

MIGRATE-N Analysis for the mtCR of *H. francisci*

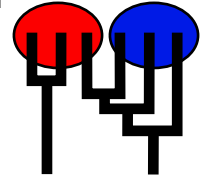
POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 4.4.4(git:) [June-1-2019]

Program started at Fri Nov 26 22:52:14 2021

Program finished at Sat Nov 27 00:01:36 2021 [Runtime:0000:01:09:22]



Options

Datatype:

DNA sequence data

Inheritance scalers in use for Thetas:

0.00

[The locus with a scaler of 1.0 used as reference]

Data set was subsampled: used a random sample of size:

30 and seed -1

Random number seed:

(with internal timer) 1650361641

Start parameters:

Theta values were generated

RANDOM start value from the prior

M values were generated

Using a percent value of the prior

Connection matrix:

m = average (average over a group of Thetas or M,

s = symmetric migration M, S = symmetric 4Nm,

0 = zero, and not estimated,

* = migration free to vary, Thetas are on diagonal

d = row population split off column population, D = split and then migration

Population	1	1	1	2	3	4	5
1 Anacapa_Island	*	*	*	*	*	*	*
1 Santa_Cruz_Isla	*	*	*	*	*	*	*
1 Catalina_Island	*	*	*	*	*	*	*
2 California_Main	*	*	*	*	*	*	*
3 Bahia_Tortugas	*	*	*	*	*	*	*

4 LSIBM * * * * * * *
 5 Las_Animas * * * * * * *

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	Θ_4	<displayed>
5	Θ_5	<displayed>
6	M _{2→1}	<displayed>
7	M _{3→1}	<displayed>
8	M _{4→1}	<displayed>
9	M _{5→1}	<displayed>
10	M _{1→2}	<displayed>
11	M _{3→2}	<displayed>
12	M _{4→2}	<displayed>
13	M _{5→2}	<displayed>
14	M _{1→3}	<displayed>
15	M _{2→3}	<displayed>
16	M _{4→3}	<displayed>
17	M _{5→3}	<displayed>
18	M _{1→4}	<displayed>
19	M _{2→4}	<displayed>
20	M _{3→4}	<displayed>
21	M _{5→4}	<displayed>
22	M _{1→5}	<displayed>
23	M _{2→5}	<displayed>
24	M _{3→5}	<displayed>
25	M _{4→5}	<displayed>

Mutation rate among loci:

Mutation rate is constant

Analysis strategy:

Bayesian inference

-Population size estimation:

Exponential Distribution

-Geneflow estimation:

Exponential Distribution

Proposal distributions for parameter

Parameter	Proposal
Theta	Metropolis sampling
M	Metropolis sampling
Divergence	Metropolis sampling
Divergence Spread	Metropolis sampling
Genealogy	Metropolis-Hastings

Prior distribution for parameter

Parameter		Prior	Minimum	Mean	Maximum	Delta	Bins	UpdateFreq
1	Theta	** Uniform	0.000000	0.000	0.000	0.000	5000	0.02000
2	Theta	** Uniform	0.000000	0.000	0.000	0.000	5000	0.02000
3	Theta	** Uniform	0.000000	0.000	0.000	0.000	5000	0.02000
4	Theta	** Uniform	0.000000	0.000	0.000	0.000	5000	0.02000
5	Theta	** Uniform	0.000000	0.000	0.000	0.000	5000	0.02000
6	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
7	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
8	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
9	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
10	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
11	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
12	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
13	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
14	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
15	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
16	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
17	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
18	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
19	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
20	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
21	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
22	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
23	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
24	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000
25	M	** Uniform	0.000000	500.0	1000.	100.0	1500	0.02000

[* * means priors were set globally]

Markov chain settings:

Number of chains	1	Long chain
Recorded steps [a]	10000	
Increment (record every x step [b])	1000	
Number of concurrent chains (replicates) [c]	1	
Visited (sampled) parameter values [a*b*c]	10000000	
Number of discard trees per chain (burn-in)	2500	

Multiple Markov chains:

Static heating scheme

1000000.00 4 chains with temperatures
3.00 1.50 1.00
Swapping interval is 1

Print options:

Data file:

migrate-n_heterodontus_input_noGI.txt

Haplotyping is turned on:	NO
Output file:	migrate-n_het_fullmodel_noGI_out
Posterior distribution raw histogram file:	bayesfile
Raw data from the MCMC run:	bayesallfile.gz
Print data:	No
Print genealogies [only some for some data type]:	None

Data summary

Data file: migrate-n_heterodontus_input_noGI.txt
 Datatype: Sequence data
 Number of loci: 1

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	HKY	[Bf:0.31 0.21 0.14 0.33, kappa=2.000]

Data set was subsampled: used a random sample of size: 30

Sites per locus

Locus	Sites
1	724

Site rate variation and probabilities:

Locus	Sublocus	Region type	Rate of change	Probability	Patch size
1	1	1	1.000	1.000	1.000

Population	Locus	Gene copies
1 Anacapa_Island	1	25
1 Santa_Cruz_Island	1	26
1 Catalina_Island	1	29
2 California_Mainland_Coast	1	131
3 Bahia_Tortugas	1	30
4 LSIBM	1	54
5 Las_Animas	1	12
Total of all populations	1	307

Subsampled dataset

Data set was subsampled randomly per population: 30 samples taken

Locus	Population	Individuals
1	Anacapa_Is	A-161 A-172 A-167 A-176 A-160 A-158 A-198 A-193 A-166 A-159 A-194 A-162 A-165 A-157 A-175 A-192 A-163 A-171 A-164 A-196 A-156 A-174 A-199 A-173 A-197
	Santa_Cruz	SC-205 SC-153 SC-90 SC-207 SC-187 SC-151 SC-209 SC-154 SC-91 SC-203 SC-206 SC-184 SC-155 SC-152 SC-188 SC-185 SC-182 SC-169 SC-204 SC-150 SC-64 SC-186 SC-183 SC-168 SC-62 SC-63
	Catalina_I	C-F34 C-F48 C-F43 C-F54 C-F28 C-F13 C-F12 C-F11 C-F197 C-F26 C-F177 C-F27 C-F32 C-F35 C-F4 C-F10 C-F16 C-F44 C-F198 C-F29 C-F8 C-F17 C-F21 C-F25 C-F45 C-F30 C-F47 C-F53 C-F14
	California	MLCA-F69 MLCA-F218 MLCA-F83 MLCA-F91 MLCA-F170 MLCA-F165 MLCA-F255 MLCA-68 MLCA-F284 MLCA-F237 MLCA-F150 MLCA-F239 MLCA-F301 MLCA-F93 MLCA-79 MLCA-F85 MLCA-F234 MLCA-F265 MLCA-F162 MLCA-F209 MLCA-F104 MLCA-F202 MLCA-F231 MLCA-F252 MLCA-F250 MLCA-F145 MLCA-111 MLCA-F176 MLCA-F241 MLCA-F147
	Bahia_Tort	BT-HF31 BT-HF48 BT-HF47 BT-HF37 BT-HF32 BT-HF46 BT-HF44 BT-HF23 BT-HF21 BT-HF41 BT-HF30 BT-HF22 BT-HF27 BT-HF38 BT-HF25 BT-HF40 BT-HF39 BT-HF24 BT-HF28 BT-HF34 BT-HF29 BT-HF19 BT-HF43 BT-HF20 BT-HF36 BT-HF35 BT-HF42 BT-HF26 BT-HF33 BT-HF45
	LSIBM	LSIBM-398 LSIBM-337 LSIBM-405 LSIBM-410 LSIBM-668 LSIBM-315 LSIBM-400 LSIBM-407 LSIBM-675 LSIBM-406 LSIBM-476 LSIBM-478 LSIBM-480 LSIBM-319 LSIBM-479 LSIBM-321 LSIBM-392 LSIBM-390 LSIBM-684 LSIBM-674 LSIBM-408 LSIBM-389 LSIBM-391 LSIBM-673 LSIBM-669 LSIBM-399 LSIBM-382 LSIBM-402 LSIBM-385 LSIBM-484
	Las_Animas	LA-647 LA-649 LA-646 LA-653 LA-652 LA-651 LA-648 LA-655 LA-650 LA-645 LA-656 LA-654

