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   Fri May 28 15:38:07 2021
Program finished at   Fri May 28 23:59:31 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) 3819900891
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	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	$M_{2\to1}$	<displayed>
24	$M_{1\to2}$	<displayed>
25	$M_{3\to2}$	<displayed>
36	$M_{2\to3}$	<displayed>
37	$M_{4\to3}$	<displayed>
48	$M_{3\to4}$	<displayed>
49	$M_{5\to4}$	<displayed>
60	$M_{4\to5}$	<displayed>
61	$M_{6\to5}$	<displayed>
72	$M_{5\to6}$	<displayed>
73	$M_{7\to6}$	<displayed>
84	$M_{6\to7}$	<displayed>
85	$M_{8\to7}$	<displayed>
96	$M_{7\to8}$	<displayed>
97	$M_{9\to8}$	<displayed>
108	$M_{8\to9}$	<displayed>
109	$M_{10\to9}$	<displayed>
120	$M_{9\to10}$	<displayed>
121	$M_{11\to10}$	<displayed>
132	$M_{10\to11}$	<displayed>
133	$M_{12\to11}$	<displayed>
144	$M_{11\to12}$	<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	10
12	1	0
Total of all populations	1	184

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	0.08001	0.18001	0.09001	0.03053
1	Θ_2	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02140
1	Θ_3	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.03026
1	Θ_4	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01451
1	Θ_5	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01578
1	Θ_6	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01736
1	Θ_7	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01375
1	Θ_8	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01444
1	Θ_9	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01325
1	Θ_{10}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02240
1	Θ_{11}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01261
1	Θ_{12}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01036
1	$M_{2->1}$	26000.0	34000.0	49000.0	60000.0	70000.0	45000.0	34123.9
1	$M_{1->2}$	2000.0	12000.0	27000.0	42000.0	56000.0	41000.0	45154.3
1	$M_{3->2}$	14000.0	28000.0	47000.0	62000.0	78000.0	59000.0	72250.1
1	$M_{2->3}$	0.0	0.0	1000.0	14000.0	18000.0	15000.0	70074.3
1	$M_{4->3}$	22000.0	36000.0	47000.0	56000.0	72000.0	49000.0	46625.8
1	$M_{3->4}$	56000.0	66000.0	79000.0	88000.0	104000.0	75000.0	54307.3
1	$M_{5->4}$	28000.0	50000.0	63000.0	76000.0	90000.0	63000.0	60332.2
1	$M_{4->5}$	52000.0	62000.0	75000.0	88000.0	124000.0	83000.0	84534.5
1	$M_{6->5}$	52000.0	60000.0	73000.0	86000.0	128000.0	73000.0	66117.1
1	$M_{5->6}$	0.0	10000.0	21000.0	28000.0	76000.0	61000.0	81710.3
1	$M_{7->6}$	0.0	2000.0	15000.0	28000.0	64000.0	25000.0	26940.7
1	$M_{6->7}$	0.0	8000.0	25000.0	38000.0	48000.0	35000.0	49010.2
1	$M_{8->7}$	32000.0	42000.0	57000.0	70000.0	84000.0	51000.0	42417.7
1	$M_{7->8}$	0.0	6000.0	19000.0	30000.0	54000.0	25000.0	24851.4
1	$M_{9->8}$	34000.0	46000.0	63000.0	76000.0	90000.0	55000.0	46908.9
1	$M_{8->9}$	4000.0	10000.0	25000.0	40000.0	76000.0	35000.0	37415.7
1	$M_{10->9}$	10000.0	20000.0	31000.0	38000.0	50000.0	33000.0	30479.5
1	$M_{9->10}$	0.0	0.0	1000.0	12000.0	50000.0	21000.0	23267.7
1	$M_{11->10}$	0.0	0.0	9000.0	20000.0	50000.0	21000.0	18047.7
1	$M_{10->11}$	0.0	24000.0	31000.0	36000.0	50000.0	27000.0	21677.2
1	$M_{12->11}$	0.0	0.0	1000.0	12000.0	22000.0	77000.0	59255.4
1	$M_{11->12}$	648000.0	826000.0	889000.0	922000.0	1000000.0	783000.0	677591.3

