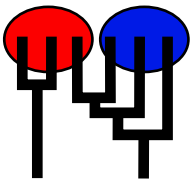
Preliminary migrate analysis of M. californianus

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 10:52:05 2021
Program finished at Tue Jun 1 16:28:18 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) 281655502
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	0.18001	0.07001	0.02672
1	Θ_{2}	0.00001	0.00001	0.01001	0.08001	0.18001	0.09001	0.03105
1	Θ_{3}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02213
1	Θ_{4}	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02542
1	Θ_{5}	0.00001	0.00001	0.01001	0.08001	0.18001	0.09001	0.02909
1	Θ_{6}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02055
1	Θ_{7}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.00166
1	Θ_{8}	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02578
1	Θ_{9}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02078
1	Θ_{10}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01376
1	Θ_{11}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01060
1	Θ_{12}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.00934
1	$M_{2->1}$	0.0	12000.0	23000.0	30000.0	42000.0	25000.0	22465.1
1	$M_{1->2}$	0.0	8000.0	19000.0	28000.0	42000.0	23000.0	18978.4
1	$M_{3->2}$	14000.0	36000.0	47000.0	60000.0	72000.0	49000.0	45947.2
1	$M_{2->3}$	0.0	0.0	13000.0	24000.0	32000.0	73000.0	77682.3
1	$M_{4->3}$	12000.0	58000.0	71000.0	86000.0	94000.0	63000.0	56828.5
1	$M_{3->4}$	14000.0	28000.0	37000.0	46000.0	60000.0	39000.0	37240.9
1	$M_{5->4}$	10000.0	48000.0	65000.0	78000.0	84000.0	59000.0	51008.8
1	$M_{4->5}$	8000.0	20000.0	31000.0	38000.0	50000.0	33000.0	30561.7
1	$M_{6->5}$	0.0	6000.0	21000.0	32000.0	72000.0	29000.0	32899.1
1	$M_{5->6}$	0.0	2000.0	15000.0	28000.0	38000.0	27000.0	31410.1
1	$M_{7->6}$	8000.0	48000.0	59000.0	68000.0	82000.0	47000.0	45842.9
1	$M_{6->7}$	0.0	26000.0	41000.0	50000.0	62000.0	37000.0	32128.8
1	$M_{8->7}$	90000.0	100000.0	117000.0	130000.0	142000.0	109000.0	90073.7
1	$M_{7->8}$	66000.0	72000.0	87000.0	98000.0	106000.0	81000.0	69166.5
1	$M_{9->8}$	18000.0	32000.0	43000.0	52000.0	66000.0	45000.0	43122.8
1	$M_{8->9}$	22000.0	30000.0	45000.0	58000.0	66000.0	55000.0	65146.9
1	$M_{10->9}$	38000.0	48000.0	65000.0	78000.0	90000.0	59000.0	45842.9
1	$M_{9->10}$	0.0	4000.0	13000.0	18000.0	30000.0	101000.0	79642.7
1	$M_{11->10}$	0.0	14000.0	27000.0	36000.0	52000.0	29000.0	27237.1
1	$M_{10->11}$	0.0	10000.0	23000.0	32000.0	48000.0	25000.0	24070.7
1	$M_{12->11}$	0.0	6000.0	21000.0	34000.0	42000.0	31000.0	63058.7
1	$M_{11->12}$	2000.0	10000.0	25000.0	38000.0	70000.0	33000.0	34367.9

