

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 18:13:47 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)

2132464641

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	$M_{2 \rightarrow 1}$	<displayed>
24	$M_{1 \rightarrow 2}$	<displayed>
25	$M_{3 \rightarrow 2}$	<displayed>
36	$M_{2 \rightarrow 3}$	<displayed>
37	$M_{4 \rightarrow 3}$	<displayed>
48	$M_{3 \rightarrow 4}$	<displayed>
49	$M_{5 \rightarrow 4}$	<displayed>
60	$M_{4 \rightarrow 5}$	<displayed>
61	$M_{6 \rightarrow 5}$	<displayed>
72	$M_{5 \rightarrow 6}$	<displayed>
73	$M_{7 \rightarrow 6}$	<displayed>
84	$M_{6 \rightarrow 7}$	<displayed>
85	$M_{8 \rightarrow 7}$	<displayed>
96	$M_{7 \rightarrow 8}$	<displayed>
97	$M_{9 \rightarrow 8}$	<displayed>
108	$M_{8 \rightarrow 9}$	<displayed>
109	$M_{10 \rightarrow 9}$	<displayed>
120	$M_{9 \rightarrow 10}$	<displayed>
121	$M_{11 \rightarrow 10}$	<displayed>
132	$M_{10 \rightarrow 11}$	<displayed>
133	$M_{12 \rightarrow 11}$	<displayed>
144	$M_{11 \rightarrow 12}$	<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