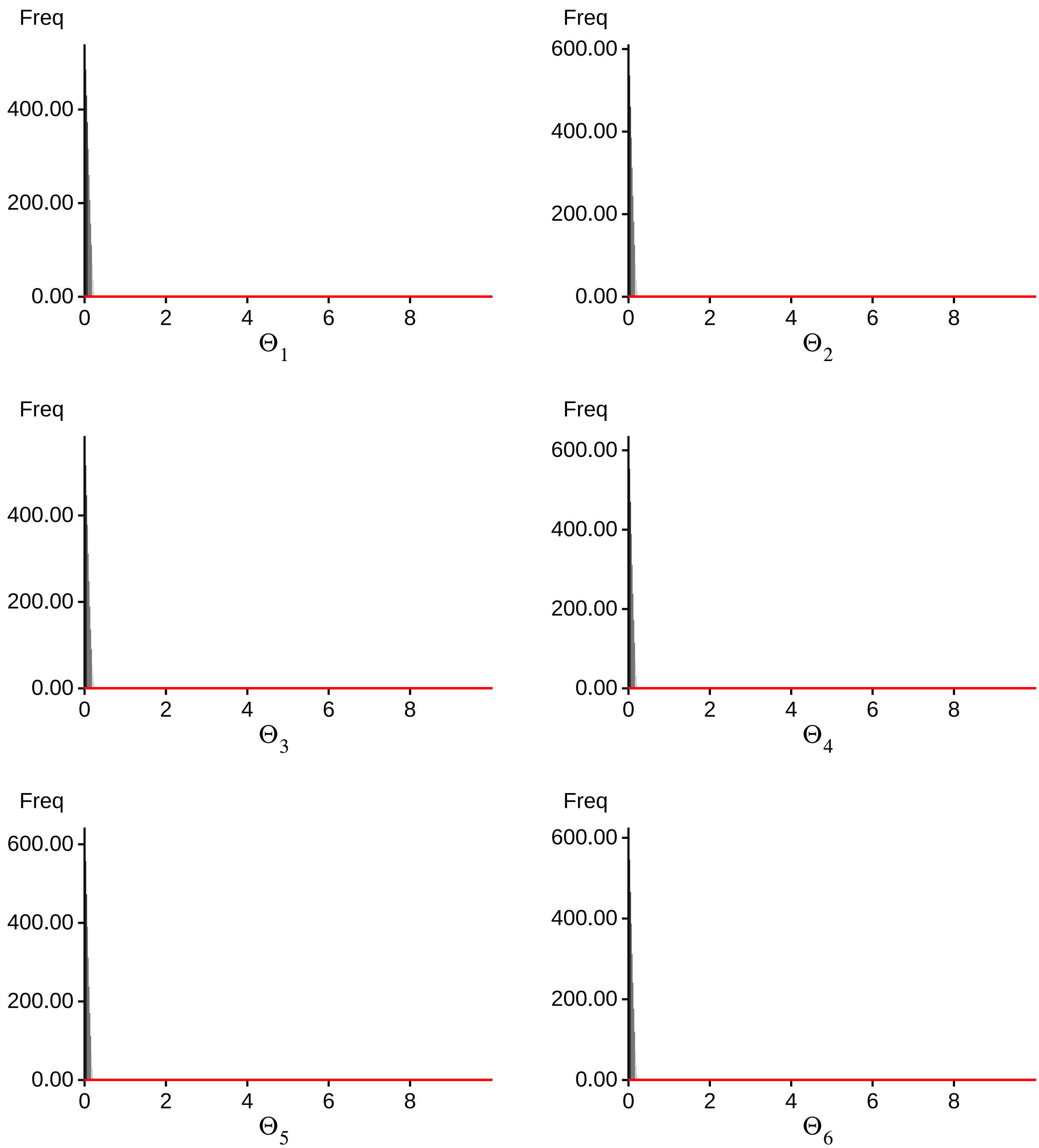
Citation suggestions:

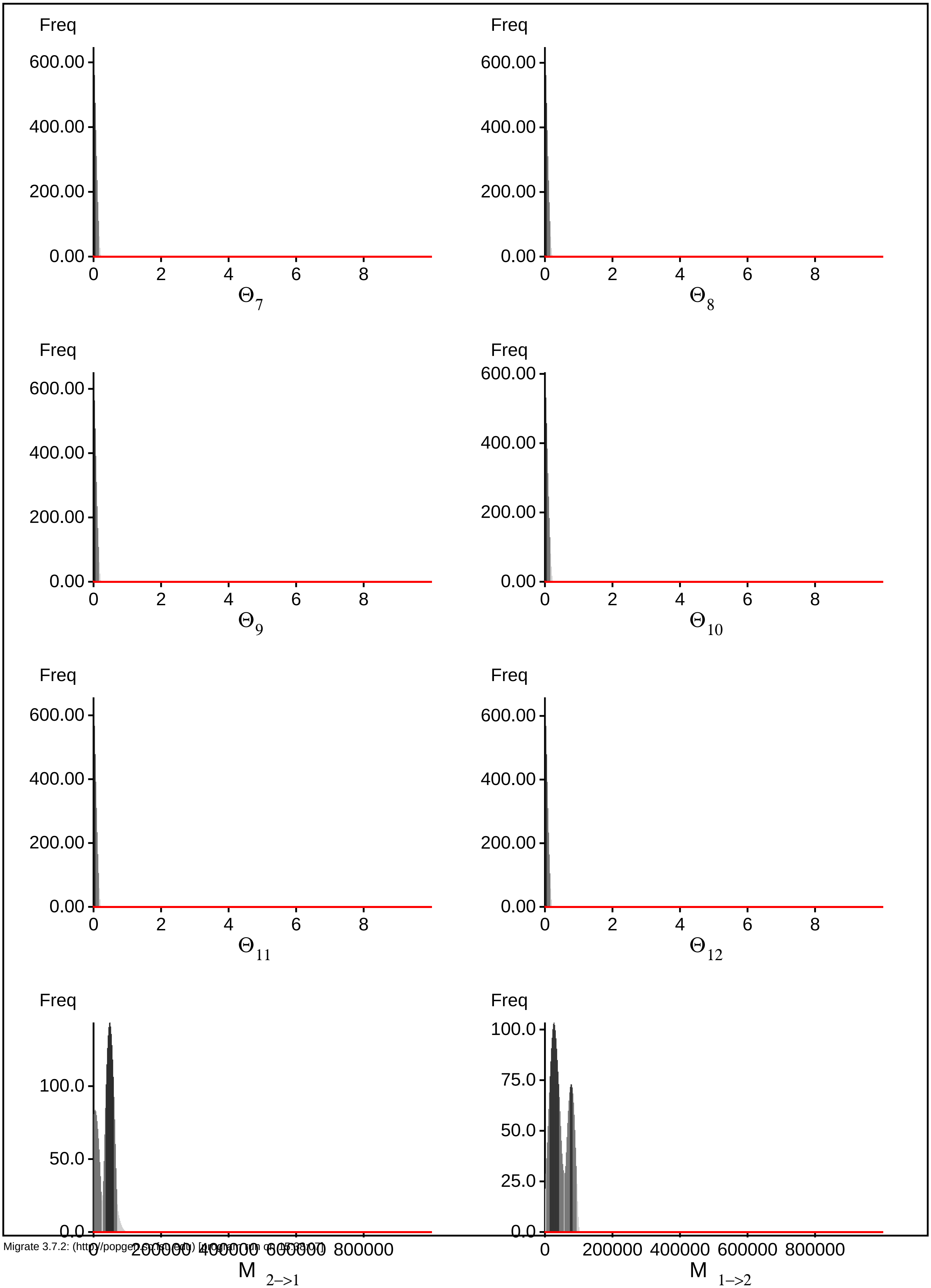
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341-345