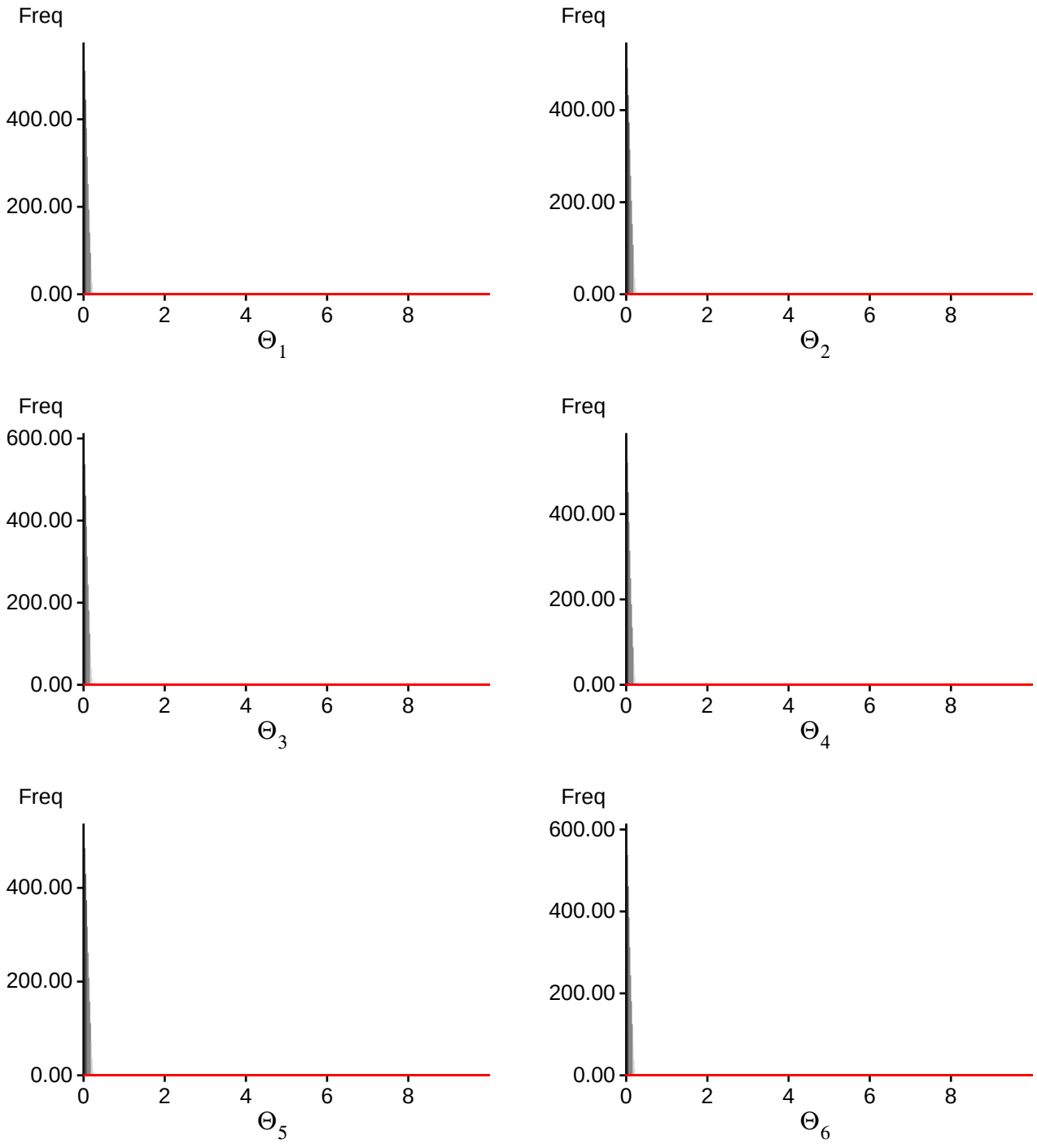
Citation suggestions:

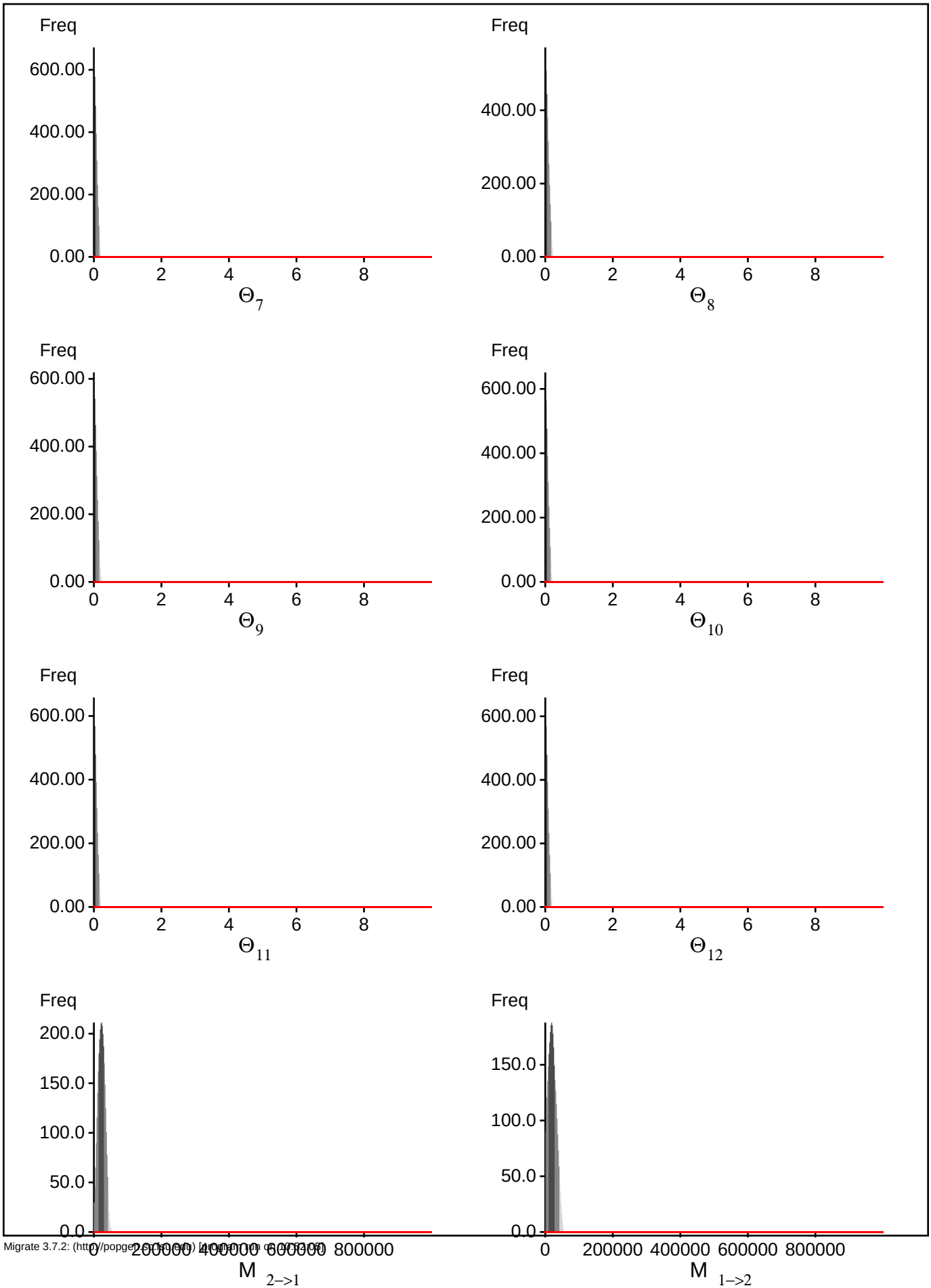
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341-345

