

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)

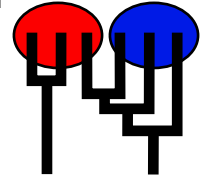
Compiled for PARALLEL computer architectures

One master and 8 compute nodes are available.

Compiled for a SYMMETRIC multiprocessors (Grandcentral)

Program started at Mon Jun 3 09:04:09 2019

Program finished at Mon Jun 3 09:32:25 2019 [Runtime:0000:00:28:16]



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) 3241586714

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→1}	<displayed>
5	M _{3→1}	<displayed>
6	M _{1→2}	<displayed>
7	M _{3→2}	<displayed>
8	M _{1→3}	<displayed>
9	M _{2→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])	100
Number of concurrent chains (replicates) [c]	1
Visited (sampled) parameter values [a*b*c]	500000
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	0.00193	0.00260	0.00470	0.00480	0.00647	0.03723	0.04150
1	Θ_2	0.00107	0.00220	0.00270	0.00333	0.00580	0.00330	0.00344
1	Θ_3	0.00287	0.00493	0.00503	0.00513	0.01953	0.01023	0.01230
1	$M_{2 \rightarrow 1}$	1863.333	1906.667	1918.333	1923.333	2216.667	1798.333	1831.828
1	$M_{3 \rightarrow 1}$	3.333	120.000	138.333	150.000	533.333	201.667	244.660
1	$M_{1 \rightarrow 2}$	156.667	440.000	451.667	453.333	853.333	498.333	520.832
1	$M_{3 \rightarrow 2}$	63.333	156.667	225.000	240.000	496.667	251.667	269.458
1	$M_{1 \rightarrow 3}$	0.000	70.000	98.333	140.000	333.333	138.333	156.974
1	$M_{2 \rightarrow 3}$	60.000	153.333	248.333	273.333	540.000	271.667	301.257
2	Θ_1	0.00127	0.00287	0.00330	0.00407	0.00947	0.00470	0.00507
2	Θ_2	0.00060	0.00147	0.00243	0.00287	0.00687	0.00303	0.00346
2	Θ_3	0.00100	0.00180	0.00270	0.00373	0.00767	0.00357	0.00410
2	$M_{2 \rightarrow 1}$	56.667	153.333	185.000	270.000	536.667	278.333	311.218
2	$M_{3 \rightarrow 1}$	0.000	0.000	5.000	30.000	136.667	38.333	50.993
2	$M_{1 \rightarrow 2}$	56.667	343.333	351.667	370.000	786.667	401.667	446.114
2	$M_{3 \rightarrow 2}$	80.000	260.000	368.333	386.667	856.667	471.667	519.954
2	$M_{1 \rightarrow 3}$	73.333	246.667	311.667	336.667	746.667	405.000	445.063
2	$M_{2 \rightarrow 3}$	30.000	173.333	205.000	296.667	686.667	335.000	382.877
3	Θ_1	0.00067	0.00133	0.00143	0.00187	0.00787	0.00283	0.00396
3	Θ_2	0.00213	0.00467	0.00623	0.00647	0.01127	0.00637	0.00682
3	Θ_3	0.00267	0.00793	0.00823	0.00913	0.02380	0.01303	0.01625
3	$M_{2 \rightarrow 1}$	256.667	510.000	541.667	576.667	1020.000	615.000	653.571
3	$M_{3 \rightarrow 1}$	266.667	576.667	611.667	623.333	1053.333	675.000	717.527
3	$M_{1 \rightarrow 2}$	10.000	50.000	75.000	113.333	233.333	108.333	117.352
3	$M_{3 \rightarrow 2}$	33.333	86.667	128.333	193.333	323.333	168.333	178.301
3	$M_{1 \rightarrow 3}$	216.667	380.000	385.000	400.000	620.000	415.000	422.327
3	$M_{2 \rightarrow 3}$	56.667	180.000	225.000	230.000	433.333	238.333	249.512
4	Θ_1	0.00173	0.00400	0.00423	0.00433	0.01207	0.00610	0.00687
4	Θ_2	0.00160	0.00280	0.00343	0.00353	0.00780	0.00430	0.00467
4	Θ_3	0.00233	0.00493	0.00537	0.00547	0.01533	0.00817	0.00920
4	$M_{2 \rightarrow 1}$	90.000	276.667	308.333	403.333	653.333	358.333	377.965
4	$M_{3 \rightarrow 1}$	0.000	0.000	5.000	36.667	140.000	41.667	50.567
4	$M_{1 \rightarrow 2}$	0.000	20.000	55.000	133.333	370.000	128.333	157.168

