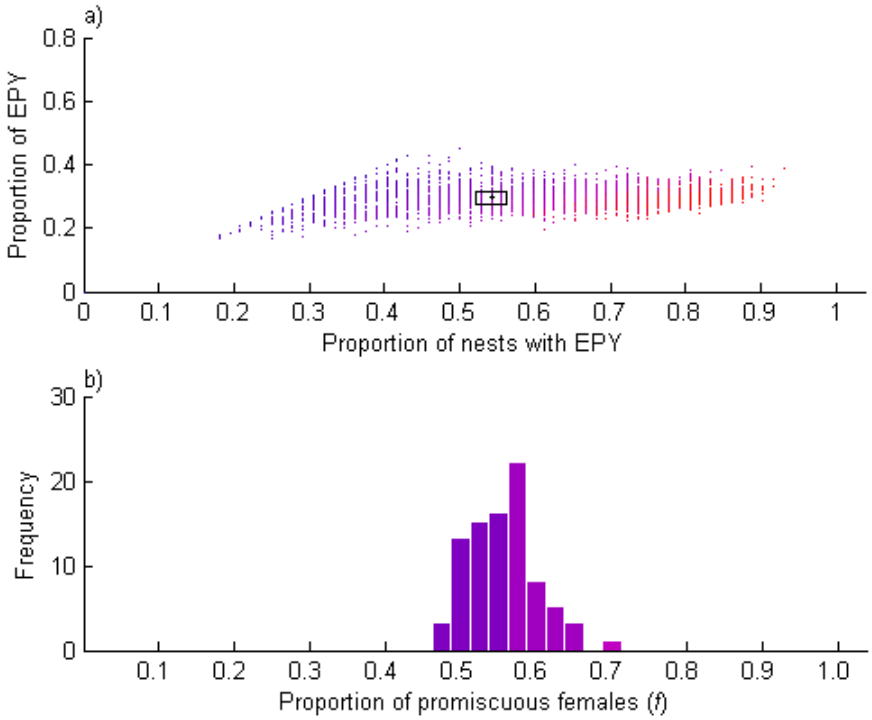


Figure 20: Population 7 from Kleven & Lifjeld, 2005.

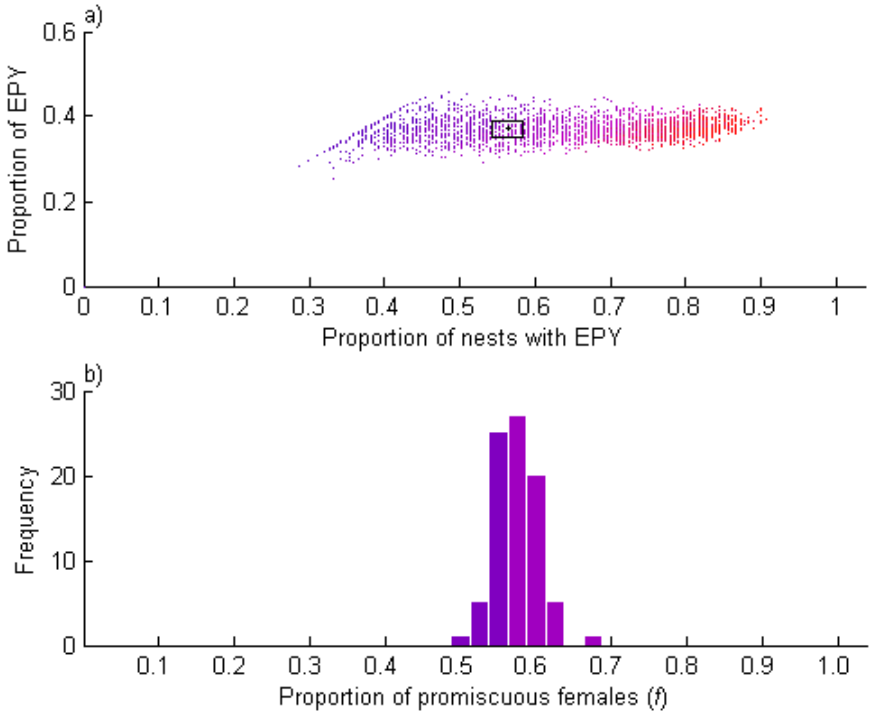


Figure 21: Population 8 from Mayer & Pasinelli, 2013.

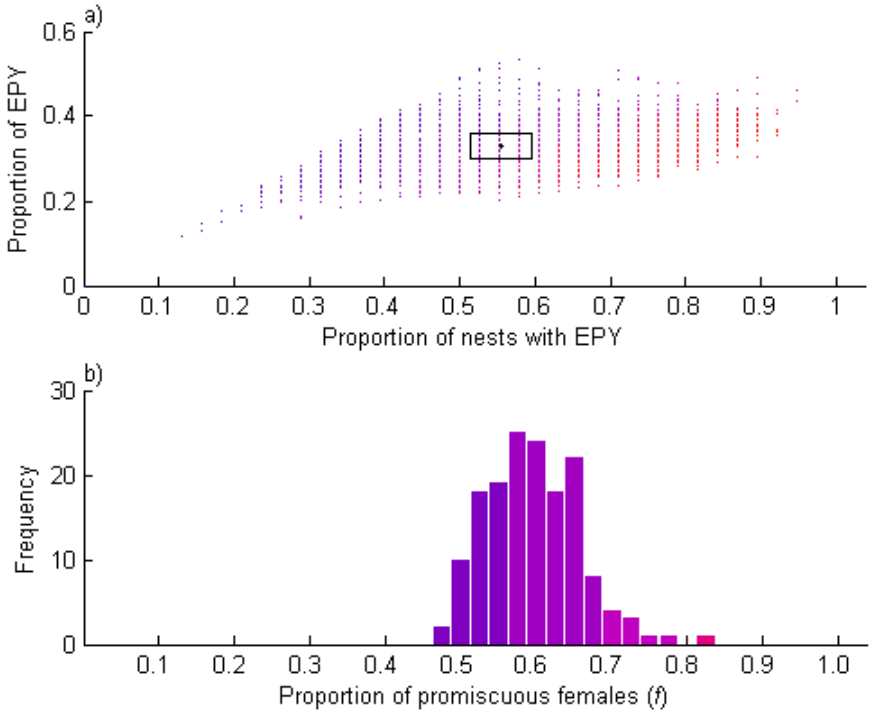


Figure 22: Population 9 from Suter *et al.*, 2009.

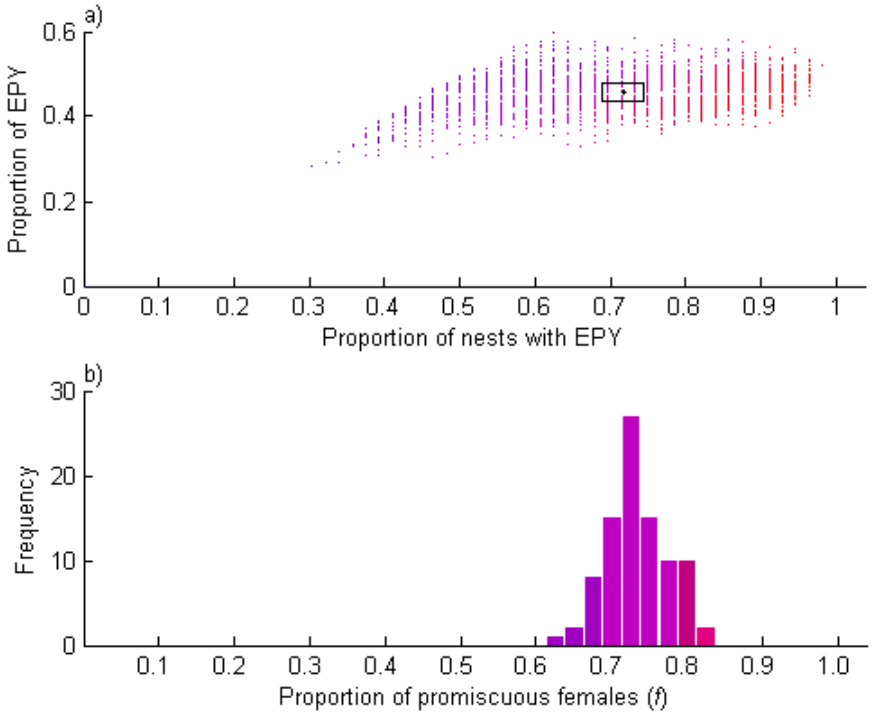


Figure 23: Population 10 from Suter *et al.*, 2009.