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

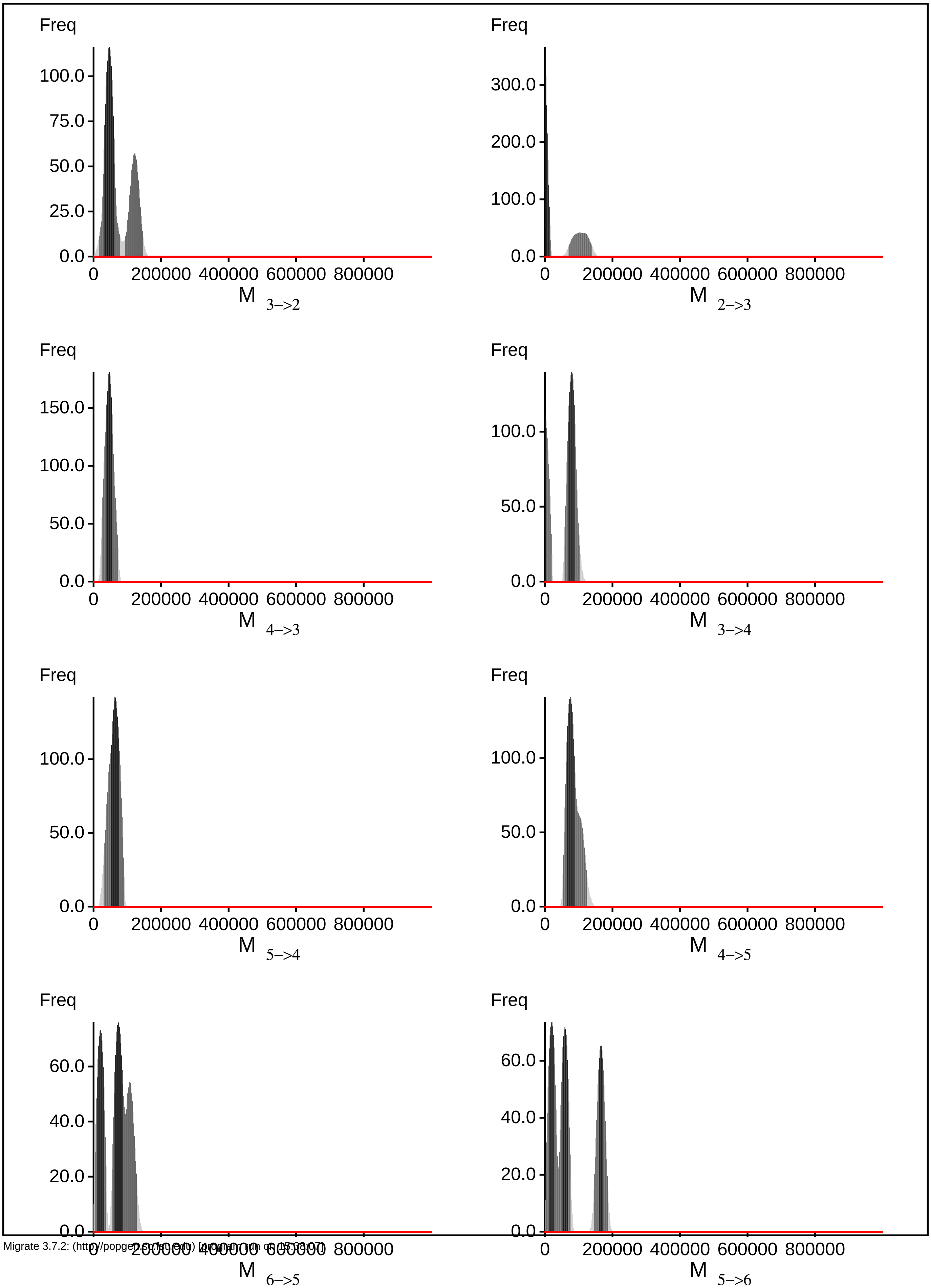
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp.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