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	Θ_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	Θ_3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	Θ_4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	Θ_5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
1	$M_{2 \rightarrow 1}$	0.0	184.0	371.0	422.0	478.7	271.0	518.7
1	$M_{3 \rightarrow 1}$	0.0	167.3	369.7	412.7	468.7	263.7	504.0
1	$M_{4 \rightarrow 1}$	12.7	231.3	379.0	452.0	524.0	301.7	592.0
1	$M_{5 \rightarrow 1}$	0.0	0.0	0.3	72.0	474.7	257.7	525.6
1	$M_{1 \rightarrow 2}$	140.0	282.7	405.0	506.7	574.7	381.7	728.5
1	$M_{3 \rightarrow 2}$	47.3	279.3	393.7	473.3	544.7	333.0	671.8
1	$M_{4 \rightarrow 2}$	0.0	0.0	0.3	66.0	416.0	250.3	493.3
1	$M_{5 \rightarrow 2}$	0.0	193.3	375.0	431.3	516.7	278.3	526.5
1	$M_{1 \rightarrow 3}$	0.0	59.3	131.0	299.3	440.0	246.3	475.9
1	$M_{2 \rightarrow 3}$	0.0	70.0	125.0	216.7	462.0	253.0	512.6
1	$M_{4 \rightarrow 3}$	114.0	293.3	399.0	486.0	570.0	369.0	720.8
1	$M_{5 \rightarrow 3}$	0.0	196.0	376.3	434.0	500.0	279.7	552.7
1	$M_{1 \rightarrow 4}$	0.0	24.7	109.7	264.7	409.3	240.3	466.5
1	$M_{2 \rightarrow 4}$	30.0	189.3	380.3	436.0	464.0	274.3	554.2
1	$M_{3 \rightarrow 4}$	36.7	292.0	397.7	480.7	556.7	347.0	704.0
1	$M_{5 \rightarrow 4}$	0.0	163.3	368.3	406.0	452.0	262.3	520.8
1	$M_{1 \rightarrow 5}$	0.0	182.0	372.3	424.7	494.0	269.7	523.2
1	$M_{2 \rightarrow 5}$	0.0	0.0	0.3	79.3	474.0	263.0	540.7
1	$M_{3 \rightarrow 5}$	44.7	278.0	393.0	474.7	555.3	337.0	664.6
1	$M_{4 \rightarrow 5}$	0.0	220.7	376.3	452.7	528.0	294.3	588.1

Citation suggestions:

Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters.

