Assessing the statistical power of genetic analyses to detect multiple mating in fishes

B. D. Neff*‡ and T. E. Pitcher†

*Department of Biology, University of Western Ontario, London, Ontario N6A 5B7, Canada and †Department of Zoology, University of Toronto, 25 Harbord St, Toronto, Ontario M5S 3G5, Canada

(Received 29 April 2002, Accepted 7 August 2002)

A single-sex model is presented that calculates the probability of detecting multiple mating (PrDM) given genetic data from the single genetic parent and a sample of its offspring. The model incorporates the effects of numbers of loci, alleles, offspring and genetic parents contributing to the multiple mating, all of which effect PrDM. The model is used to determine the actual number of loci and offspring that are required to detect multiply mated broods with high probability (80 and 95%). For example, if two sires contribute with equal fertilization success to multiply mated broods, then only 10 offspring and one locus with seven equally common alleles are required to ensure that 80% of multiple mated broods are detected. Ninety-five per cent of multiple mated broods can be detected with 10 offspring and five loci with four equally common alleles. The utility of the model is demonstrated with biological examples addressing geographic variation in multiple paternity among natural populations of guppies Poecilia reticulata and mosquitofish Gambusia holbrooki.

© 2002 The Fisheries Society of the British Isles. Published by Elsevier Science Ltd. All rights reserved.

Key words: multiple mating; statistical power; paternity; reproductive skew; genetic markers; Poeciliidae.

INTRODUCTION

The recent discovery that multiple mating is prevalent in the animal kingdom has revolutionized the study of mating systems (Reynolds, 1996; Krebs & Davies, 1997; Jennions & Petrie, 2000). Multiple mating occurs when individuals of one sex mate with more than one individual of the opposite sex (Reynolds, 1996; Neff et al., 2000a). Multiple mating can affect the intensity of natural and sexual selection (Andersson, 1994; Fleming & Gross, 1994; Evans & Magurran, 1999a, 2001), effective population sizes (Sugg & Chesser, 1994; Martinez et al., 2000), and genetic variability and introgression (Moran & Garcia-Vazquez, 1998; Baer & Schmid-Hempel, 1999). In this paper the frequency of multiple mating is defined as the proportion of broods in the population that contain genes from at least two males (or alternatively two females). The term mating is used to reflect fertilization and not just copulation. Frequency differs from the degree of multiple mating, which instead calculates the number of individuals genetically contributing to each brood. The frequency of multiple mating is critical to understanding the evolution of mating systems and to the conservation of endangered populations (Kelly et al., 1999; Kichler et al., 1999; Zane et al., 1999; Moore & Ball, 2002).

‡Author to whom correspondence should be addressed. Tel.: +1 519 850-2532; fax: +1 519 661-2014; email: bneff@uwo.ca
Generally, the frequency of multiple mating can be inferred from the degree of multiple mating, for example from the proportion of broods with two or more sires, and several models have been developed to calculate degree (Levine et al., 1980; Pedersen & Boomsma, 1999; DeWoody et al., 2000a,b; Lexer et al., 2000). Large numbers of loci, however, can be required to accurately calculate the likely number of sires per brood. Alternatively, the frequency can be more efficiently calculated by analysing all broods simultaneously and adjusting the observed proportion of multiply mated broods (i.e. broods having three or more paternal alleles at a locus) by a correction factor (Milkman & Zeitler, 1974; Cobbs, 1977; Williams & Evarts, 1989; Kichler et al., 1999). The correction factor accounts for the possibility that less than three paternal alleles are detected in a multiply mated brood, which can occur when two sires share a common genotype at the loci used in the analysis or when a sire’s offspring are not sampled from the brood (due to random sampling error). The correction factor thus represents the probability of detecting a multiple mating (PrDM). PrDM depends on four parameters: (1) the number of loci; (2) the number of alleles (and their frequency); (3) the number of offspring analysed from the brood; (4) the number of genetic parents (and their reproductive skew). No model yet exists, however, that simultaneously accounts for all of these factors when calculating PrDM. The aim of the present study was to develop a model to calculate PrDM that incorporated any combination of these four parameters.