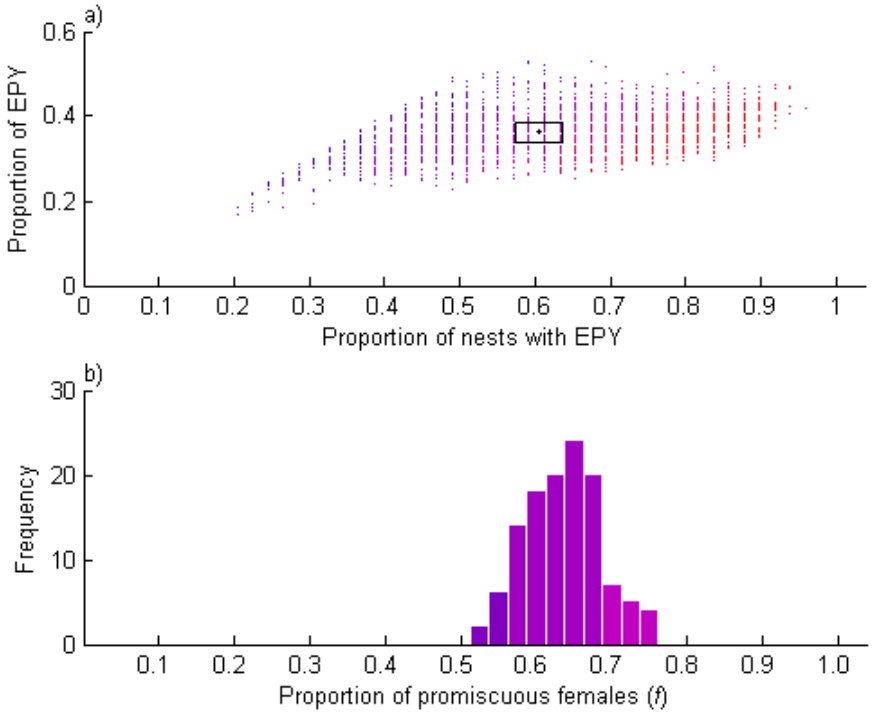


Figure 24: Population 11 from Suter *et al.*, 2009.

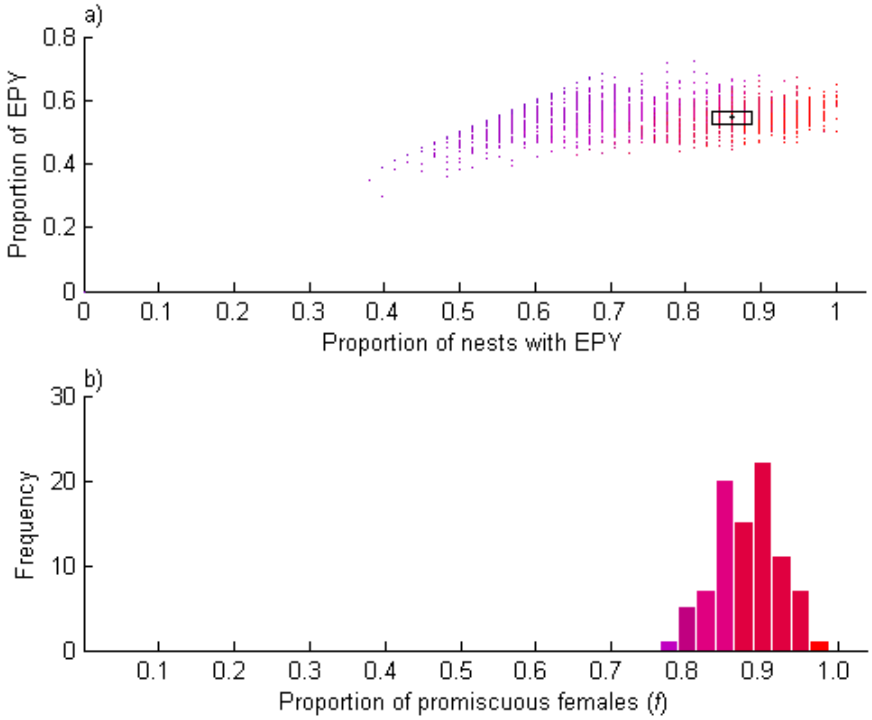


Figure 25: Population 12 from Dixon *et al.*, 1994.

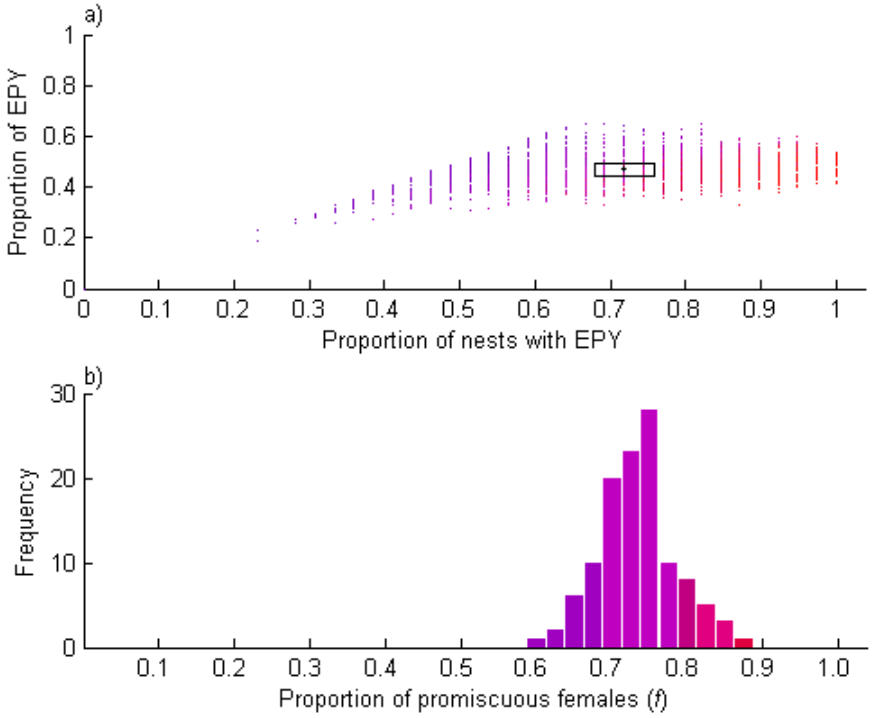


Figure 26: Population 13 from Dunn *et al.*, 1994.

