

Current protocols example dataset

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 4.4.3(git:) [March-21-2019]

Using Intel AVX (Advanced Vector Extensions)

Compiled for PARALLEL computer architectures

One master and 8 compute nodes are available.

Compiled for a SYMMETRIC multiprocessors (Grandcentral)

Program started at Mon Jun 3 08:46:14 2019

Program finished at Mon Jun 3 08:49:31 2019 [Runtime:0000:00:03:17]



Options

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) 3296668800

Start parameters:

Theta values were generated Using a percent value of 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	2	3
1 Arbon_1	*	*	*
2 Berg_2	*	*	*
3 Chur_3	*	*	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	$M_{2 \rightarrow 1}$	<displayed>
5	$M_{3 \rightarrow 1}$	<displayed>
6	$M_{1 \rightarrow 2}$	<displayed>
7	$M_{3 \rightarrow 2}$	<displayed>
8	$M_{1 \rightarrow 3}$	<displayed>
9	$M_{2 \rightarrow 3}$	<displayed>

Mutation rate among loci:

Mutation rate is constant for all loci

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	0.050	0.100	0.010	1500	0.05556
2	Theta **	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
3	Theta **	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
4	M **	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
5	M **	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
6	M **	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
7	M **	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
8	M **	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
9	M **	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556

[* * means priors were set globally]

Markov chain settings:

Long chain

Number of chains	1
Recorded steps [a]	5000
Increment (record every x step [b])	10
Number of concurrent chains (replicates) [c]	1
Visited (sampled) parameter values [a*b*c]	50000
Number of discard trees per chain (burn-in)	5000

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: infile
Haplotyping is turned on: NO
Output file: outfile_short
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: infile
 Datatype: Haplotype data
 Number of loci: 10

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Jukes-Cantor	[Basefreq: =0.25]
2	1	Jukes-Cantor	[Basefreq: =0.25]
3	1	Jukes-Cantor	[Basefreq: =0.25]
4	1	Jukes-Cantor	[Basefreq: =0.25]
5	1	Jukes-Cantor	[Basefreq: =0.25]
6	1	Jukes-Cantor	[Basefreq: =0.25]
7	1	Jukes-Cantor	[Basefreq: =0.25]
8	1	Jukes-Cantor	[Basefreq: =0.25]
9	1	Jukes-Cantor	[Basefreq: =0.25]
10	1	Jukes-Cantor	[Basefreq: =0.25]

Sites per locus

Locus	Sites
1	1000
2	1000
3	1000
4	1000
5	1000
6	1000
7	1000
8	1000
9	1000
10	1000

Site rate variation and probabilities:

Locus	Sublocus	Region type	Rate of change	Probability	Patch size
1	1	1	1.000	1.000	1.000
2	1	1	1.000	1.000	1.000
3	1	1	1.000	1.000	1.000
4	1	1	1.000	1.000	1.000
5	1	1	1.000	1.000	1.000
6	1	1	1.000	1.000	1.000

7	1	1	1.000	1.000	1.000		
8	1	1	1.000	1.000	1.000		
9	1	1	1.000	1.000	1.000		
10	1	1	1.000	1.000	1.000		
Population			Locus		Gene copies		
					data	(missing)	
1 Arbon_1			1		10		
			2		10		
			3		10		
			4		10		
			5		10		
			6		10		
			7		10		
			8		10		
			9		10		
			10		10		
2 Berg_2			1		10		
			2		10		
			3		10		
			4		10		
			5		10		
			6		10		
			7		10		
			8		10		
			9		10		
			10		10		
3 Chur_3			1		10		
			2		10		
			3		10		
			4		10		
			5		10		
			6		10		
			7		10		
			8		10		
			9		10		
			10		10		
Total of all populations			1		30	(0)	
			2		30	(0)	
			3		30	(0)	
			4		30	(0)	
			5		30	(0)	
			6		30	(0)	
			7		30	(0)	
			8		30	(0)	
			9		30	(0)	