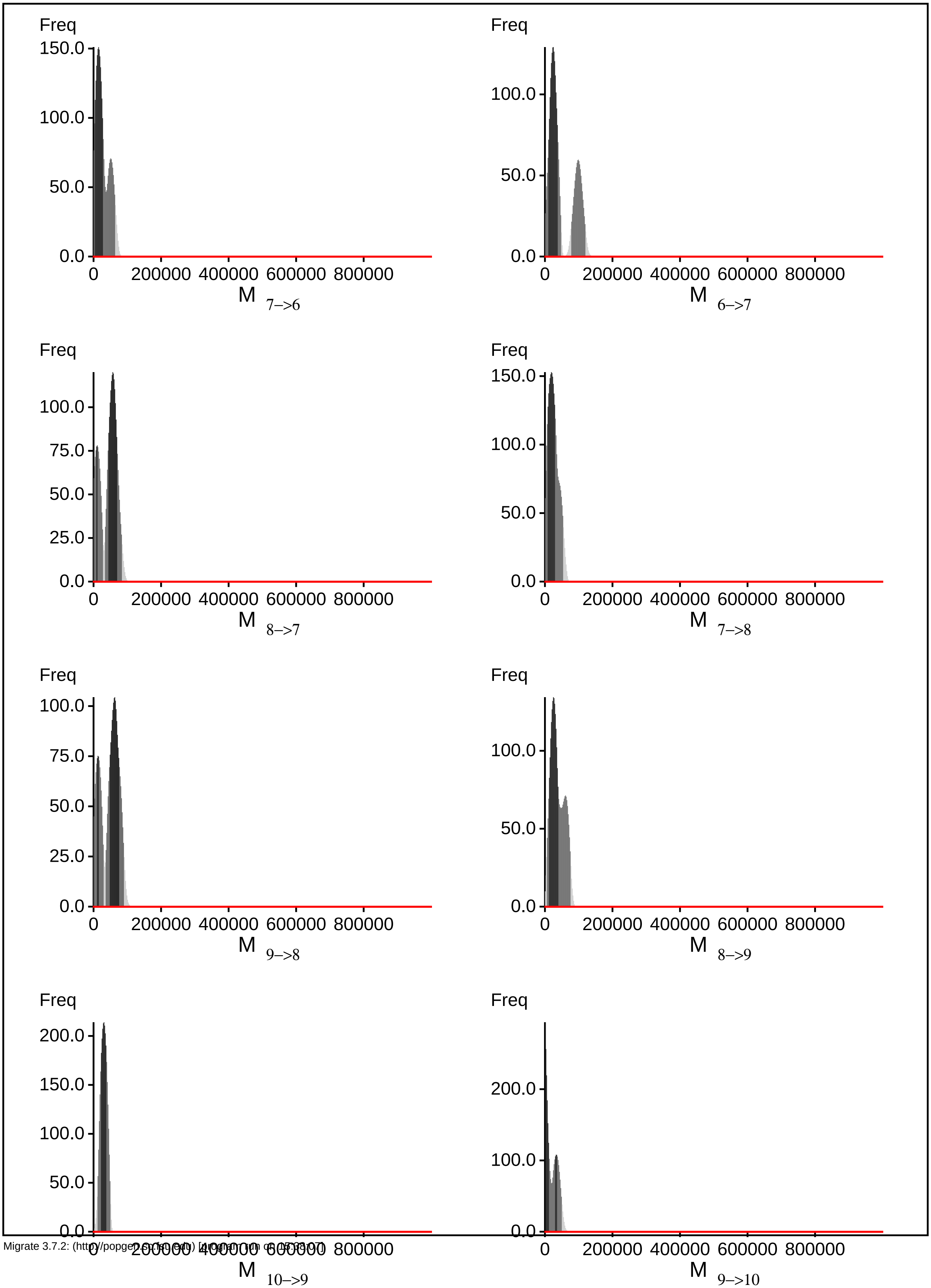
Bayesian Analysis: Posterior distribution over all loci