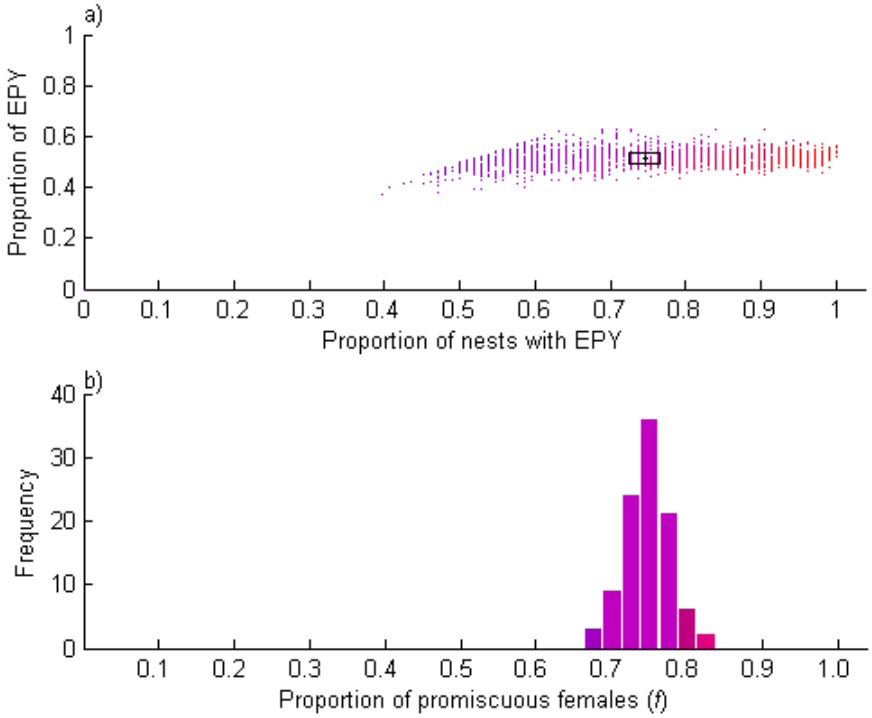


Figure 27: Population 14 from Conrad *et al.*, 2001.

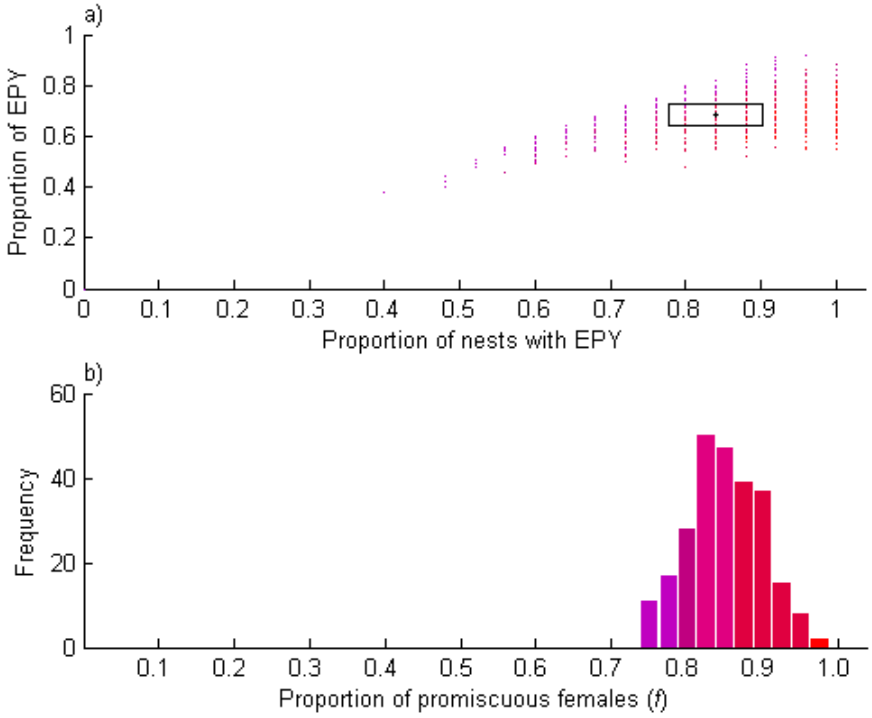


Figure 28: Population 15 from Barber *et al.*, 1996.

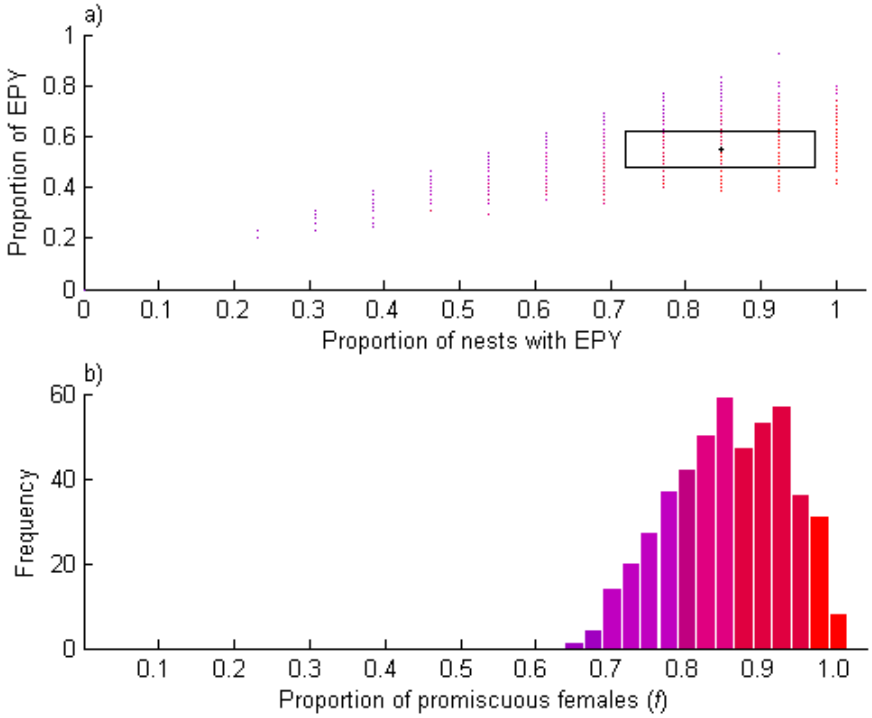


Figure 29: Population 16 from Conrad *et al.*, 2001.

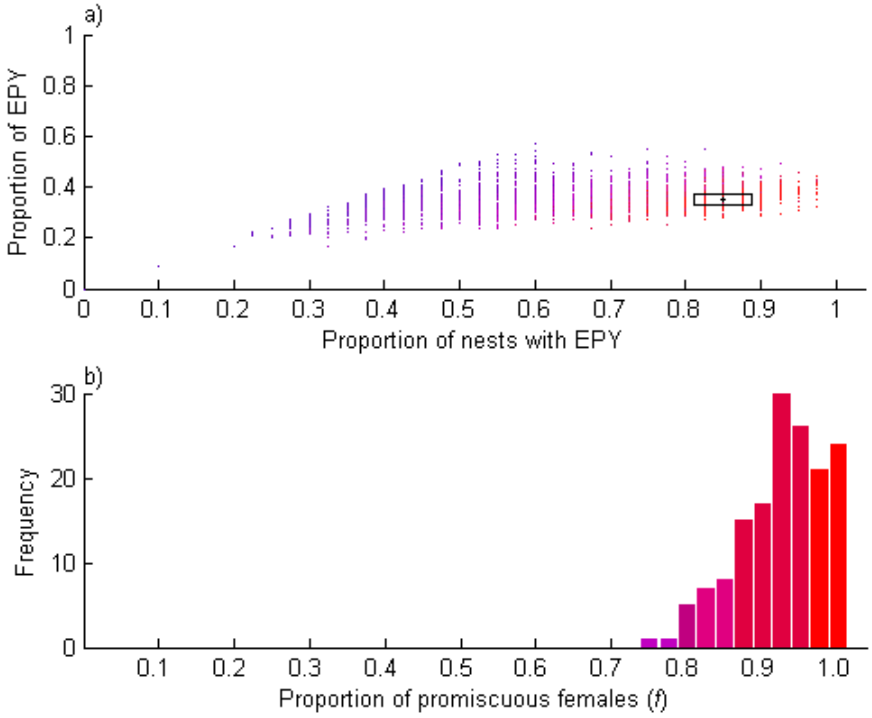


Figure 30: Population 17 from O'Brien & Dawson, 2007.

