

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)

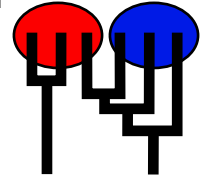
Compiled for PARALLEL computer architectures

One master and 8 compute nodes are available.

Compiled for a SYMMETRIC multiprocessors (Grandcentral)

Program started at Sat Jun 1 22:17:15 2019

Program finished at Sat Jun 1 22:17:59 2019 [Runtime:0000:00:00:44]



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) 2281898755

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])	1
Number of concurrent chains (replicates) [c]	1
Visited (sampled) parameter values [a*b*c]	5000
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_tooshort
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)

10

30

(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.01033	0.01040	0.01050	0.01053	0.01060	0.01123	0.01400
1	Θ_2	0.00560	0.00560	0.00570	0.00573	0.00573	0.00630	0.00737
1	Θ_3	0.01673	0.01673	0.01683	0.01687	0.01687	0.01690	0.02755
1	$M_{2 \rightarrow 1}$	260.000	263.333	271.667	276.667	276.667	348.333	550.870
1	$M_{3 \rightarrow 1}$	136.667	143.333	148.333	153.333	153.333	228.333	281.306
1	$M_{1 \rightarrow 2}$	970.000	970.000	975.000	976.667	976.667	801.667	829.079
1	$M_{3 \rightarrow 2}$	390.000	400.000	405.000	406.667	406.667	775.000	747.822
1	$M_{1 \rightarrow 3}$	903.333	903.333	908.333	910.000	910.000	785.000	785.286
1	$M_{2 \rightarrow 3}$	1130.000	1136.667	1141.667	1150.000	1156.667	1021.667	909.869
2	Θ_1	0.00147	0.00147	0.00157	0.00160	0.00180	0.00383	0.00455
2	Θ_2	0.01040	0.01060	0.01070	0.01073	0.01073	0.00863	0.00975
2	Θ_3	0.00227	0.00260	0.00270	0.00273	0.00280	0.00350	0.00458
2	$M_{2 \rightarrow 1}$	956.667	956.667	961.667	963.333	963.333	841.667	861.592
2	$M_{3 \rightarrow 1}$	150.000	160.000	165.000	166.667	170.000	221.667	242.710
2	$M_{1 \rightarrow 2}$	303.333	310.000	318.333	320.000	320.000	345.000	351.972
2	$M_{3 \rightarrow 2}$	373.333	373.333	378.333	380.000	380.000	448.333	479.579
2	$M_{1 \rightarrow 3}$	580.000	580.000	585.000	586.667	590.000	575.000	547.421
2	$M_{2 \rightarrow 3}$	436.667	440.000	448.333	453.333	466.667	448.333	457.439
3	Θ_1	0.00580	0.00600	0.00610	0.00613	0.00620	0.00617	0.00663
3	Θ_2	0.00947	0.00960	0.00970	0.00973	0.00973	0.01057	0.01203
3	Θ_3	0.00667	0.00667	0.00683	0.00687	0.00687	0.00890	0.01207
3	$M_{2 \rightarrow 1}$	653.333	653.333	658.333	660.000	660.000	661.667	665.026
3	$M_{3 \rightarrow 1}$	193.333	193.333	198.333	200.000	200.000	281.667	270.723
3	$M_{1 \rightarrow 2}$	353.333	353.333	358.333	363.333	366.667	498.333	516.508
3	$M_{3 \rightarrow 2}$	306.667	306.667	315.000	316.667	316.667	285.000	299.549
3	$M_{1 \rightarrow 3}$	1453.333	1453.333	1458.333	1460.000	1460.000	1461.667	1431.583
3	$M_{2 \rightarrow 3}$	866.667	866.667	875.000	876.667	886.667	821.667	838.429
4	Θ_1	0.00627	0.00640	0.00650	0.00653	0.00673	0.00730	0.00823
4	Θ_2	0.00660	0.00660	0.00670	0.00673	0.00673	0.00350	0.00438
4	Θ_3	0.00427	0.00427	0.00437	0.00440	0.00440	0.01157	0.01341
4	$M_{2 \rightarrow 1}$	220.000	223.333	231.667	233.333	233.333	275.000	292.167
4	$M_{3 \rightarrow 1}$	416.667	426.667	431.667	433.333	433.333	465.000	508.004
4	$M_{1 \rightarrow 2}$	560.000	576.667	588.333	590.000	606.667	688.333	764.990