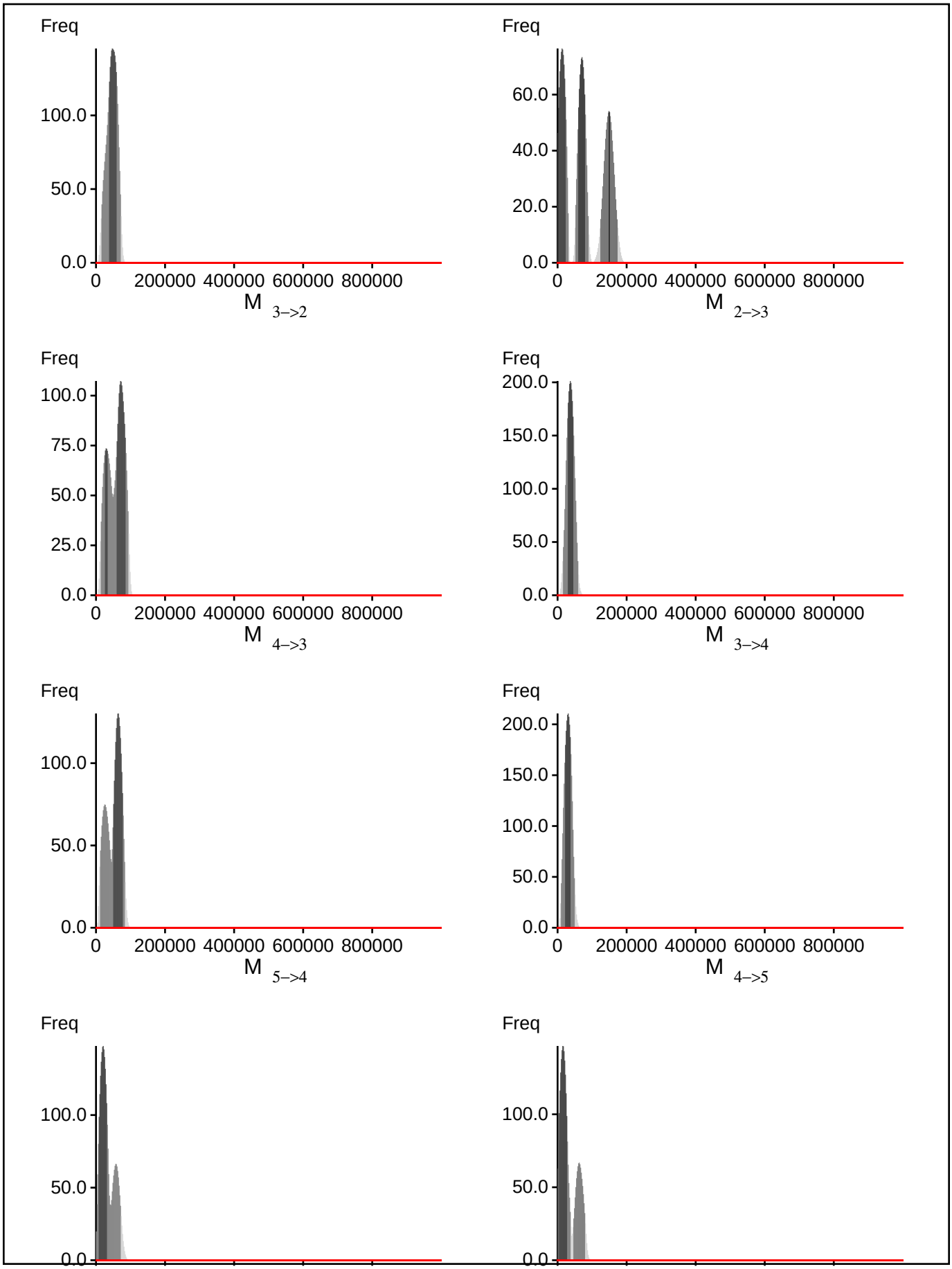
Beerli P., 2007. Estimation of the population scaled mutation rate from microsatellite data, Genetics, 177:1967-1968.