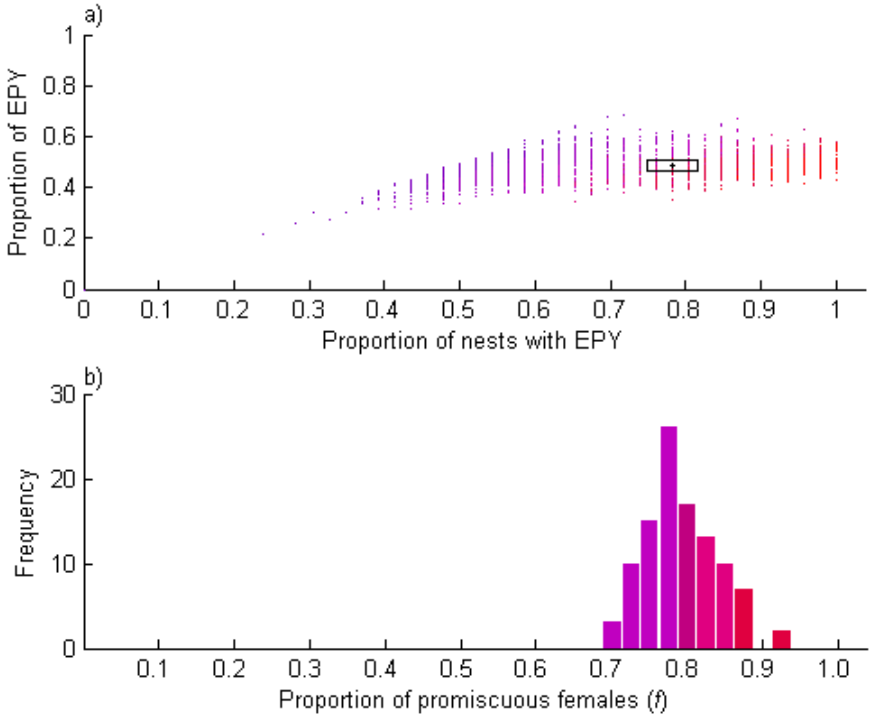


Figure 31: Population 18 from Whittingham *et al.*, 2006.

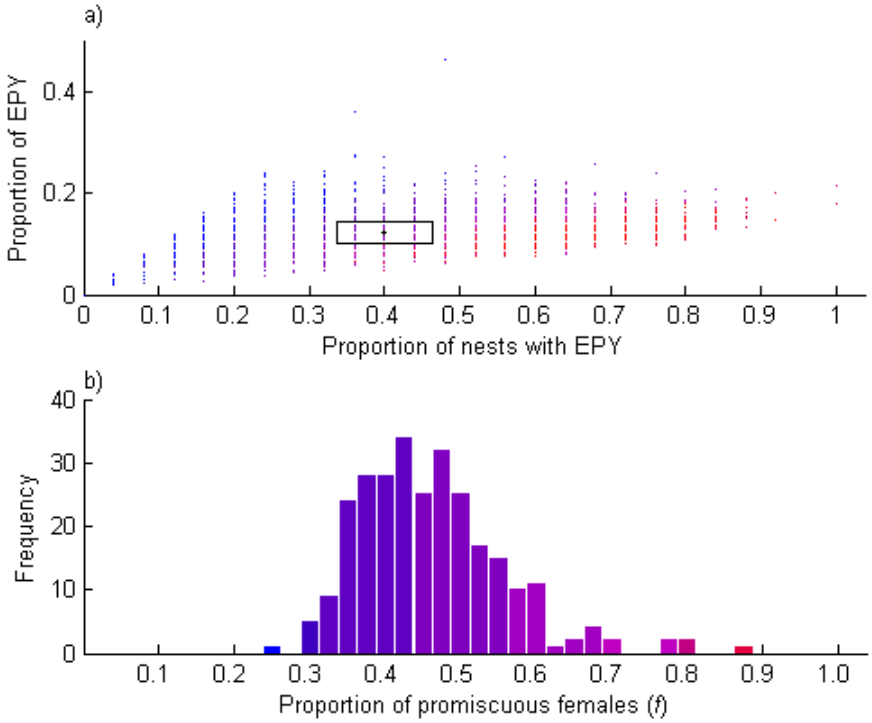


Figure 32: Population 19 from Charmantier & Blondel, 2003.

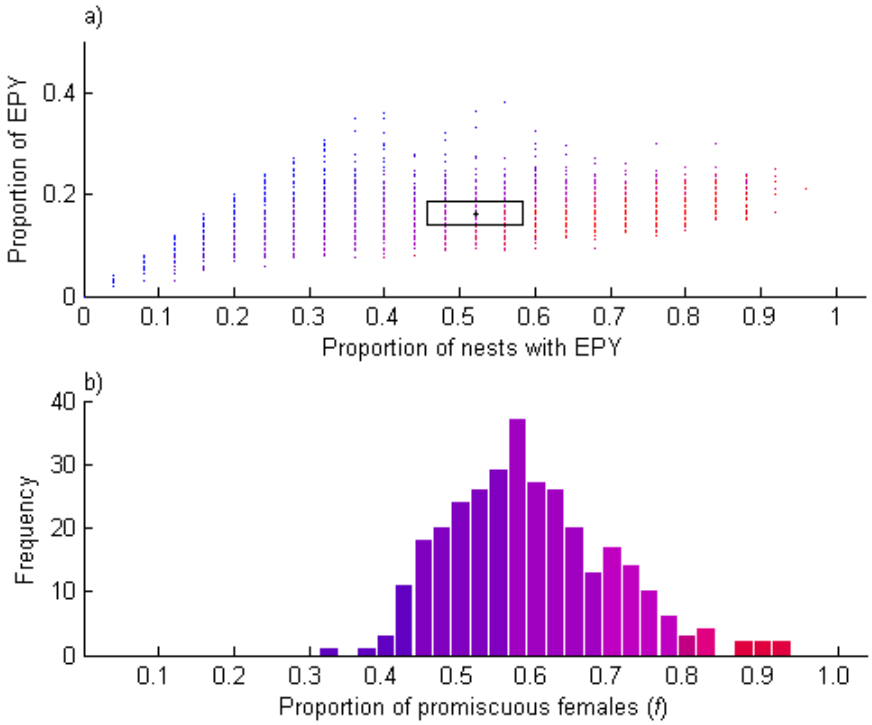


Figure 33: Population 20 from Charmantier & Blondel, 2003.

