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   Wed Jun 2 13:54:44 2021
Program finished at   Wed Jun 2 21:23:29 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)   2409929486 |

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	*	s	0	0	0	0	0	0	0	0	0	0
2 Bamfiel	s	*	s	0	0	0	0	0	0	0	0	0
3 PortRen	0	s	*	s	0	0	0	0	0	0	0	0
4 WalkOnB	0	0	s	*	s	0	0	0	0	0	0	0
5 BodegaH	0	0	0	s	*	s	0	0	0	0	0	0
6 Davenpo	0	0	0	0	s	*	s	0	0	0	0	0
7 VistaDe	0	0	0	0	0	s	*	s	0	0	0	0
8 HazardR	0	0	0	0	0	0	s	*	s	0	0	0
9 Refugio	0	0	0	0	0	0	0	s	*	s	0	0
10 Carpint	0	0	0	0	0	0	0	0	s	*	s	0

11 WhitePo	0	0	0	0	0	0	0	0	0	s	*	s	
12 LaJolla	0	0	0	0	0	0	0	0	0	s	*		

Order of parameters:

1	Θ_1		<displayed>
2	Θ_2		<displayed>
3	Θ_3		<displayed>
4	Θ_4		<displayed>
5	Θ_5		<displayed>
6	Θ_6		<displayed>
7	Θ_7		<displayed>
8	Θ_8		<displayed>
9	Θ_9		<displayed>
10	Θ_{10}		<displayed>
11	Θ_{11}		<displayed>
12	Θ_{12}		<displayed>
13	$M_{2\rightarrow1}$ =	$M_{2\rightarrow1}$ [s]	<displayed>
24	$M_{1\rightarrow2}$ =	$M_{2\rightarrow1}$ [s]	
25	$M_{3\rightarrow2}$ =	$M_{3\rightarrow2}$ [s]	<displayed>
36	$M_{2\rightarrow3}$ =	$M_{3\rightarrow2}$ [s]	
37	$M_{4\rightarrow3}$ =	$M_{4\rightarrow3}$ [s]	<displayed>
48	$M_{3\rightarrow4}$ =	$M_{4\rightarrow3}$ [s]	
49	$M_{5\rightarrow4}$ =	$M_{5\rightarrow4}$ [s]	<displayed>
60	$M_{4\rightarrow5}$ =	$M_{5\rightarrow4}$ [s]	
61	$M_{6\rightarrow5}$ =	$M_{6\rightarrow5}$ [s]	<displayed>
72	$M_{5\rightarrow6}$ =	$M_{6\rightarrow5}$ [s]	
73	$M_{7\rightarrow6}$ =	$M_{7\rightarrow6}$ [s]	<displayed>
84	$M_{6\rightarrow7}$ =	$M_{7\rightarrow6}$ [s]	
85	$M_{8\rightarrow7}$ =	$M_{8\rightarrow7}$ [s]	<displayed>
96	$M_{7\rightarrow8}$ =	$M_{8\rightarrow7}$ [s]	
97	$M_{9\rightarrow8}$ =	$M_{9\rightarrow8}$ [s]	<displayed>
108	$M_{8\rightarrow9}$ =	$M_{9\rightarrow8}$ [s]	
109	$M_{10\rightarrow9}$ =	$M_{10\rightarrow9}$ [s]	<displayed>
120	$M_{9\rightarrow10}$ =	$M_{10\rightarrow9}$ [s]	
121	$M_{11\rightarrow10}$ =	$M_{11\rightarrow10}$ [s]	<displayed>
132	$M_{10\rightarrow11}$ =	$M_{11\rightarrow10}$ [s]	
133	$M_{12\rightarrow11}$ =	$M_{12\rightarrow11}$ [s]	<displayed>
144	$M_{11\rightarrow12}$ =	$M_{12\rightarrow11}$ [s]	