10

30

(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00727	0.00727	0.00737	0.00740	0.00887	0.01070	0.01408
1	Θ_2	0.00153	0.00247	0.00297	0.00353	0.00573	0.00383	0.00399
1	Θ_3	0.00780	0.00793	0.00817	0.00827	0.01267	0.01090	0.01247
1	$M_{2 \rightarrow 1}$	426.667	483.333	498.333	503.333	676.667	515.000	524.940
1	$M_{3 \rightarrow 1}$	93.333	250.000	275.000	283.333	286.667	218.333	223.236
1	$M_{1 \rightarrow 2}$	160.000	246.667	251.667	256.667	306.667	365.000	377.173
1	$M_{3 \rightarrow 2}$	30.000	103.333	108.333	110.000	276.667	235.000	279.275
1	$M_{1 \rightarrow 3}$	50.000	126.667	141.667	146.667	276.667	181.667	194.512
1	$M_{2 \rightarrow 3}$	0.000	83.333	101.667	103.333	223.333	108.333	119.258
2	Θ_1	0.00073	0.00147	0.00170	0.00220	0.00500	0.00277	0.00324
2	Θ_2	0.00133	0.00320	0.00330	0.00347	0.00653	0.00410	0.00494
2	Θ_3	0.00207	0.00493	0.00503	0.00540	0.00907	0.00530	0.00597
2	$M_{2 \rightarrow 1}$	223.333	223.333	228.333	230.000	230.000	678.333	759.407
2	$M_{3 \rightarrow 1}$	126.667	163.333	171.667	173.333	173.333	271.667	349.879
2	$M_{1 \rightarrow 2}$	186.667	186.667	195.000	196.667	213.333	358.333	400.560
2	$M_{3 \rightarrow 2}$	326.667	346.667	351.667	353.333	366.667	568.333	598.809
2	$M_{1 \rightarrow 3}$	283.333	310.000	318.333	320.000	336.667	458.333	538.005
2	$M_{2 \rightarrow 3}$	513.333	516.667	521.667	526.667	530.000	315.000	377.675
3	Θ_1	0.00180	0.00273	0.00283	0.00300	0.00393	0.00317	0.00388
3	Θ_2	0.00280	0.00667	0.00683	0.00687	0.00853	0.00603	0.00638
3	Θ_3	0.01107	0.01193	0.01203	0.01207	0.01233	0.01230	0.01645
3	$M_{2 \rightarrow 1}$	626.667	643.333	648.333	653.333	670.000	728.333	770.803
3	$M_{3 \rightarrow 1}$	290.000	293.333	298.333	300.000	340.000	421.667	450.893
3	$M_{1 \rightarrow 2}$	273.333	393.333	411.667	433.333	633.333	428.333	446.048
3	$M_{3 \rightarrow 2}$	0.000	3.333	8.333	43.333	130.000	61.667	73.575
3	$M_{1 \rightarrow 3}$	320.000	340.000	348.333	353.333	406.667	425.000	446.826
3	$M_{2 \rightarrow 3}$	6.667	120.000	128.333	133.333	333.333	175.000	195.075
4	Θ_1	0.00160	0.00240	0.00257	0.00320	0.00593	0.00390	0.00430
4	Θ_2	0.00193	0.00313	0.00330	0.00340	0.00700	0.00583	0.00622
4	Θ_3	0.00267	0.00527	0.00557	0.00567	0.01040	0.00670	0.00732
4	$M_{2 \rightarrow 1}$	256.667	360.000	365.000	366.667	453.333	365.000	390.979
4	$M_{3 \rightarrow 1}$	463.333	503.333	508.333	516.667	616.667	518.333	531.756
4	$M_{1 \rightarrow 2}$	243.333	336.667	341.667	346.667	473.333	398.333	414.542