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
4	$M_{3 \rightarrow 2}$	30.000	203.333	261.667	263.333	620.000	335.000	362.789
4	$M_{1 \rightarrow 3}$	50.000	160.000	218.333	266.667	566.667	275.000	327.104
4	$M_{2 \rightarrow 3}$	0.000	6.667	18.333	46.667	156.667	58.333	72.771
5	Θ_1	0.00173	0.00467	0.00550	0.00687	0.02280	0.01957	0.03133
5	Θ_2	0.00113	0.00213	0.00277	0.00360	0.00640	0.00337	0.00360
5	Θ_3	0.00240	0.00660	0.00670	0.00713	0.02113	0.01083	0.01633
5	$M_{2 \rightarrow 1}$	503.333	740.000	745.000	750.000	1456.667	988.333	1023.866
5	$M_{3 \rightarrow 1}$	46.667	113.333	145.000	153.333	436.667	255.000	276.532
5	$M_{1 \rightarrow 2}$	166.667	426.667	468.333	483.333	773.333	471.667	491.998
5	$M_{3 \rightarrow 2}$	0.000	20.000	48.333	80.000	246.667	85.000	104.149
5	$M_{1 \rightarrow 3}$	93.333	270.000	335.000	460.000	896.667	538.333	754.445
5	$M_{2 \rightarrow 3}$	0.000	0.000	1.667	70.000	253.333	71.667	99.711
6	Θ_1	0.00087	0.00193	0.00390	0.00527	0.01013	0.00510	0.00783
6	Θ_2	0.00067	0.00233	0.00310	0.00360	0.00880	0.00397	0.00962
6	Θ_3	0.00587	0.01220	0.01230	0.01273	0.02347	0.01443	0.01585
6	$M_{2 \rightarrow 1}$	0.000	0.000	18.333	113.333	576.667	1045.000	1225.586
6	$M_{3 \rightarrow 1}$	0.000	0.000	28.333	143.333	626.667	201.667	308.545
6	$M_{1 \rightarrow 2}$	0.000	56.667	98.333	213.333	790.000	515.000	1075.760
6	$M_{3 \rightarrow 2}$	0.000	0.000	5.000	90.000	523.333	151.667	240.317
6	$M_{1 \rightarrow 3}$	0.000	0.000	1.667	46.667	320.000	55.000	115.403
6	$M_{2 \rightarrow 3}$	0.000	0.000	1.667	33.333	203.333	35.000	76.770
7	Θ_1	0.00087	0.00253	0.00290	0.00407	0.00853	0.00397	0.00518
7	Θ_2	0.00187	0.00720	0.00757	0.00820	0.01500	0.01377	0.02282
7	Θ_3	0.00460	0.00700	0.00810	0.00840	0.02253	0.01243	0.01533
7	$M_{2 \rightarrow 1}$	346.667	446.667	451.667	460.000	1216.667	825.000	887.735
7	$M_{3 \rightarrow 1}$	0.000	0.000	38.333	66.667	336.667	121.667	153.827
7	$M_{1 \rightarrow 2}$	406.667	606.667	615.000	616.667	1666.667	1065.000	1091.213
7	$M_{3 \rightarrow 2}$	13.333	150.000	185.000	193.333	510.000	215.000	247.285
7	$M_{1 \rightarrow 3}$	0.000	0.000	5.000	76.667	326.667	108.333	150.780
7	$M_{2 \rightarrow 3}$	16.667	270.000	291.667	296.667	663.333	305.000	341.476
8	Θ_1	0.00100	0.00233	0.00263	0.00387	0.00760	0.00383	0.00433
8	Θ_2	0.00133	0.00360	0.00390	0.00433	0.01060	0.00523	0.00651
8	Θ_3	0.00353	0.00633	0.00657	0.00680	0.01807	0.00983	0.01094
8	$M_{2 \rightarrow 1}$	16.667	143.333	155.000	180.000	586.667	271.667	312.811
8	$M_{3 \rightarrow 1}$	6.667	160.000	168.333	170.000	456.667	208.333	250.053
8	$M_{1 \rightarrow 2}$	0.000	63.333	141.667	183.333	523.333	205.000	241.402
8	$M_{3 \rightarrow 2}$	186.667	386.667	415.000	493.333	1010.000	581.667	651.338