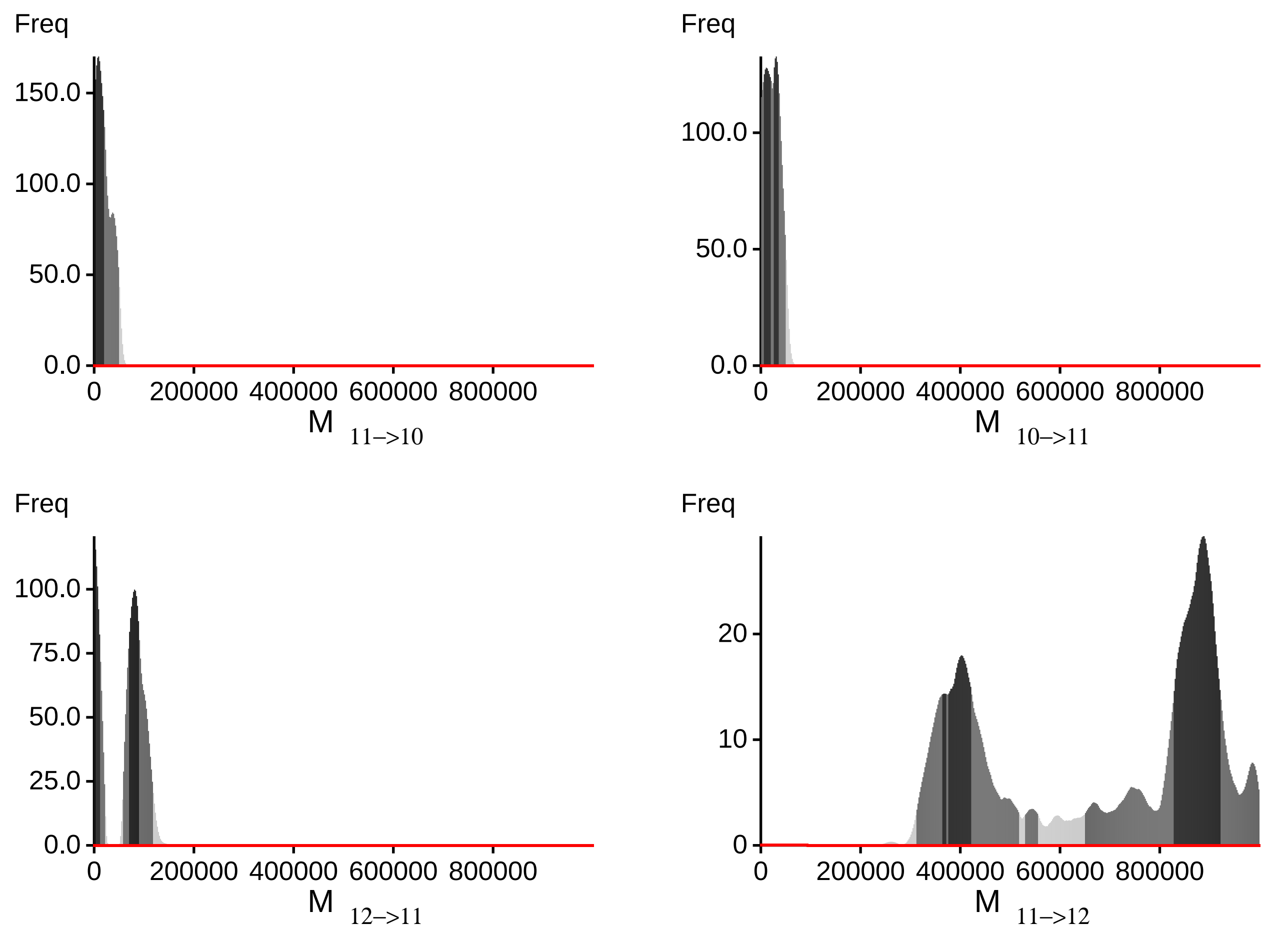




Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_7
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_8
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_9
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_{10}
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_{11}
Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ_{12}
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 2->1
Freq
100.0
75.0
50.0
25.0
0.0
0
200000
400000
600000
800000
M 1->2