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	Θ_1	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02667
1	Θ_2	0.00001	0.00001	0.01001	0.08001	0.18001	0.09001	0.03165
1	Θ_3	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02688
1	Θ_4	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02019
1	Θ_5	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01219
1	Θ_6	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01330
1	Θ_7	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01261
1	Θ_8	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.00982
1	Θ_9	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.00742
1	Θ_{10}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01344
1	Θ_{11}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01398
1	Θ_{12}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01365
1	$M_{2\to1}$	18000.0	26000.0	41000.0	54000.0	62000.0	51000.0	58128.4
1	$M_{1\to2}$	18000.0	26000.0	41000.0	54000.0	62000.0	51000.0	58128.4
1	$M_{3\to2}$	0.0	6000.0	23000.0	34000.0	48000.0	35000.0	53409.7
1	$M_{2\to3}$	0.0	6000.0	23000.0	34000.0	48000.0	35000.0	53409.7
1	$M_{4\to3}$	50000.0	58000.0	75000.0	86000.0	96000.0	69000.0	52756.5
1	$M_{3\to4}$	50000.0	58000.0	75000.0	86000.0	96000.0	69000.0	52756.5
1	$M_{5\to4}$	12000.0	22000.0	39000.0	52000.0	64000.0	49000.0	58955.4
1	$M_{4\to5}$	12000.0	22000.0	39000.0	52000.0	64000.0	49000.0	58955.4
1	$M_{6\to5}$	56000.0	64000.0	79000.0	92000.0	98000.0	73000.0	59832.2
1	$M_{5\to6}$	56000.0	64000.0	79000.0	92000.0	98000.0	73000.0	59832.2
1	$M_{7\to6}$	2000.0	12000.0	29000.0	42000.0	54000.0	39000.0	43258.3
1	$M_{6\to7}$	2000.0	12000.0	29000.0	42000.0	54000.0	39000.0	43258.3
1	$M_{8\to7}$	6000.0	12000.0	27000.0	38000.0	44000.0	35000.0	41267.6
1	$M_{7\to8}$	6000.0	12000.0	27000.0	38000.0	44000.0	35000.0	41267.6
1	$M_{9\to8}$	0.0	0.0	13000.0	26000.0	36000.0	27000.0	31632.0
1	$M_{8\to9}$	0.0	0.0	13000.0	26000.0	36000.0	27000.0	31632.0
1	$M_{10\to9}$	4000.0	14000.0	29000.0	42000.0	78000.0	39000.0	40559.2
1	$M_{9\to10}$	4000.0	14000.0	29000.0	42000.0	78000.0	39000.0	40559.2
1	$M_{11\to10}$	0.0	12000.0	21000.0	30000.0	42000.0	23000.0	21433.9
1	$M_{10\to11}$	0.0	12000.0	21000.0	30000.0	42000.0	23000.0	21433.9
1	$M_{12\to11}$	10000.0	24000.0	33000.0	42000.0	54000.0	35000.0	33790.2
1	$M_{11\to12}$	10000.0	24000.0	33000.0	42000.0	54000.0	35000.0	33790.2

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.