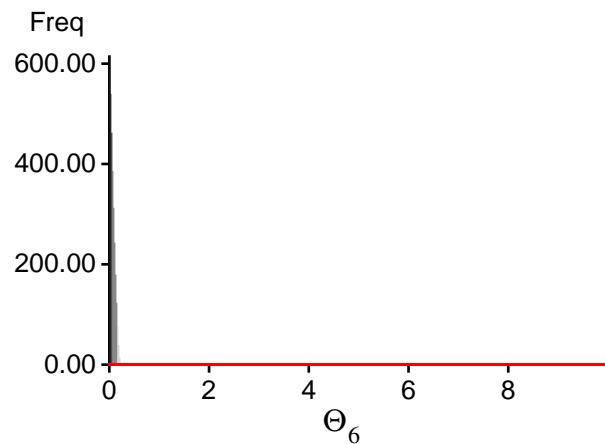
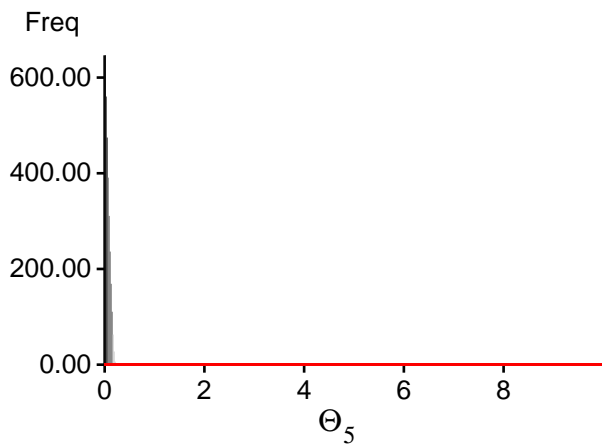
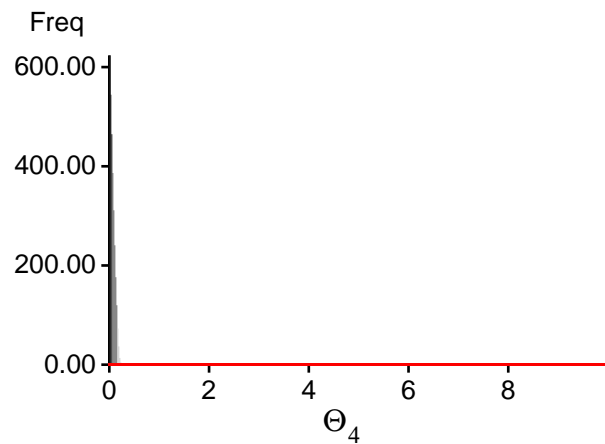
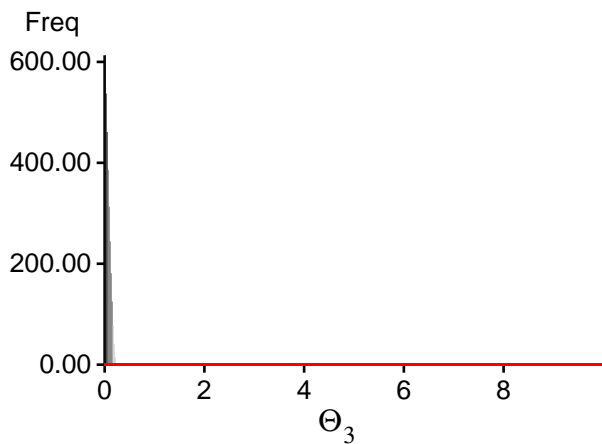
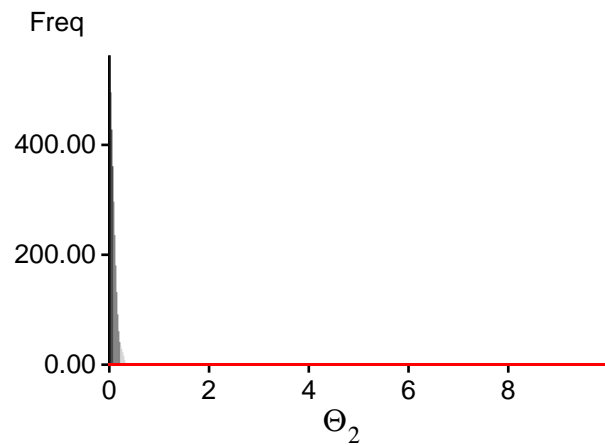
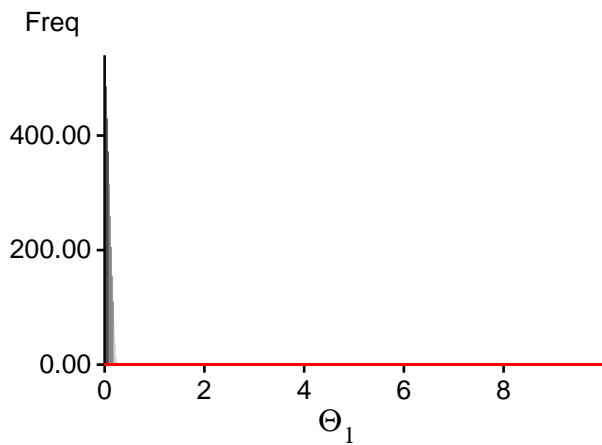
Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00001	0.00001	0.01001	0.08001	0.18001	0.09001	0.03191
1	Θ_2	0.00001	0.00001	0.01001	0.08001	0.22001	0.09001	0.05153
1	Θ_3	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02069
1	Θ_4	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01860
1	Θ_5	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01487
1	Θ_6	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.02081
1	Θ_7	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01908
1	Θ_8	0.00001	0.00001	0.01001	0.06001	0.18001	0.07001	0.02305
1	Θ_9	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01584
1	Θ_{10}	0.00001	0.00001	0.01001	0.06001	0.16001	0.07001	0.01418
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	0.01062
1	$M_{2 \rightarrow 1}$	0.0	4000.0	23000.0	38000.0	50000.0	35000.0	48489.7
1	$M_{1 \rightarrow 2}$	0.0	0.0	3000.0	12000.0	54000.0	45000.0	48244.8
1	$M_{3 \rightarrow 2}$	44000.0	66000.0	81000.0	94000.0	106000.0	69000.0	53793.8
1	$M_{2 \rightarrow 3}$	78000.0	86000.0	101000.0	116000.0	156000.0	99000.0	92185.4
1	$M_{4 \rightarrow 3}$	0.0	36000.0	49000.0	62000.0	74000.0	41000.0	36422.1
1	$M_{3 \rightarrow 4}$	0.0	0.0	15000.0	30000.0	42000.0	31000.0	40882.0
1	$M_{5 \rightarrow 4}$	24000.0	72000.0	85000.0	92000.0	102000.0	69000.0	64898.2
1	$M_{4 \rightarrow 5}$	56000.0	70000.0	87000.0	118000.0	142000.0	115000.0	171898.2
1	$M_{6 \rightarrow 5}$	0.0	0.0	11000.0	26000.0	72000.0	53000.0	78215.1
1	$M_{5 \rightarrow 6}$	8000.0	22000.0	35000.0	56000.0	90000.0	47000.0	48107.8
1	$M_{7 \rightarrow 6}$	0.0	0.0	11000.0	22000.0	34000.0	23000.0	49304.4
1	$M_{6 \rightarrow 7}$	0.0	6000.0	23000.0	36000.0	44000.0	33000.0	46516.4
1	$M_{8 \rightarrow 7}$	0.0	2000.0	11000.0	18000.0	28000.0	91000.0	71368.0
1	$M_{7 \rightarrow 8}$	12000.0	34000.0	53000.0	62000.0	74000.0	49000.0	45614.7
1	$M_{9 \rightarrow 8}$	0.0	2000.0	17000.0	28000.0	66000.0	27000.0	28549.7
1	$M_{8 \rightarrow 9}$	6000.0	12000.0	33000.0	50000.0	66000.0	47000.0	67669.2
1	$M_{10 \rightarrow 9}$	20000.0	36000.0	49000.0	58000.0	78000.0	51000.0	48813.0
1	$M_{9 \rightarrow 10}$	0.0	0.0	7000.0	18000.0	26000.0	63000.0	57361.2
1	$M_{11 \rightarrow 10}$	84000.0	94000.0	105000.0	116000.0	170000.0	107000.0	92601.9
1	$M_{10 \rightarrow 11}$	28000.0	72000.0	85000.0	98000.0	112000.0	77000.0	73167.8
1	$M_{12 \rightarrow 11}$	50000.0	78000.0	93000.0	106000.0	118000.0	77000.0	66389.8
1	$M_{11 \rightarrow 12}$	0.0	8000.0	21000.0	32000.0	124000.0	61000.0	61156.9

