

migrate-n work sheet

Plan

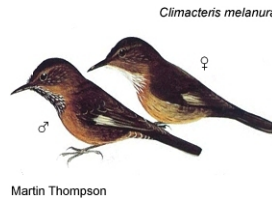
1. Introduction, very short demo
2. Hands-On
3. Evaluation of results on blackboard

A reminder: if you want to keep results (they are written by default into a file called `outfile` and `outfile.pdf`), save it using either `cp outfile yourfavoredfilename` or `mv outfile yourfavoredfilename` or change the name of the output file in the menu.

Dataset

you can find them in <http://workshop.molecularrevolution.org/software/lamarc/>

- `infile.treecreeper`: a two-population dataset with 2 loci.



- `infile.msat`: an example of a microsatellite data set (simulated data)
- `infile.3pop`: 3 populations of the treecreepers, 2 loci. 1

Hands-on

1. Run the 2-population treecreeper data data set in MIGRATE-N (Bayes-mode)
 - (a) Check where the data set is located (and remember the name)
 - (b) Start MIGRATE-N from the commandline in the terminal
 - (c) Explore the menus (especially the STRATEGY, INPUT/OUTPUT)
 - (d) In the STRATEGY submenu, switch to Bayesian inference
 - (e) Adjust the number of long chains to 1.
 - (f) Check that the prior is uniform and in the range of 0.000001 to 0.1 for Θ (Theta) and 0 to 2000 for M (the migration parameter). You need to specify 3 numbers for the uniform distribution : minimum maximum interval, the interval is best set around 1/10 of the difference of maximum and minimum.
 - (g) Back out to the main menu and SAVE the parameter file (“Write a parmfile”)

- (h) Run (this should take a couple of seconds), inspect the print to the screen, it shows effective sample size, acceptance ratios (if you use Slice sampling instead of Metropolis-Hastings for the parameter changes, then the acceptance ratio will be 1.0).
- (i) Look at the output: there is a TEXT file and a PDF file. On Macs, you can open the PDF file from the commandline with `open filename`. Check out the posterior distribution.
- (j) Start MIGRATE-N or (on Macs) try out FASTMIGRATE-N (FASTMIGRATE-N is a shell script that calls the parallel version of MIGRATE-N, you should not see any difference in the menu except that it shows in the title that this is the parallel version. It will run about twice as fast for this 2 locus data set, but will only write a minimal report to the screen.)
- (k) Adjust the number of long chains to 1, change the number of recorded long steps to 50000 and then run (will take a couple of minutes).
- (l) Write the value of the mode for the combined estimates of Θ_2 and $M_{1 \rightarrow 2}$ on the Blackboard (towards the end of the demo we will tally the numbers up and look at the variance of the estimates).

2. Likelihood ratio tests using the ML module in MIGRATE-N; use `infile.treecreeper`

- (a) Rename the file called parmfile (for example with `mv parmfile parmfile.bayes`)
- (b) Start MIGRATE-N (check the STRATEGY menu, it should show Maximum Likelihood as the inference strategy)
- (c) Shorten the **samples** of long and short chains by a factor of 10.
- (d) Go back to the main menu, go to INPUT/OUTPUT menu and from there to the likelihood ratio test menu. The input is rather complicated: enter the number of populations (2). A possible test is the test whether the migration rates are the same. MIGRATE-N uses * for parameters that are not of interest in the test. The default model for two population is organized in a matrix with Θ on the diagonal, the parameters are organized like this:

$$\begin{array}{|c|c|} \hline \Theta_1 & M_{21} \\ \hline M_{12} & \Theta_1 \\ \hline \end{array} \rightarrow \begin{array}{|c|c|c|c|} \hline \Theta_1 & M_{21} & M_{12} & \Theta_1 \\ \hline \end{array}$$

Possible tests are: are population sizes the same, are migration rates the same, are migration rate zero in one direction,

This can be expressed as the following strings:

Same sizes: enter **m * * m**

Same rates: enter *** m m ***

Same rates and sizes: enter **m m m m**

Direction: enter *** 0 * ***

You can enter any number of different tests one after another. Type **end** once your done. Without the 'end' the program will keep asking to enter hypotheses.

- (e) Run

- (f) If the run above was done in short time, redo the analysis with 10x higher values (increase the increment for the long and short chains).
- (g) We will discuss the LRT once all reached this point (keep your output files, LRT is not yet implemented in the PDF files!).

3. Explore **AFTER** you run the worksheet for **Lamar**

- Replace the Uniform prior with an Exponential prior (Bayes)
- Heating (ML or Bayes)
- Replication (Bayes or ML) **Use fastmigrate-n**
- Use `infile.3pop` Add the custom migration matrix: in the menu **Parameter** adjust submenu number 7 under the parameter menu: The custom migration matrix is somewhat difficult to enter: I typically draw a migration matrix on paper, for this case this would look like this:

$$\begin{array}{rcccl}
 & \textit{Doomagee} & & * & * & 0 \\
 \textit{DouglasHotSprings} & & & * & * & 0 \\
 \textit{FitzroyCrossing} & & 0 & * & * &
 \end{array}$$

Linearized this will look like this:

$$\begin{array}{cccccccc}
 * & * & 0 & * & * & 0 & 0 & * & *
 \end{array}$$

- 4. Use a two-population dataset. In the Input/Output section turn on the **Event histogram** option (use the *All events, record every event, set the timeinterval to 0.001* and also turn on the **skyline plot**).