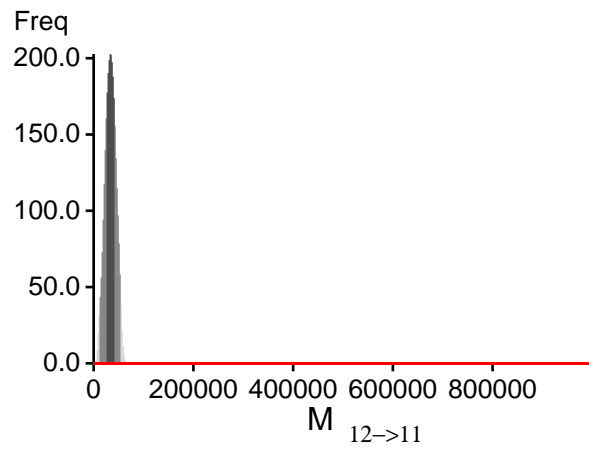# *Bayesian Analysis: Posterior distribution over all loci*
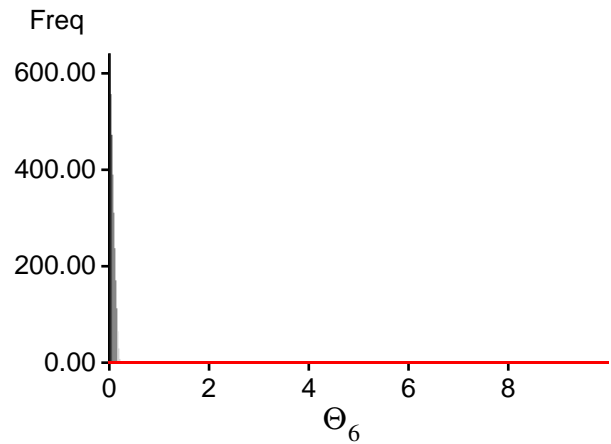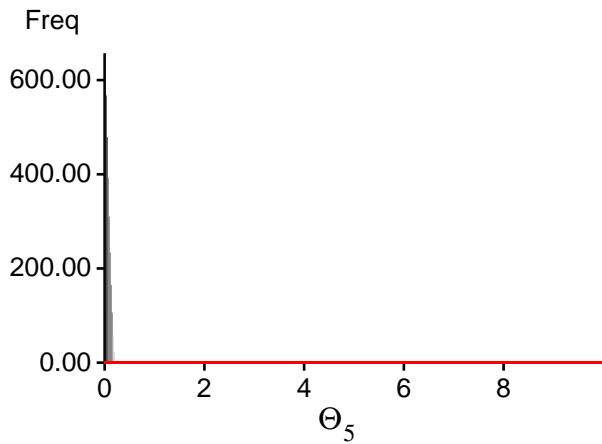