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 18:26:07 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) 2829207646
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	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	$M_{2->1}$	<displayed>
24	$M_{1->2}$	<displayed>
25	$M_{3->2}$	<displayed>
36	$M_{2->3}$	<displayed>
37	$M_{4->3}$	<displayed>
48	$M_{3->4}$	<displayed>
49	$M_{5->4}$	<displayed>
60	$M_{4->5}$	<displayed>
61	$M_{6->5}$	<displayed>
72	$M_{5->6}$	<displayed>
73	$M_{7->6}$	<displayed>
84	$M_{6->7}$	<displayed>
85	$M_{8->7}$	<displayed>
96	$M_{7->8}$	<displayed>
97	$M_{9->8}$	<displayed>
108	$M_{8->9}$	<displayed>
109	$M_{10->9}$	<displayed>
120	$M_{9->10}$	<displayed>
121	$M_{11->10}$	<displayed>
132	$M_{10->11}$	<displayed>
133	$M_{12->11}$	<displayed>
144	$M_{11->12}$	<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	Θ_1	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01692
1	Θ_2	0.00001	0.00001	0.01001	0.08001	0.20001	0.09001	0.04325
1	Θ_3	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01954
1	Θ_4	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02338
1	Θ_5	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01542
1	Θ_6	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02423
1	Θ_7	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01155
1	Θ_8	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02396
1	Θ_9	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02120
1	Θ_{10}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01190
1	Θ_{11}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01651
1	Θ_{12}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.00954
1	$M_{2->1}$	20000.0	30000.0	45000.0	56000.0	94000.0	57000.0	56686.8
