

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)

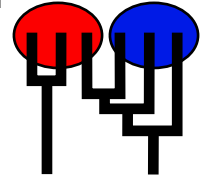
Compiled for PARALLEL computer architectures

One master and 8 compute nodes are available.

Compiled for a SYMMETRIC multiprocessors (Grandcentral)

Program started at Mon Jun 3 08:49:32 2019

Program finished at Mon Jun 3 08:51:15 2019 [Runtime:0000:00:01:43]



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) 301250715

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	*	0	0
2 Berg_2	*	*	0
3 Chur_3	0	*	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	$M_{1 \rightarrow 2}$	<displayed>
5	$M_{2 \rightarrow 3}$	<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	0.10000
2	Theta	** Uniform	0.000000	0.050	0.100	0.010	1500	0.10000
3	Theta	** Uniform	0.000000	0.050	0.100	0.010	1500	0.10000
4	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.10000
5	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.10000

[* * 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

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

Print options:

Data file:

infile

Haplotyping is turned on:	NO
Output file:	outfile_model1
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	0.00380	0.00853	0.00870	0.00880	0.01260	0.00890	0.00975
1	Θ_2	0.00113	0.00267	0.00337	0.00353	0.00740	0.00377	0.00451
1	Θ_3	0.00767	0.00853	0.00910	0.00947	0.01360	0.01270	0.01418
1	$M_{1 \rightarrow 2}$	26.667	90.000	101.667	103.333	360.000	221.667	319.825
1	$M_{2 \rightarrow 3}$	0.000	26.667	38.333	40.000	150.000	68.333	83.789
2	Θ_1	0.00253	0.00407	0.00417	0.00440	0.01080	0.00837	0.00877
2	Θ_2	0.00140	0.00180	0.00197	0.00200	0.00653	0.00383	0.00426
2	Θ_3	0.00140	0.00280	0.00290	0.00300	0.00613	0.00410	0.00450
2	$M_{1 \rightarrow 2}$	43.333	173.333	185.000	210.000	760.000	385.000	488.090
2	$M_{2 \rightarrow 3}$	0.000	36.667	45.000	46.667	370.000	191.667	243.807
3	Θ_1	0.00273	0.00453	0.00510	0.00520	0.00973	0.00563	0.00610
3	Θ_2	0.00533	0.00873	0.00897	0.00927	0.00980	0.01170	0.01886
3	Θ_3	0.00527	0.00933	0.00943	0.01020	0.01633	0.01123	0.01239
3	$M_{1 \rightarrow 2}$	160.000	333.333	341.667	356.667	500.000	345.000	362.658
3	$M_{2 \rightarrow 3}$	0.000	36.667	41.667	86.667	173.333	81.667	90.003
4	Θ_1	0.00660	0.01227	0.01237	0.01253	0.01640	0.01137	0.01184
4	Θ_2	0.00060	0.00207	0.00217	0.00220	0.00840	0.00363	0.00477
4	Θ_3	0.00127	0.00260	0.00270	0.00287	0.00727	0.00490	0.00564
4	$M_{1 \rightarrow 2}$	303.333	426.667	431.667	433.333	583.333	758.333	942.238
4	$M_{2 \rightarrow 3}$	66.667	326.667	335.000	336.667	556.667	405.000	491.599
5	Θ_1	0.00393	0.00653	0.00663	0.00700	0.01013	0.00677	0.00723
5	Θ_2	0.00093	0.00287	0.00317	0.00333	0.01287	0.00597	0.00950
5	Θ_3	0.00813	0.00920	0.00937	0.00940	0.01453	0.01217	0.01385
5	$M_{1 \rightarrow 2}$	530.000	760.000	765.000	766.667	1326.667	948.333	995.364
5	$M_{2 \rightarrow 3}$	33.333	93.333	98.333	103.333	106.667	108.333	137.440
6	Θ_1	0.01007	0.01600	0.01610	0.01613	0.01620	0.01410	0.01421
6	Θ_2	0.00093	0.00127	0.00157	0.00167	0.00787	0.00583	0.00797
6	Θ_3	0.01467	0.01467	0.01490	0.01507	0.01573	0.02463	0.03151
6	$M_{1 \rightarrow 2}$	110.000	576.667	598.333	600.000	903.333	601.667	799.533
6	$M_{2 \rightarrow 3}$	610.000	613.333	625.000	630.000	630.000	758.333	1129.741

