# *Preliminary migrate analysis of M. californianus*

MIGRATION RATE AND POPULATION SIZE ESTIMATION
using the coalescent and maximum likelihood or Bayesian inference
Migrate-n version 3.7.2 [April-12-18]
Program started at Tue Jun 1 13:45:41 2021
Program finished at Tue Jun 1 18:20:38 2021

## *Options*

Datatype: DNA sequence data

Inheritance scalers in use for Thetas:
All loci use an inheritance scaler of 1.0
[The locus with a scaler of 1.0 used as reference]
Random number seed: (with internal timer) 1635060950
Start parameters:

Theta values were generated from guessed values
Theta = 0.01000
M values were generated from guessed values
M-matrix:
100000.00 [all are the same]
Connection type matrix:
where m = average (average over a group of Thetas or M,
s = symmetric M, S = symmetric 4Nm, 0 = zero, and not estimated,
* = free to vary, Thetas are on diagonal

| Population | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 ElfinCo | * | * | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 Bamfiel | * | * | * | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 PortRen | 0 | * | * | * | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 WalkOnB | 0 | 0 | * | * | * | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 BodegaH | 0 | 0 | 0 | * | * | * | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 Davenpo | 0 | 0 | 0 | 0 | * | * | * | 0 | 0 | 0 | 0 | 0 |
| 7 VistaDe | 0 | 0 | 0 | 0 | 0 | * | * | * | 0 | 0 | 0 | 0 |
| 8 HazardR | 0 | 0 | 0 | 0 | 0 | 0 | * | * | * | 0 | 0 | 0 |
| 9 Refugio | 0 | 0 | 0 | 0 | 0 | 0 | 0 | * | * | * | 0 | 0 |
| 10 Carpint | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | * | * | * | 0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 WhitePo | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | * | * | * |
| 12 LaJolla | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | * | * |

Order of parameters:

| | | |
|---|---|---|
| 1 | $\Theta_1$ | <displayed> |
| 2 | $\Theta_2$ | <displayed> |
| 3 | $\Theta_3$ | <displayed> |
| 4 | $\Theta_4$ | <displayed> |
| 5 | $\Theta_5$ | <displayed> |
| 6 | $\Theta_6$ | <displayed> |
| 7 | $\Theta_7$ | <displayed> |
| 8 | $\Theta_8$ | <displayed> |
| 9 | $\Theta_9$ | <displayed> |
| 10 | $\Theta_{10}$ | <displayed> |
| 11 | $\Theta_{11}$ | <displayed> |
| 12 | $\Theta_{12}$ | <displayed> |
| 13 | $M_{2\text{->}1}$ | <displayed> |
| 24 | $M_{1\text{->}2}$ | <displayed> |
| 25 | $M_{3\text{->}2}$ | <displayed> |
| 36 | $M_{2\text{->}3}$ | <displayed> |
| 37 | $M_{4\text{->}3}$ | <displayed> |
| 48 | $M_{3\text{->}4}$ | <displayed> |
| 49 | $M_{5\text{->}4}$ | <displayed> |
| 60 | $M_{4\text{->}5}$ | <displayed> |
| 61 | $M_{6\text{->}5}$ | <displayed> |
| 72 | $M_{5\text{->}6}$ | <displayed> |
| 73 | $M_{7\text{->}6}$ | <displayed> |
| 84 | $M_{6\text{->}7}$ | <displayed> |
| 85 | $M_{8\text{->}7}$ | <displayed> |
| 96 | $M_{7\text{->}8}$ | <displayed> |
| 97 | $M_{9\text{->}8}$ | <displayed> |
| 108 | $M_{8\text{->}9}$ | <displayed> |
| 109 | $M_{10\text{->}9}$ | <displayed> |
| 120 | $M_{9\text{->}10}$ | <displayed> |
| 121 | $M_{11\text{->}10}$ | <displayed> |
| 132 | $M_{10\text{->}11}$ | <displayed> |
| 133 | $M_{12\text{->}11}$ | <displayed> |
| 144 | $M_{11\text{->}12}$ | <displayed> |

Mutation rate among loci: Mutation rate is constant

Analysis strategy: Bayesian inference

Proposal distributions for parameter

| Parameter | Proposal |
|---|---|
| Theta | Metropolis sampling |
| M | Slice sampling |

Prior distribution for parameter

| Parameter | Prior | Minimum | Mean* | Maximum | Delta | Bins |
|---|---|---|---|---|---|---|
| Theta | Exp window | 0.000010 | 0.010000 | 10.000000 | 1.000000 | 500 |
| M | Exp window | 0.000100 | 100000.000000 | 1000000.000000 | 100000.000000 | 500 |

| Markov chain settings: | Long chain |
|---|---|
| Number of chains | 1 |
| Recorded steps [a] | 1000 |
| Increment (record every x step [b] | 100 |
| Number of concurrent chains (replicates) [c] | 3 |
| Visited (sampled) parameter values [a*b*c] | 300000 |
| Number of discard trees per chain (burn-in) | 1000 |