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
4	$M_{3 \rightarrow 2}$	443.333	446.667	455.000	456.667	470.000	455.000	457.547
4	$M_{1 \rightarrow 3}$	466.667	466.667	471.667	473.333	486.667	435.000	448.424
4	$M_{2 \rightarrow 3}$	246.667	246.667	251.667	253.333	253.333	338.333	369.072
5	Θ_1	0.00567	0.00573	0.00583	0.00587	0.00587	0.01097	0.01380
5	Θ_2	0.00213	0.00213	0.00223	0.00227	0.00227	0.00850	0.00925
5	Θ_3	0.00580	0.00580	0.00590	0.00593	0.00600	0.00597	0.00664
5	$M_{2 \rightarrow 1}$	456.667	460.000	465.000	466.667	466.667	468.333	481.635
5	$M_{3 \rightarrow 1}$	1036.667	1036.667	1045.000	1046.667	1046.667	1015.000	986.194
5	$M_{1 \rightarrow 2}$	623.333	623.333	628.333	630.000	630.000	821.667	854.719
5	$M_{3 \rightarrow 2}$	583.333	586.667	591.667	593.333	596.667	838.333	861.368
5	$M_{1 \rightarrow 3}$	590.000	603.333	608.333	613.333	620.000	665.000	671.668
5	$M_{2 \rightarrow 3}$	486.667	490.000	495.000	496.667	496.667	645.000	676.558
6	Θ_1	0.00520	0.00520	0.00530	0.00540	0.00540	0.00637	0.00694
6	Θ_2	0.00420	0.00420	0.00430	0.00433	0.00433	0.00630	0.00712
6	Θ_3	0.00913	0.00920	0.00930	0.00933	0.00933	0.01517	0.01634
6	$M_{2 \rightarrow 1}$	156.667	173.333	181.667	186.667	186.667	205.000	223.540
6	$M_{3 \rightarrow 1}$	713.333	713.333	718.333	720.000	720.000	741.667	745.610
6	$M_{1 \rightarrow 2}$	1883.333	1883.333	1888.333	1890.000	1890.000	1218.333	1319.793
6	$M_{3 \rightarrow 2}$	763.333	763.333	768.333	770.000	770.000	745.000	764.236
6	$M_{1 \rightarrow 3}$	600.000	600.000	605.000	606.667	606.667	788.333	778.676
6	$M_{2 \rightarrow 3}$	436.667	436.667	445.000	446.667	446.667	455.000	459.238
7	Θ_1	0.00153	0.00153	0.00163	0.00173	0.00200	0.00203	0.00215
7	Θ_2	0.00627	0.00627	0.00637	0.00640	0.00640	0.01150	0.01395
7	Θ_3	0.02080	0.02080	0.02090	0.02093	0.02100	0.02297	0.02604
7	$M_{2 \rightarrow 1}$	350.000	366.667	371.667	373.333	383.333	365.000	369.136
7	$M_{3 \rightarrow 1}$	1083.333	1103.333	1108.333	1110.000	1110.000	931.667	958.047
7	$M_{1 \rightarrow 2}$	700.000	700.000	708.333	720.000	720.000	681.667	677.401
7	$M_{3 \rightarrow 2}$	236.667	236.667	241.667	243.333	243.333	638.333	646.188
7	$M_{1 \rightarrow 3}$	333.333	333.333	338.333	340.000	340.000	341.667	323.787
7	$M_{2 \rightarrow 3}$	376.667	386.667	395.000	396.667	396.667	455.000	460.072
8	Θ_1	0.00627	0.00627	0.00637	0.00640	0.00640	0.01963	0.03665
8	Θ_2	0.00427	0.00440	0.00450	0.00453	0.00467	0.00697	0.00745
8	Θ_3	0.00813	0.00813	0.00823	0.00827	0.00827	0.01363	0.01670
8	$M_{2 \rightarrow 1}$	316.667	316.667	321.667	323.333	333.333	368.333	398.255
8	$M_{3 \rightarrow 1}$	176.667	176.667	181.667	183.333	186.667	481.667	475.184
8	$M_{1 \rightarrow 2}$	606.667	606.667	611.667	613.333	613.333	781.667	774.378
8	$M_{3 \rightarrow 2}$	140.000	140.000	145.000	146.667	146.667	418.333	416.197

