

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 23:06:15 2019

Program finished at Sat Jun 1 23:10:15 2019 [Runtime:0000:00:04:00]



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) 575434568

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)	

10

30

(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00393	0.00393	0.00403	0.00407	0.00407	0.01083	0.01309
1	Θ_2	0.00100	0.00207	0.00250	0.00333	0.00480	0.00283	0.00314
1	Θ_3	0.00480	0.00500	0.00517	0.00520	0.00520	0.00863	0.01432
1	$M_{2 \rightarrow 1}$	363.333	466.667	471.667	473.333	580.000	491.667	509.455
1	$M_{3 \rightarrow 1}$	450.000	500.000	505.000	506.667	530.000	508.333	526.212
1	$M_{1 \rightarrow 2}$	636.667	700.000	705.000	723.333	820.000	798.333	809.626
1	$M_{3 \rightarrow 2}$	73.333	206.667	211.667	220.000	343.333	245.000	277.491
1	$M_{1 \rightarrow 3}$	713.333	896.667	905.000	906.667	980.000	915.000	945.626
1	$M_{2 \rightarrow 3}$	256.667	300.000	305.000	306.667	346.667	398.333	413.172
2	Θ_1	0.00080	0.00207	0.00257	0.00287	0.00567	0.00317	0.00352
2	Θ_2	0.00280	0.00420	0.00430	0.00453	0.00847	0.00550	0.00577
2	Θ_3	0.00160	0.00213	0.00223	0.00227	0.00360	0.00297	0.00372
2	$M_{2 \rightarrow 1}$	1013.333	1023.333	1028.333	1030.000	1036.667	881.667	901.936
2	$M_{3 \rightarrow 1}$	340.000	343.333	348.333	350.000	453.333	515.000	533.579
2	$M_{1 \rightarrow 2}$	210.000	353.333	361.667	363.333	446.667	331.667	330.744
2	$M_{3 \rightarrow 2}$	123.333	183.333	188.333	196.667	323.333	241.667	248.022
2	$M_{1 \rightarrow 3}$	176.667	286.667	301.667	303.333	410.000	365.000	398.995
2	$M_{2 \rightarrow 3}$	1160.000	1170.000	1175.000	1176.667	1183.333	1148.333	1174.980
3	Θ_1	0.00207	0.00327	0.00343	0.00387	0.00760	0.00457	0.00469
3	Θ_2	0.00220	0.00227	0.00237	0.00240	0.00240	0.01457	0.02551
3	Θ_3	0.01387	0.01387	0.01410	0.01413	0.01473	0.02370	0.03000
3	$M_{2 \rightarrow 1}$	613.333	746.667	751.667	763.333	913.333	858.333	974.823
3	$M_{3 \rightarrow 1}$	226.667	393.333	398.333	413.333	560.000	425.000	432.184
3	$M_{1 \rightarrow 2}$	1810.000	1816.667	1821.667	1823.333	1840.000	2018.333	2025.747
3	$M_{3 \rightarrow 2}$	700.000	720.000	725.000	730.000	733.333	801.667	880.950
3	$M_{1 \rightarrow 3}$	1650.000	1650.000	1658.333	1660.000	1686.667	1231.667	1233.581
3	$M_{2 \rightarrow 3}$	66.667	100.000	108.333	116.667	116.667	435.000	474.896
4	Θ_1	0.00187	0.00380	0.00397	0.00427	0.00920	0.00510	0.00598
4	Θ_2	0.00153	0.00413	0.00423	0.00440	0.00793	0.00483	0.00542
4	Θ_3	0.00640	0.00640	0.00683	0.00693	0.00760	0.00857	0.01003
4	$M_{2 \rightarrow 1}$	613.333	663.333	671.667	676.667	680.000	465.000	496.879
4	$M_{3 \rightarrow 1}$	1260.000	1273.333	1278.333	1283.333	1296.667	928.333	953.415
4	$M_{1 \rightarrow 2}$	760.000	773.333	788.333	790.000	863.333	708.333	696.890