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
4	$M_{3 \rightarrow 2}$	253.333	380.000	388.333	406.667	680.000	471.667	484.429
4	$M_{1 \rightarrow 3}$	70.000	266.667	271.667	273.333	360.000	235.000	254.255
4	$M_{2 \rightarrow 3}$	230.000	430.000	445.000	453.333	556.667	448.333	522.662
5	Θ_1	0.00353	0.00540	0.00550	0.00560	0.00947	0.00623	0.00666
5	Θ_2	0.00167	0.00167	0.00197	0.00200	0.00380	0.00630	0.01165
5	Θ_3	0.00387	0.00560	0.00570	0.00573	0.00900	0.00997	0.01441
5	$M_{2 \rightarrow 1}$	33.333	116.667	121.667	123.333	220.000	125.000	158.593
5	$M_{3 \rightarrow 1}$	236.667	260.000	268.333	283.333	363.333	291.667	306.653
5	$M_{1 \rightarrow 2}$	2046.667	2056.667	2061.667	2070.000	2086.667	1845.000	1842.756
5	$M_{3 \rightarrow 2}$	273.333	296.667	301.667	303.333	310.000	288.333	339.769
5	$M_{1 \rightarrow 3}$	313.333	403.333	408.333	410.000	720.000	501.667	544.737
5	$M_{2 \rightarrow 3}$	210.000	250.000	258.333	260.000	556.667	445.000	472.524
6	Θ_1	0.00140	0.00140	0.00150	0.00153	0.00153	0.05097	0.04943
6	Θ_2	0.00220	0.00313	0.00323	0.00333	0.00647	0.00423	0.00447
6	Θ_3	0.00853	0.00920	0.00930	0.00940	0.00940	0.01663	0.02296
6	$M_{2 \rightarrow 1}$	1163.333	1170.000	1175.000	1176.667	1180.000	1488.333	1536.492
6	$M_{3 \rightarrow 1}$	126.667	200.000	208.333	210.000	306.667	221.667	241.268
6	$M_{1 \rightarrow 2}$	33.333	96.667	105.000	106.667	210.000	248.333	271.992
6	$M_{3 \rightarrow 2}$	113.333	156.667	168.333	173.333	330.000	248.333	284.518
6	$M_{1 \rightarrow 3}$	306.667	310.000	325.000	326.667	343.333	398.333	380.990
6	$M_{2 \rightarrow 3}$	0.000	0.000	11.667	20.000	63.333	88.333	180.709
7	Θ_1	0.00287	0.00653	0.00663	0.00673	0.00793	0.00590	0.00586
7	Θ_2	0.03087	0.03087	0.03097	0.03100	0.03107	0.03983	0.04405
7	Θ_3	0.01967	0.01980	0.01990	0.01993	0.02020	0.01810	0.02090
7	$M_{2 \rightarrow 1}$	150.000	283.333	288.333	296.667	406.667	308.333	318.341
7	$M_{3 \rightarrow 1}$	206.667	403.333	415.000	416.667	523.333	418.333	464.189
7	$M_{1 \rightarrow 2}$	1750.000	1763.333	1768.333	1770.000	1773.333	1965.000	2007.169
7	$M_{3 \rightarrow 2}$	0.000	6.667	11.667	13.333	76.667	191.667	252.949
7	$M_{1 \rightarrow 3}$	256.667	256.667	261.667	263.333	286.667	418.333	439.310
7	$M_{2 \rightarrow 3}$	73.333	206.667	211.667	226.667	343.333	245.000	271.251
8	Θ_1	0.00127	0.00173	0.00197	0.00207	0.00300	0.00430	0.00639
8	Θ_2	0.00393	0.00553	0.00590	0.00620	0.00667	0.00663	0.00702
8	Θ_3	0.00353	0.00433	0.00443	0.00453	0.00460	0.00750	0.00813
8	$M_{2 \rightarrow 1}$	1000.000	1016.667	1021.667	1023.333	1056.667	988.333	1098.717
8	$M_{3 \rightarrow 1}$	150.000	170.000	175.000	190.000	200.000	248.333	294.493
8	$M_{1 \rightarrow 2}$	76.667	83.333	88.333	96.667	193.333	158.333	173.696
8	$M_{3 \rightarrow 2}$	13.333	66.667	81.667	110.000	220.000	111.667	118.797