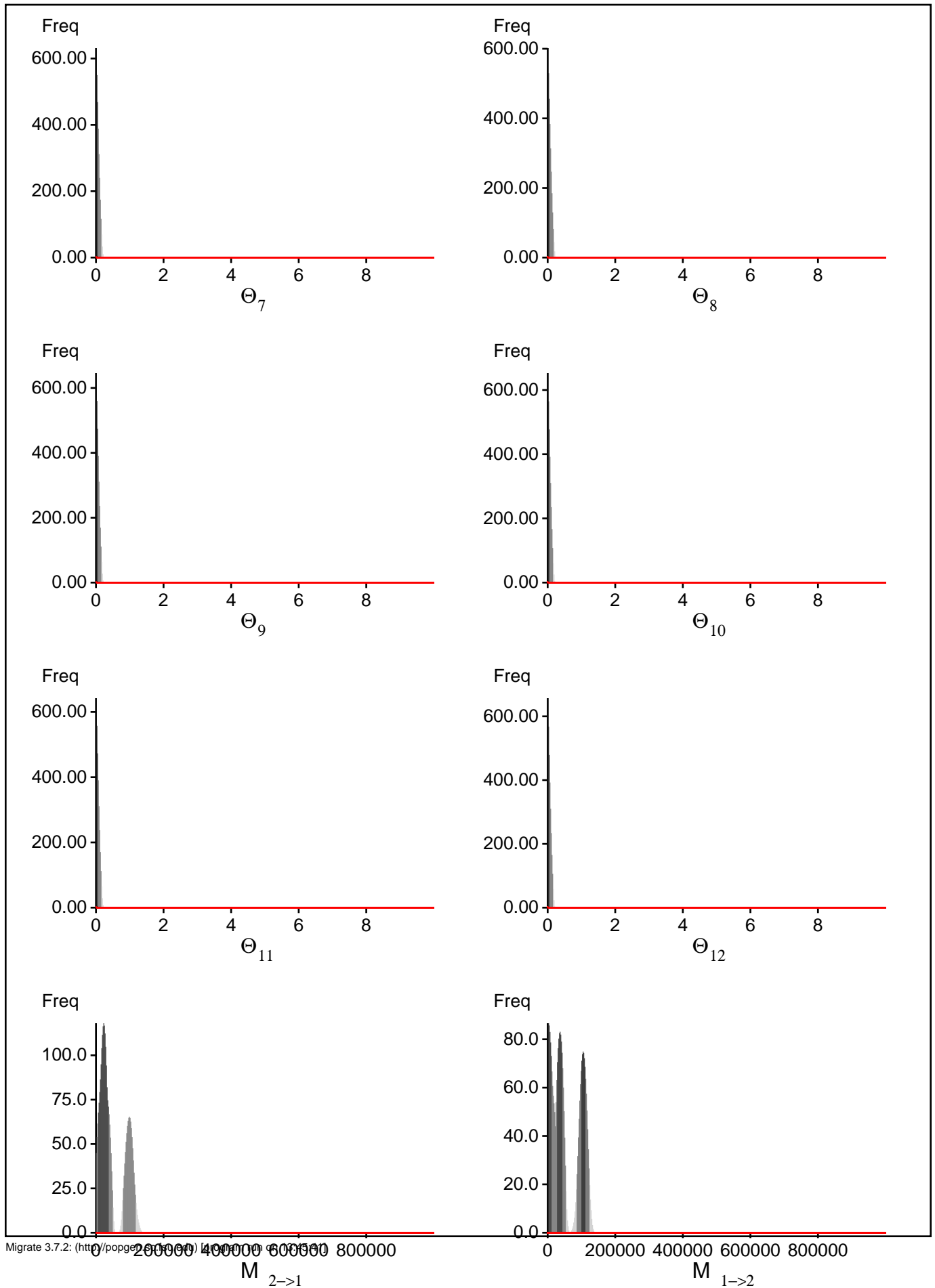
Citation suggestions:

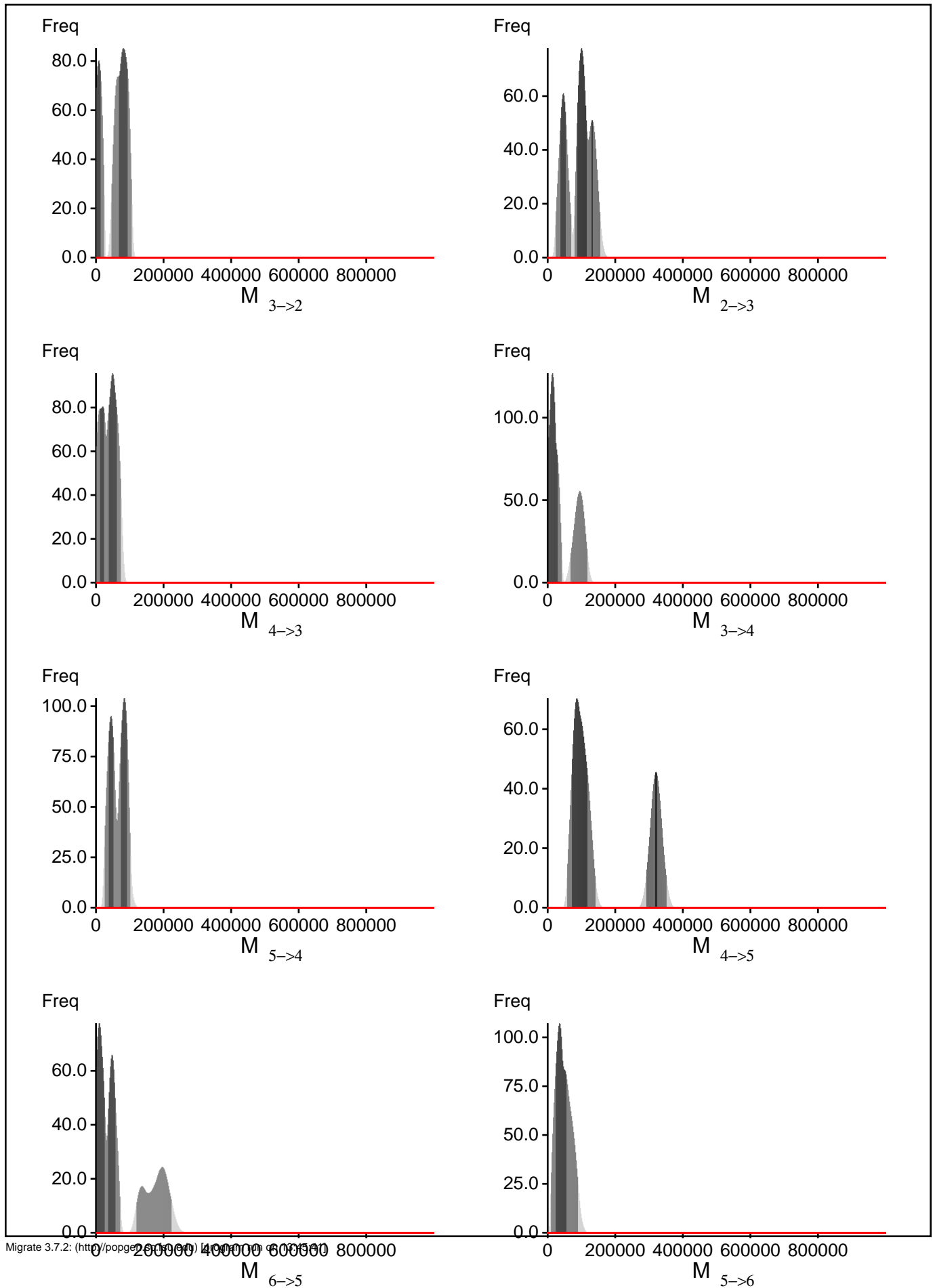
- Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345
- Beerli P., 2007. Estimation of the population scaled mutation rate from microsatellite data, *Genetics*, 177:1967-1968.
- Beerli P., 2009. How to use MIGRATE or why 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.