1	$M_{1->2}$	2000.0	14000.0	29000.0	42000.0	52000.0	39000.0	44449.6
1	$M_{3->2}$	0.0	10000.0	21000.0	30000.0	42000.0	25000.0	20136.8
1	$M_{2->3}$	26000.0	34000.0	53000.0	66000.0	80000.0	63000.0	72503.7
1	$M_{4->3}$	6000.0	14000.0	29000.0	42000.0	82000.0	37000.0	40490.6
1	$M_{3->4}$	6000.0	18000.0	31000.0	40000.0	58000.0	33000.0	31264.4
1	$M_{5->4}$	0.0	12000.0	33000.0	38000.0	52000.0	29000.0	24174.8
1	$M_{4->5}$	76000.0	88000.0	105000.0	126000.0	166000.0	133000.0	161888.3
1	$M_{6->5}$	74000.0	88000.0	107000.0	124000.0	146000.0	99000.0	86200.6
1	$M_{5->6}$	0.0	0.0	13000.0	24000.0	30000.0	23000.0	48429.4
1	$M_{7->6}$	0.0	0.0	5000.0	18000.0	48000.0	19000.0	15091.8
1	$M_{6->7}$	20000.0	28000.0	45000.0	64000.0	120000.0	85000.0	99615.7
1	$M_{8->7}$	44000.0	48000.0	63000.0	78000.0	178000.0	113000.0	111500.9
1	$M_{7->8}$	42000.0	52000.0	67000.0	82000.0	94000.0	61000.0	52744.9