Freq
M 3–>2
M 2–>3
M 4–>3
M 3–>4
M 5–>4
M 4–>5
M 6–>5
M 5–>6
0 200000 400000 600000 800000


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


Freq
150.0
100.0
50.0
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
100.0
75.0
50.0
25.0
0.0
0
200000 400000 600000 800000
M 12–>11
Freq
20
10
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	-2145.214806	(1a)
	-2072.646892	(1b)
Harmonic mean	-1828.556608	(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	1090/4418	0.24672
Θ_2	1899/4488	0.42313
Θ_3	1298/4441	0.29228
Θ_4	504/4347	0.11594
Θ_5	1456/4386	0.33197
Θ_6	778/4418	0.17610
Θ_7	1754/4418	0.39701
Θ_8	1474/4403	0.33477
Θ_9	1404/4434	0.31664
Θ_{10}	1281/4412	0.29034
Θ_{11}	1862/4535	0.41058
Θ_{12}	3573/4344	0.82251
$M_{2\to1}$	4373/4373	1.00000
$M_{1\to2}$	4442/4442	1.00000
$M_{3\to2}$	4347/4347	1.00000
$M_{2\to3}$	4458/4458	1.00000
$M_{4\to3}$	4439/4439	1.00000
$M_{3\to4}$	4386/4386	1.00000
$M_{5\to4}$	4437/4437	1.00000
$M_{4\to5}$	4447/4447	1.00000
$M_{6\to5}$	4536/4536	1.00000
$M_{5\to6}$	4499/4499	1.00000
$M_{7\to6}$	4482/4482	1.00000
$M_{6\to7}$	4434/4434	1.00000
$M_{8\to7}$	4358/4358	1.00000
$M_{7\to8}$	4440/4440	1.00000
$M_{9\to8}$	4407/4407	1.00000
$M_{8\to9}$	4488/4488	1.00000
$M_{10\to9}$	4444/4444	1.00000
$M_{9\to10}$	4379/4379	1.00000
$M_{11\to10}$	4391/4391	1.00000
$M_{10\to11}$	4415/4415	1.00000
$M_{12\to11}$	4417/4417	1.00000
$M_{11\to12}$	4476/4476	1.00000
Genealogies	33187/149461	0.22204

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.80413	386.29
Θ_2	0.69339	581.47
Θ_3	0.75688	442.61
Θ_4	0.94953	78.51
Θ_5	0.79393	360.64
Θ_6	0.87543	204.89
Θ_7	0.62821	715.17
Θ_8	0.70481	527.07
Θ_9	0.76357	406.78
Θ_{10}	0.76781	404.94
Θ_{11}	0.69773	586.40
Θ_{12}	0.34302	1552.22
$M_{2\to1}$	0.79272	356.77
$M_{1\to2}$	0.85995	227.94
$M_{3\to2}$	0.76483	405.88
$M_{2\to3}$	0.77874	392.12
$M_{4\to3}$	0.71033	513.27
$M_{3\to4}$	0.80964	322.31
$M_{5\to4}$	0.83452	277.06
$M_{4\to5}$	0.83898	264.16
$M_{6\to5}$	0.79059	358.74
$M_{5\to6}$	0.72603	477.13
$M_{7\to6}$	0.82474	289.84
$M_{6\to7}$	0.74542	458.85
$M_{8\to7}$	0.79897	346.36
$M_{7\to8}$	0.84363	259.24
$M_{9\to8}$	0.75046	438.75
$M_{8\to9}$	0.73934	453.69
$M_{10\to9}$	0.77133	393.54
$M_{9\to10}$	0.79784	338.46
$M_{11\to10}$	0.74317	446.34
$M_{10\to11}$	0.70196	549.06
$M_{12\to11}$	0.73650	462.68
$M_{11\to12}$	0.67666	607.61
Ln[Prob(D|G)]	0.98133	28.27

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.

Param 144 (Locus 1): Upper prior boundary seems too low!