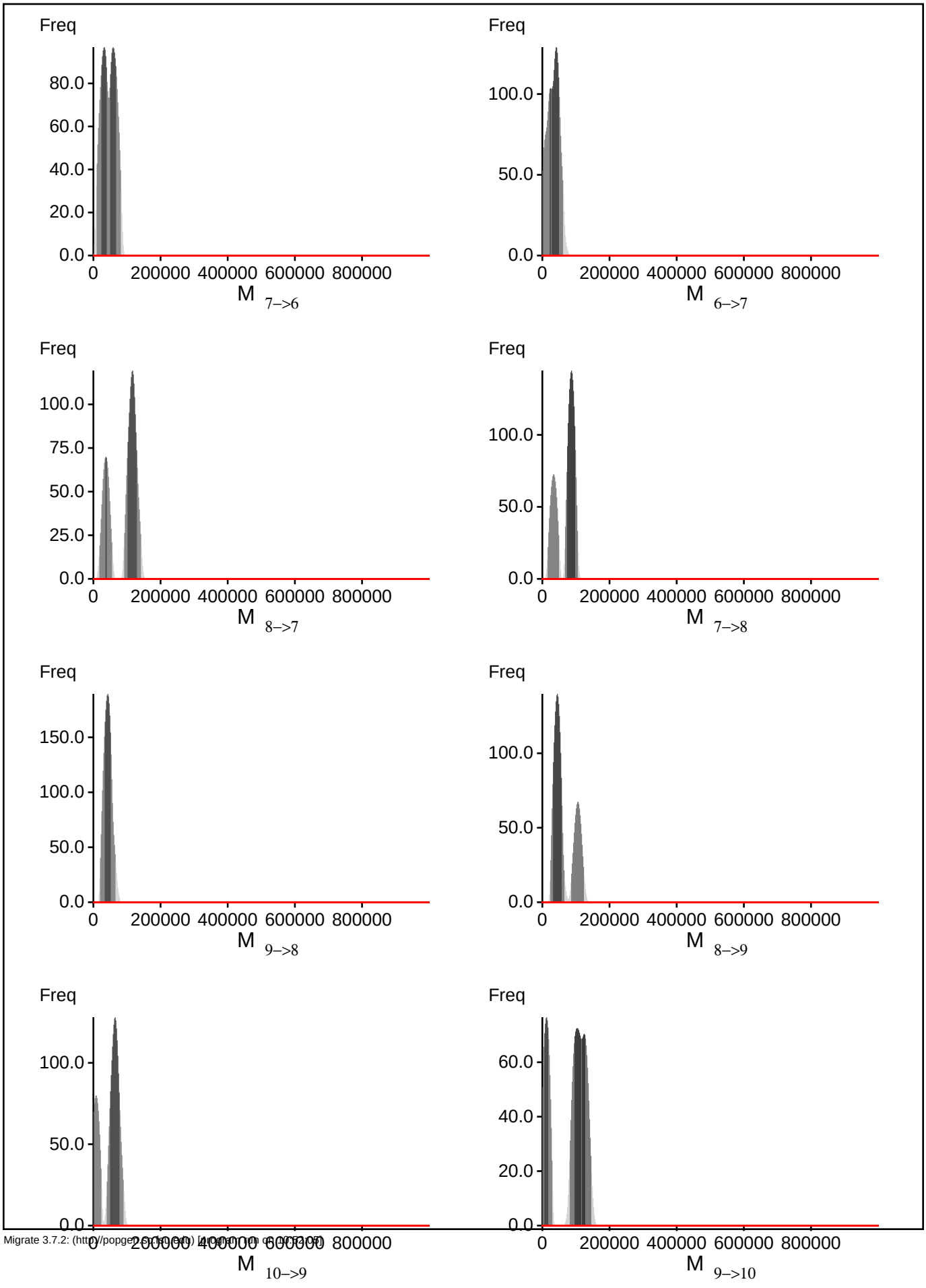
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp.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