1	$M_{9->8}$	0.0	0.0	5000.0	16000.0	66000.0	51000.0	49109.8
1	$M_{8->9}$	8000.0	18000.0	33000.0	44000.0	74000.0	39000.0	39424.6
1	$M_{10->9}$	0.0	0.0	1000.0	14000.0	18000.0	15000.0	35143.2
1	$M_{9->10}$	10000.0	64000.0	81000.0	96000.0	118000.0	75000.0	67914.2
1	$M_{11->10}$	0.0	2000.0	21000.0	40000.0	150000.0	39000.0	56044.7
1	$M_{10->11}$	4000.0	12000.0	29000.0	42000.0	54000.0	37000.0	45369.6
1	$M_{12->11}$	0.0	0.0	1000.0	14000.0	32000.0	15000.0	22912.4
1	$M_{11->12}$	54000.0	66000.0	95000.0	138000.0	504000.0	133000.0	201419.2

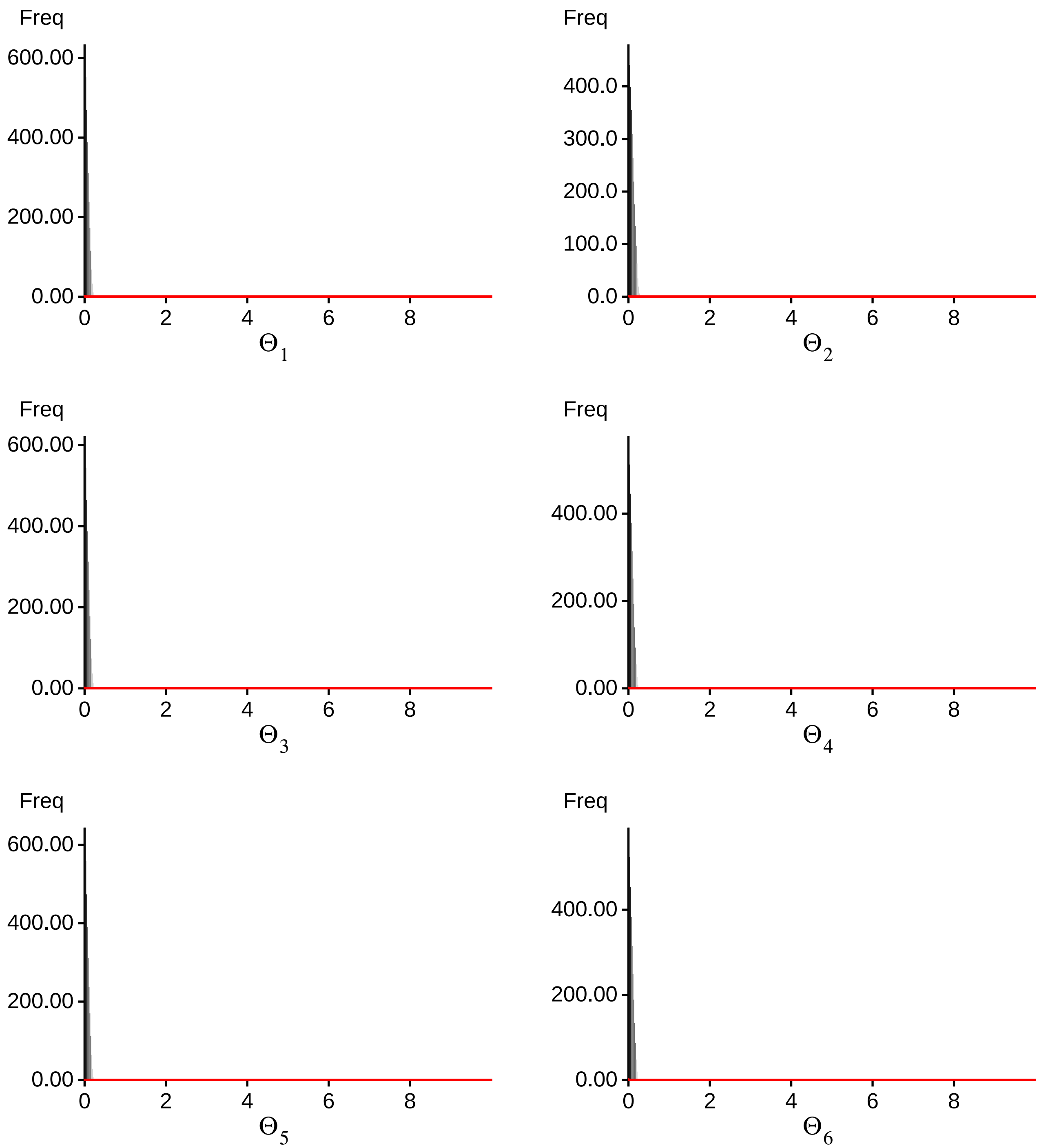
Citation suggestions:

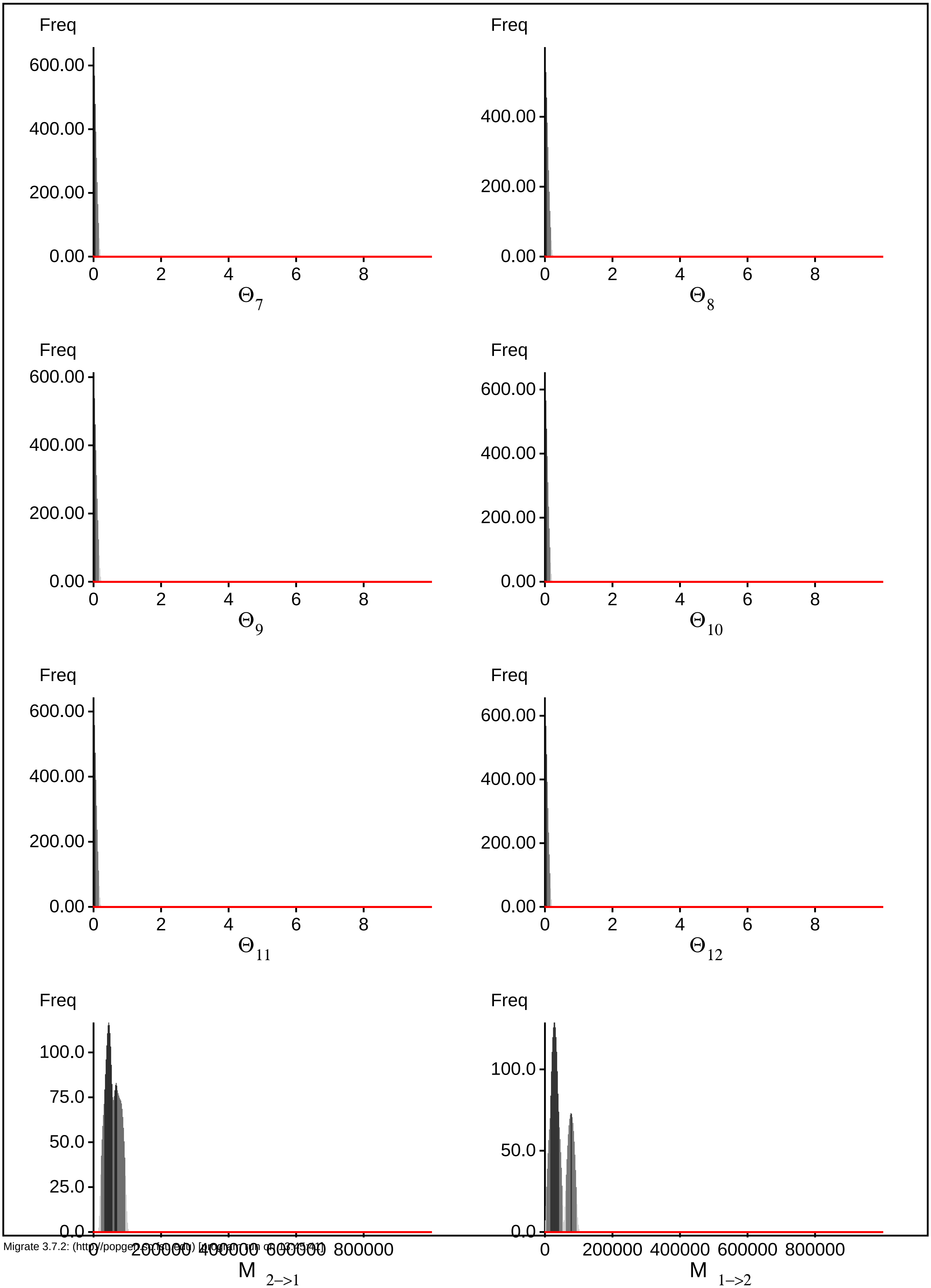
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341-345

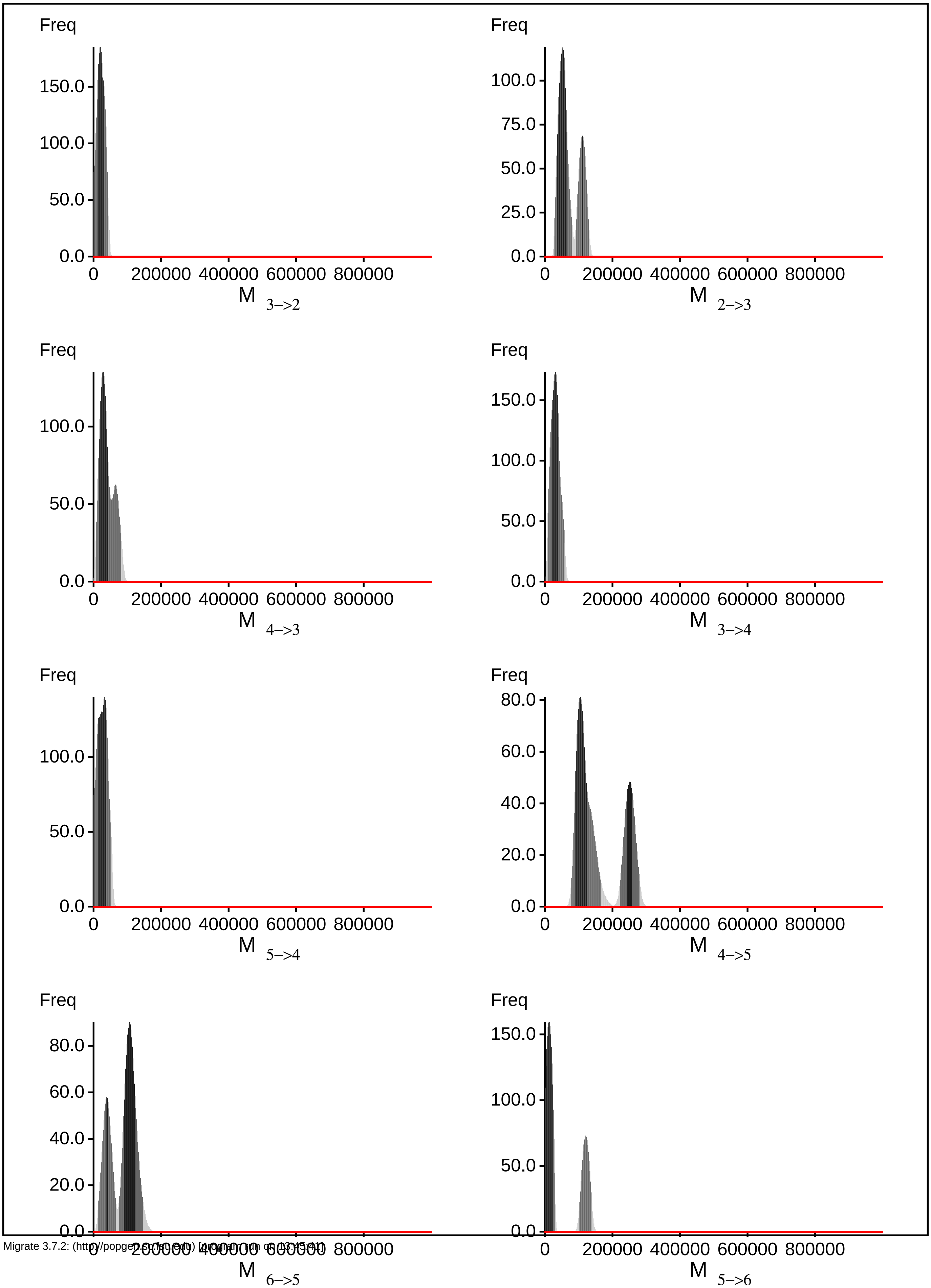
Beerli P., 2007. Estimation of the population scaled mutation rate from microsatellite data, Genetics, 177:1967-1968.

