EffMot
Brief description

EffMot estimates the effective number of mothers and its confidence interval based on a data sample collected from seeds. The confidence interval is calculated using the parametric bootstrap method.

Preparing input data files

EffMot requires information about haplotype frequencies in megagametophytes and embryos. The input files can be prepared in two formats, as described below.

1. Haplotype frequency input file.

Format:

N_{HM}		*[number of examined haplotypes in megagametophytes (integer)]*
V_{HM}		*[number of haplotype variants in megagametophytes (integer)]*
i_{HM}	f_{HM}	*[i_{HM} - haplotype variant ID (integer),* *f_{HM} – haplotype variant frequency (real)]*
…		
N_{HE}		*[number of examined haplotypes in embryos (integer)]*
V_{HE}		*[number of haplotype variants in embryos (integer)]*
i_{HE}	f_{HE}	*[i_{HE} - haplotype variant ID (integer),* *f_{HE} – haplotype variant frequency (real)]*
…		

e.g.:

100	
19	
1	0.002
2	0.003
3	0.011
4	0.005
….	
18	0.001
19	0.002
95	
35	
1	0.002
2	0.003
3	0.011
4	0.005
….	
34	0.001
35	0.002

There are two major blocks in this file. The first one contains data from the megagametophytes, and the second one – from the embryos. Each block begins with two lines corresponding to the number of all examined haplotypes (first line) and the number of identified haplotype variants (second line). Next, a tab or space-delimited list of pairs haplotype ID – haplotype frequency follows. The IDs should be consecutive numbers. The number of data points should correspond to the number of variants declared in the second line of the block.

2. Raw haplotype input file.

Format:

N_{LM}	*[number of loci examined in megagametophytes]*
N_{HM}	*[number of haplotypes found in megagametophytes]*
H_1 H_2 … H_{NLM}	*[list of alleles in the first haplotype]*
…	
N_{LE}	*[number of loci examined in embryos]*
N_{HE}	*[number of haplotypes found in emryos]*
H_1 H_2 … H_{NLE}	*[list of alleles in the first haplotype]*
…	

e.g.

```
5
100
112 134 124 122 108
110 130 124 122 108
112 134 124 122 108
112 133 124 122 108
…
110 119 134 122 108
5
87
110 130 124 122 108
112 134 124 122 108
110 132 124 122 108
112 134 124 121 108
…
110 130 124 122 108
```

Similarly as in haplotype frequency input file, there are two data blocks here. The first one contains the data from megagametophytes, and the second one from embryos. In each block the first line is the number of loci, and the second line is the number of all examined haplotypes. The lines that follow represent haplotypes: in each line as many tab/space-delimited numbers are expected as many loci were declared in the first line of the block. Each number represents an allele. Number of the lines must match the number of haplotypes declared in the second line of the block. Any number is integer.

Please, note that both files contain only numbers and they must be separated with either tab or space. In the case of real numbers, a period has to be used as a decimal separator. In the case

of the raw haplotype input file, any missing data should be recorded as '-1'. The haplotypes with data gaps will be excluded from the analysis.

To load the input data file, select the '*Data*' menu. After the data is successfully loaded, a message will show on the screen and the '*Show data*' menu will become enabled. This menu allows the user to display the data from the file on the screen. If the input file contains the raw haplotype data, the haplotypes along with their frequencies, as estimated by the program, will be shown.

Running the program

After the input data is successfully loaded, the '*Run*' menu will become enabled. The default values for simulation parameters (the number of iterations and the confidence level) can be changed using the '*Options*' menu.

Please, note that the computation time is proportional to the number of iterations. Please select this number accordingly to the resources at hand.

To start the analysis, select the '*Run*' menu. The results will be output to the screen, e.g.:

```
Female haplotype identity: 0.050096
Male haplotype identity: 0.016066

Effective number of mothers: 28.913329

Simulation parameters:
        Number of iteration: 1000
        Confidence level: 0.95
0.95-confidence interval:
        [ 22.739828 ; 33.228956 ]
```

Simulations can be run repeatedly. All results can be saved to a file.