In the model, mating systems with single-sex multiple mating are considered. Single-sex multiple mating occurs when there is multiple mating by only one sex. Here, multiple mating is defined at the level of the brood (Neff et al., 2000a). Therefore, single-sex multiple mating gives rise to broods that contain the genes of either a single female that has mated with multiple males or a single male that has mated with multiple females. As such, all of the young within a brood are either full- or half-sibs. The model requires single-locus genetic data, such as microsatellites or allozymes, from the known genetic parent and a sample of the brood. A genetic sample from the breeding population is also required to estimate allele frequencies.

The model is tested with two biological examples based on published data. The first example addresses geographic variation in multiple paternity among wild populations of guppies Poecilia reticulata (Peters) (Kelly et al., 1999). Guppies experience two distinct predation regimes. Guppies are considered to be in high predation locales when they co-exist with the piscivorous cichlid Crenicichla alta Eigenmann and low predation locales when they co-exist with the gape-limited cyprinodontid Rivulus hartii (Boulenger), which has a limited impact only on juvenile guppies (Farr, 1975; Houde, 1997). As the costs of mating increases due to increased predation, female guppies are less selective in their choice of mates (Godin & Briggs, 1996; Gong & Gibson, 1996) and males alter their mating tactics to include fewer conspicuous courtship displays and more sneak copulations (Luyten & Liley, 1985; Houde & Endler, 1990). Therefore, it might be expected that multiply sired broods are more common in high predation populations (Kelly et al., 1999; Houde, 1988; Evans & Magurran, 1999a, b, 2000). The second example calculates and compares the frequency of multiple mating in two wild populations of the related mosquitofish Gambusia holbrooki Girard from Fisher Pond and Fire Pond (Zane et al., 1999). Previous analysis of
multiple mating in other populations of mosquitofish found that c. 55–65% of broods were multiply sired (Chesser et al., 1984; Greene & Brown, 1991). Genetic data from the populations studied by Zane et al. (1999), however, suggest that genetic variation is relatively low. Thus, if females mate multiply to increase genetic variation in their sperm supply (Jennions & Petrie, 2000), then the frequency of multiple mating should be higher (>65%) in Fisher Pond and Fire Pond.

MATERIAL AND METHODS

THE MODEL

The model builds on those developed by Cobbs (1977) and Kichler et al. (1999). It calculates the probability that a multiple mating is detected (given that there is a multiple mating) using a Monte Carlo simulation (Manly, 1997). While it might be desirable to generate a statistic directly, this is exceedingly difficult when multiple loci are used (Harshman & Clarke, 1998). The model requires four parameters: (1) the number of loci; (2) the number of alleles and their frequencies within the breeding population; (3) the number of offspring analysed from the brood; (4) the number of mates (e.g. sires) and their relative contributions (i.e. skew in fertilization success within the brood). The genotype of a known genetic parent can also be an input for individual broods. When estimating \( Pr_{DM} \) for a population or set of independent broods it may be omitted and instead randomly generated by the model based on the population allele frequency data. All of the parameters influence \( Pr_{DM} \) and must therefore be accounted for in the analysis (Cobbs, 1977). Because the number of mates and their reproductive skew will probably be unknown methods to estimate them are discussed below. Also, Harshman & Clark (1998) develop a model based on a Monte Carlo simulation to directly estimate reproductive skew from genetic data, assuming that the probability that a female remates follows a Poisson distribution and the skew follows a geometric distribution.