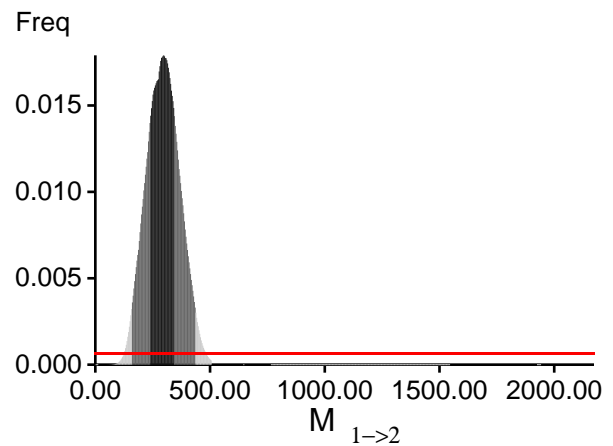
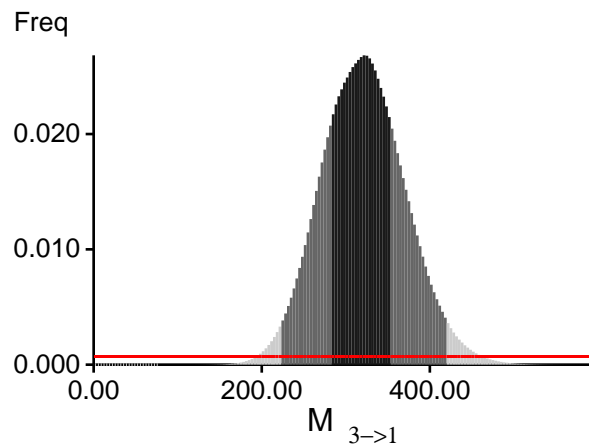
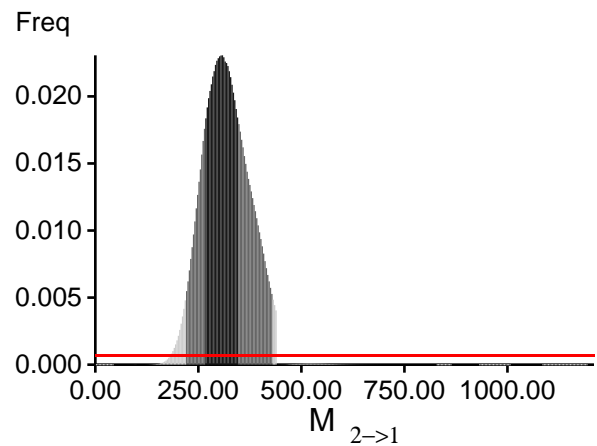
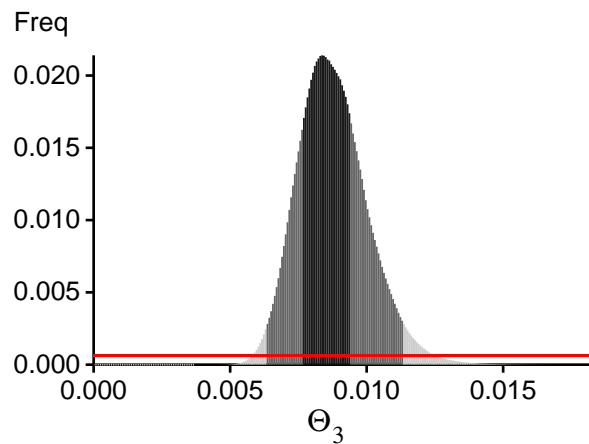
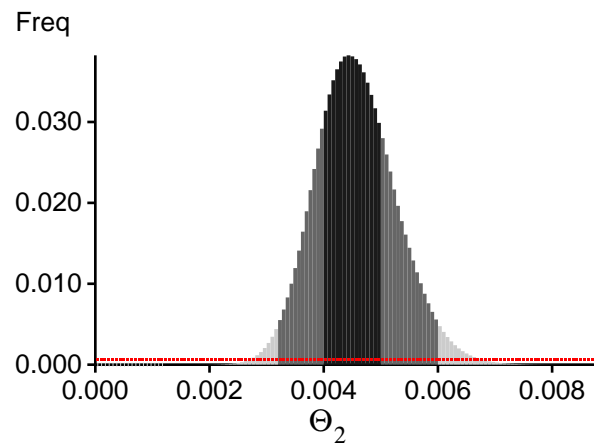
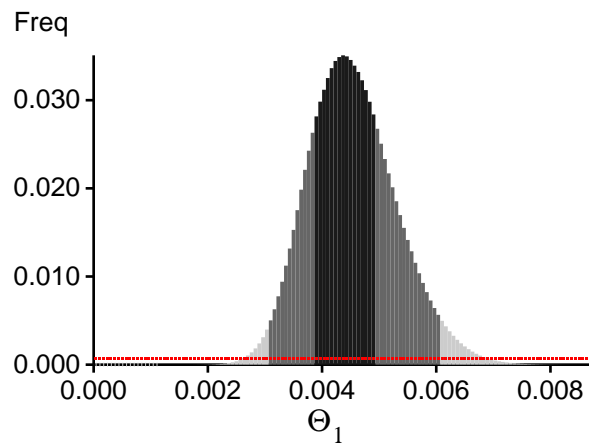
Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$M_{1 \rightarrow 3}$	3.333	50.000	68.333	80.000	183.333	115.000	129.795
8	$M_{2 \rightarrow 3}$	110.000	140.000	148.333	153.333	420.000	258.333	274.737
9	Θ_1	0.00493	0.00500	0.00510	0.00513	0.00720	0.00723	0.00760
9	Θ_2	0.00187	0.00220	0.00243	0.00260	0.00560	0.00350	0.00371
9	Θ_3	0.01293	0.01293	0.01310	0.01313	0.01340	0.03503	0.04057
9	$M_{2 \rightarrow 1}$	110.000	133.333	138.333	140.000	463.333	265.000	279.469
9	$M_{3 \rightarrow 1}$	316.667	400.000	408.333	410.000	776.667	538.333	552.684
9	$M_{1 \rightarrow 2}$	3.333	60.000	65.000	96.667	233.333	108.333	191.355
9	$M_{3 \rightarrow 2}$	283.333	380.000	418.333	433.333	623.333	461.667	509.549
9	$M_{1 \rightarrow 3}$	750.000	876.667	881.667	883.333	1050.000	911.667	929.284
9	$M_{2 \rightarrow 3}$	176.667	303.333	308.333	356.667	496.667	358.333	367.622
10	Θ_1	0.00133	0.00260	0.00317	0.00320	0.00580	0.00370	0.00384
10	Θ_2	0.00413	0.00507	0.00523	0.00527	0.00720	0.00837	0.01049
10	Θ_3	0.01480	0.01493	0.01517	0.01560	0.01580	0.02510	0.03487
10	$M_{2 \rightarrow 1}$	0.000	86.667	95.000	96.667	463.333	255.000	407.066
10	$M_{3 \rightarrow 1}$	153.333	446.667	468.333	470.000	560.000	391.667	388.714
10	$M_{1 \rightarrow 2}$	156.667	166.667	171.667	173.333	173.333	635.000	618.239
10	$M_{3 \rightarrow 2}$	46.667	116.667	141.667	146.667	180.000	405.000	437.389
10	$M_{1 \rightarrow 3}$	320.000	320.000	325.000	336.667	630.000	521.667	549.346
10	$M_{2 \rightarrow 3}$	66.667	83.333	91.667	93.333	96.667	1165.000	1073.379
All	Θ_1	0.00300	0.00380	0.00437	0.00493	0.00607	0.00457	0.00454
All	Θ_2	0.00313	0.00393	0.00443	0.00500	0.00600	0.00457	0.00457
All	Θ_3	0.00627	0.00760	0.00837	0.00940	0.01133	0.00877	0.00878
All	$M_{2 \rightarrow 1}$	216.667	263.333	308.333	346.667	430.000	318.333	317.643
All	$M_{3 \rightarrow 1}$	220.000	280.000	321.667	353.333	420.000	325.000	321.411
All	$M_{1 \rightarrow 2}$	156.667	236.667	298.333	343.333	436.667	298.333	296.410
All	$M_{3 \rightarrow 2}$	126.667	166.667	198.333	226.667	300.000	208.333	208.477
All	$M_{1 \rightarrow 3}$	130.000	206.667	248.333	250.000	250.000	225.000	226.434
All	$M_{2 \rightarrow 3}$	153.333	203.333	231.667	273.333	340.000	245.000	244.741