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
4	$M_{3 \rightarrow 2}$	306.667	436.667	441.667	443.333	543.333	541.667	580.793
4	$M_{1 \rightarrow 3}$	303.333	353.333	365.000	376.667	493.333	345.000	351.482
4	$M_{2 \rightarrow 3}$	643.333	663.333	668.333	676.667	686.667	641.667	608.687
5	Θ_1	0.01140	0.01167	0.01177	0.01180	0.01180	0.03383	0.03982
5	Θ_2	0.00073	0.00127	0.00183	0.00200	0.00340	0.00203	0.00207
5	Θ_3	0.00587	0.00587	0.00603	0.00627	0.00907	0.01237	0.01953
5	$M_{2 \rightarrow 1}$	1823.333	1823.333	1828.333	1830.000	1830.000	1365.000	1400.584
5	$M_{3 \rightarrow 1}$	600.000	616.667	621.667	623.333	633.333	501.667	520.627
5	$M_{1 \rightarrow 2}$	1033.333	1043.333	1051.667	1060.000	1090.000	1198.333	1228.057
5	$M_{3 \rightarrow 2}$	500.000	503.333	511.667	513.333	526.667	481.667	511.674
5	$M_{1 \rightarrow 3}$	973.333	1100.000	1108.333	1113.333	1190.000	1101.667	1117.216
5	$M_{2 \rightarrow 3}$	500.000	503.333	508.333	510.000	633.333	615.000	659.295
6	Θ_1	0.00147	0.00287	0.00330	0.00373	0.00660	0.00383	0.00402
6	Θ_2	0.00407	0.00407	0.00437	0.00440	0.00513	0.00550	0.00844
6	Θ_3	0.01033	0.01093	0.01117	0.01120	0.01180	0.01770	0.02052
6	$M_{2 \rightarrow 1}$	260.000	263.333	271.667	273.333	293.333	258.333	274.199
6	$M_{3 \rightarrow 1}$	293.333	293.333	298.333	300.000	390.000	481.667	523.351
6	$M_{1 \rightarrow 2}$	786.667	793.333	798.333	800.000	820.000	1008.333	1037.483
6	$M_{3 \rightarrow 2}$	310.000	396.667	421.667	436.667	673.333	415.000	422.484
6	$M_{1 \rightarrow 3}$	3.333	23.333	35.000	43.333	136.667	91.667	104.811
6	$M_{2 \rightarrow 3}$	243.333	303.333	311.667	313.333	360.000	415.000	469.782
7	Θ_1	0.00253	0.00333	0.00370	0.00380	0.00580	0.00437	0.00472
7	Θ_2	0.01553	0.01553	0.01563	0.01567	0.01567	0.02763	0.03454
7	Θ_3	0.00520	0.00733	0.00763	0.00827	0.01053	0.01223	0.01733
7	$M_{2 \rightarrow 1}$	206.667	296.667	315.000	316.667	423.333	405.000	419.000
7	$M_{3 \rightarrow 1}$	236.667	243.333	248.333	250.000	266.667	501.667	517.626
7	$M_{1 \rightarrow 2}$	1450.000	1470.000	1478.333	1480.000	1546.667	1611.667	1624.353
7	$M_{3 \rightarrow 2}$	350.000	350.000	358.333	360.000	476.667	465.000	496.364
7	$M_{1 \rightarrow 3}$	90.000	270.000	275.000	276.667	283.333	325.000	352.457
7	$M_{2 \rightarrow 3}$	423.333	520.000	525.000	533.333	543.333	831.667	1028.442
8	Θ_1	0.00293	0.00400	0.00423	0.00440	0.00940	0.00637	0.00714
8	Θ_2	0.00440	0.00567	0.00577	0.00580	0.00580	0.00670	0.00853
8	Θ_3	0.00313	0.00667	0.00703	0.00720	0.01107	0.00790	0.00883
8	$M_{2 \rightarrow 1}$	20.000	116.667	121.667	123.333	166.667	101.667	108.209
8	$M_{3 \rightarrow 1}$	133.333	200.000	208.333	210.000	396.667	268.333	273.681
8	$M_{1 \rightarrow 2}$	260.000	310.000	315.000	320.000	323.333	518.333	540.921
8	$M_{3 \rightarrow 2}$	120.000	433.333	438.333	440.000	450.000	328.333	350.321