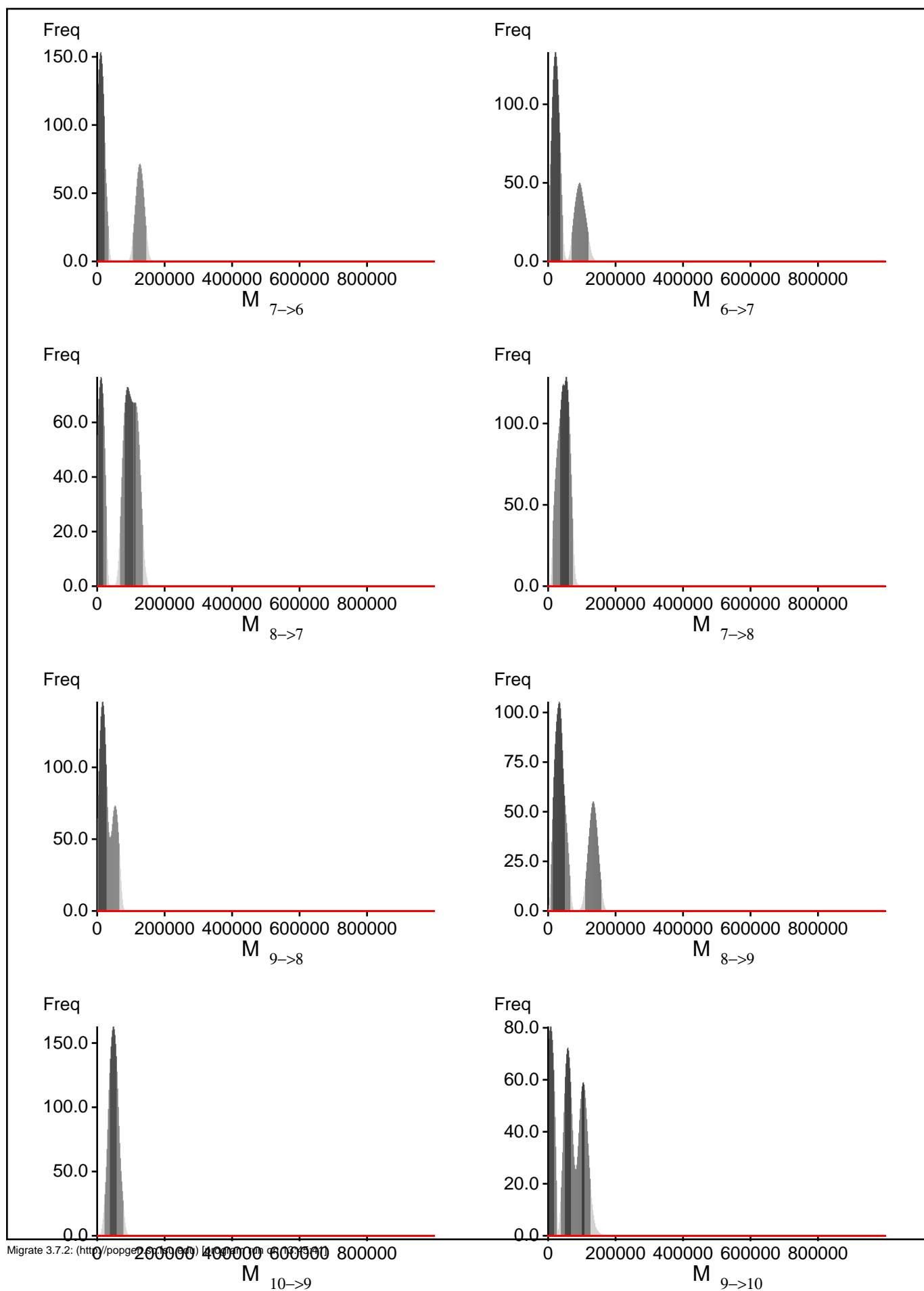
Bayesian Analysis: Posterior distribution over all loci