The simulation (Fig. 1) generates genotypes for a genetic mother (or father) and its \( m \) mates based on the genetic polymorphism data. Each of the \( b \) offspring within the brood is then generated by selecting one of the \( m \) mates according to their relative fertilization success (skew). For example, if two mates contribute to a brood with relative success of 75 and 25\%, then it is three times more likely that the first mate will produce any particular offspring and on average the first mate will produce 75\% of the offspring within a brood. One of the two alleles from each parent at each locus is randomly assigned (with equal probability) to the offspring. The genotypes of the \( b \) offspring are then analysed to determine the number of different alleles from the mates. The multiple mating is detected when more than two different alleles from the mates at any locus are identified in the \( b \) offspring (Birdsall & Nash, 1973; Barry et al., 1992; Baker et al., 1999). The multiple mating could also be detected when only two paternal alleles are detected, but the frequency of the two alleles in the offspring differ from the expected 1:1 ratio under Mendelian inheritance and assuming that there is only a single heterozygous sire (Barry et al., 1992). This latter possibility, however, is not considered because it adds little resolution to the analysis when the loci used have more than two common alleles (as is the case for most microsatellites) and when only a limited number of offspring are analysed (Ochando et al., 1996). In cases where a mate’s allele is ambiguous, for example when the known genetic parent and offspring are heterozygous with the same alleles, the minimum number of alleles is calculated. The entire process is repeated for a total of 10 000 broods and \( Pr_{DM} \) is calculated as the proportion of these broods that a multiple mating (at least three paternal or maternal alleles) is detected.

MULTIPLE MATING IN GUPPIES

Kelly et al. (1999) used two microsatellite loci to estimate the proportion of broods with multiple paternity from 10 wild guppy populations (five populations from each of high and low predation locales). They reported the results of an analysis of covariance
(ANCOVA) test which controls only for differences in the number of offspring analysed per brood and showed that the predation regime significantly affects the frequency of multiple paternity in guppies ($F_{1,7} = 8.31, P = 0.024$). The loci, however, were more polymorphic in the high predation populations which will increase the chance of detecting multiple matings and could lead to spurious differences in the frequency of multiple mating among populations.

For each population $PrDM$ was calculated using the mean and s.d. in the number of offspring analysed per brood reported in Kelly et al. (1999) and the genetic polymorphism data reported in Kelly (1999). Because the number of sires contributing to each brood was unknown six situations were considered: (1) two sires with equal fertilization success (mean paternity of 50% for each male) for broods from all populations; (2) two sires with skewed success (66.7 and 33.3%) for all populations; (3) three sires with equal success (33.3% for each male) for broods from all populations; (4) three sires with skewed success (57, 28.5 and 14.5%) for all populations; (5) two sires with skewed success (66.7 and
33·3%) for populations from the high predation locales and three sires with skewed success (57, 28·5 and 14·5%) for populations from the low predation locales; (6) two sires with skewed success (66·7 and 33·3%) for populations from the low predation locales and three sires with skewed success (57, 28·5 and 14·5%) for populations from the high predation locales. The first four situations have no difference in the number of sires between populations from the two predation regimes, while the last two situations consider differing numbers. The six situations cover a variety of multiple matings and probably encompass the actual distribution of fertilization success in these populations (Houde, 1997; Kelly et al., 1999). For each of the six situations ANCOVA was conducted with the observed frequency of multiply mated broods as the dependent variable, predation regime (low v. high) as the independent variable and PrDM as the covariate. Both the frequency and PrDM were first arcsine transformed to restore normality to the data (Zar, 1999).

MULTIPLE MATING IN MOSQUITOFISH

Zane et al. (1999) used three microsatellite loci to estimate the proportion of broods with multiple paternity from two wild mosquitofish populations, Fisher Pond and Fire Pond. The two ponds are located c. 10 km apart at the Savannah River site near Aiken, South Carolina, U.S.A. They are believed to be genetically distinct (Zane et al., 1999). Zane et al. (1999) found that based on direct paternal allele counts in each brood the frequency of multiple mating in Fisher Pond was 84% and in Fire Pond it was 88%. Because the combined exclusion probability across the three loci was only 0·84 in each population, they concluded that their frequency estimates were conservative (i.e. minimal estimates). They were unable, however, to calculate the actual accuracy of their estimates.