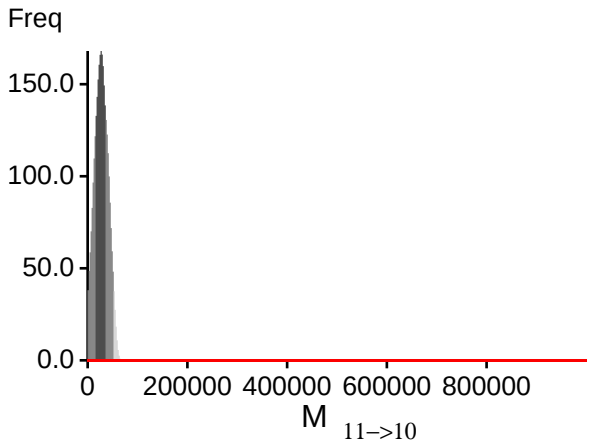
Bayesian Analysis: Posterior distribution over all loci

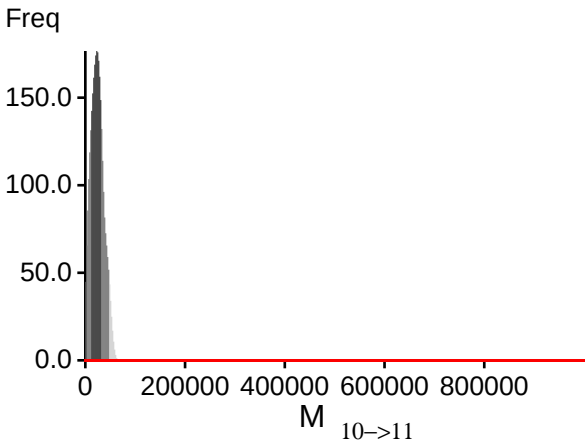


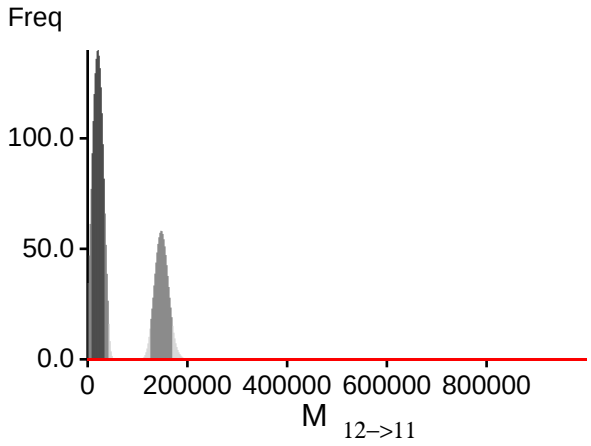

Freq
600.00
400.00
200.00
0.00
0
2
4
6
8
Θ7
Θ8
Θ9
Θ10
Θ11
Θ12
200.0
150.0
100.0
50.0
0.0
200000
400000
600000
800000
M 2–>1
M 1–>2