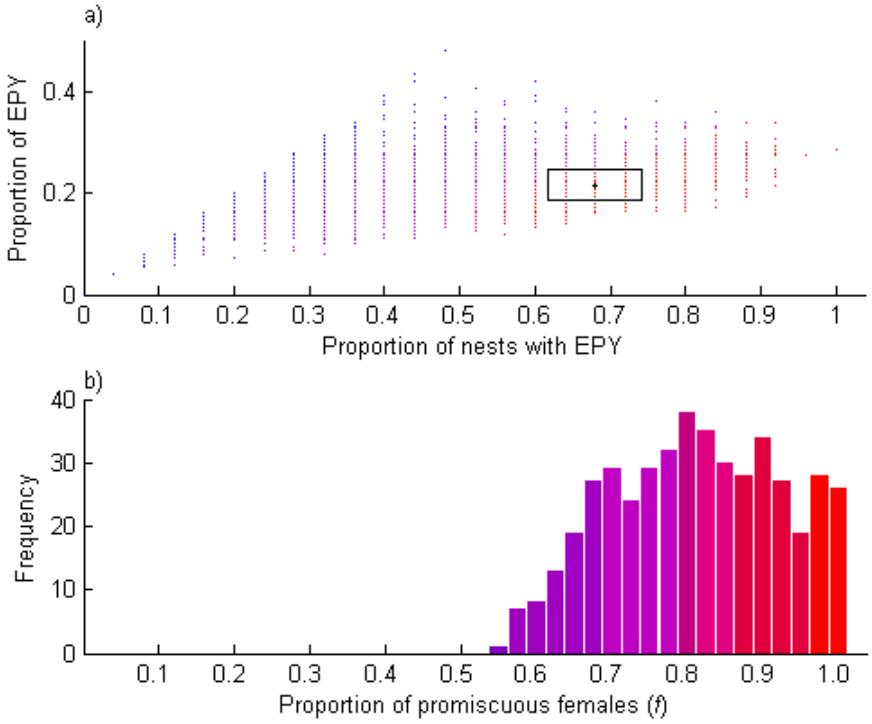


Figure 34: Population 21 from Charmantier & Blondel, 2003.

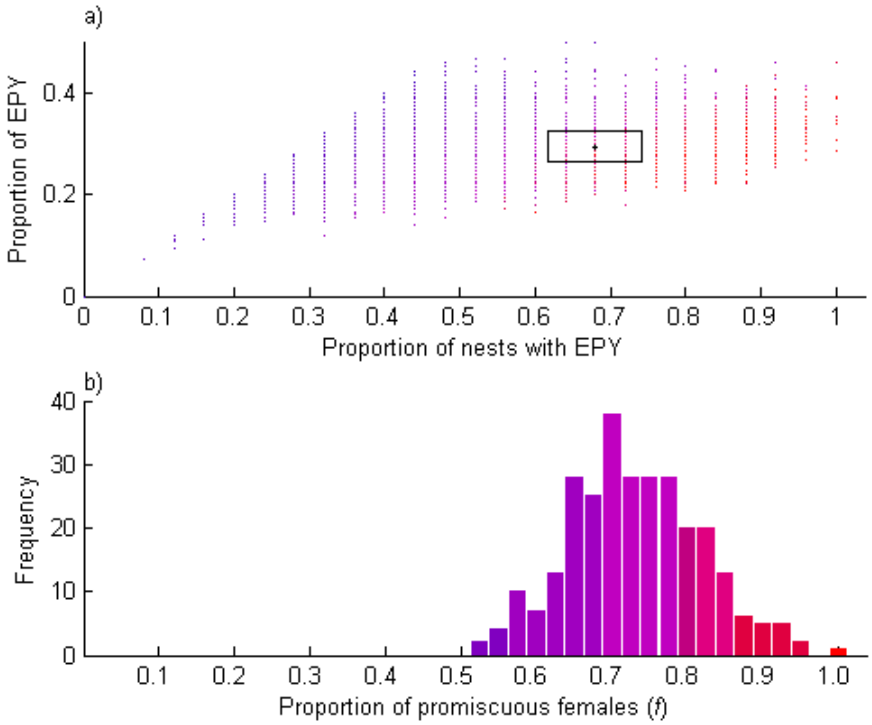


Figure 35: Population 22 from Charmantier & Blondel, 2003.

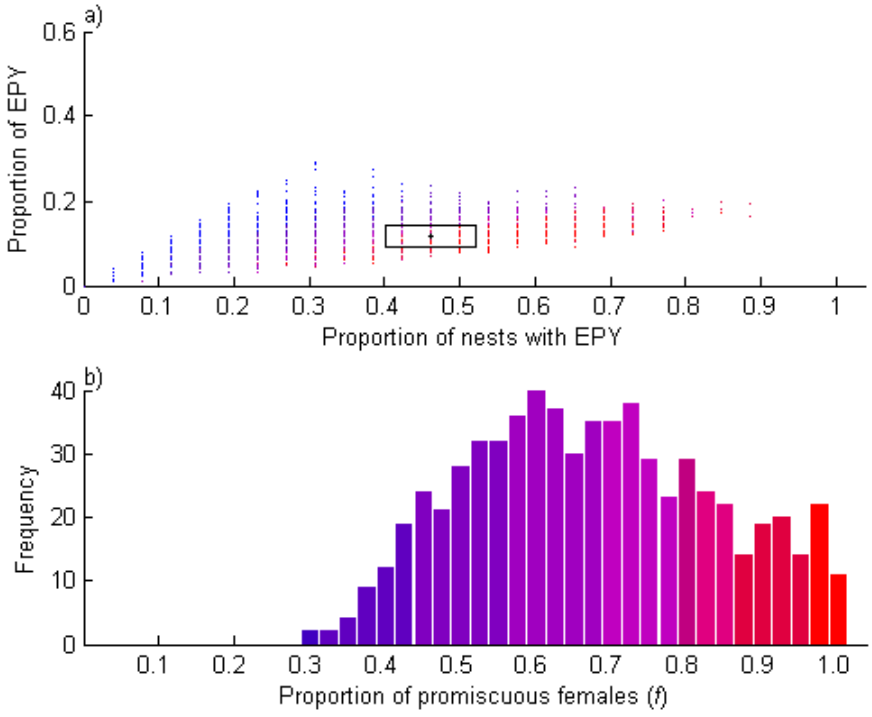


Figure 36: Population 23 from García-Navas *et al.*, 2013.

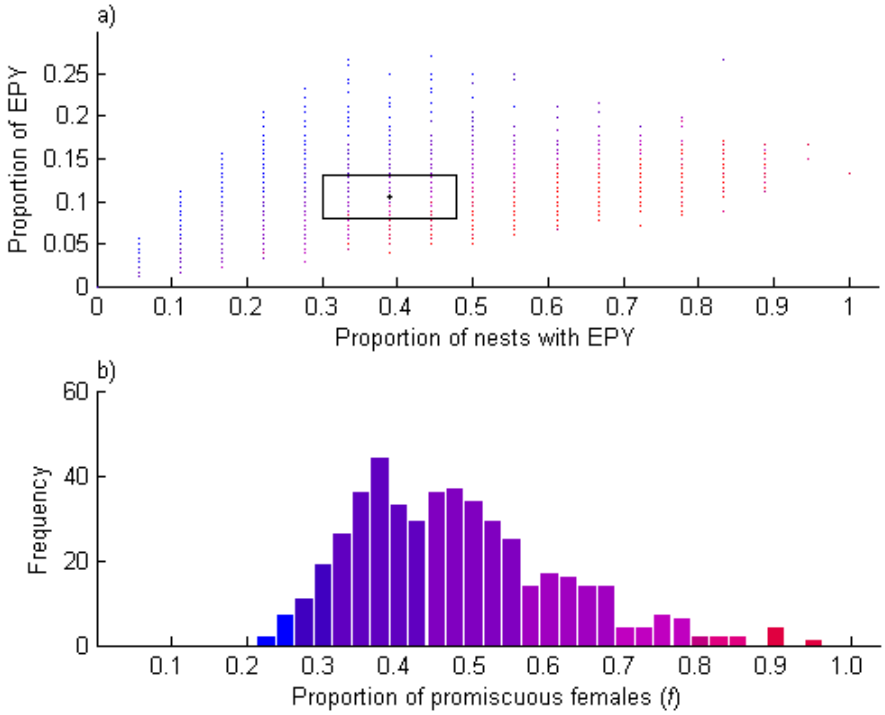


Figure 37: Population 24 from Krokene *et al.*, 1998.