For each population PrDM was calculated using the mean number of offspring and the genetic data reported in Zane et al. (1999). Here two situations were considered for the number of sires contributing to multiply mated broods: (1) two sires with equal fertilization success (50% each); (2) two sires with skewed success (66·7 and 33·3%). For each of the two situations, PrDM was used to adjust the observed frequency of multiple mating according to the following calculation: actual frequency of multiple mating=(observed frequency of multiple mating)(PrDM)−1.

RESULTS

THE MODEL

Increasing the numbers of loci, alleles, offspring and genetic parents increases PrDM (Fig. 2 and Table I). Increasing either the number of loci used to genotype the individuals or using loci with more alleles or with frequencies that are less skewed (i.e. more effective alleles; Neff et al. 2000b) increases genetic polymorphism. In turn, it is more likely that each genetic parent is heterozygous and contains different alleles and therefore PrDM increases. Loci with less skewed allele frequencies are considerably more valuable than loci with highly skewed frequencies. For example, when allele frequencies are skewed, nearly twice as many alleles are required to obtain equivalent PrDM values as loci with even allele frequencies (Table I).

It is more likely that a multiple mating is detected when the known parent is homozygous (unpubl. data). PrDM, however, increases with locus polymorphism (number of alleles) even though polymorphism decreases the probability that the known parent is homozygous. This is attributed to the predominant positive effect of polymorphism on the probability that the known parent’s mates are heterozygous and possess different alleles, which are more likely to be detected in the sample of offspring.
Increasing the number of offspring analysed from a brood increases the probability that the alleles from each genetic parent are sampled and therefore increases $PrDM$ [Fig. 2(b)]. Finally, increasing the number of parents that genetically contribute to the brood, or equalizing the relative contribution of each parent, increases the probability that more than two different alleles are detected within the brood and therefore increases $PrDM$ [Fig. 2(c)]. Generally, however, the number of sires contributing to a brood and their reproductive skew will be a property of the mating system and out of the hands of the researcher. Nevertheless, all else equal, mating systems with more sires and lower skew will have greater statistical power to detect multiple matings.

While it requires relatively few loci, alleles and offspring to obtain $PrDM$ values of 80%, it requires considerably more to obtain $PrDM$ values of 95%. For example, assuming that two sires contribute to multiply mated broods with equal success, a $PrDM$ value of 80% can be obtained by analysing 10 offspring with one locus with seven equally common alleles, or two loci with four equally common alleles [Fig. 2(a)]. To obtain a $PrDM$ value of 95%, 10 offspring would have to be analysed with one locus with 15 equally common alleles or two loci with seven equally common alleles. Thus, loci with approximately twice as many
alleles are required to raise \( PrDM \) from 80 to 95%. Nevertheless, with microsatellite genetic markers, \( PrDM \) values of 95% should be obtainable.

### MULTIPLE MATING IN GUPPIES

The mean \( PrDM \) for the low predation populations ranged from 0.31–0.45 and for high predation populations it ranged from 0.62–0.80 depending on the reproductive skew situation. Thus, the probability of detecting a multiple mating in low predation populations was considerably lower than in the high predation populations. Subsequently, it was found that \( PrDM \) had a highly significant effect and explained most of the variation in the observed frequency of multiple mating among the populations (80–99%\(^\text{a}\); Table II). The significance of predation intensity, however, was dependent on the number of sires. Provided that the low predation populations had the same or more sires contributing to each multiply mated brood as compared to the high predation populations, there was a small but significant effect of predation on the frequency of multiple mating. If the high predation populations had more sires contributing to each multiply mated brood, however, then there was no significant effect of predation (Table II).

---

### Table I. \( PrDM \) values for combinations of numbers of loci, offspring, and alleles.

Multiply mated broods were assumed to be the product of either two sires with fertilization success of two-thirds and one-third (first \( PrDM \) number) or three sires with equal success (one-third each; second number).