Multiple Markov chains:

Static heating scheme — 4 chains with temperatures

100000.00 3.00 1.50 1.00

Swapping interval is 1

| Print options: | |
|---|---|
| Data file: | ../../mcalifornianus_210528.mig |
| Output file: | outfile.txt |
| Posterior distribution raw histogram file: | bayesfile |
| Print data: | No |
| Print genealogies [only some for some data type]: | None |

# *Data summary*

| | |
|---|---|
| Datatype: | Sequence data |
| Number of loci: | 1 |

| Population | Locus | Gene copies |
|---|---|---|
| 1 ElfinCo | 1 | 19 |
| 2 Bamfiel | 1 | 23 |
| 3 PortRen | 1 | 15 |
| 4 WalkOnB | 1 | 16 |
| 5 BodegaH | 1 | 7 |
| 6 Davenpo | 1 | 17 |
| 7 VistaDe | 1 | 19 |
| 8 HazardR | 1 | 23 |
| 9 Refugio | 1 | 16 |
| 10 Carpint | 1 | 19 |
| 11 WhitePo | 1 | 11 |
| 12 LaJolla | 1 | 8 |
| Total of all populations | 1 | 193 |

# *Bayesian Analysis: Posterior distribution table*

| Locus | Parameter | 2.5% | 25.0% | Mode | 75.0% | 97.5% | Median | Mean |
|---|---|---|---|---|---|---|---|---|
| 1 | $\Theta_1$ | 0.00001 | 0.00001 | 0.01001 | 0.08001 | 0.20001 | 0.09001 | 0.03684 |
| 1 | $\Theta_2$ | 0.00001 | 0.00001 | 0.01001 | 0.08001 | 0.20001 | 0.09001 | 0.03361 |
| 1 | $\Theta_3$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.01734 |
| 1 | $\Theta_4$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.02044 |
| 1 | $\Theta_5$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.01496 |
| 1 | $\Theta_6$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.00883 |
| 1 | $\Theta_7$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.18001 | 0.07001 | 0.02431 |
| 1 | $\Theta_8$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.18001 | 0.07001 | 0.02608 |
| 1 | $\Theta_9$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.18001 | 0.07001 | 0.02627 |
| 1 | $\Theta_{10}$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.00573 |
| 1 | $\Theta_{11}$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.01015 |
| 1 | $\Theta_{12}$ | 0.00001 | 0.00001 | 0.01001 | 0.06001 | 0.16001 | 0.07001 | 0.00824 |
| 1 | $M_{2->1}$ | 52000.0 | 64000.0 | 79000.0 | 92000.0 | 106000.0 | 73000.0 | 57379.1 |
| 1 | $M_{1->2}$ | 2000.0 | 10000.0 | 25000.0 | 36000.0 | 46000.0 | 33000.0 | 40256.9 |
| 1 | $M_{3->2}$ | 0.0 | 16000.0 | 33000.0 | 40000.0 | 50000.0 | 29000.0 | 23404.1 |
| 1 | $M_{2->3}$ | 4000.0 | 58000.0 | 71000.0 | 78000.0 | 94000.0 | 53000.0 | 50311.1 |
| 1 | $M_{4->3}$ | 34000.0 | 48000.0 | 61000.0 | 72000.0 | 112000.0 | 67000.0 | 67936.5 |
| 1 | $M_{3->4}$ | 0.0 | 8000.0 | 23000.0 | 34000.0 | 66000.0 | 31000.0 | 30476.6 |
| 1 | $M_{5->4}$ | 30000.0 | 38000.0 | 53000.0 | 64000.0 | 120000.0 | 99000.0 | 106631.7 |
| 1 | $M_{4->5}$ | 40000.0 | 88000.0 | 103000.0 | 116000.0 | 174000.0 | 101000.0 | 102938.8 |
| 1 | $M_{6->5}$ | 34000.0 | 46000.0 | 67000.0 | 86000.0 | 102000.0 | 81000.0 | 126415.2 |
| 1 | $M_{5->6}$ | 10000.0 | 52000.0 | 69000.0 | 82000.0 | 90000.0 | 61000.0 | 54625.4 |
| 1 | $M_{7->6}$ | 46000.0 | 76000.0 | 101000.0 | 114000.0 | 130000.0 | 95000.0 | 90724.2 |
| 1 | $M_{6->7}$ | 42000.0 | 62000.0 | 81000.0 | 92000.0 | 104000.0 | 63000.0 | 53666.0 |
| 1 | $M_{8->7}$ | 64000.0 | 78000.0 | 89000.0 | 98000.0 | 112000.0 | 91000.0 | 89440.0 |
| 1 | $M_{7->8}$ | 32000.0 | 40000.0 | 55000.0 | 68000.0 | 104000.0 | 63000.0 | 65871.1 |
| 1 | $M_{9->8}$ | 66000.0 | 74000.0 | 91000.0 | 104000.0 | 114000.0 | 85000.0 | 76359.2 |
| 1 | $M_{8->9}$ | 22000.0 | 56000.0 | 73000.0 | 86000.0 | 94000.0 | 67000.0 | 61410.5 |
| 1 | $M_{10->9}$ | 8000.0 | 18000.0 | 39000.0 | 56000.0 | 68000.0 | 51000.0 | 70936.2 |
| 1 | $M_{9->10}$ | 38000.0 | 54000.0 | 65000.0 | 82000.0 | 104000.0 | 73000.0 | 71023.9 |
| 1 | $M_{11->10}$ | 32000.0 | 78000.0 | 95000.0 | 110000.0 | 122000.0 | 89000.0 | 81429.0 |
| 1 | $M_{10->11}$ | 6000.0 | 14000.0 | 27000.0 | 40000.0 | 82000.0 | 65000.0 | 76879.8 |
| 1 | $M_{12->11}$ | 48000.0 | 66000.0 | 83000.0 | 92000.0 | 114000.0 | 85000.0 | 82224.9 |
| 1 | $M_{11->12}$ | 0.0 | 2000.0 | 15000.0 | 28000.0 | 34000.0 | 97000.0 | 92670.4 |