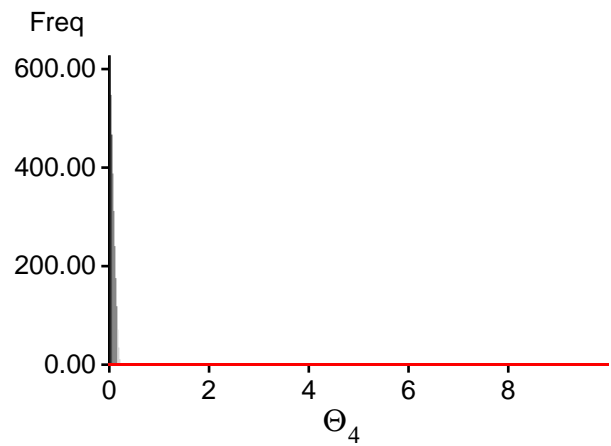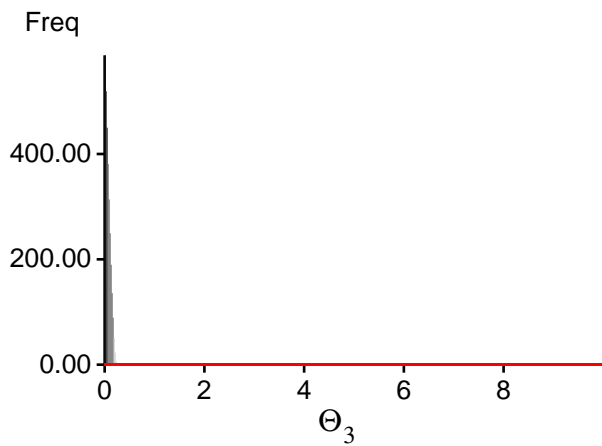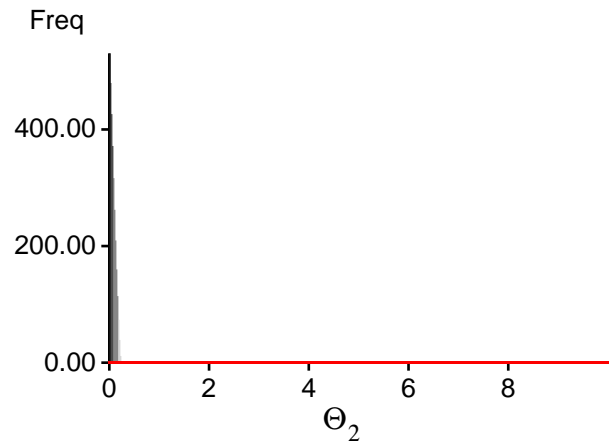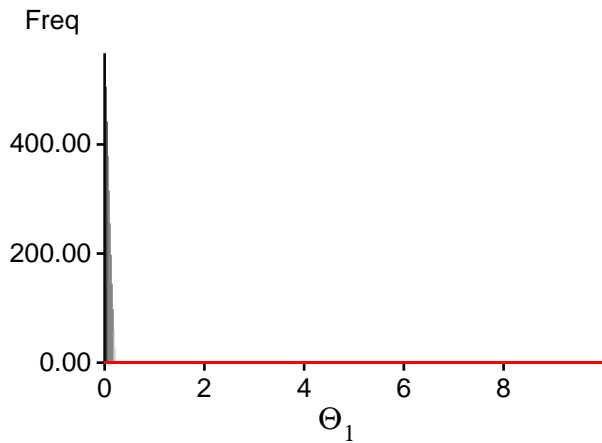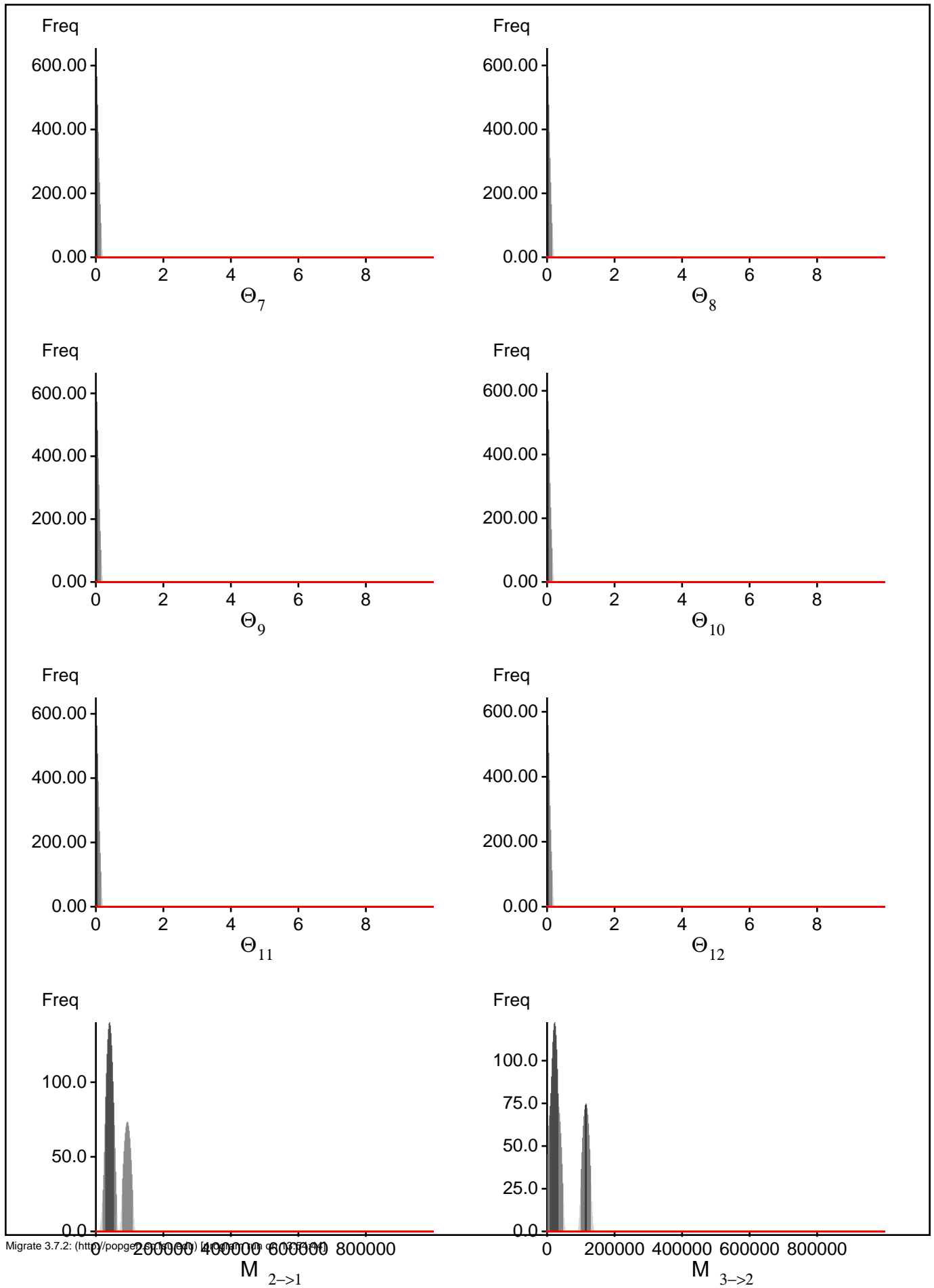
Freq