Bioinformatics 22:341-345

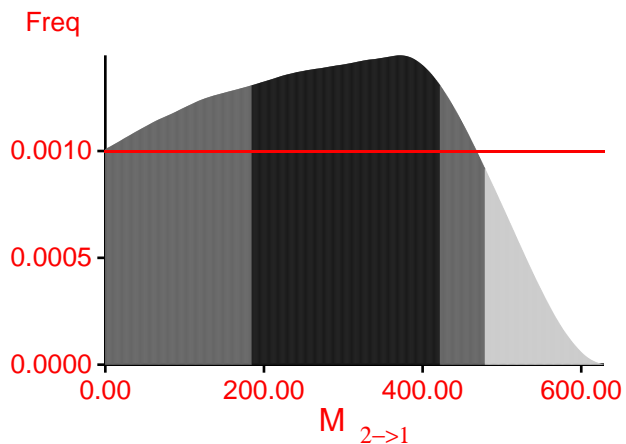
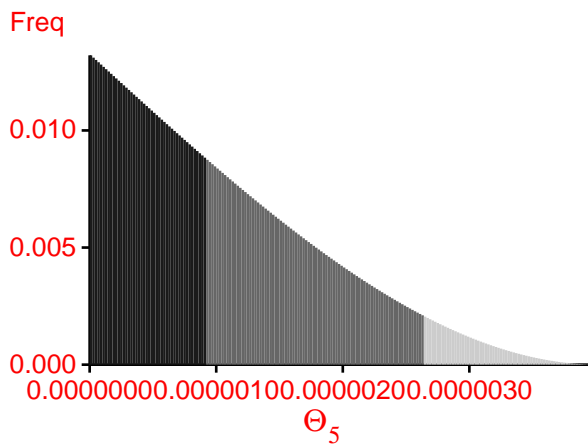
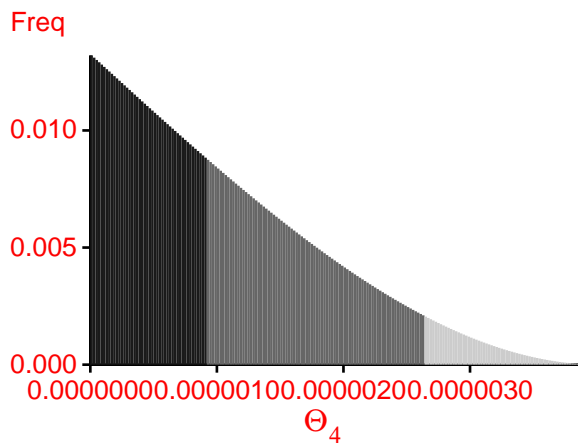
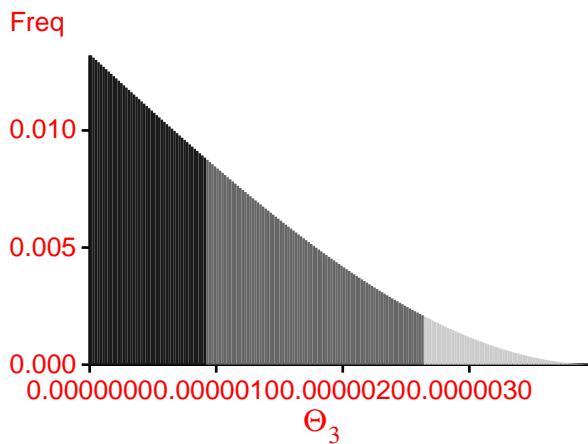
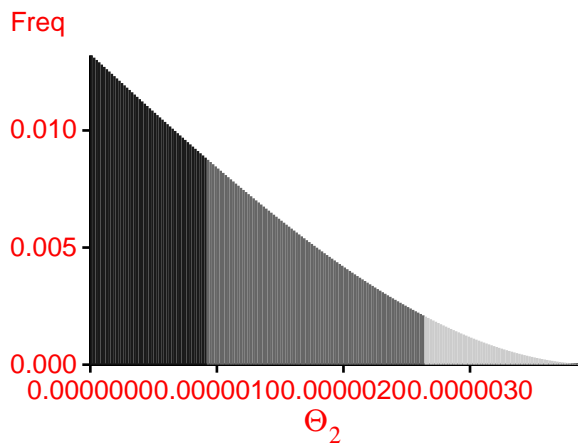
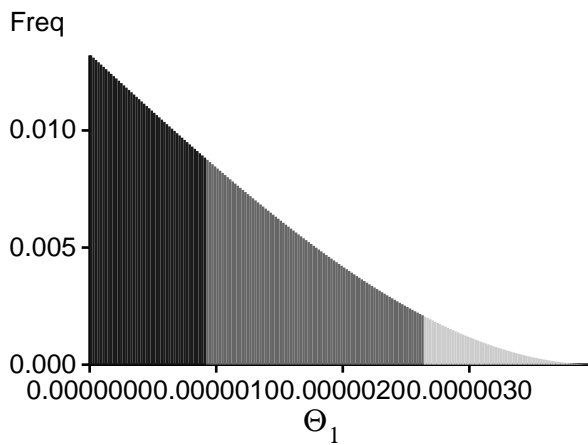
Beerli P., 2007. Estimation of the population scaled mutation rate from microsatellite data,