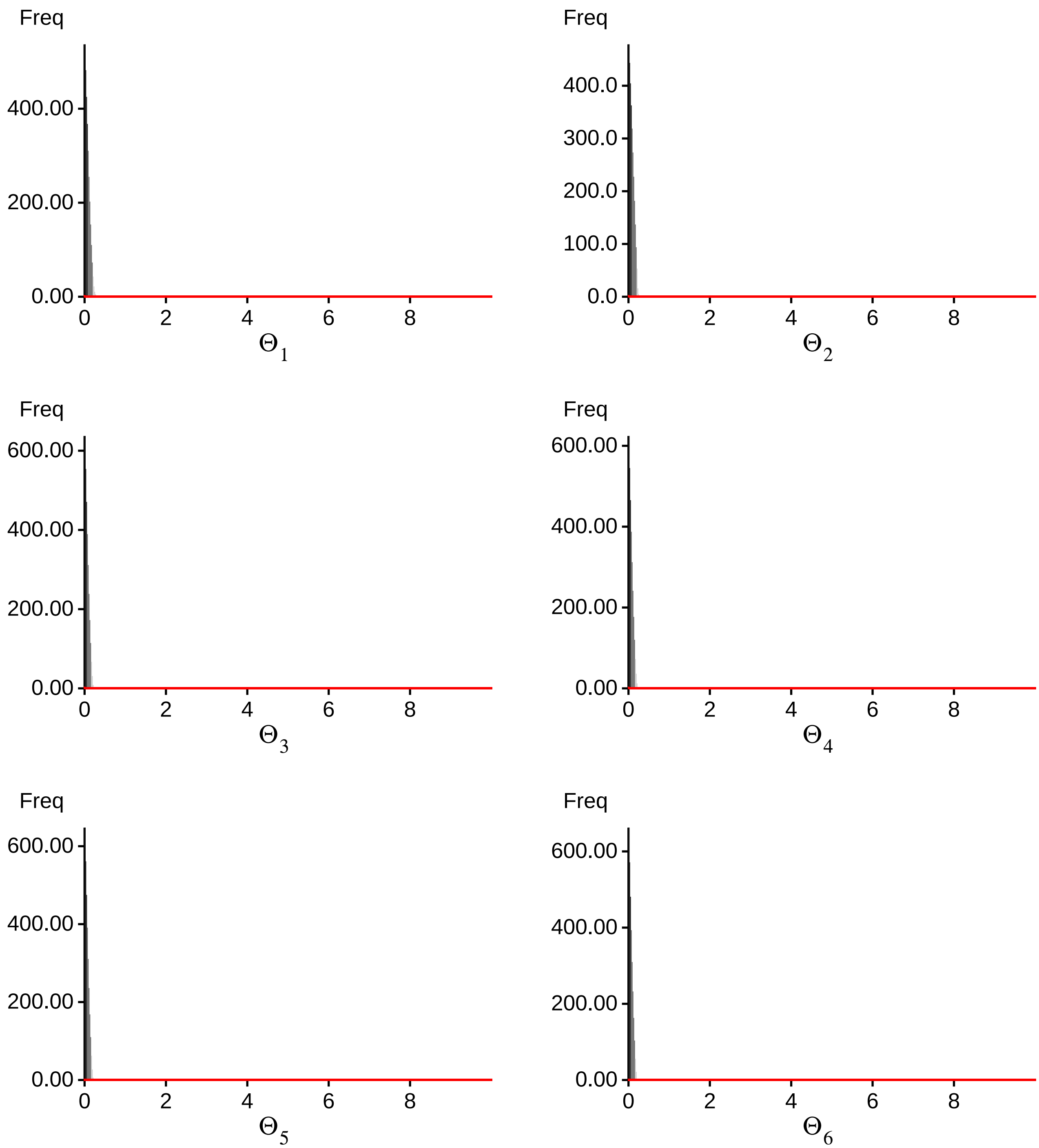
Citation suggestions:

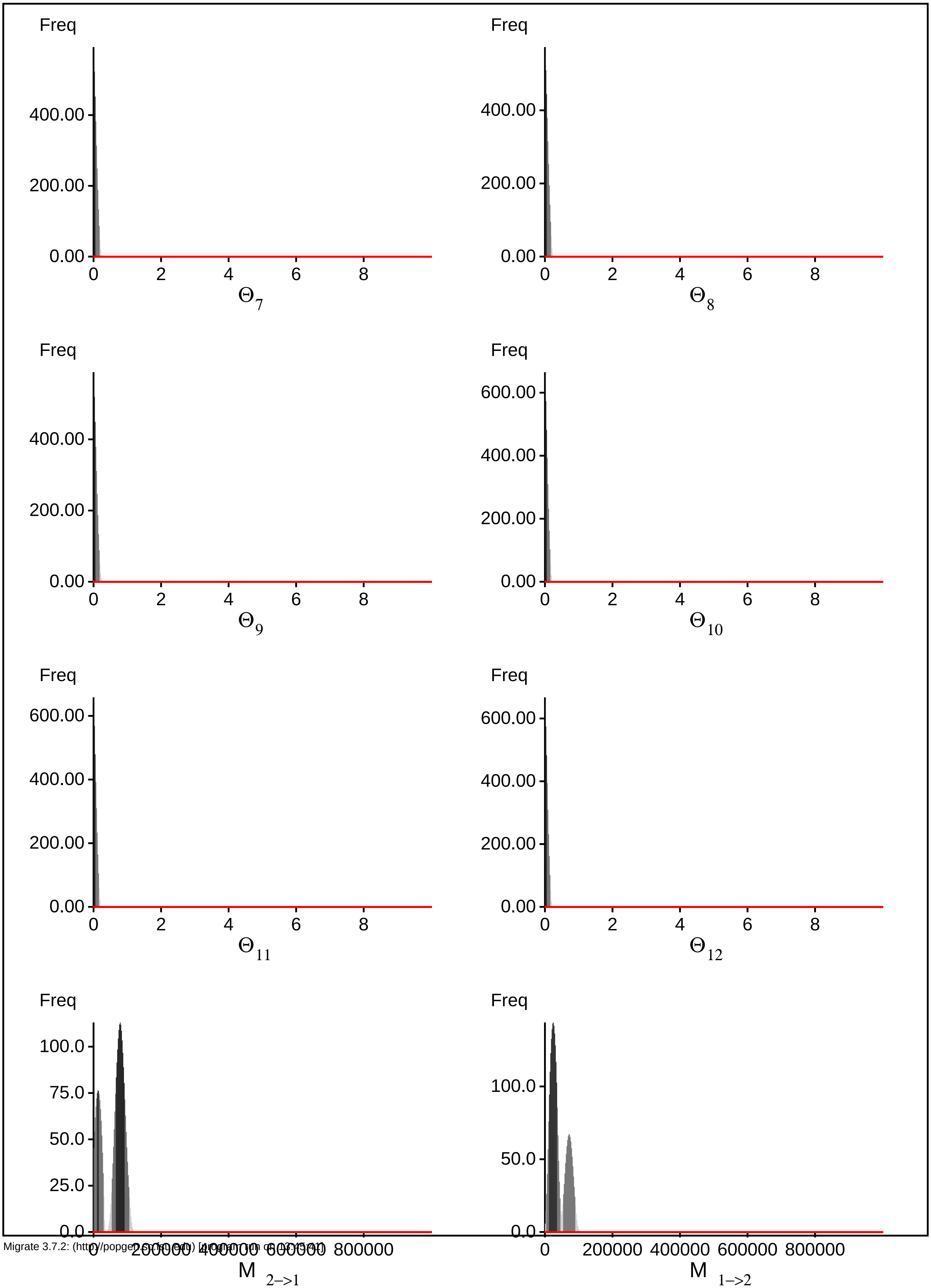
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341-345