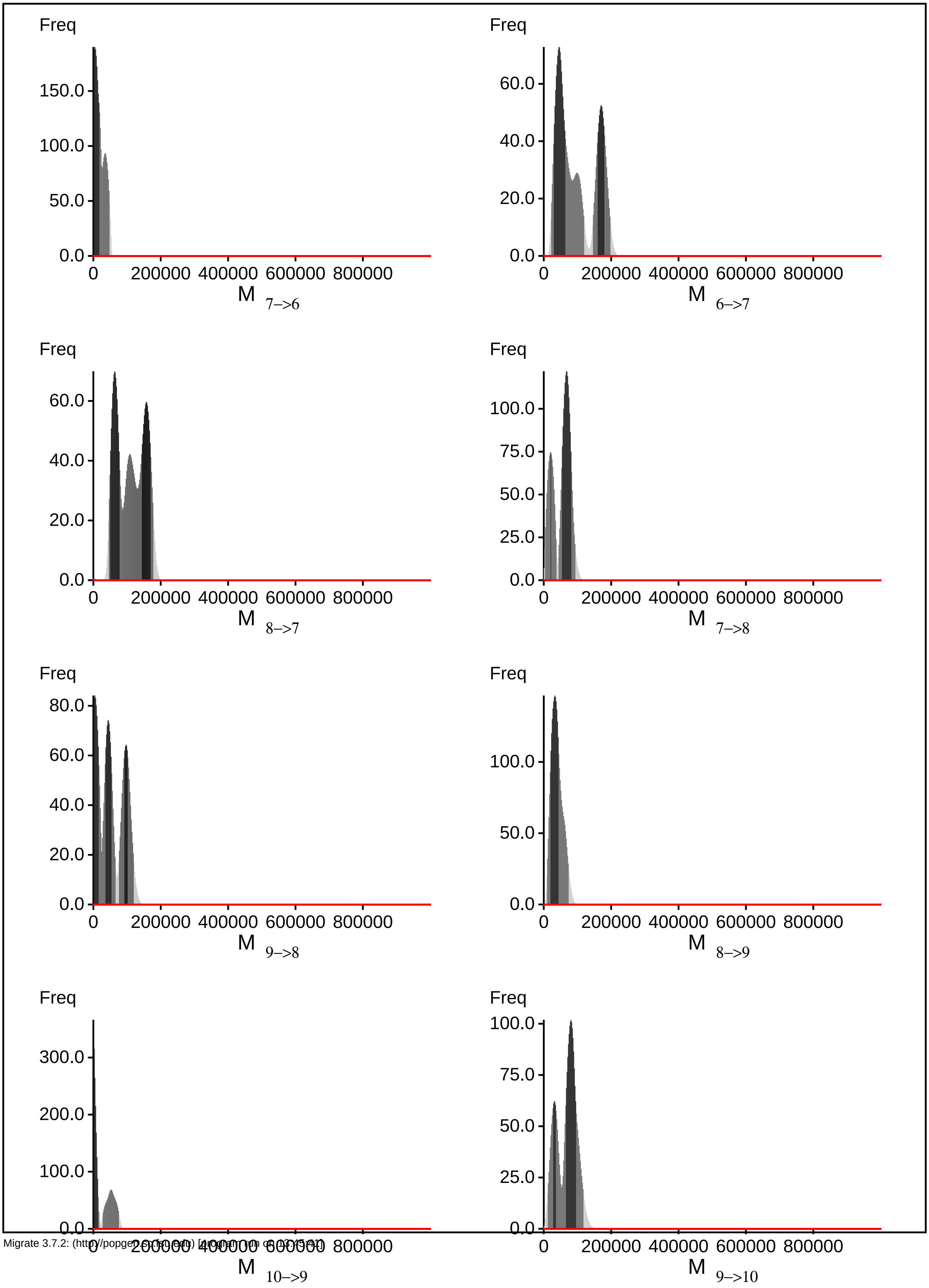
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp.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

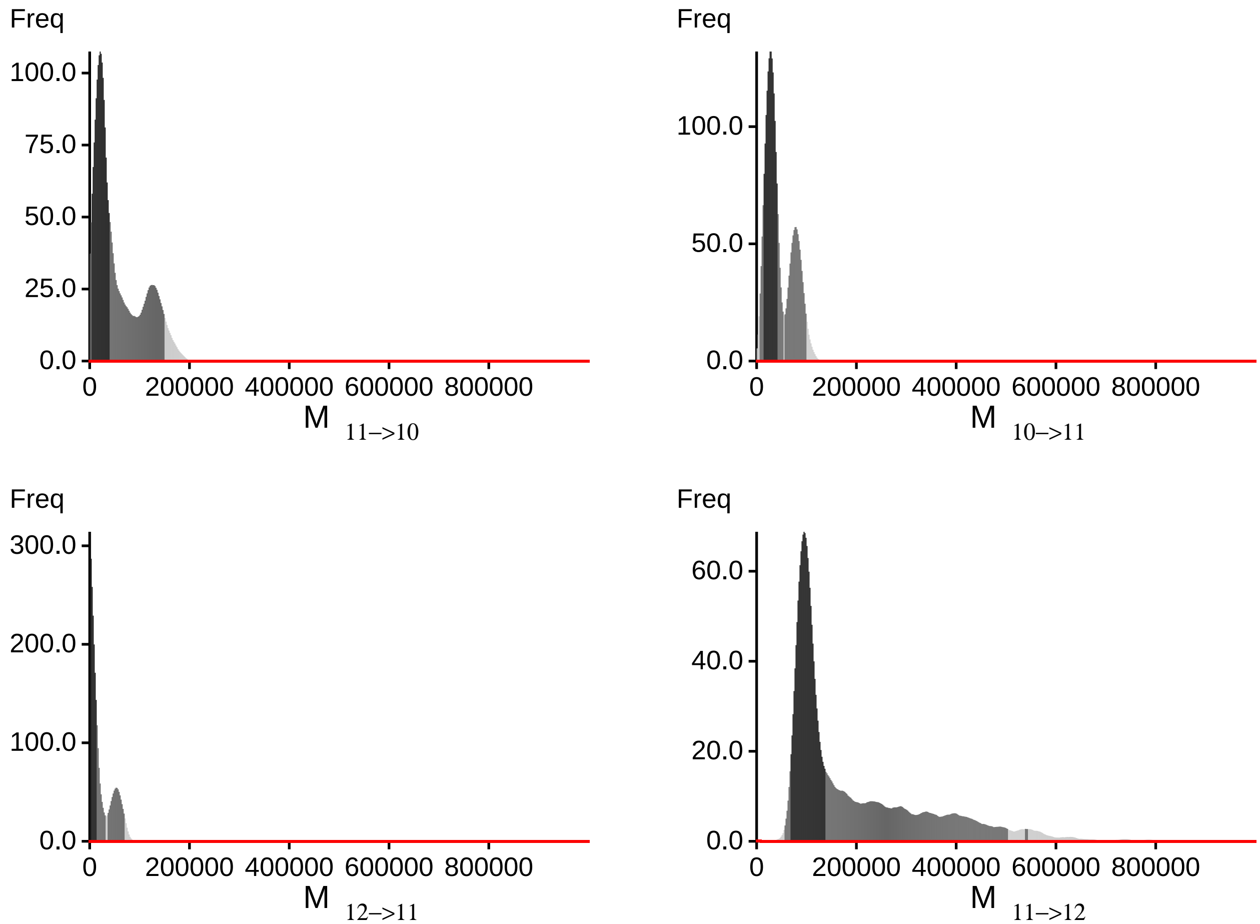
Bayesian Analysis: Posterior distribution over all loci




Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_7
Θ_8
Θ_9
Θ_{10}
Θ_{11}
Θ_{12}
100.0
75.0
50.0
25.0
0.0
200000
400000
600000
800000
$M_{2\to1}$
$M_{1\to2}$


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


Freq
150.0
100.0
50.0
0.0
0 200000 400000 600000 800000
M 7–>6
Freq
60.0
40.0
20.0
0.0
0 200000 400000 600000 800000
M 6–>7
Freq
60.0
40.0
20.0
0.0
0 200000 400000 600000 800000
M 8–>7
Freq
100.0
75.0
50.0
25.0
0.0
0 200000 400000 600000 800000
M 7–>8
Freq
80.0
60.0
40.0
20.0
0.0
0 200000 400000 600000 800000
M 9–>8
Freq
100.0
50.0
0.0
0 200000 400000 600000 800000
M 8–>9
Freq
300.0
200.0
100.0
0.0
0 200000 400000 600000 800000
M 10–>9
Freq
100.0
75.0
50.0
25.0
0.0
0 200000 400000 600000 800000
M 9–>10


Freq
100.0
75.0
50.0
25.0
0.0
0
200000
400000
600000
800000
M 11–>10
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 10–>11
Freq
300.0
200.0
100.0