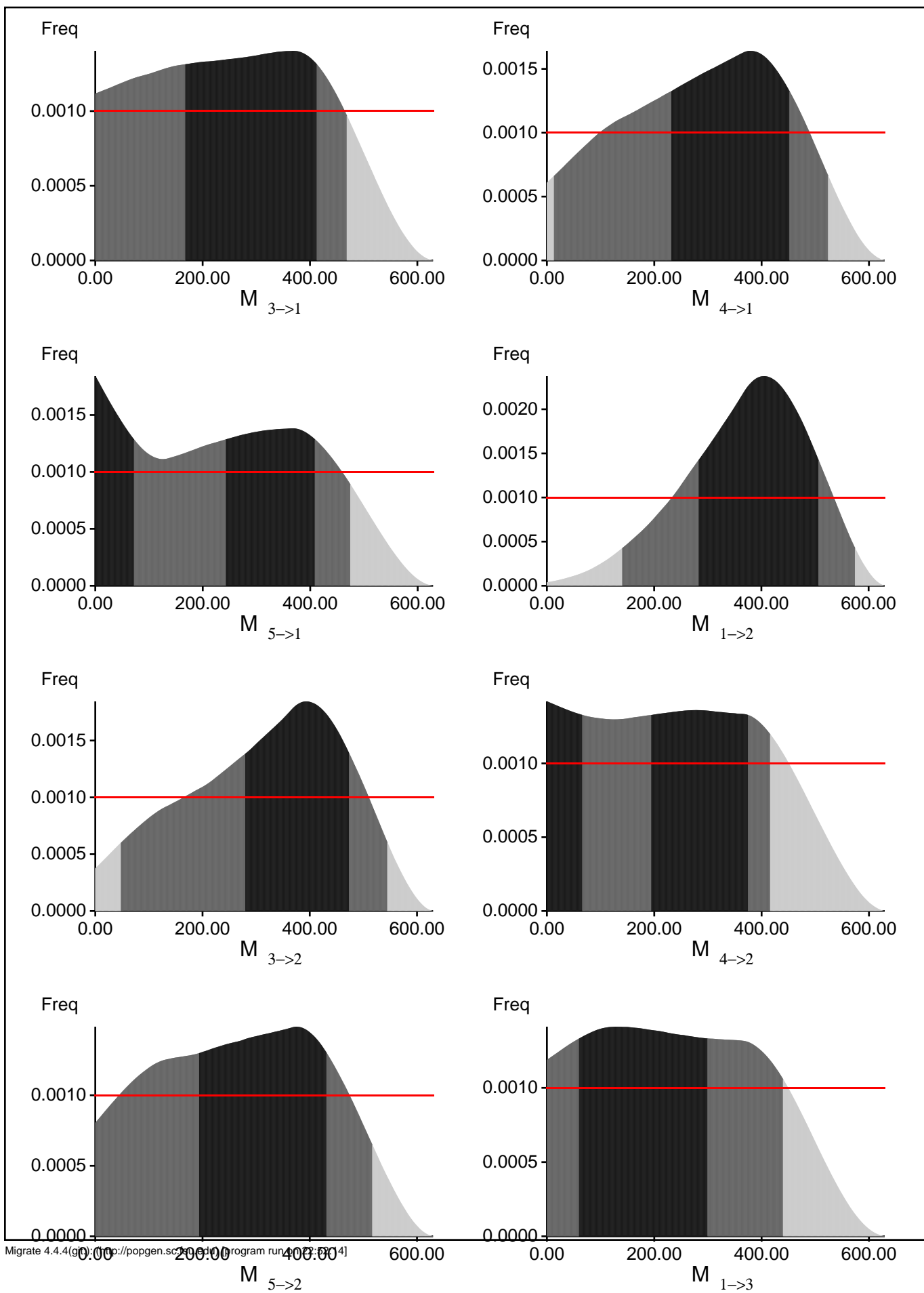
Genetics, 177:1967-1968.