<table>
<thead>
<tr>
<th>Loci</th>
<th>Offspring</th>
<th>Even</th>
<th>Alleles</th>
<th>Skewed*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>0.13</td>
<td>0.26</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.21</td>
<td>0.42</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.28</td>
<td>0.49</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.47</td>
<td>0.74</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.40</td>
<td>0.61</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.66</td>
<td>0.87</td>
<td>0.94</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.24</td>
<td>0.44</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.38</td>
<td>0.66</td>
<td>0.78</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.48</td>
<td>0.74</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.72</td>
<td>0.93</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.63</td>
<td>0.85</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.88</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.47</td>
<td>0.72</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.68</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.80</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.95</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.92</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Skewed allele frequencies were calculated such that each subsequent allele had 50% the frequency of the previous allele. As an example, for three alleles the frequencies would be 0.571, 0.286 and 0.143.
TABLE II. Analysis of variance of the frequency of multiple mating in guppies between high and low predation regimes. Six possible situations were considered for the number of sires and their relative paternity (equal or skewed) within the broods. The proportion of the explained variation, per cent in parentheses, was partitioned between Predation and the covariate PrDM

<table>
<thead>
<tr>
<th>Number of sires</th>
<th>Low predation (n=5)</th>
<th>High predation (n=5)</th>
<th>Source</th>
<th>$r^2$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Two equal Two equal</td>
<td>Model 0·92</td>
<td>$&lt;0·001$</td>
<td>Predation (7) 0·047</td>
<td>$&lt;0·001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$PrDM$ (93)</td>
<td></td>
<td>$&lt;0·001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Two skewed Two skewed</td>
<td>Model 0·92</td>
<td>$&lt;0·001$</td>
<td>Predation (7) 0·047</td>
<td>$&lt;0·001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$PrDM$ (93)</td>
<td></td>
<td>$&lt;0·001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Three equal Three equal</td>
<td>Model 0·92</td>
<td>$&lt;0·001$</td>
<td>Predation (7) 0·055</td>
<td>$&lt;0·001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$PrDM$ (93)</td>
<td></td>
<td>$&lt;0·001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Three skewed Three skewed</td>
<td>Model 0·92</td>
<td>$&lt;0·001$</td>
<td>Predation (6) 0·054</td>
<td>$&lt;0·001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$PrDM$ (94)</td>
<td></td>
<td>$&lt;0·001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Three skewed Two skewed</td>
<td>Model 0·92</td>
<td>$&gt;0·001$</td>
<td>Predicted (20) 0·005</td>
<td>$&lt;0·001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$PrDM$ (80)</td>
<td></td>
<td>$&lt;0·001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Two skewed Three skewed</td>
<td>Model 0·93</td>
<td>$&lt;0·001$</td>
<td>Predation (1) 0·395</td>
<td>$&lt;0·001$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$PrDM$ (99)</td>
<td></td>
<td>$&lt;0·001$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MULTIPLE MATING IN MOSQUITOFISH
If two sires contributed equally to multiply mated broods, then $PrDM$ in Fisher Pond and Fire Pond would be 0·975 and 0·982, respectively. If two sires contributed with skewed success, then $PrDM$ would be 0·967 and 0·974 in the two respective populations. Thus, with only three loci, Zane et al. (1999) had high statistical power to detect multiply mated broods. Based on the calculation of $PrDM$, the actual frequencies of multiple mating in Fisher Pond and Fire Pond were 86·2 and 89·6%, respectively, for two sires with equal success and 86·9 and 90·3% for two sires with skewed success. Thus, consistent with the conclusions of Zane et al. (1999), the frequency of multiple mating in the two populations is very similar.

DISCUSSION
In this paper a model is developed to calculate the statistical power of genetic analyses to detect multiple mating. The model incorporates the number of loci, number of alleles and their frequencies, number of offspring analysed from each brood, and number of genetic sires and their reproductive skew. The probability of detecting a multiple mating, and thus statistical power, increases with the number of loci and offspring analysed, and the number of genetic parents
contributing to the brood (Fig. 2). It decreases as the population frequency of the alleles carried by the single genetic parent increases (and if the parent is heterozygous) and as reproductive skew among the parents increases (Fig. 2). Analogous results have been found in the context of genetic parentage analysis and population assignments (Bernatchez & Duchesne, 2000; Neff et al., 2000b).