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	-2215.275423	(1a)
	-2140.436243	(1b)
Harmonic mean	-1851.391046	(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	407/4457	0.09132
Θ_2	295/4369	0.06752
Θ_3	1392/4456	0.31239
Θ_4	349/4364	0.07997
Θ_5	2663/4527	0.58825
Θ_6	938/4462	0.21022
Θ_7	2510/4373	0.57398
Θ_8	1406/4486	0.31342
Θ_9	1421/4361	0.32584
Θ_{10}	2111/4487	0.47047
Θ_{11}	2145/4394	0.48817
Θ_{12}	2612/4539	0.57546
$M_{2\to1}$	4447/4447	1.00000
$M_{1\to2}$	4334/4334	1.00000
$M_{3\to2}$	4394/4394	1.00000
$M_{2\to3}$	4352/4352	1.00000
$M_{4\to3}$	4348/4348	1.00000
$M_{3\to4}$	4337/4337	1.00000
$M_{5\to4}$	4419/4419	1.00000
$M_{4\to5}$	4393/4393	1.00000
$M_{6\to5}$	4282/4282	1.00000
$M_{5\to6}$	4371/4371	1.00000
$M_{7\to6}$	4381/4381	1.00000
$M_{6\to7}$	4421/4421	1.00000
$M_{8\to7}$	4433/4433	1.00000
$M_{7\to8}$	4439/4439	1.00000
$M_{9\to8}$	4405/4405	1.00000
$M_{8\to9}$	4254/4254	1.00000
$M_{10\to9}$	4470/4470	1.00000
$M_{9\to10}$	4443/4443	1.00000
$M_{11\to10}$	4361/4361	1.00000
$M_{10\to11}$	4387/4387	1.00000
$M_{12\to11}$	4419/4419	1.00000
$M_{11\to12}$	4444/4444	1.00000
Genealogies	32452/150191	0.21607