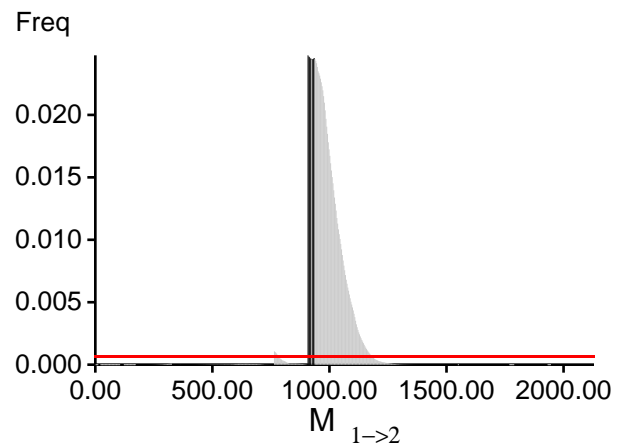
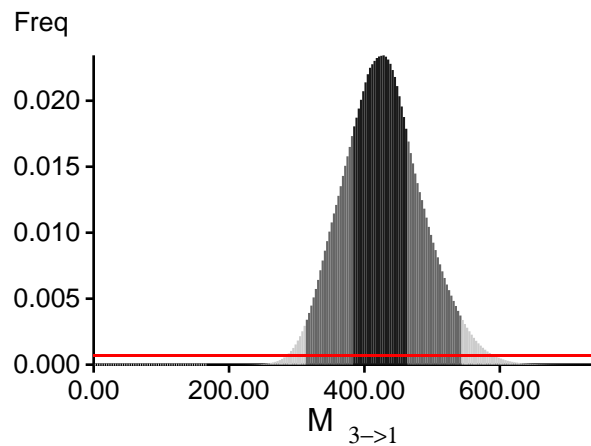
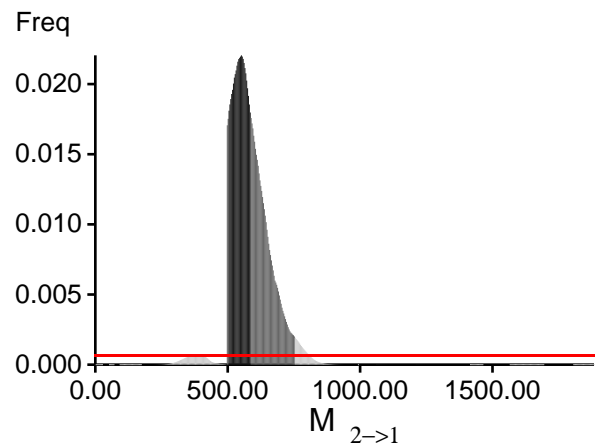
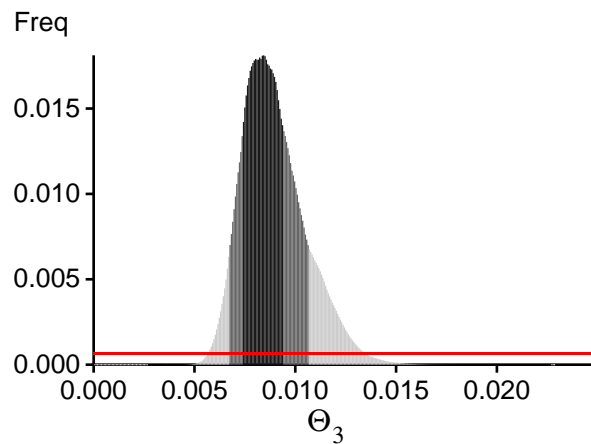
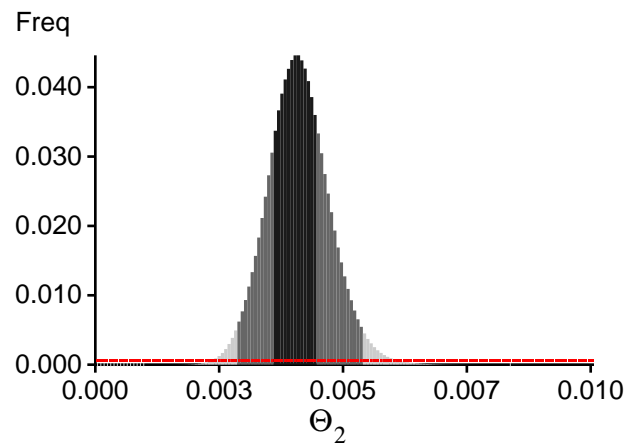
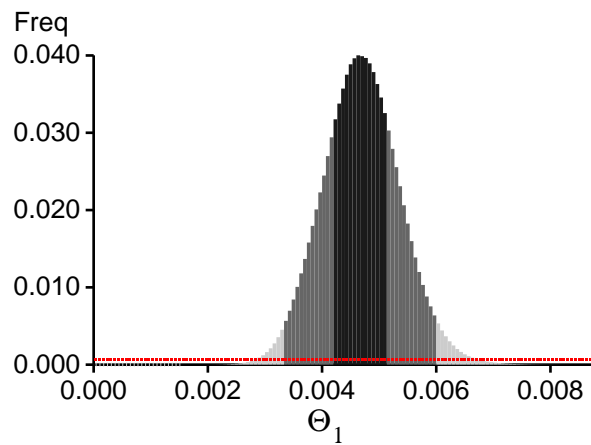
Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$M_{1 \rightarrow 3}$	256.667	373.333	385.000	390.000	486.667	398.333	414.439
8	$M_{2 \rightarrow 3}$	130.000	246.667	251.667	256.667	370.000	301.667	315.363
9	Θ_1	0.00320	0.00360	0.00370	0.00373	0.00660	0.00603	0.00761
9	Θ_2	0.00487	0.00507	0.00517	0.00520	0.00520	0.00737	0.00749
9	Θ_3	0.02287	0.02287	0.02310	0.02313	0.02327	0.05117	0.04929
9	$M_{2 \rightarrow 1}$	370.000	423.333	431.667	446.667	736.667	595.000	626.282
9	$M_{3 \rightarrow 1}$	93.333	250.000	291.667	306.667	663.333	381.667	451.876
9	$M_{1 \rightarrow 2}$	126.667	216.667	231.667	233.333	326.667	235.000	242.343
9	$M_{3 \rightarrow 2}$	16.667	73.333	85.000	86.667	120.000	85.000	86.687
9	$M_{1 \rightarrow 3}$	13.333	83.333	98.333	100.000	310.000	155.000	167.369
9	$M_{2 \rightarrow 3}$	370.000	496.667	508.333	523.333	783.333	575.000	591.058
10	Θ_1	0.00193	0.00347	0.00357	0.00387	0.00687	0.00437	0.00454
10	Θ_2	0.00327	0.00333	0.00343	0.00347	0.00353	0.00617	0.00734
10	Θ_3	0.00633	0.00947	0.00957	0.00960	0.01340	0.01657	0.02109
10	$M_{2 \rightarrow 1}$	776.667	840.000	851.667	853.333	880.000	751.667	775.356
10	$M_{3 \rightarrow 1}$	323.333	603.333	608.333	613.333	790.000	571.667	594.526
10	$M_{1 \rightarrow 2}$	700.000	703.333	718.333	720.000	730.000	875.000	925.641
10	$M_{3 \rightarrow 2}$	660.000	660.000	665.000	666.667	673.333	811.667	806.720
10	$M_{1 \rightarrow 3}$	1313.333	1313.333	1318.333	1320.000	1330.000	1205.000	1220.713
10	$M_{2 \rightarrow 3}$	193.333	250.000	255.000	260.000	286.667	221.667	245.395
All	Θ_1	0.00327	0.00413	0.00463	0.00513	0.00600	0.00477	0.00469
All	Θ_2	0.00280	0.00353	0.00410	0.00447	0.00540	0.00417	0.00411
All	Θ_3	0.00667	0.00733	0.00843	0.00940	0.01067	0.00883	0.00897
All	$M_{2 \rightarrow 1}$	493.333	500.000	551.667	586.667	753.333	578.333	585.277
All	$M_{3 \rightarrow 1}$	310.000	380.000	428.333	463.333	543.333	428.333	427.592
All	$M_{1 \rightarrow 2}$	903.333	903.333	908.333	923.333	923.333	978.333	984.800
All	$M_{3 \rightarrow 2}$	156.667	220.000	248.333	293.333	393.333	265.000	267.915
All	$M_{1 \rightarrow 3}$	163.333	190.000	238.333	240.000	240.000	421.667	459.176
All	$M_{2 \rightarrow 3}$	263.333	413.333	501.667	540.000	603.333	461.667	448.535