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$M_{1 \rightarrow 3}$	6.667	60.000	78.333	100.000	243.333	98.333	113.197
8	$M_{2 \rightarrow 3}$	53.333	123.333	155.000	200.000	353.333	191.667	204.437
9	Θ_1	0.00167	0.00280	0.00430	0.00753	0.01960	0.00843	0.01703
9	Θ_2	0.00087	0.00233	0.00243	0.00247	0.00607	0.00290	0.00313
9	Θ_3	0.00493	0.00913	0.01010	0.01013	0.02380	0.01303	0.01548
9	$M_{2 \rightarrow 1}$	560.000	850.000	865.000	883.333	1650.000	1118.333	1165.575
9	$M_{3 \rightarrow 1}$	0.000	20.000	88.333	133.333	403.333	151.667	187.728
9	$M_{1 \rightarrow 2}$	243.333	666.667	695.000	726.667	1206.667	728.333	789.970
9	$M_{3 \rightarrow 2}$	0.000	43.333	71.667	113.333	313.333	108.333	136.848
9	$M_{1 \rightarrow 3}$	0.000	40.000	111.667	146.667	270.000	145.000	190.081
9	$M_{2 \rightarrow 3}$	0.000	3.333	11.667	46.667	130.000	45.000	59.659
10	Θ_1	0.00133	0.00260	0.00303	0.00427	0.00887	0.00457	0.00500
10	Θ_2	0.00220	0.00407	0.00417	0.00420	0.01313	0.00690	0.00847
10	Θ_3	0.00467	0.00813	0.00857	0.00967	0.01927	0.01183	0.01518
10	$M_{2 \rightarrow 1}$	10.000	100.000	151.667	170.000	430.000	218.333	249.095
10	$M_{3 \rightarrow 1}$	60.000	210.000	231.667	286.667	506.667	278.333	304.934
10	$M_{1 \rightarrow 2}$	156.667	356.667	431.667	450.000	810.000	468.333	497.325
10	$M_{3 \rightarrow 2}$	23.333	183.333	188.333	190.000	456.667	198.333	220.416
10	$M_{1 \rightarrow 3}$	0.000	16.667	41.667	96.667	376.667	125.000	157.937
10	$M_{2 \rightarrow 3}$	76.667	296.667	305.000	310.000	646.667	321.667	377.543
All	Θ_1	0.00273	0.00353	0.00410	0.00460	0.00560	0.00423	0.00420
All	Θ_2	0.00253	0.00327	0.00377	0.00420	0.00507	0.00390	0.00383
All	Θ_3	0.00587	0.00720	0.00817	0.00900	0.01093	0.00837	0.00836
All	$M_{2 \rightarrow 1}$	340.000	433.333	495.000	546.667	656.667	501.667	500.249
All	$M_{3 \rightarrow 1}$	40.000	80.000	108.333	136.667	193.333	118.333	117.147
All	$M_{1 \rightarrow 2}$	183.333	243.333	285.000	320.000	396.667	291.667	291.489
All	$M_{3 \rightarrow 2}$	106.667	153.333	181.667	213.333	273.333	191.667	189.358
All	$M_{1 \rightarrow 3}$	73.333	116.667	148.333	176.667	240.000	155.000	155.327
All	$M_{2 \rightarrow 3}$	43.333	80.000	105.000	130.000	183.333	111.667	111.009