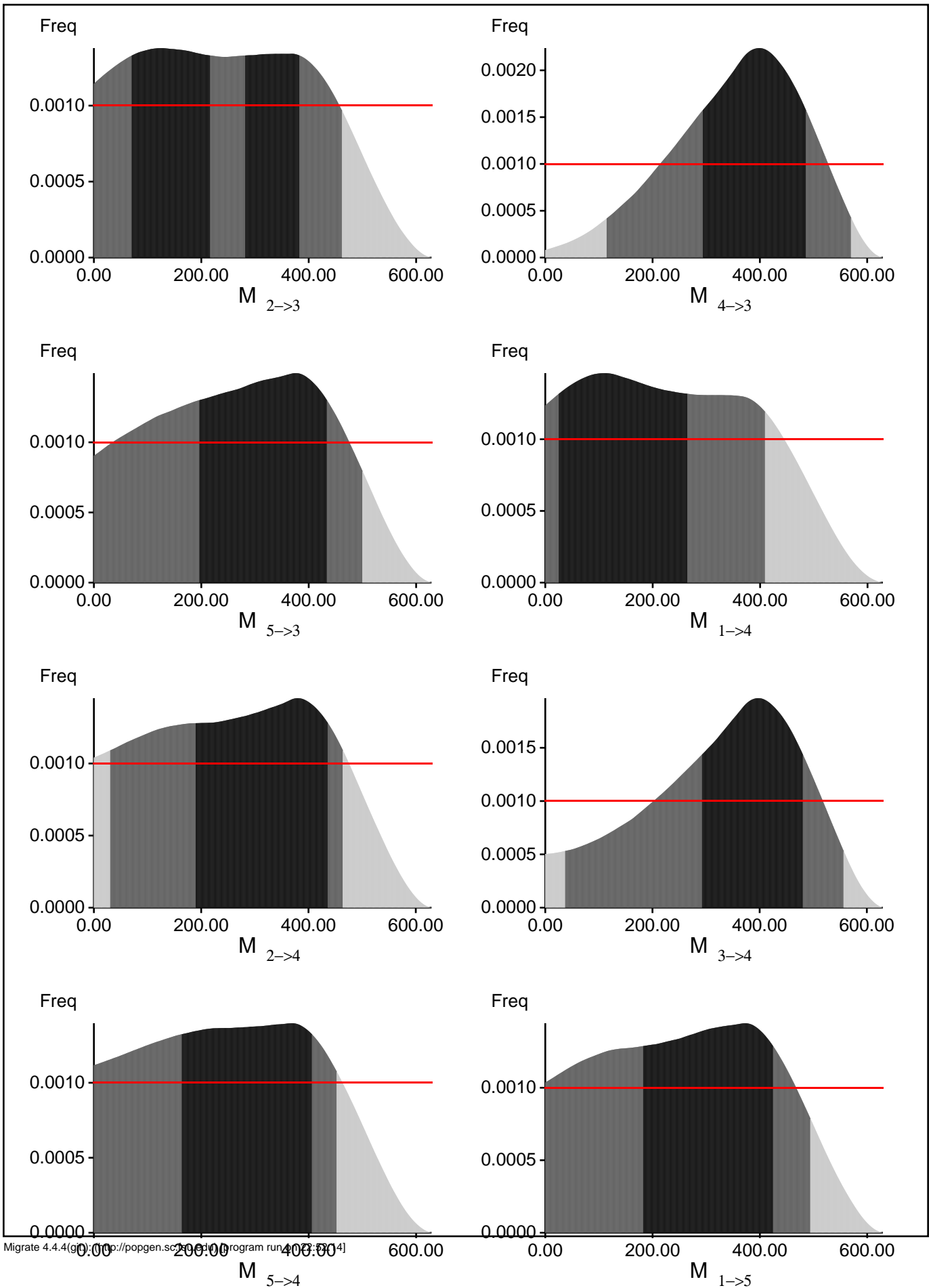
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use?