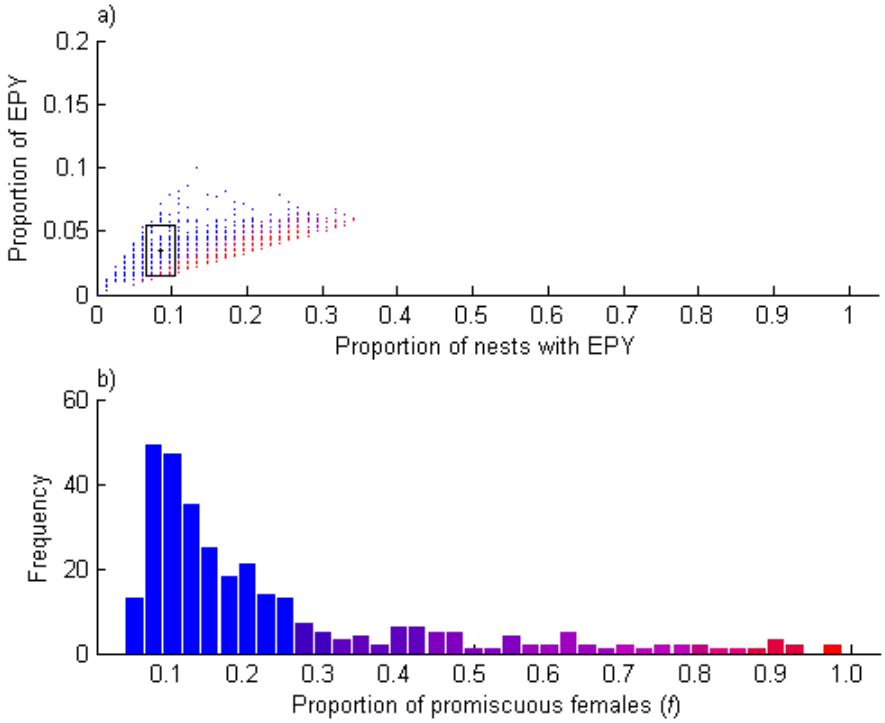


Figure 38: Population 25 from Verboven & Mateman, 1997.

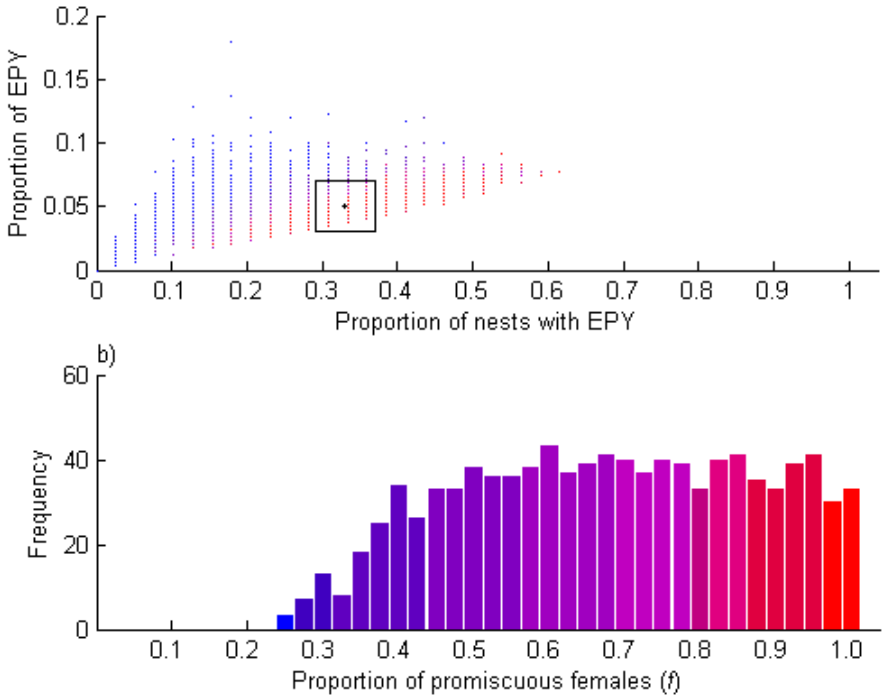


Figure 39: Population 26 from Strohbach *et al.*, 1998.

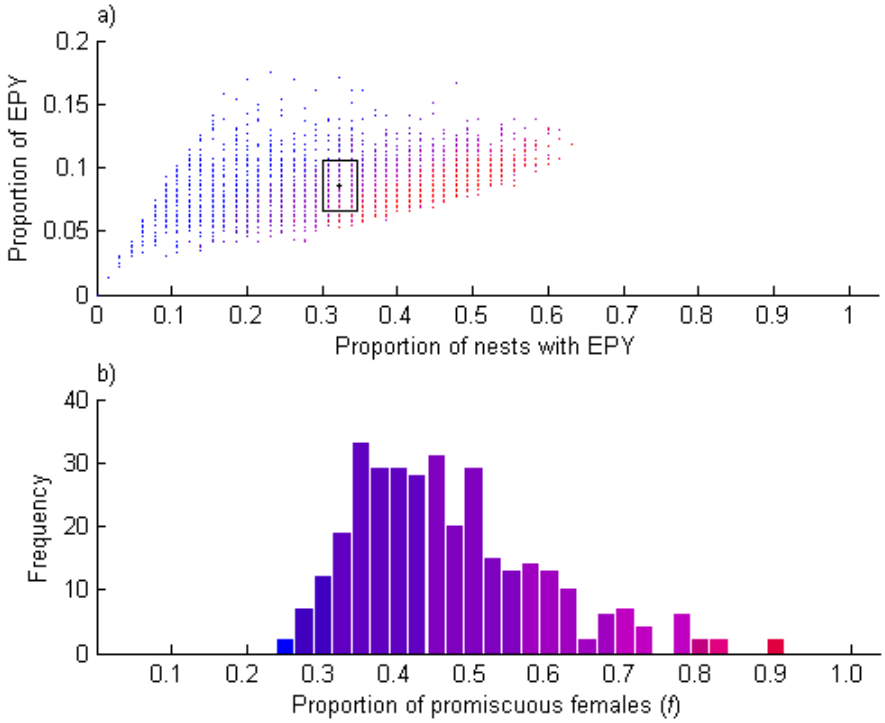


Figure 40: Population 27 from Lubjuhn *et al.*, 1999.

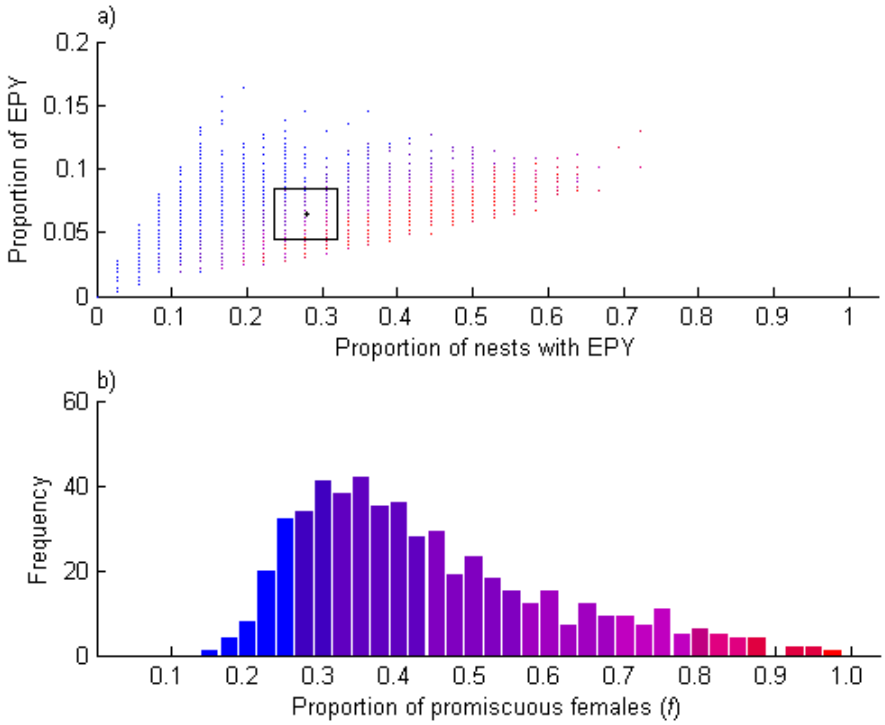


Figure 41: Population 28 from Lubjuhn *et al.*, 1999.