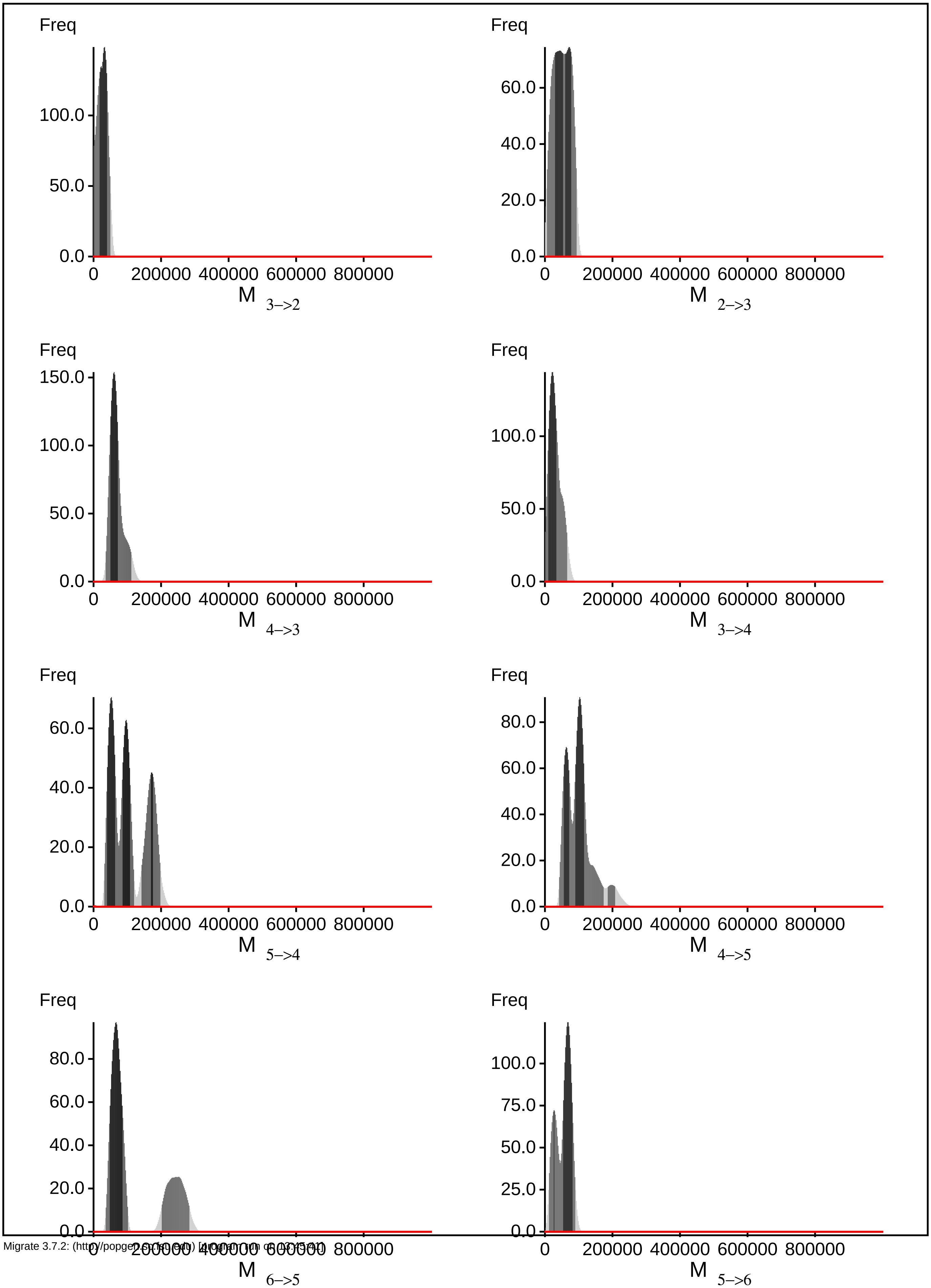
Beerli P., 2007. Estimation of the population scaled mutation rate from microsatellite data, Genetics, 177:1967-1968.