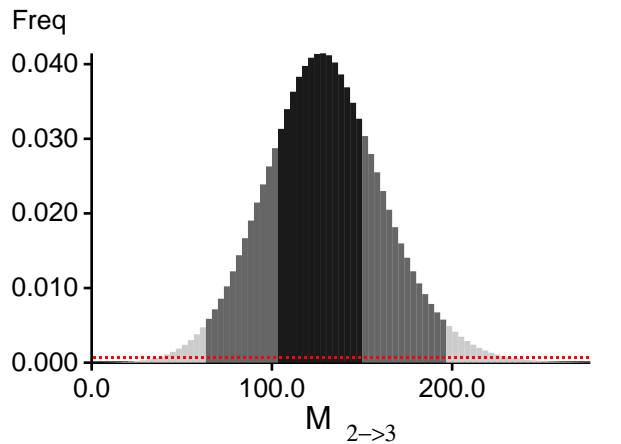
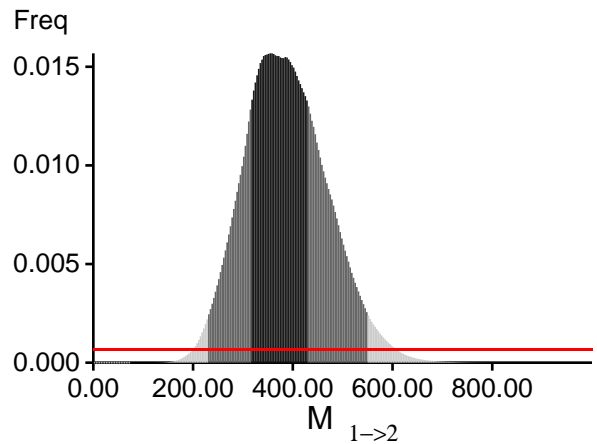
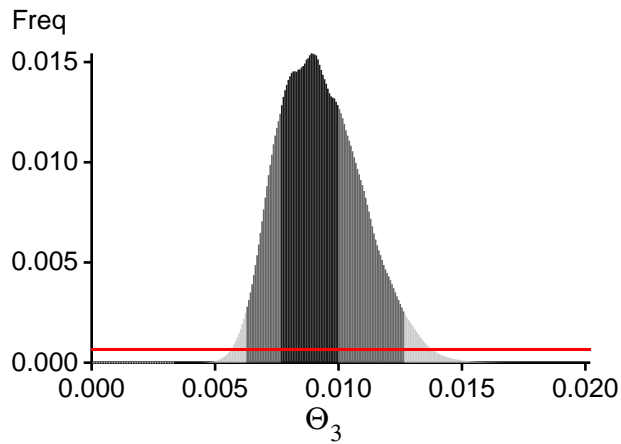
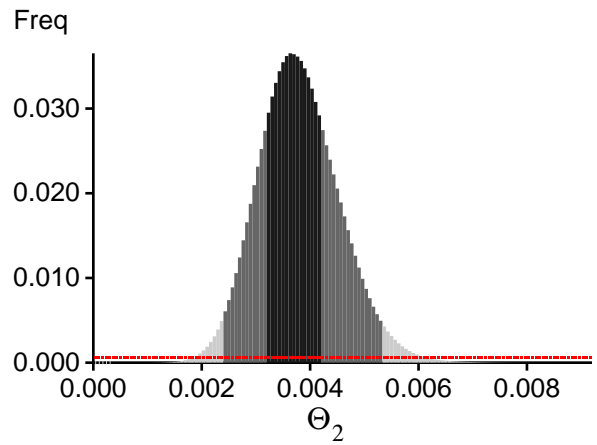
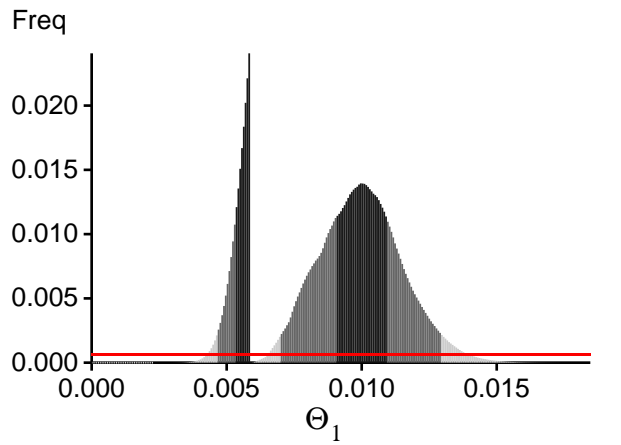
7	Θ_1	0.01660	0.01680	0.01690	0.01707	0.01967	0.01490	0.01497
7	Θ_2	0.00453	0.00547	0.00563	0.00573	0.00573	0.01783	0.02647
7	Θ_3	0.00700	0.01400	0.01410	0.01413	0.02007	0.01463	0.01631
7	$M_{1 \rightarrow 2}$	403.333	413.333	428.333	433.333	433.333	1348.333	1551.646
7	$M_{2 \rightarrow 3}$	0.000	10.000	41.667	46.667	160.000	68.333	78.107
<hr/>								
8	Θ_1	0.01273	0.01813	0.01863	0.01880	0.02487	0.01950	0.02028
8	Θ_2	0.00060	0.00160	0.00170	0.00180	0.00687	0.00330	0.00377
8	Θ_3	0.01427	0.01427	0.01443	0.01447	0.01447	0.05410	0.05358
8	$M_{1 \rightarrow 2}$	666.667	753.333	765.000	766.667	840.000	1285.000	1430.500
8	$M_{2 \rightarrow 3}$	2000.000	2000.000	2005.000	2006.667	2010.000	2461.667	2577.540
<hr/>								
9	Θ_1	0.01627	0.01927	0.01943	0.01947	0.02060	0.01597	0.01679
9	Θ_2	0.00107	0.00167	0.00177	0.00193	0.00300	0.00343	0.00540
9	Θ_3	0.00913	0.01147	0.01223	0.01233	0.01340	0.02277	0.02832
9	$M_{1 \rightarrow 2}$	96.667	96.667	111.667	116.667	120.000	838.333	1048.513
9	$M_{2 \rightarrow 3}$	23.333	366.667	371.667	376.667	643.333	335.000	367.523
<hr/>								
10	Θ_1	0.00227	0.00420	0.00430	0.00467	0.01440	0.00843	0.00929
10	Θ_2	0.00153	0.00620	0.00643	0.00700	0.01180	0.00737	0.00847
10	Θ_3	0.01033	0.01193	0.01217	0.01227	0.01227	0.02043	0.02747