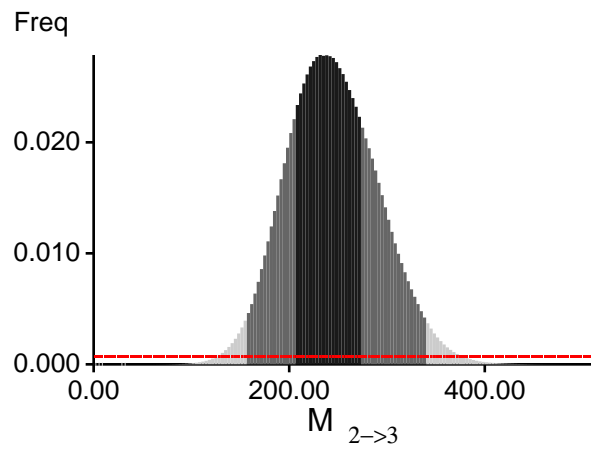
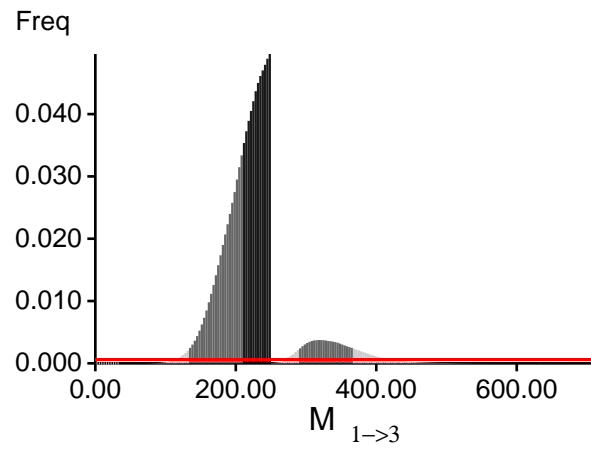
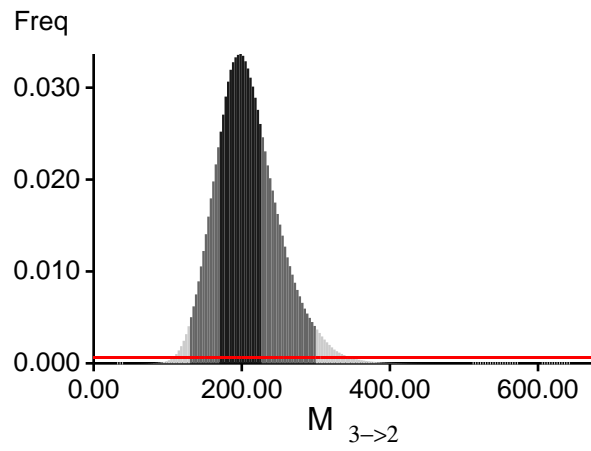
Citation suggestions:

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

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



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]

Locus	Raw thermodynamic score(1a)	Bezier approximation score(1b)	Harmonic mean(2)
1	-3442.08	-2489.15	-2303.90
2	-2885.06	-2129.38	-1982.45
3	-3107.43	-2277.52	-2122.70
4	-3667.13	-2474.43	-2240.63
5	-3195.72	-2356.72	-2194.89
6	-3141.65	-2459.70	-2316.30
7	-3064.77	-2346.53	-2209.78
8	-3640.68	-2508.14	-2268.01
9	-3630.36	-2516.64	-2284.46
10	-3146.89	-2314.83	-2164.65
All	-32800.40	-23751.68	-21966.41

(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

[Scaling factor = 121.358912]

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	13210/27808	0.47504
Θ_2	12202/28057	0.43490
Θ_3	13398/27627	0.48496
$M_{2 \rightarrow 1}$	10999/27912	0.39406
$M_{3 \rightarrow 1}$	11476/27837	0.41226
$M_{1 \rightarrow 2}$	12952/27749	0.46676
$M_{3 \rightarrow 2}$	11849/27953	0.42389
$M_{1 \rightarrow 3}$	13343/27691	0.48185
$M_{2 \rightarrow 3}$	12621/27526	0.45851
Genealogies	31802/249840	0.12729

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.92842	2589.50
Θ_2	0.90583	3485.37
Θ_3	0.90944	3241.36
$M_{2 \rightarrow 1}$	0.90478	3296.95
$M_{3 \rightarrow 1}$	0.91214	3080.56
$M_{1 \rightarrow 2}$	0.90898	3142.62
$M_{3 \rightarrow 2}$	0.89438	3903.26
$M_{1 \rightarrow 3}$	0.90901	3441.72
$M_{2 \rightarrow 3}$	0.92372	2885.91
Genealogies	0.89911	3831.78

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