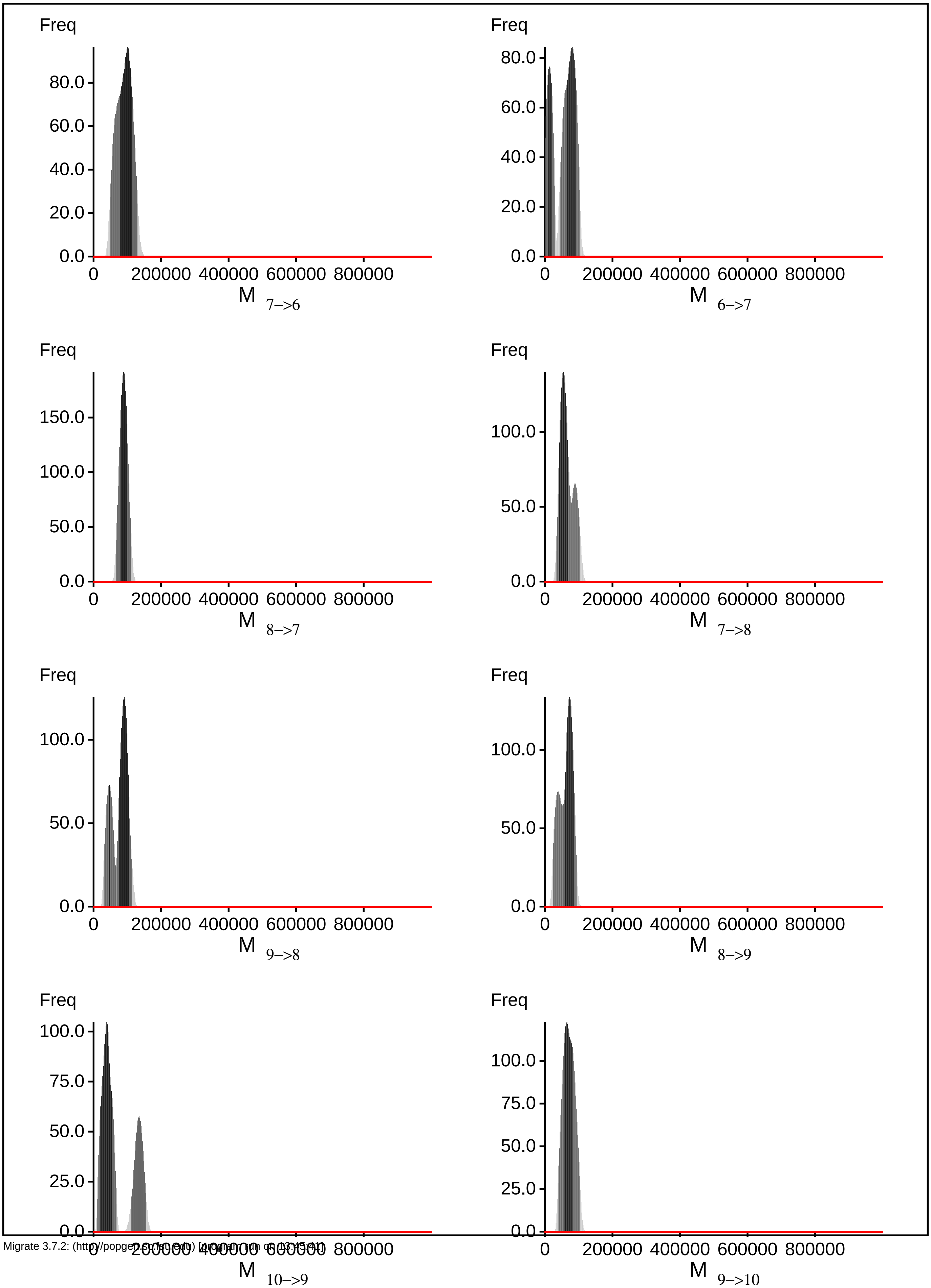
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp. 42-79.

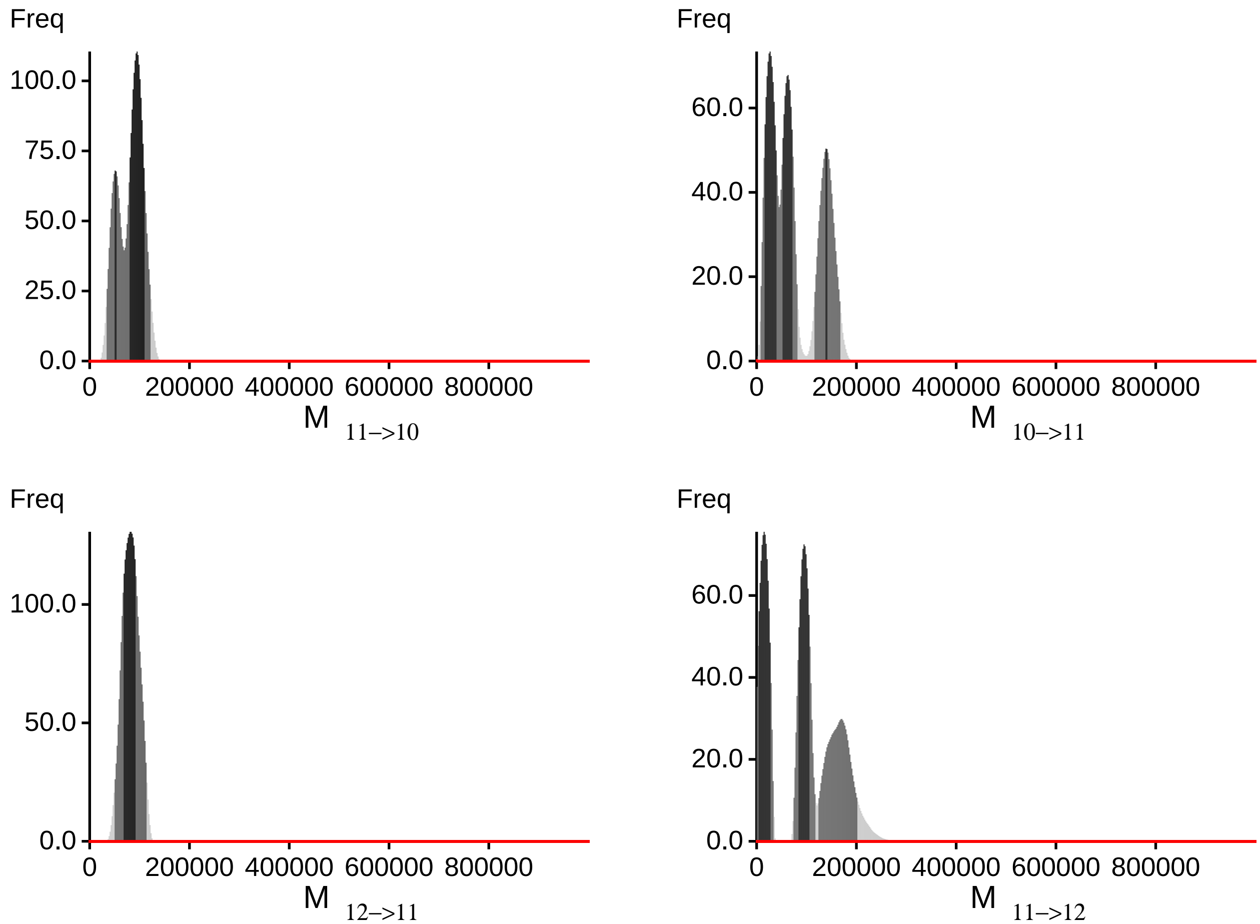
## *Bayesian Analysis: Posterior distribution over all loci*