10	$M_{1 \rightarrow 2}$	90.000	150.000	158.333	160.000	426.667	275.000	391.160
10	$M_{2 \rightarrow 3}$	190.000	283.333	291.667	296.667	300.000	568.333	696.624
<hr/>								
All	Θ_1	0.00460	0.00527	0.00583	0.00587	0.00587	0.00950	0.00902
All	Θ_2	0.00233	0.00313	0.00363	0.00420	0.00533	0.00383	0.00380
All	Θ_3	0.00620	0.00760	0.00890	0.01000	0.01267	0.00923	0.00928
All	$M_{1 \rightarrow 2}$	226.667	313.333	355.000	430.000	550.000	385.000	387.844
All	$M_{2 \rightarrow 3}$	60.000	100.000	128.333	150.000	196.667	131.667	129.360

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	-3156.22	-2460.43	-2303.30
2	-2689.81	-2083.29	-1973.09
3	-2824.73	-2230.35	-2130.34
4	-3531.59	-2474.19	-2284.86
5	-3049.19	-2340.45	-2202.95
6	-2956.46	-2419.16	-2316.06
7	-3016.86	-2327.47	-2214.53
8	-3148.76	-2393.97	-2259.90
9	-3174.16	-2437.71	-2297.71
10	-2829.60	-2265.35	-2158.27
All	-30332.36	-23387.34	-22095.98

(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 = 45.020785]

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	25671/49990	0.51352
Θ_2	22676/50141	0.45224
Θ_3	25296/49798	0.50797
$M_{1 \rightarrow 2}$	26865/50166	0.53552
$M_{2 \rightarrow 3}$	20799/50022	0.41580
Genealogies	32671/249883	0.13075

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.86498	4705.99
Θ_2	0.88750	4235.57
Θ_3	0.81074	7199.17
$M_{1 \rightarrow 2}$	0.91771	2999.59
$M_{2 \rightarrow 3}$	0.79799	8105.37
Genealogies	0.86574	5087.60

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