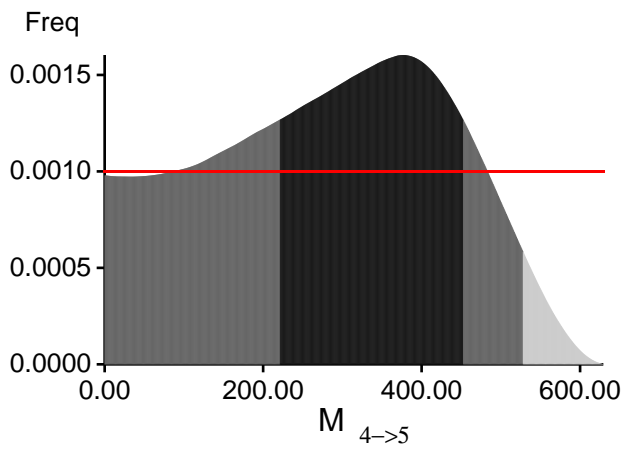
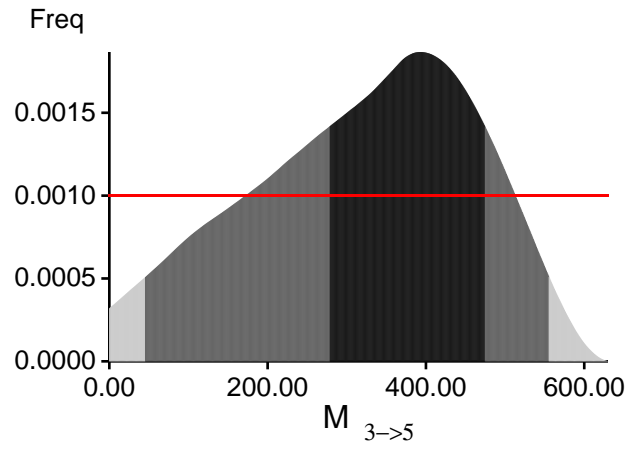
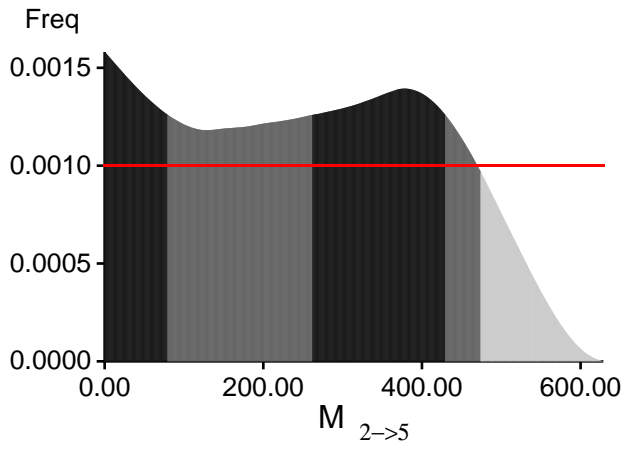
In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp. 42-79.

Bayesian Analysis: Posterior distribution for locus 1









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	-1583.599226	(1a)
	-1305.628533	(1b)
Harmonic mean	-1167.115774	(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	103974/200484	0.51861
Θ_2	61795/199893	0.30914
Θ_3	126228/199811	0.63174
Θ_4	66763/200000	0.33381
Θ_5	66135/199987	0.33070
M _{2→1}	169683/199518	0.85046
M _{3→1}	172219/200533	0.85881
M _{4→1}	163577/199919	0.81822
M _{5→1}	169872/200014	0.84930
M _{1→2}	144550/199991	0.72278
M _{3→2}	149369/199315	0.74941
M _{4→2}	179524/199909	0.89803
M _{5→2}	176838/199647	0.88575
M _{1→3}	179365/200473	0.89471
M _{2→3}	170578/199965	0.85304
M _{4→3}	146172/200767	0.72807
M _{5→3}	171165/200394	0.85414
M _{1→4}	179068/199906	0.89576
M _{2→4}	170539/199536	0.85468
M _{3→4}	146242/200136	0.73071
M _{5→4}	176543/200261	0.88156
M _{1→5}	181251/200976	0.90185
M _{2→5}	177216/200008	0.88604
M _{3→5}	156473/199661	0.78369
M _{4→5}	171306/200415	0.85476
Genealogies	1814302/4998481	0.36297

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.00279	9943.29
Θ_2	0.02958	9424.40
Θ_3	0.06078	8853.20
Θ_4	0.03569	9309.86
Θ_5	0.06513	8776.23
M _{2→1}	0.18361	6896.73
M _{3→1}	0.14673	7440.23
M _{4→1}	0.20387	6612.49
M _{5→1}	0.15826	7266.61
M _{1→2}	-0.00568	10113.26
M _{3→2}	0.21476	6463.50
M _{4→2}	0.06558	8768.32
M _{5→2}	0.09351	8288.90
M _{1→3}	0.08950	8356.19
M _{2→3}	0.17067	7083.50
M _{4→3}	0.06915	8705.59
M _{5→3}	0.15886	7257.58
M _{1→4}	0.07707	8568.02
M _{2→4}	0.14699	7436.17
M _{3→4}	0.16310	7194.74
M _{5→4}	0.11463	7942.32
M _{1→5}	0.09043	8340.62
M _{2→5}	0.10955	8024.45
M _{3→5}	0.13429	7631.36
M _{4→5}	0.11879	7875.64
Genealogies	0.66559	2007.59

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