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$M_{1 \rightarrow 3}$	253.333	253.333	265.000	266.667	266.667	481.667	566.640
8	$M_{2 \rightarrow 3}$	166.667	166.667	185.000	186.667	186.667	255.000	274.219
9	Θ_1	0.00727	0.00740	0.00750	0.00753	0.00767	0.01383	0.01566
9	Θ_2	0.00420	0.00440	0.00450	0.00453	0.00467	0.00457	0.00478
9	Θ_3	0.00760	0.00760	0.00783	0.00800	0.00800	0.00997	0.01033
9	$M_{2 \rightarrow 1}$	996.667	1006.667	1011.667	1016.667	1016.667	978.333	975.158
9	$M_{3 \rightarrow 1}$	340.000	340.000	345.000	346.667	346.667	365.000	389.054
9	$M_{1 \rightarrow 2}$	466.667	480.000	485.000	486.667	493.333	565.000	584.715
9	$M_{3 \rightarrow 2}$	463.333	473.333	478.333	480.000	480.000	488.333	511.617
9	$M_{1 \rightarrow 3}$	476.667	476.667	481.667	483.333	483.333	551.667	574.447
9	$M_{2 \rightarrow 3}$	316.667	320.000	325.000	326.667	326.667	341.667	354.239
10	Θ_1	0.00167	0.00173	0.00183	0.00187	0.00253	0.00303	0.00343
10	Θ_2	0.00547	0.00547	0.00557	0.00560	0.00560	0.00730	0.00770
10	Θ_3	0.00967	0.00967	0.00977	0.00980	0.01000	0.01310	0.01364
10	$M_{2 \rightarrow 1}$	550.000	556.667	561.667	566.667	573.333	641.667	640.547
10	$M_{3 \rightarrow 1}$	596.667	596.667	601.667	606.667	606.667	635.000	650.198
10	$M_{1 \rightarrow 2}$	640.000	653.333	658.333	663.333	663.333	661.667	634.353
10	$M_{3 \rightarrow 2}$	660.000	660.000	665.000	666.667	666.667	738.333	736.578
10	$M_{1 \rightarrow 3}$	290.000	303.333	311.667	313.333	313.333	331.667	346.615
10	$M_{2 \rightarrow 3}$	440.000	440.000	445.000	446.667	446.667	648.333	647.149
All	Θ_1	0.00427	0.00493	0.00563	0.00613	0.00680	0.00583	0.00595
All	Θ_2	0.00407	0.00527	0.00603	0.00653	0.00773	0.00597	0.00594
All	Θ_3	0.00587	0.00667	0.00723	0.00727	0.00727	0.00837	0.00823
All	$M_{2 \rightarrow 1}$	330.000	373.333	405.000	426.667	483.333	405.000	402.263
All	$M_{3 \rightarrow 1}$	196.667	223.333	261.667	263.333	263.333	391.667	408.152
All	$M_{1 \rightarrow 2}$	510.000	563.333	598.333	636.667	710.000	615.000	628.570
All	$M_{3 \rightarrow 2}$	393.333	436.667	491.667	543.333	686.667	525.000	534.004
All	$M_{1 \rightarrow 3}$	390.000	496.667	548.333	556.667	556.667	508.333	500.834
All	$M_{2 \rightarrow 3}$	386.667	430.000	458.333	490.000	550.000	468.333	467.986