200.0

150.0

100.0

50.0

0.0

0   200000 400000 600000 800000

$M_{12->11}$
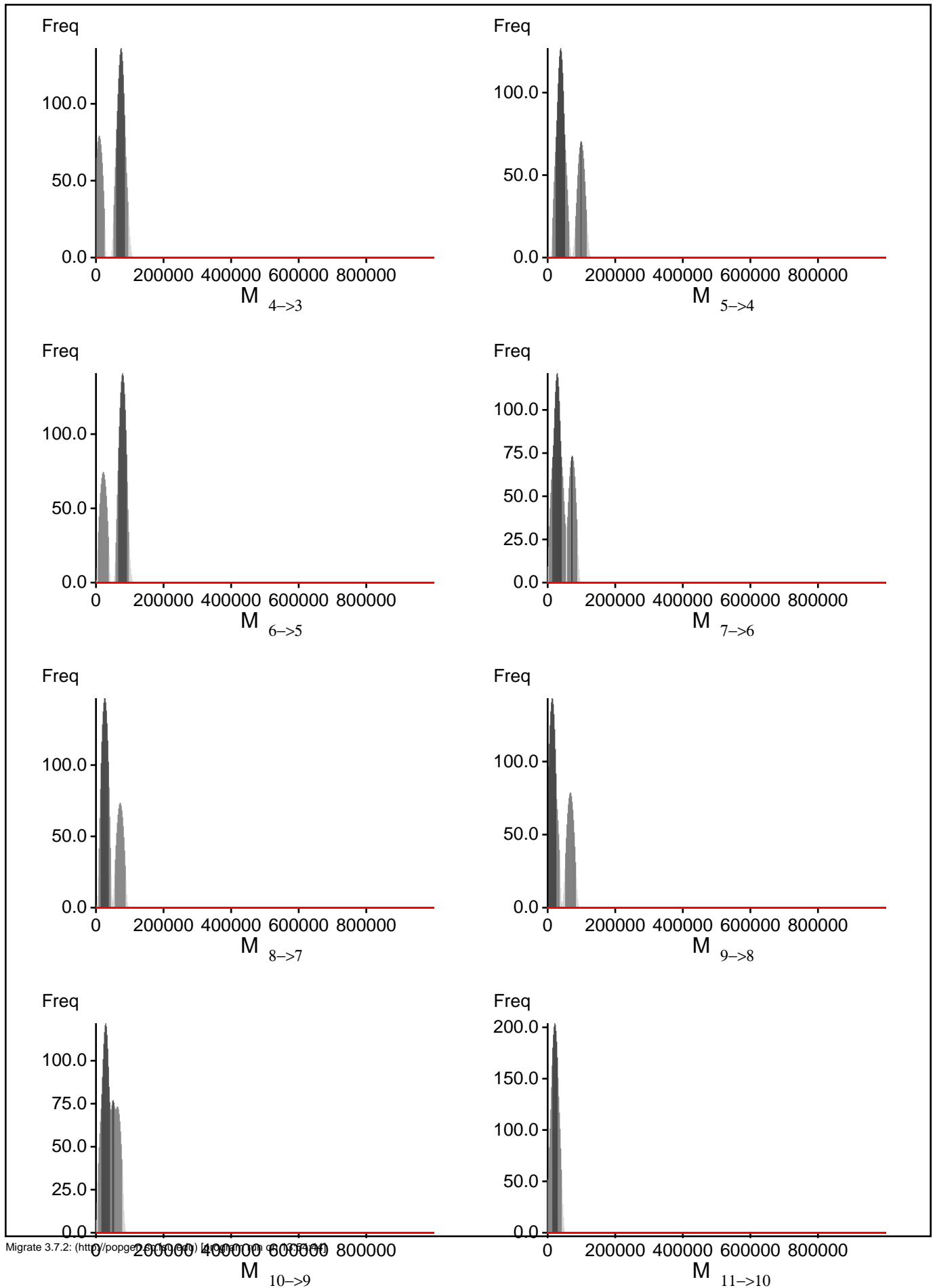
Log-Probability of the data given the model (marginal likelihood)