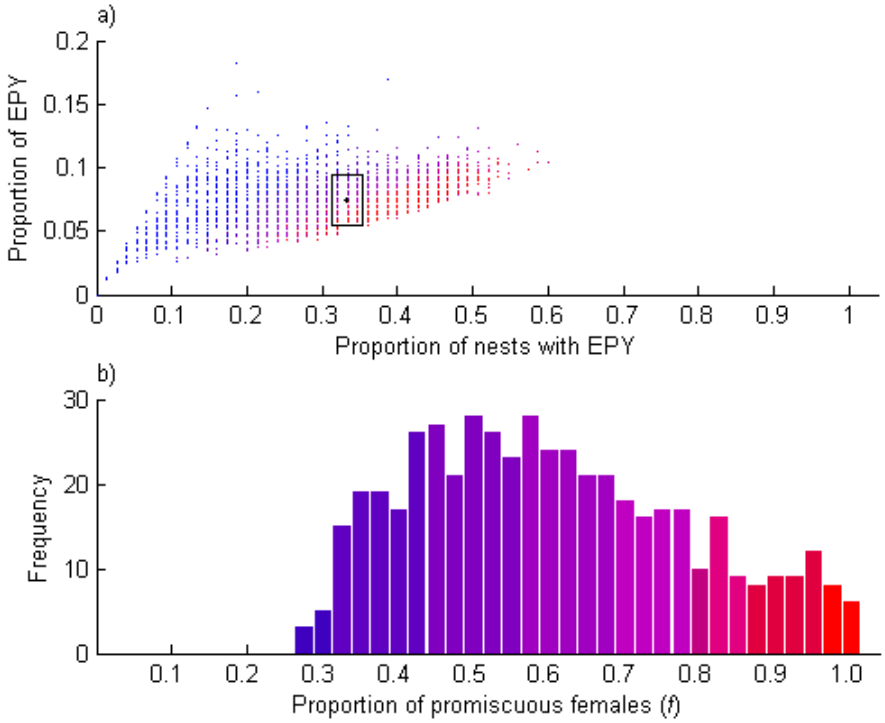
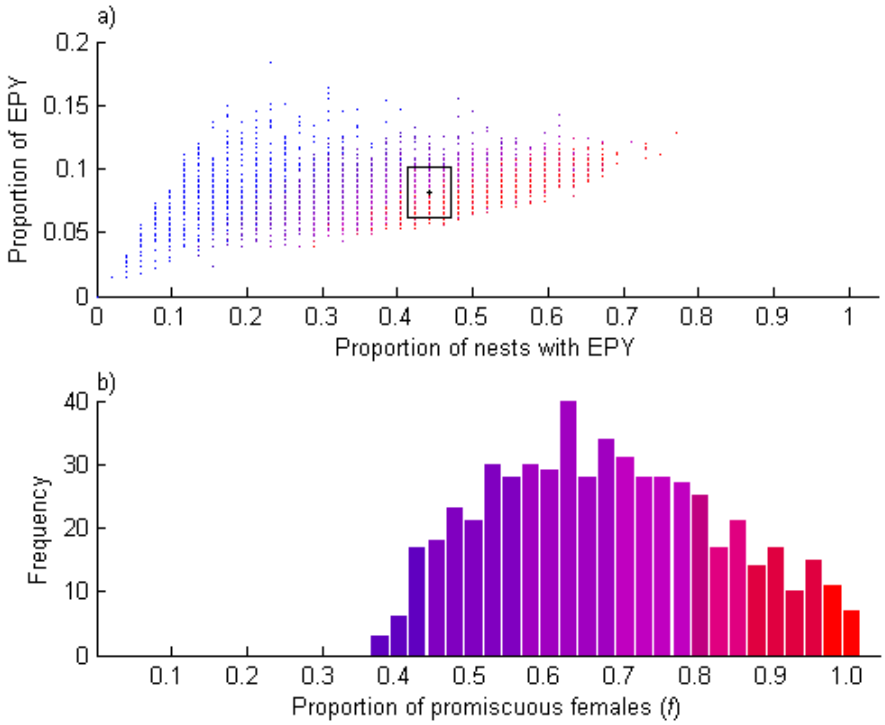


Figure 42: Population 29 from Lubjuhn *et al.*, 1999.



Figur 43: Population 30 from Lubjuhn *et al.*, 1999.