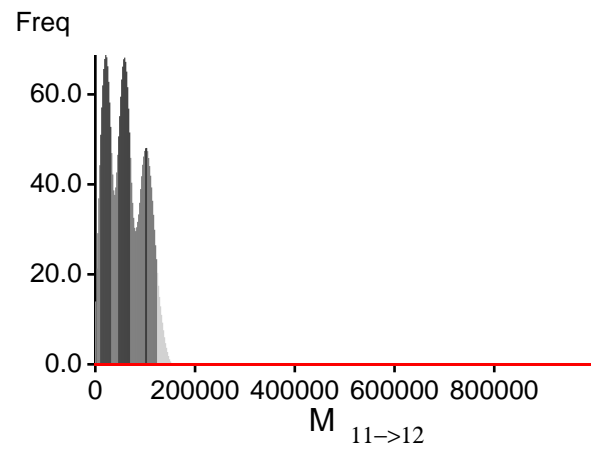
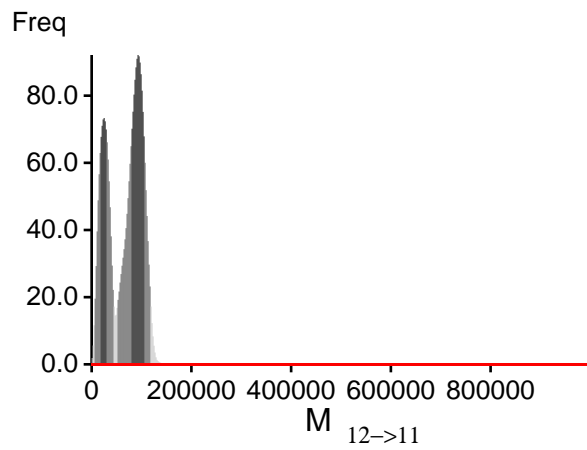
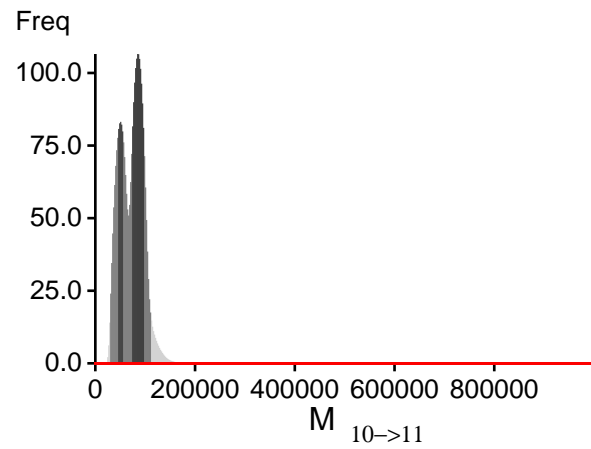
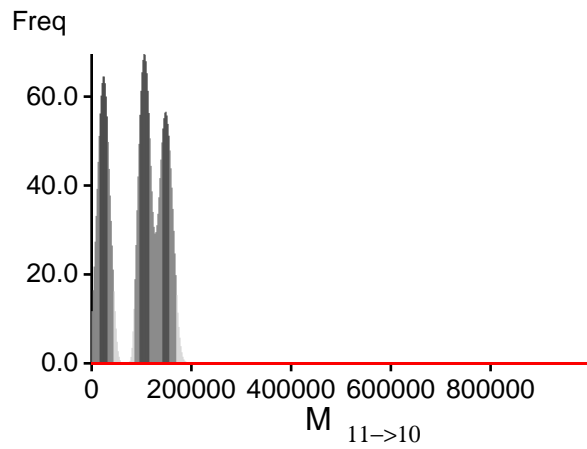