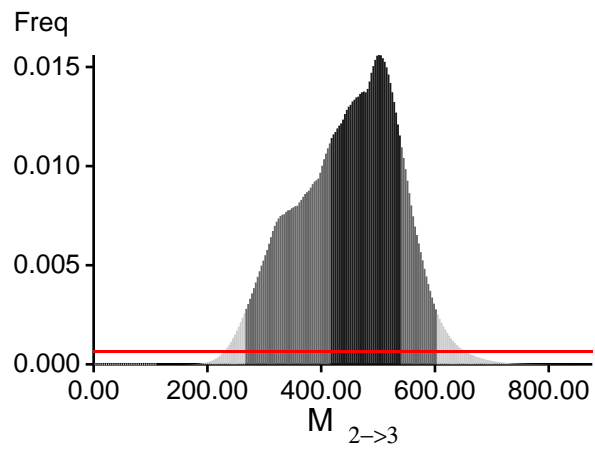
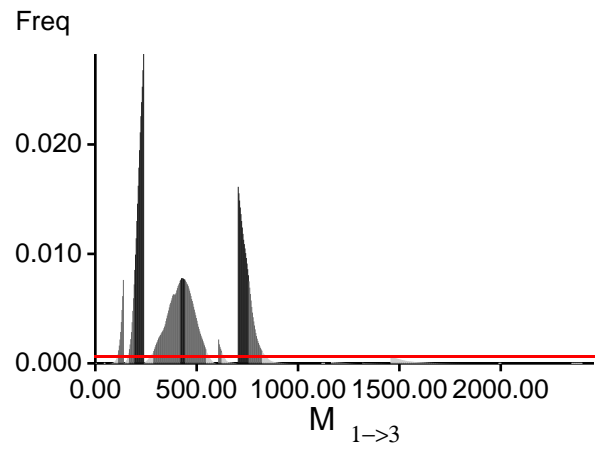
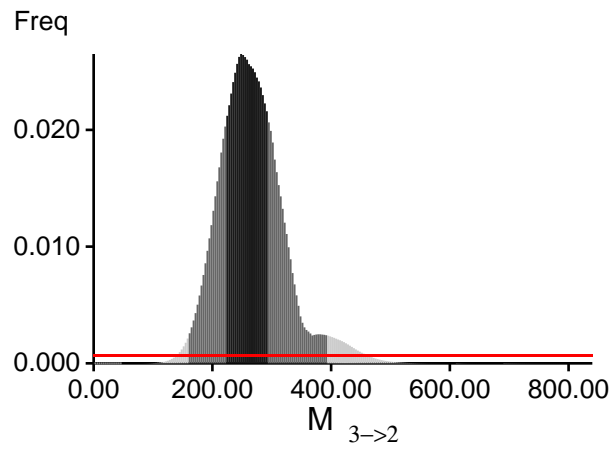
Citation suggestions:

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	-3456.04	-2499.07	-2300.40
2	-2915.19	-2123.89	-1963.74
3	-3145.27	-2280.73	-2119.38
4	-3698.76	-2481.84	-2238.90
5	-3235.26	-2361.44	-2197.79
6	-3171.63	-2452.19	-2308.93
7	-3142.57	-2355.65	-2212.10
8	-3491.00	-2442.12	-2238.21
9	-3615.51	-2495.63	-2275.61
10	-3141.42	-2320.21	-2164.90
All	-32923.77	-23723.89	-21931.07

(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 = 88.879896]

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	13953/27740	0.50299
Θ_2	11755/28107	0.41822
Θ_3	14490/27850	0.52029
$M_{2 \rightarrow 1}$	11965/27710	0.43179
$M_{3 \rightarrow 1}$	12831/27945	0.45915
$M_{1 \rightarrow 2}$	12370/27814	0.44474
$M_{3 \rightarrow 2}$	10835/27962	0.38749
$M_{1 \rightarrow 3}$	13897/27398	0.50723
$M_{2 \rightarrow 3}$	12414/28033	0.44284
Genealogies	29388/249441	0.11782

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.92616	2766.93
Θ_2	0.92363	2755.57
Θ_3	0.93079	2283.92
$M_{2 \rightarrow 1}$	0.91649	2920.51
$M_{3 \rightarrow 1}$	0.90795	3238.09
$M_{1 \rightarrow 2}$	0.89553	4102.90
$M_{3 \rightarrow 2}$	0.89701	4072.23
$M_{1 \rightarrow 3}$	0.90678	3511.07
$M_{2 \rightarrow 3}$	0.91087	3498.36
Genealogies	0.88251	4180.30

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