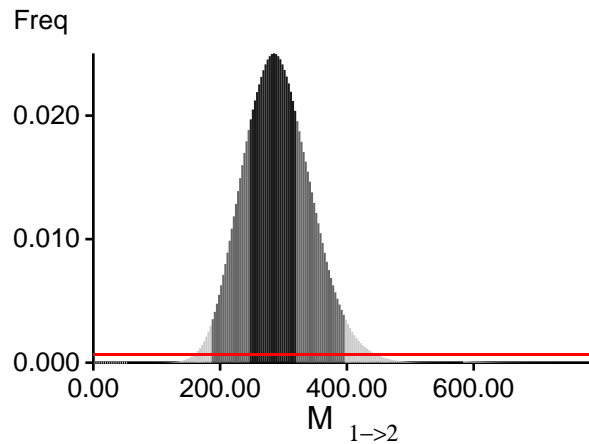
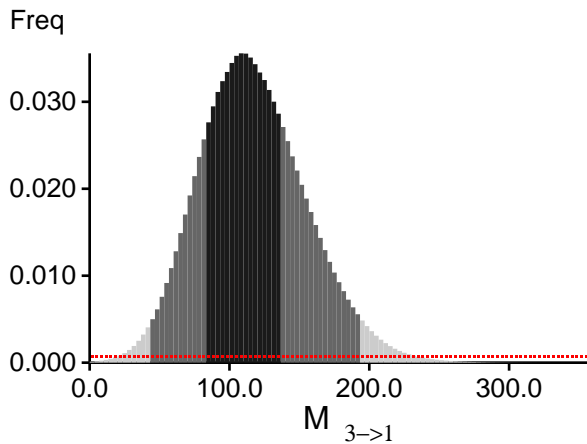
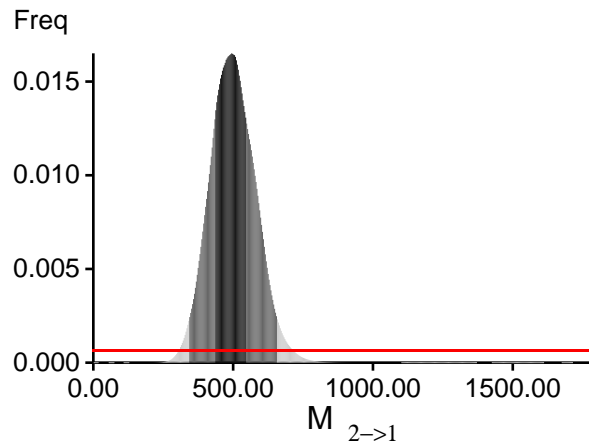
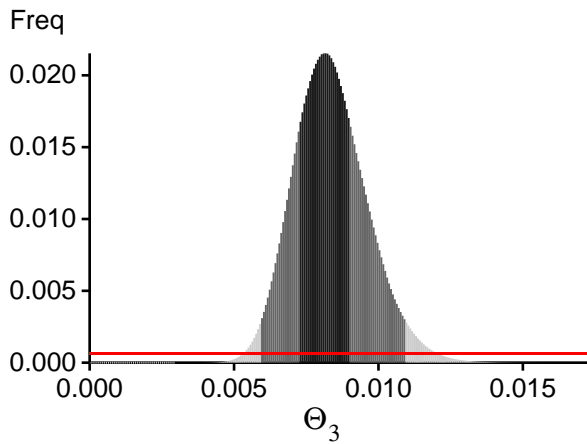
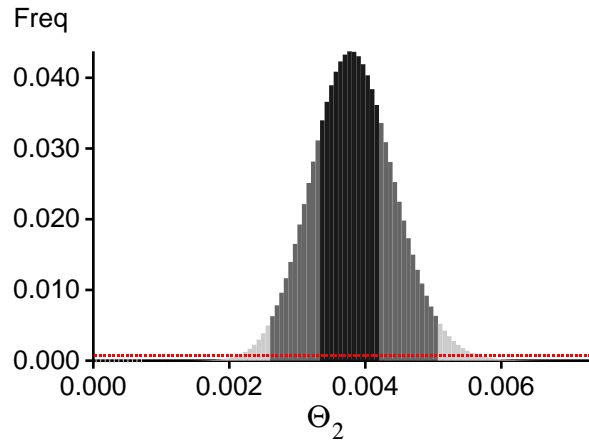
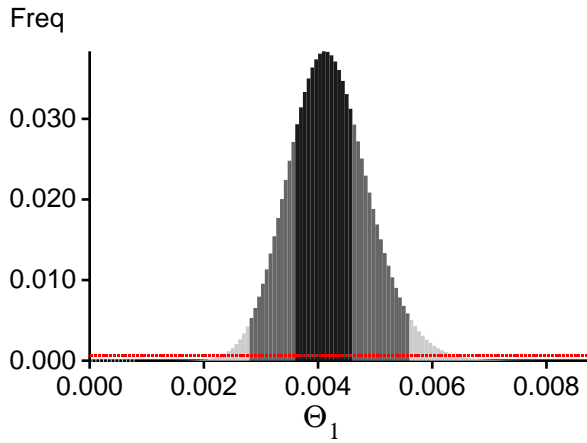
Citation suggestions:

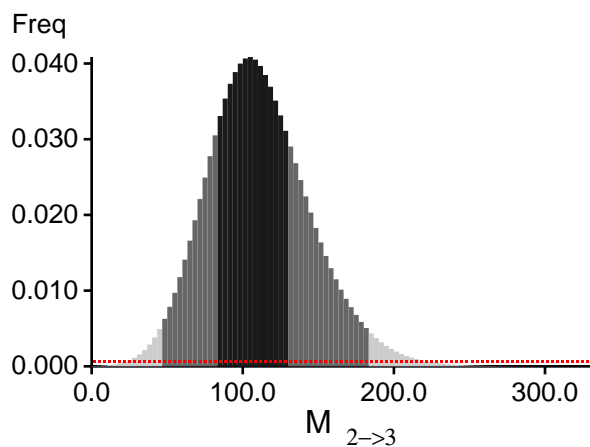
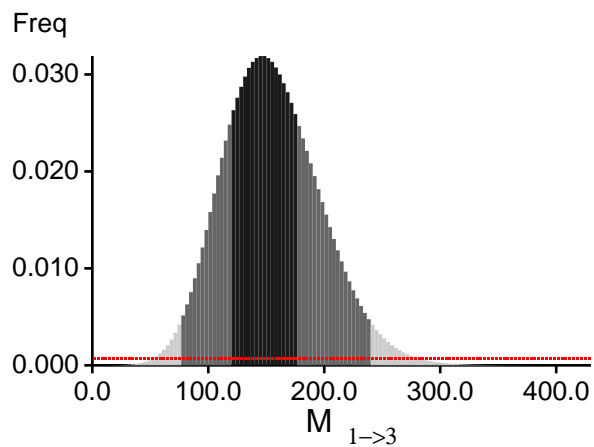
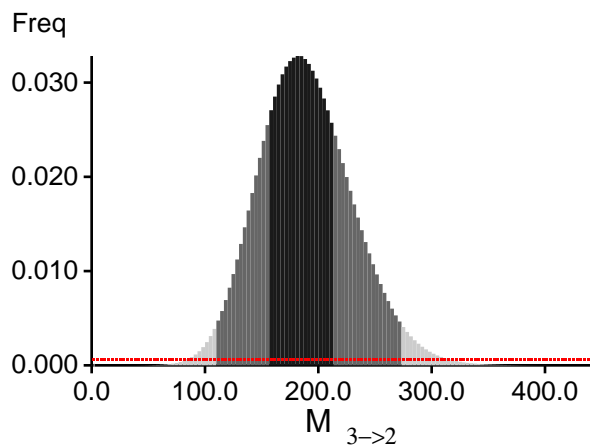
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345

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 = 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]

Locus	Raw thermodynamic score(1a)	Bezier approximation score(1b)	Harmonic mean(2)
1	-3297.22	-2536.18	-2328.09
2	-2737.51	-2123.98	-1985.22
3	-2880.00	-2291.96	-2111.46
4	-3332.85	-2480.85	-2247.19
5	-3010.33	-2372.53	-2191.43
6	-3220.12	-2607.88	-2345.55
7	-3016.40	-2366.59	-2227.83
8	-3401.84	-2506.22	-2246.40
9	-3451.11	-2533.27	-2328.66
10	-2990.85	-2329.84	-2172.53
All	-31205.30	-24016.38	-22051.44

(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 = 132.934704]

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	144547/277917	0.52011
Θ_2	104849/278071	0.37706
Θ_3	115318/277826	0.41507
M _{2→1}	110301/277878	0.39694
M _{3→1}	102447/278804	0.36745
M _{1→2}	114087/277231	0.41152
M _{3→2}	141436/278552	0.50775
M _{1→3}	111944/277719	0.40308
M _{2→3}	102223/277901	0.36784
Genealogies	338925/2498101	0.13567

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.73779	11410.03
Θ_2	0.67651	14435.59
Θ_3	0.59141	18746.24
$M_{2 \rightarrow 1}$	0.67070	14382.03
$M_{3 \rightarrow 1}$	0.54913	21533.51
$M_{1 \rightarrow 2}$	0.64699	16243.54
$M_{3 \rightarrow 2}$	0.59130	18090.58
$M_{1 \rightarrow 3}$	0.73298	11234.67
$M_{2 \rightarrow 3}$	0.66842	13750.45
Genealogies	0.83134	6574.27

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