Preliminary migrate analysis of *M. californianus* CO1 haplotypes for Evolution 2 -- 8









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

Use this value for Bayes factor calculations:

$BF = \text{Exp}[\ln(\text{Prob}(D \mid \text{thisModel}) - \ln(\text{Prob}(D \mid \text{otherModel}))]$

or as $LBF = 2 (\ln(\text{Prob}(D \mid \text{thisModel}) - \ln(\text{Prob}(D \mid \text{otherModel})))$

shows the support for thisModel]

Method	$\ln(\text{Prob}(D \mid \text{Model}))$	Notes
Thermodynamic integration	-2223.583852	(1a)
	-2132.036586	(1b)
Harmonic mean	-1899.860783	(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	790/4427	0.17845
Θ_2	675/4426	0.15251
Θ_3	1660/4501	0.36881
Θ_4	914/4411	0.20721
Θ_5	2069/4402	0.47001
Θ_6	1701/4442	0.38294
Θ_7	1477/4409	0.33500
Θ_8	1297/4518	0.28707
Θ_9	2027/4405	0.46016
Θ_{10}	2107/4428	0.47584
Θ_{11}	1886/4382	0.43040
Θ_{12}	1582/4359	0.36293
M _{2→1}	4455/4455	1.00000
M _{1→2}	4322/4322	1.00000
M _{3→2}	4463/4463	1.00000
M _{2→3}	4353/4353	1.00000
M _{4→3}	4453/4453	1.00000
M _{3→4}	4345/4345	1.00000
M _{5→4}	4465/4465	1.00000
M _{4→5}	4418/4418	1.00000
M _{6→5}	4268/4268	1.00000
M _{5→6}	4424/4424	1.00000
M _{7→6}	4322/4322	1.00000
M _{6→7}	4353/4353	1.00000
M _{8→7}	4340/4340	1.00000
M _{7→8}	4568/4568	1.00000
M _{9→8}	4451/4451	1.00000
M _{8→9}	4464/4464	1.00000
M _{10→9}	4464/4464	1.00000
M _{9→10}	4425/4425	1.00000
M _{11→10}	4511/4511	1.00000
M _{10→11}	4374/4374	1.00000
M _{12→11}	4366/4366	1.00000
M _{11→12}	4472/4472	1.00000
Genealogies	36367/149814	0.24275

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.86050	234.74
Θ_2	0.89197	184.14
Θ_3	0.68306	618.35
Θ_4	0.84798	281.63
Θ_5	0.68851	563.98
Θ_6	0.73554	478.33
Θ_7	0.73853	454.11
Θ_8	0.85276	246.93
Θ_9	0.63521	680.30
Θ_{10}	0.59303	778.21
Θ_{11}	0.65717	656.96
Θ_{12}	0.71906	534.08
$M_{2 \rightarrow 1}$	0.79018	357.91
$M_{1 \rightarrow 2}$	0.75481	433.39
$M_{3 \rightarrow 2}$	0.78869	355.70
$M_{2 \rightarrow 3}$	0.85457	237.25
$M_{4 \rightarrow 3}$	0.74590	447.23
$M_{3 \rightarrow 4}$	0.83075	277.50
$M_{5 \rightarrow 4}$	0.79331	364.10
$M_{4 \rightarrow 5}$	0.77594	388.68
$M_{6 \rightarrow 5}$	0.87903	197.94
$M_{5 \rightarrow 6}$	0.82348	292.28
$M_{7 \rightarrow 6}$	0.70133	539.16
$M_{6 \rightarrow 7}$	0.86796	220.49
$M_{8 \rightarrow 7}$	0.79326	346.71
$M_{7 \rightarrow 8}$	0.74694	442.78
$M_{9 \rightarrow 8}$	0.75348	459.85
$M_{8 \rightarrow 9}$	0.85396	240.78
$M_{10 \rightarrow 9}$	0.76559	410.07
$M_{9 \rightarrow 10}$	0.72038	487.55
$M_{11 \rightarrow 10}$	0.79539	342.82
$M_{10 \rightarrow 11}$	0.69921	531.29
$M_{12 \rightarrow 11}$	0.81893	307.66
$M_{11 \rightarrow 12}$	0.88255	188.45
$\text{Ln}[\text{Prob}(D G)]$	0.97774	33.78

Potential Problems

This section reports potential problems with your run, but such reporting is often not very accurate. With 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 whether 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