Freq
400.00
200.00
0.00
0
2
4
6
8
Θ7
Freq
400.00
200.00
0.00
0
2
4
6
8
Θ8
Freq
400.00
200.00
0.00
0
2
4
6
8
Θ9
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ10
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ11
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ12
Freq
100.0
75.0
50.0
25.0
0.0
0
200000
400000
600000
800000
M 2->1
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 1->2


Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 3–>2
Freq
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 2–>3
Freq
150.0
100.0
50.0
0.0
0
200000
400000
600000
800000
M 4–>3
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 3–>4
Freq
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 5–>4
Freq
80.0
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 4–>5
Freq
80.0
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 6–>5
Freq
100.0
75.0
50.0
25.0
0.0
0
200000
400000
600000
800000
M 5–>6


Freq
80.0
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 7–>6
Freq
80.0
60.0
40.0
20.0
0.0
M 6–>7
Freq
150.0
100.0
50.0
0.0
M 8–>7
Freq
100.0
50.0
0.0
M 7–>8
Freq
100.0
50.0
0.0
M 9–>8
Freq
100.0
50.0
0.0
M 8–>9
Freq
100.0
75.0
50.0
25.0
0.0
M 10–>9
Freq
100.0
75.0
50.0
25.0
0.0
M 9–>10


Freq
100.0
75.0
50.0
25.0
0.0
0
200000
400000
600000
800000
M 11–>10
Freq
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 10–>11
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 12–>11
Freq
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 11–>12

## *Log-Probability of the data given the model (marginal likelihood)*

Use this value for Bayes factor calculations:
BF = Exp[ ln(Prob(D | thisModel) - ln( Prob( D | otherModel)
or as LBF = 2 (ln(Prob(D | thisModel) - ln( Prob( D | otherModel))
shows the support for thisModel]

| Method | ln(Prob(D\|Model)) | Notes |
|---|---|---|
| Thermodynamic integration | -2215.773818 | (1a) |
| | -2147.116214 | (1b) |
| Harmonic mean | -1848.243092 | (2) |

(1a, 1b and 2) are approximations to the marginal likelihood, make sure that the program run long enough!
(1a, 1b) and (2) should give similar results, in principle.
But (2) is overestimating the likelihood, it is presented for historical reasons and should not be used
(1a, 1b) needs heating with chains that span a temperature range of 1.0 to at least 100,000.
(1b) is using a Bezier-curve to get better approximations for runs with low number of heated chains

Citation suggestions:

Beerli P. and M. Palczewski, 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations, Genetics, 185: 313-326.