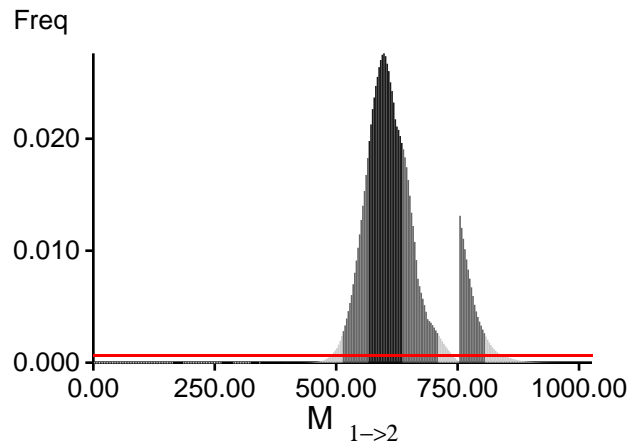
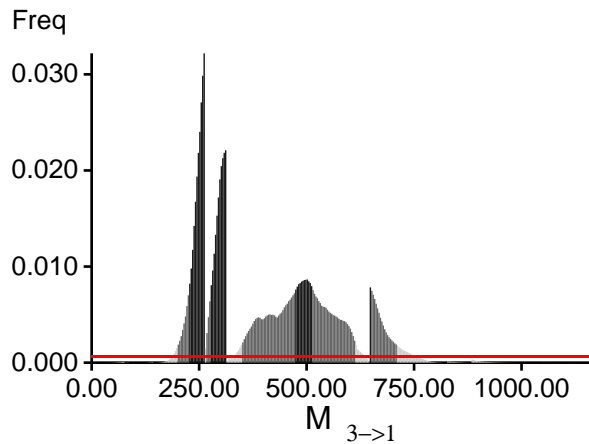
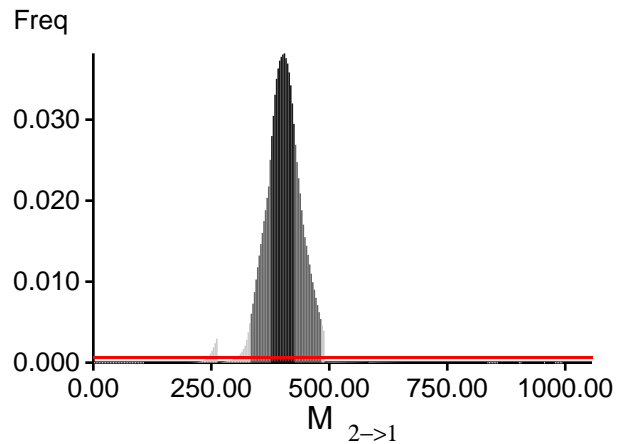
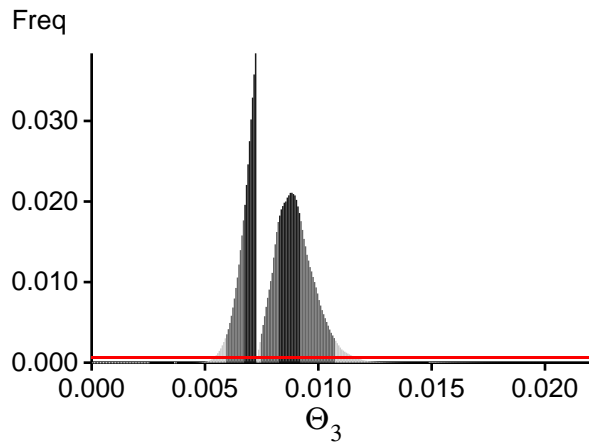
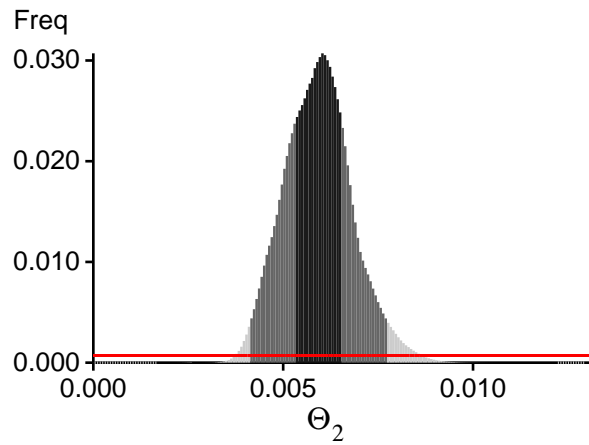
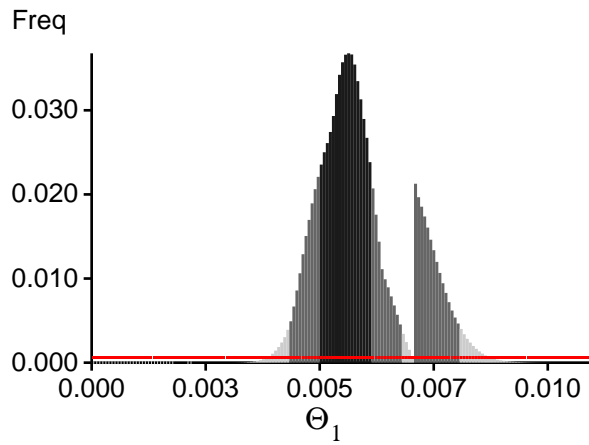
Citation suggestions:

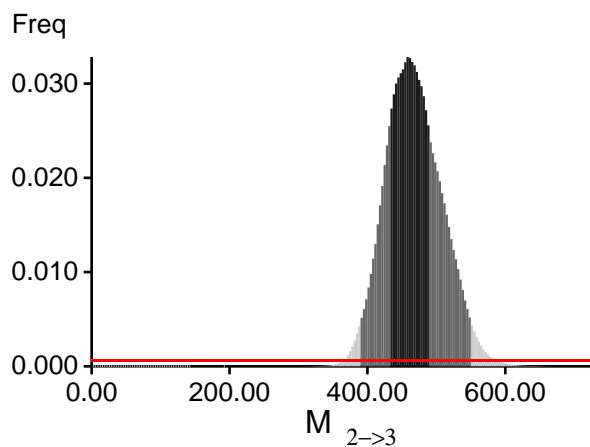
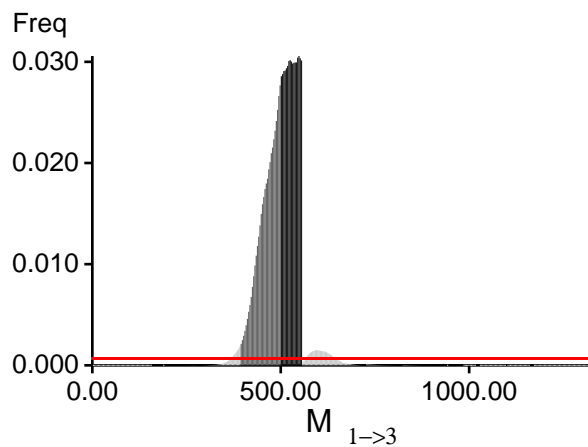
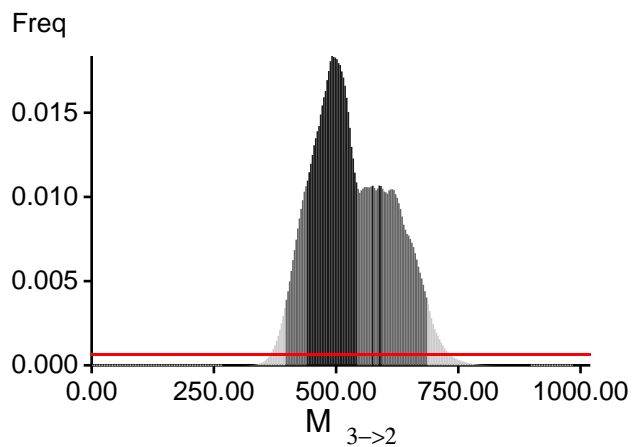
Berli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345

Berli 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	-3429.29	-2463.62	-2295.97
2	-2753.44	-2090.37	-1960.61
3	-2989.48	-2247.13	-2117.13
4	-3581.44	-2459.57	-2247.16
5	-3141.57	-2347.36	-2216.19
6	-3116.08	-2422.71	-2301.01
7	-2991.44	-2325.34	-2202.43
8	-3454.49	-2416.63	-2232.81
9	-3486.49	-2463.50	-2276.54
10	-2971.21	-2278.34	-2157.30
All	-31809.66	-23409.30	-21901.87

(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 = 105.268124]

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	1379/2801	0.49232
Θ_2	1512/2818	0.53655
Θ_3	1823/2857	0.63808
M _{2→1}	1432/2727	0.52512
M _{3→1}	1363/2700	0.50481
M _{1→2}	1764/2827	0.62398
M _{3→2}	1591/2717	0.58557
M _{1→3}	1433/2732	0.52452
M _{2→3}	1530/2791	0.54819
Genealogies	2606/25030	0.10412

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.99193	261.30
Θ_2	0.99016	362.12
Θ_3	0.99319	291.04
M _{2→1}	0.98354	623.20
M _{3→1}	0.98738	459.91
M _{1→2}	0.98744	458.52
M _{3→2}	0.98800	446.11
M _{1→3}	0.98560	558.16
M _{2→3}	0.98610	518.42
Genealogies	0.97066	1027.84

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