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.95232	73.81
Θ_2	0.95817	64.73
Θ_3	0.78138	370.76
Θ_4	0.95648	67.48
Θ_5	0.56366	913.55
Θ_6	0.89194	172.32
Θ_7	0.57569	812.30
Θ_8	0.76792	414.16
Θ_9	0.74877	462.11
Θ_{10}	0.64660	658.20
Θ_{11}	0.60489	768.11
Θ_{12}	0.46633	1098.96
$M_{2\rightarrow 1}$	0.81362	310.86
$M_{1\rightarrow 2}$	0.79492	353.85
$M_{3\rightarrow 2}$	0.81198	311.41
$M_{2\rightarrow 3}$	0.80012	346.12
$M_{4\rightarrow 3}$	0.89367	169.84
$M_{3\rightarrow 4}$	0.75985	412.16
$M_{5\rightarrow 4}$	0.79797	338.45
$M_{4\rightarrow 5}$	0.71986	495.57
$M_{6\rightarrow 5}$	0.87899	193.51
$M_{5\rightarrow 6}$	0.74141	451.29
$M_{7\rightarrow 6}$	0.69806	553.96
$M_{6\rightarrow 7}$	0.82493	289.00
$M_{8\rightarrow 7}$	0.69933	535.88
$M_{7\rightarrow 8}$	0.73596	458.91
$M_{9\rightarrow 8}$	0.77728	379.65
$M_{8\rightarrow 9}$	0.75003	430.21
$M_{10\rightarrow 9}$	0.64773	649.85
$M_{9\rightarrow 10}$	0.74296	442.20
$M_{11\rightarrow 10}$	0.87091	208.03
$M_{10\rightarrow 11}$	0.74374	443.05
$M_{12\rightarrow 11}$	0.83120	277.01
$M_{11\rightarrow 12}$	0.72791	473.51
Ln[Prob(D|G)]	0.97649	35.79

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