Generally, only a few loci and samples are required to obtain high statistical power to detect multiple matings. For example, when two sires contribute with equal success to a brood, only 10 offspring (per brood) and one locus with seven equally common alleles are required to ensure that 80% of multiply mated broods are detected. $PrDM$ values of 95% can be obtained with 10 offspring and two loci with seven equally common alleles or five loci with four equally common alleles. When allele frequencies are skewed, considerably more loci are required to obtain equivalent statistical power (Table I). Thus, future studies may be better to focus on loci with fewer, evenly distributed alleles. Nevertheless, with genetic markers like microsatellites, high statistical power should be feasible in many studies.

The effectiveness of the model has been demonstrated by analysing the effect of predation intensity on the frequency of multiple mating among wild guppy populations (Kelly et al., 1999). There were two important consequences of $PrDM$ on the analysis. First, there was a higher chance of detecting multiple matings in the high predation populations because of the increased genetic polymorphism (i.e. number of alleles at each locus) in these populations. Most (80–99%) of the explained variation in the frequency of multiple mating was attributed to differences in $PrDM$ (Table II). As such, the observed difference in frequencies of multiple mating between high (64·2% of broods) and low (24·6% of broods) predation populations observed by Kelly et al. (1999) may be attributed to both biological and statistical effects. Second, the frequency of multiple mating was significantly higher in the high predation populations as compared to the low predation populations, unless the number of sires contributing to a multiply mated brood was greater in the high predation populations. In this latter case there would be a difference in the degree of multiple mating, but not the frequency. Both results are interesting and biologically important.

In the second example on mosquitofish, it has been shown that with only three microsatellite loci Zane et al. (1999) had high statistical power to detect multiple matings in their two wild populations. Assuming that only two males contributed to multiply mated broods with equal or marginally skewed fertilization success, >96% of such broods should have been detected. Interestingly, $PrDM$ was considerably higher than the parentage exclusion probability (=84%). Thus, exclusion probabilities do not accurately reflect the probability of detecting multiple matings. Incorporating $PrDM$ into the observed frequency of multiple mating, it was concluded that the actual frequency of multiple mating was very similar in the two populations (Fisher Pond and Fire Pond), ranging between 86 and 90%. These estimates are substantially higher than estimates of 55–65% from other populations. Because the Fisher Pond and Fire Pond populations have lower genetic variability, the increased rate of multiple mating in these populations might reflect females attempting to increase the genetic diversity in their sperm supplies.
In conclusion, this study highlights a general point for studies assessing the frequency of multiple mating: \(PrDM\) directly affects the proportion of multiple matings that are detected and therefore should be included in statistical analyses. The accuracy of statistical analyses can be maximized by ensuring that \(PrDM\) is not only as high as possible, but also equal between groups that are compared (e.g. the high and low predation populations in guppies or the two populations of mosquitofish). Equating \(PrDM\) between groups may involve increasing the number of loci for groups that have low genetic polymorphism or increasing the number of offspring analysed from each brood. The model can be used with polymorphism data collected from a preliminary screening of the groups to determine the optimal approach to achieve a desired \(PrDM\). It is also possible that the Monte Carlo simulation developed here could be incorporated into a Bayesian framework to provide estimates of the frequency of multiple mating. The software is available as a downloadable executable file (http://publish.uwo.ca/~bneff/software.htm PrDM).

We are grateful to P. Fu, C. Kelly, N. Mrosovsky, J. Olden, J. Reynolds, D. Reznick, H. Rodd and two anonymous reviewers for their comments and helpful discussion. We also thank the authors of the two papers from which we extracted the multiple mating data. This work was supported by the Natural Science and Engineering Council of Canada.

References