6.2. MATLAB script

The version of the MATLAB program was R2013a.

```
close all;
clear all;

[DataFromExcel, TextFromExcel] = xlsread('\\ustaoset.uib.no\dse024\Settings\Desktop\Feltstudier
EPP data.xlsx','Sheet1','O82:Z84');
%Columns:
%1: Clutch size (c)
%2: SamplePopSize (Npop)
%3: EPP in pop. (p)
%4: Proportion nests with EPY
%5: Proportion EPY in nests with at least 1 EPY

Nests = round(10000); %Total number of nests
p_cat_max = 10; %Number of p-values
f_cat_max = 40; %Number of f-values
SamplePopSize = 200; %Number of nests in each population sample
N_Populations = floor(Nests/SamplePopSize); %Number of population samples
MaxClutchSize = round(10); %Maximum clutch size
FocalClutchSize = 10; %Clutch size being focused on
UseRealData = -1;
%UseRealData > 0: Particular population from excel file
%UseRealData = 0: Array (no data from excel file)
%UseRealData = -1: All selected populations from excel file

if (UseRealData > 0);  %A value of zero means full array with parameter combinations. A value of
1 or higher is a specific pop. from table above.
p_cat_max = 1;
p = DataFromExcel(UseRealData,3);
N_Populations = 100;
MaxClutchSize = round(DataFromExcel(UseRealData,1));
```

```
FocalClutchSize = round(DataFromExcel(UseRealData,1));
SamplePopSize = DataFromExcel(UseRealData,2);
NestsWithEPYValue = DataFromExcel(UseRealData,4);
MeanEPYInEPYNestsValue = DataFromExcel(UseRealData,6);
Nests = SamplePopSize*N_Populations;
elseif (UseRealData == -1);
p_cat_max = length(DataFromExcel(:,1)); %#ok<*NASGU>
N_Populations = 100;
end

%Find whether an egg has EP sire or social father
if (UseRealData >= 0);
EPYorWPY(1:MaxClutchSize,1:Nests,1:f_cat_max,1:p_cat_max) = round(0); %Initialize with social
father as father
EPYsInNest(1:MaxClutchSize,1:Nests,1:f_cat_max,1:p_cat_max) = round(0); %The number of EPYs in
nests
NestHasEPY(1:MaxClutchSize,1:Nests,1:f_cat_max,1:p_cat_max) = round(0); %Matrix of nests that
contain at least one EPY
EPYDistr(1:MaxClutchSize+1,1:MaxClutchSize,1:f_cat_max,1:p_cat_max) = 0.; %Distribution of number
of EPYs in nests
Frequency_EPYsInSubPop(1:MaxClutchSize+1,1:N_Populations,1:f_cat_max,1:p_cat_max) = 0.; %The
number of EPYs in nests in population samples
EPYNests(1:MaxClutchSize,1:3,1:f_cat_max,1:p_cat_max)= round(0); %Proportion of nests with EPY
end

ContourPlotData(1:2,1:f_cat_max,1:p_cat_max) = -2.;
ContourPlotData2(1:2,1:p_cat_max,1:f_cat_max) = -2.;
ContourPlotData3(1:2,1:N_Populations,1:f_cat_max,1:p_cat_max) = 0.;
pvalueEPYNests(1:f_cat_max,1:p_cat_max) = 0.;
pvalues = 1/p_cat_max:(1/p_cat_max):1;
fvalues = 1/f_cat_max:(1/f_cat_max):1;
pvalueEPYsInNest(1:f_cat_max,1:p_cat_max) = 0.;
fvalueWPYsInNest(1:p_cat_max,1:f_cat_max) = 0.;
fvalueMeanEPYsInNest(1:N_Populations,1:f_cat_max,1:p_cat_max) = 0.;
Frequency_WPYsInSubPop(1:f_cat_max,1:N_Populations,1:p_cat_max) = 0.;
Frequency_WPYsInTotalPop(1:p_cat_max,1:f_cat_max) = 0.;
M = zeros(p_cat_max,f_cat_max);

for p_cat = 1:p_cat_max;
if (UseRealData == 0)
p = p_cat/p_cat_max;
elseif (UseRealData == -1);
p = DataFromExcel(p_cat,3);
MaxClutchSize = round(DataFromExcel(p_cat,1));
FocalClutchSize = round(DataFromExcel(p_cat,1));
SamplePopSize = DataFromExcel(p_cat,2);
NestsWithEPYValue = DataFromExcel(p_cat,4);
MeanEPYInEPYNestsValue = DataFromExcel(p_cat,6);
Nests = SamplePopSize*N_Populations;

xUpperLimit_p(p_cat) = NestsWithEPYValue + max(1.5*(1/(SamplePopSize-1)),0.02); %Upper limit x
xLowerLimit_p(p_cat) = NestsWithEPYValue - max(1.5*(1/(SamplePopSize-1)),0.02); %Lower limit x
yUpperLimit_p(p_cat) = p + max(1.5*(3/(SamplePopSize*FocalClutchSize-1)),0.02); %Upper limit y
yLowerLimit_p(p_cat) = p - max(1.5*(3/(SamplePopSize*FocalClutchSize-1)),0.02); %Lower limit y

clear('EPYorWPY','EPYsInNest','NestHasEPY','EPYDistr','Frequency_EPYsInSubPop','EPYNests');
EPYorWPY(1:MaxClutchSize,1:Nests,1:f_cat_max,1:p_cat_max) = round(0); %Initialize with social
father as father
EPYsInNest(1:MaxClutchSize,1:Nests,1:f_cat_max,1:p_cat_max) = round(0); %The number of EPYs in
nests
NestHasEPY(1:MaxClutchSize,1:Nests,1:f_cat_max,1:p_cat_max) = round(0); %Matrix of nests that
contain EPY
EPYDistr(1:MaxClutchSize+1,1:MaxClutchSize,1:f_cat_max,1:p_cat_max) = 0.; %Distribution of number
of EPYs in nests
Frequency_EPYsInSubPop(1:MaxClutchSize+1,1:N_Populations,1:f_cat_max,1:p_cat_max) = 0.; %The
number of EPYs in nests in population samples
EPYNests(1:MaxClutchSize,1:3,1:f_cat_max,1:p_cat_max)= round(0);
end
for f_cat = 1:f_cat_max;
f = f_cat/f_cat_max;
if (p/f > 1);
```

```
ContourPlotData(:,f_cat,p_cat) = -1.;
ContourPlotData2(:,p_cat,f_cat) = -1.;
else
for Nest = 1:Nests;
RandomValue = rand();
if (RandomValue <= f); %This is a female with EPC behaviour
for Egg = 1:MaxClutchSize;
if (rand() <= (p/f));
EPYorWPY(Egg,Nest,f_cat,p_cat) = round(1); %This egg is EPY given female EPC behaviour that is
variable between nests but constant within nests
end
end
end
for ClutchSize = 1:MaxClutchSize;
EPYsInNest(ClutchSize,Nest,f_cat,p_cat) = sum(EPYorWPY(1:ClutchSize,Nest,f_cat,p_cat)); %Number
of EPYs in nest at ClutchSize 'Egg'
if (EPYsInNest(ClutchSize,Nest,f_cat,p_cat) > 0);
NestHasEPY(ClutchSize,Nest,f_cat,p_cat) = round(1); %Nest has at least one EPY
end
end
end
ContourPlotData(1,f_cat,p_cat) = sum(NestHasEPY(FocalClutchSize,:,f_cat,p_cat))/Nests;
%Proportion of nests with EPY
ContourPlotData(2,f_cat,p_cat) = sum(EPYsInNest(FocalClutchSize,:,f_cat,p_cat))/Nests; %Mean # of
EPYs in nests
ContourPlotData2(1,p_cat,f_cat) = 1-sum(NestHasEPY(FocalClutchSize,:,f_cat,p_cat))/Nests;
%Proportion of nests without EPY
ContourPlotData2(2,p_cat,f_cat) =
sum(EPYsInNest(FocalClutchSize,:,f_cat,p_cat))/sum(NestHasEPY(FocalClutchSize,:,f_cat,p_cat));
%Mean # of EPY in nests with at least 1 EPY

%Find distribution of EPY over nests
ClutchSizeAxis = 1:MaxClutchSize;
for ClutchSize = 1:MaxClutchSize;
EPYDistr(:,ClutchSize,f_cat,p_cat) =
hist(EPYsInNest(ClutchSize,:,f_cat,p_cat),0:MaxClutchSize)/Nests; %Frequency distribution of EPY
offspring per nest in population
end
for ClutchSize = 1:MaxClutchSize
EPYNests(ClutchSize,1,f_cat,p_cat) = sum(NestHasEPY(ClutchSize,:,f_cat,p_cat))/Nests; %Proportion
of nests with EPY
EPYNests(ClutchSize,2,f_cat,p_cat) = 1.96*sqrt(var(NestHasEPY(ClutchSize,:,f_cat,p_cat)));
EPYNests(ClutchSize,3,f_cat,p_cat) = 1-(1-p)^ClutchSize; %Analytic solution
end
end
if (UseRealData == -1);
for Pop = 1:N_Populations
if (p/f <= 1);
StartNest = 1 + (Pop-1)*SamplePopSize;
EndNest = StartNest + SamplePopSize - 1;
ContourPlotData3(1,Pop,f_cat,p_cat) =
sum(NestHasEPY(FocalClutchSize,StartNest:EndNest,f_cat,p_cat))/SamplePopSize; %Proportion of
nests with EPY
ContourPlotData3(2,Pop,f_cat,p_cat) =
sum(EPYsInNest(FocalClutchSize,StartNest:EndNest,f_cat,p_cat))/(FocalClutchSize*SamplePopSize);
%Proportion of EPY
x = ContourPlotData3(1,Pop,f_cat,p_cat);
y = ContourPlotData3(2,Pop,f_cat,p_cat);
if (x >= xLowerLimit_p(p_cat) && x <= xUpperLimit_p(p_cat)); %Upper and lower limit in x-axis
direction
if (y >= yLowerLimit_p(p_cat) && y <= yUpperLimit_p(p_cat)); %Upper and lower limit in y-axis
direction
M(p_cat,f_cat) = M(p_cat,f_cat) + 1;
end
end
end
end
end
end %f_cat
end %p_cat
```