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	-2255.256579	(1a)
	-2167.397084	(1b)
Harmonic mean	-1863.146678	(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
Θ_1	834/4382	0.19032
Θ_2	687/4409	0.15582
Θ_3	962/4401	0.21859
Θ_4	1336/4429	0.30165
Θ_5	1274/4429	0.28765
Θ_6	720/4381	0.16435
Θ_7	1353/4414	0.30652
Θ_8	762/4397	0.17330
Θ_9	1026/4375	0.23451
Θ_{10}	1380/4327	0.31893
Θ_{11}	1370/4408	0.31080
Θ_{12}	850/4524	0.18789
$M_{2->1}$	4394/4394	1.00000
$M_{1->2}$	4394/4394	1.00000
$M_{3->2}$	4370/4370	1.00000
$M_{2->3}$	4370/4370	1.00000
$M_{4->3}$	4395/4395	1.00000
$M_{3->4}$	4395/4395	1.00000
$M_{5->4}$	4350/4350	1.00000
$M_{4->5}$	4350/4350	1.00000
$M_{6->5}$	4466/4466	1.00000
$M_{5->6}$	4466/4466	1.00000
$M_{7->6}$	4498/4498	1.00000
$M_{6->7}$	4498/4498	1.00000
$M_{8->7}$	4447/4447	1.00000
$M_{7->8}$	4447/4447	1.00000
$M_{9->8}$	4423/4423	1.00000
$M_{8->9}$	4423/4423	1.00000
$M_{10->9}$	4295/4295	1.00000
$M_{9->10}$	4295/4295	1.00000
$M_{11->10}$	4550/4550	1.00000
$M_{10->11}$	4550/4550	1.00000
$M_{12->11}$	4340/4340	1.00000
$M_{11->12}$	4340/4340	1.00000
Genealogies	31919/150005	0.21279

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.85442	239.19
Θ_2	0.90101	161.37
Θ_3	0.85613	243.48
Θ_4	0.80996	318.39
Θ_5	0.80054	334.56
Θ_6	0.90970	145.25
Θ_7	0.85345	237.17
Θ_8	0.89965	162.20
Θ_9	0.82843	285.42
Θ_{10}	0.78457	365.09
Θ_{11}	0.80603	322.43
Θ_{12}	0.89122	178.21
$M_{2\to1}$	0.80308	335.11
$M_{1\to2}$	0.80308	335.11
$M_{3\to2}$	0.77040	393.76
$M_{2\to3}$	0.77040	393.76
$M_{4\to3}$	0.77108	393.24
$M_{3\to4}$	0.77108	393.24
$M_{5\to4}$	0.83577	270.67
$M_{4\to5}$	0.83577	270.67
$M_{6\to5}$	0.81164	317.65
$M_{5\to6}$	0.81164	317.65
$M_{7\to6}$	0.82742	283.96
$M_{6\to7}$	0.82742	283.96
$M_{8\to7}$	0.82428	292.91
$M_{7\to8}$	0.82428	292.91
$M_{9\to8}$	0.79984	336.99
$M_{8\to9}$	0.79984	336.99
$M_{10\to9}$	0.86470	218.35
$M_{9\to10}$	0.86470	218.35
$M_{11\to10}$	0.81607	306.27
$M_{10\to11}$	0.81607	306.27
$M_{12\to11}$	0.87860	194.88
$M_{11\to12}$	0.87860	194.88
Ln[Prob(D|G)]	0.97622	36.24

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