

A C++ program for assessing the frequency of multiple mating in nature

Bryan D. Neff^{*‡}, Trevor E. Pitcher[†], and Joe Repka[§]

Department of Biology, Biological & Geological Sciences Building, University of Western Ontario, London N6A 5B7, Canada (*bneff@uwo.ca*)
Phone: (519) 850-2532; Fax: (519) 661-2014

[†]Department of Zoology, University of Toronto, 25 Harbord St, Toronto, Ontario M5S 3G5, Canada (*tpitcher@zoo.utoronto.ca*)

[§]Department of Mathematics, University of Toronto, 100 St. George St, Toronto, Ontario M5S 3G3, Canada (*repka@math.toronto.edu*)

July 20, 2009

Description

This program calculates the frequency of multiple mating in a population based on a sample of nests. The program outputs the probability distribution associated with each f_{mm} from 0 to 100% in increments of 1%. From the distribution any CI can be calculated. The expected f_{mm} is also calculated. The model assumes single-sex multiple mating (i.e., polygyny or polyandry) and therefore all offspring in a brood are either full-sibs or half-sibs. The criterion used to detect the multiple mating is three or more paternal alleles in the set offspring. The program is written in C++ for PC-type computers and is an executable file (*fmm.exe*). It is run by double clicking the icon or using the 'Run' command line.

Directions

Before running the program four data files must be set-up. This is most easily done using Microsoft Excel (or equivalent). Each file must be saved as a **tab delimited** file. The following data must be entered into each spreadsheet:

(1) *loci.txt*

number of loci

for each locus: allele frequencies

NOTES: Exactly 40 allele frequencies must be entered for each locus. If there is more than 40 alleles, alleles must be binned. If there are fewer than 40 alleles, simply enter 0 for the remaining alleles. The 40 allele frequencies must sum to 1.0. Each locus is entered on a separate line. A maximum of 25 loci can be used. The execution of the program becomes increasingly slow with each additional loci used.

The following is a sample spreadsheet for loci.txt. The example assumes two loci, the first has 10 alleles with even frequencies and the second has 5 with skewed frequencies.

```

2
0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0 0 0 0 0 0 0 0 0 0 0 0
0.5 0.3 0.1 0.05 0.05 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

```

(2) genetics.txt

number of nests

for each nest: number of loci used; number of offspring analyzed; number of paternal alleles observed at each locus; genetic mother’s genotype at each locus (alleles are numbered 1 through 40)

NOTES: For some nests not all loci may be used. In this case the loci must be ordered according to their use. For example, if a maximum of three loci are used then the first locus must be used in all nests, the second locus must be used by all nests with two or more loci, and the third by all nests with three loci. Thus, the program does not allow only the first and third loci to be used for a given nest. Allowable combinations are 1 or 1 and 2 or 1, 2 and 3. The maximum number of nests that can be analyzed is 200 and the maximum number of offspring from any one nest is 100.

The following is a sample spreadsheet for genetics.txt. The example assumes a maximum of two loci are used and five nests are analyzed. As an example, the second row of data in the spreadsheet (first nest) implies that two loci were used, 10 offspring were analyzed, two and three paternal alleles were observed at the first and second loci, respectively, and the mother was homozygous for the first allele at both loci. Allele ‘1’ corresponds to the first allele frequency in loci.txt. The last row (fifth nest) implies that only one locus was used (the first locus), 15 offspring were analyzed, one paternal allele was observed and the mother was heterozygous for the first and second alleles.

```

5
2      10      2      3      1      1      1      1
2      10      2      2      2      2      2      3
2      10      3      3      6      8      4      4
2      10      1      1      3      4      1      1
1      15      1      1      2

```

(3) skew.txt

mean paternity for each male in multiply mated nests

NOTES: Exactly 20 paternity values must be entered and they must sum to 1.0. If fewer than 20 males contribute to a multiply mated brood then enter zero for non-contributing males.

The following is a sample spreadsheet for skew.txt. The example assumes that three males contribute to a multiply mated brood with mean paternities of 0.50, 0.30 and 0.20.

0.50 0.30 0.20 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

(4) prior.txt

prior probability for each possible frequency of multiple mating ranging from 0 to 1 (i.e. 100%) in increments of 1%.

NOTES: Exactly 101 probability values must be entered and they must sum to 1.0. Each number can be entered on a separate line (row) in the spreadsheet. The first number corresponds to the prior probability of a frequency of multiple mating of 0, the second to a frequency of 0.01 (1%), and the 101st to a frequency of 1 (100%). For a uniform prior probability distribution enter 0.0099 for the first value and then 0.09901 for the remaining 100 values.

The data files must be titled as listed above and must be saved in the same file folder (directory) as the C++ program. If any difficulties occur (and are detected) while the program is running, an error message will appear on the screen. Because of the nature of a Monte Carlo simulation (used by the program), some minor fluctuations occur in the probabilities reported each time the program is run with the same data.

The output is displayed on the screen and written to the file fmmresults.txt. This file can be viewed using Excel or other similar programs. The file contains three columns of data. The first is the frequency of multiple mating (range 0 to 100). The second is the probability of each f_{mm} value calculated based on the genetic data (files 1-4 above). The third is a cumulative probability distribution. The expected f_{mm} value is also reported at the bottom. The cumulative distribution can be used to calculate a confidence interval. For example, the 95% CI is calculated by finding the f_{mm} values corresponding to 0.025 and 0.975.