## *Acceptance ratios for all parameters and the genealogies*

| Parameter | Accepted changes | Ratio |
|---|---|---|
| $\Theta_1$ | 977/4539 | 0.21525 |
| $\Theta_2$ | 643/4405 | 0.14597 |
| $\Theta_3$ | 2333/4392 | 0.53119 |
| $\Theta_4$ | 1730/4395 | 0.39363 |
| $\Theta_5$ | 2258/4396 | 0.51365 |
| $\Theta_6$ | 1369/4326 | 0.31646 |
| $\Theta_7$ | 1107/4394 | 0.25193 |
| $\Theta_8$ | 960/4306 | 0.22294 |
| $\Theta_9$ | 573/4388 | 0.13058 |
| $\Theta_{10}$ | 732/4403 | 0.16625 |
| $\Theta_{11}$ | 1440/4505 | 0.31964 |
| $\Theta_{12}$ | 1371/4432 | 0.30934 |
| $M_{2->1}$ | 4256/4256 | 1.00000 |
| $M_{1->2}$ | 4405/4405 | 1.00000 |
| $M_{3->2}$ | 4409/4409 | 1.00000 |
| $M_{2->3}$ | 4501/4501 | 1.00000 |
| $M_{4->3}$ | 4435/4435 | 1.00000 |
| $M_{3->4}$ | 4384/4384 | 1.00000 |
| $M_{5->4}$ | 4433/4433 | 1.00000 |
| $M_{4->5}$ | 4373/4373 | 1.00000 |
| $M_{6->5}$ | 4448/4448 | 1.00000 |
| $M_{5->6}$ | 4352/4352 | 1.00000 |
| $M_{7->6}$ | 4461/4461 | 1.00000 |
| $M_{6->7}$ | 4552/4552 | 1.00000 |
| $M_{8->7}$ | 4408/4408 | 1.00000 |
| $M_{7->8}$ | 4449/4449 | 1.00000 |
| $M_{9->8}$ | 4389/4389 | 1.00000 |
| $M_{8->9}$ | 4317/4317 | 1.00000 |
| $M_{10->9}$ | 4369/4369 | 1.00000 |
| $M_{9->10}$ | 4381/4381 | 1.00000 |
| $M_{11->10}$ | 4500/4500 | 1.00000 |
| $M_{10->11}$ | 4462/4462 | 1.00000 |
| $M_{12->11}$ | 4420/4420 | 1.00000 |
| $M_{11->12}$ | 4436/4436 | 1.00000 |
| Genealogies | 33231/149979 | 0.22157 |

## *MCMC-Autocorrelation and Effective MCMC Sample Size*

| Parameter | Autocorrelation | Effective Sampe Size |
|---|---|---|
| $\Theta_1$ | 0.84160 | 270.94 |
| $\Theta_2$ | 0.88565 | 184.43 |
| $\Theta_3$ | 0.58274 | 820.63 |
| $\Theta_4$ | 0.76043 | 408.70 |
| $\Theta_5$ | 0.63980 | 705.33 |
| $\Theta_6$ | 0.72382 | 480.64 |
| $\Theta_7$ | 0.83246 | 276.09 |
| $\Theta_8$ | 0.88476 | 187.84 |
| $\Theta_9$ | 0.90499 | 153.74 |
| $\Theta_{10}$ | 0.87697 | 207.68 |
| $\Theta_{11}$ | 0.81898 | 311.43 |
| $\Theta_{12}$ | 0.78908 | 419.27 |
| $M_{2->1}$ | 0.71929 | 492.60 |
| $M_{1->2}$ | 0.73055 | 468.57 |
| $M_{3->2}$ | 0.74876 | 432.94 |
| $M_{2->3}$ | 0.71545 | 498.62 |
| $M_{4->3}$ | 0.75136 | 433.06 |
| $M_{3->4}$ | 0.80557 | 335.32 |
| $M_{5->4}$ | 0.73009 | 477.93 |
| $M_{4->5}$ | 0.75679 | 433.25 |
| $M_{6->5}$ | 0.76870 | 393.95 |
| $M_{5->6}$ | 0.79989 | 338.73 |
| $M_{7->6}$ | 0.77009 | 391.01 |
| $M_{6->7}$ | 0.73034 | 480.28 |
| $M_{8->7}$ | 0.77361 | 387.11 |
| $M_{7->8}$ | 0.78180 | 377.37 |
| $M_{9->8}$ | 0.80157 | 337.52 |
| $M_{8->9}$ | 0.84502 | 251.87 |
| $M_{10->9}$ | 0.81018 | 315.70 |
| $M_{9->10}$ | 0.67713 | 591.27 |
| $M_{11->10}$ | 0.77834 | 394.17 |
| $M_{10->11}$ | 0.85435 | 235.61 |
| $M_{12->11}$ | 0.79305 | 345.92 |
| $M_{11->12}$ | 0.85337 | 239.26 |
| Ln[Prob(D\|G)] | 0.97325 | 40.64 |

# *Potential Problems*

This section reports potential problems with your run, but such reporting is often not very accurate. Whith many parameters in a multilocus analysis, it is very common that some parameters for some loci will not be very informative, triggering suggestions (for example to increase the prior range) that are not sensible. This suggestion tool will improve with time, therefore do not blindly follow its suggestions. If some parameters are flagged, inspect the tables carefully and judge wether an action is required. For example, if you run a Bayesian inference with sequence data, for macroscopic species there is rarely the need to increase the prior for Theta beyond 0.1; but if you use microsatellites it is rather common that your prior distribution for Theta should have a range from 0.0 to 100 or more. With many populations (>3) it is also very common that some migration routes are estimated poorly because the data contains little or no information for that route. Increasing the range will not help in such situations, reducing number of parameters may help in such situations.

No warning was recorded during the run