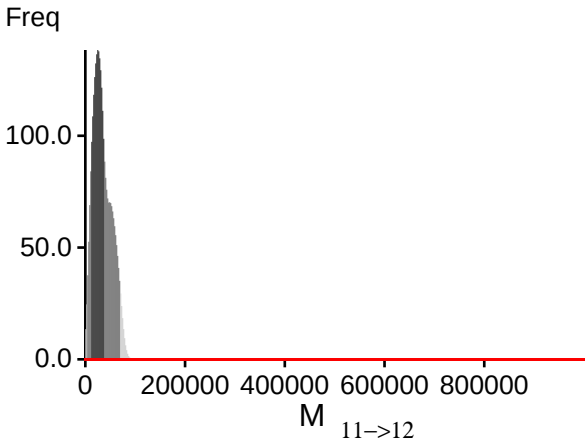

Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 3–>2
Freq
60.0
40.0
20.0
0.0
0
200000
400000
600000
800000
M 2–>3
Freq
100.0
75.0
50.0
25.0
0.0
0
200000
400000
600000
800000
M 4–>3
Freq
200.0
150.0
100.0
50.0
0.0
0
200000
400000
600000
800000
M 3–>4
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 5–>4
Freq
200.0
150.0
100.0
50.0
0.0
0
200000
400000
600000
800000
M 4–>5
Freq
100.0
50.0
0.0
0
200000
400000
600000
800000
M 6–>5
Freq
100.0
50.0
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
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
100.0
50.0
0.0
0
200000
400000
600000
800000
M
7–>8
Freq
150.0
100.0
50.0
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
100.0
50.0
0.0
0
200000
400000
600000
800000
M
10–>9
Freq
60.0
40.0
20.0
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
150.0
100.0
50.0
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
100.0
50.0
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	-2206.290263	(1a)
	-2134.796093	(1b)
Harmonic mean	-1848.074345	(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	1201/4415	0.27203
Θ_2	974/4437	0.21952
Θ_3	1435/4369	0.32845
Θ_4	1358/4412	0.30780
Θ_5	1006/4440	0.22658
Θ_6	1481/4406	0.33613
Θ_7	254/4464	0.05690
Θ_8	871/4361	0.19972
Θ_9	1286/4458	0.28847
Θ_{10}	1903/4407	0.43181
Θ_{11}	1534/4428	0.34643
Θ_{12}	1461/4332	0.33726
$M_{2->1}$	4411/4411	1.00000
$M_{1->2}$	4348/4348	1.00000
$M_{3->2}$	4321/4321	1.00000
$M_{2->3}$	4514/4514	1.00000
$M_{4->3}$	4381/4381	1.00000
$M_{3->4}$	4447/4447	1.00000
$M_{5->4}$	4403/4403	1.00000
$M_{4->5}$	4319/4319	1.00000
$M_{6->5}$	4419/4419	1.00000
$M_{5->6}$	4448/4448	1.00000
$M_{7->6}$	4421/4421	1.00000
$M_{6->7}$	4282/4282	1.00000
$M_{8->7}$	4453/4453	1.00000
$M_{7->8}$	4377/4377	1.00000
$M_{9->8}$	4377/4377	1.00000
$M_{8->9}$	4446/4446	1.00000
$M_{10->9}$	4504/4504	1.00000
$M_{9->10}$	4396/4396	1.00000
$M_{11->10}$	4419/4419	1.00000
$M_{10->11}$	4398/4398	1.00000
$M_{12->11}$	4461/4461	1.00000
$M_{11->12}$	4349/4349	1.00000
Genealogies	34031/150177	0.22661

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.79748	347.79
Θ_2	0.85053	253.87
Θ_3	0.75719	453.79
Θ_4	0.78077	389.32
Θ_5	0.80167	351.89
Θ_6	0.76921	429.35
Θ_7	0.95746	65.95
Θ_8	0.86239	224.57
Θ_9	0.81742	304.42
Θ_{10}	0.63496	678.89
Θ_{11}	0.74385	441.70
Θ_{12}	0.75634	458.19
$M_{2->1}$	0.79574	343.23
$M_{1->2}$	0.81524	306.79
$M_{3->2}$	0.80089	336.09
$M_{2->3}$	0.66143	626.43
$M_{4->3}$	0.70952	523.68
$M_{3->4}$	0.80662	324.19
$M_{5->4}$	0.68667	568.78
$M_{4->5}$	0.74635	441.43
$M_{6->5}$	0.77892	374.54
$M_{5->6}$	0.76708	399.34
$M_{7->6}$	0.78306	382.51
$M_{6->7}$	0.82905	280.51
$M_{8->7}$	0.81497	308.49
$M_{7->8}$	0.74564	440.49
$M_{9->8}$	0.74949	448.15
$M_{8->9}$	0.77975	391.83
$M_{10->9}$	0.84581	252.08
$M_{9->10}$	0.74828	435.96
$M_{11->10}$	0.65039	635.42
$M_{10->11}$	0.74003	453.14
$M_{12->11}$	0.78454	372.19
$M_{11->12}$	0.89215	174.61
Ln[Prob(D|G)]	0.98168	27.71

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