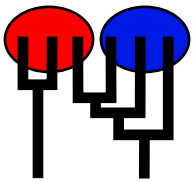
Current protocols example dataset

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 23:24:18 2019
Program finished at Sat Jun 1 23:25:54 2019 [Runtime:0000:00:01:36]



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) 2678086089
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	$M_{1->2}$	<displayed>
5	$M_{2->3}$	<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	0.10000
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	500.0	1000.	100.0	1500	0.10000
5	M	**	Uniform	0.000000	500.0	1000.	100.0	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	0.00493	0.00747	0.00770	0.00867	0.01387	0.00923	0.00970
1	Θ_2	0.00387	0.00527	0.00597	0.00607	0.01453	0.00823	0.00936
1	Θ_3	0.00547	0.00920	0.00937	0.00940	0.01360	0.01323	0.01501
1	$M_{1->2}$	17.333	57.333	58.333	61.333	108.000	83.667	110.355
1	$M_{2->3}$	6.667	58.667	60.333	60.667	94.667	73.000	87.674
2	Θ_1	0.00393	0.00653	0.00703	0.00707	0.01287	0.00743	0.00817
2	Θ_2	0.00120	0.00240	0.00263	0.00267	0.00853	0.00390	0.00454
2	Θ_3	0.00160	0.00460	0.00477	0.00493	0.00833	0.00517	0.00575
2	$M_{1->2}$	114.000	118.000	119.000	120.000	164.000	164.333	174.179
2	$M_{2->3}$	1.333	44.000	45.000	47.333	132.667	65.667	79.726
3	Θ_1	0.00353	0.00660	0.00670	0.00673	0.01013	0.00670	0.00696
3	Θ_2	0.00273	0.00307	0.00317	0.00320	0.00333	0.01383	0.02175
3	Θ_3	0.00540	0.01027	0.01043	0.01060	0.01787	0.01230	0.01389
3	$M_{1->2}$	222.000	253.333	256.333	256.667	267.333	357.000	374.187
3	$M_{2->3}$	4.667	33.333	34.333	34.667	36.000	41.000	50.068
4	Θ_1	0.00207	0.00300	0.00343	0.00347	0.00680	0.00450	0.00485
4	Θ_2	0.00093	0.00173	0.00223	0.00227	0.00660	0.00397	0.00460
4	Θ_3	0.00240	0.00633	0.00650	0.00660	0.00900	0.00657	0.00729
4	$M_{1->2}$	113.333	119.333	120.333	122.000	122.000	179.000	208.768
4	$M_{2->3}$	75.333	98.667	101.000	101.333	101.333	162.333	189.629
5	Θ_1	0.00300	0.00320	0.00343	0.00347	0.00633	0.01063	0.01087
5	Θ_2	0.00193	0.00573	0.00583	0.00593	0.01313	0.00737	0.00905
5	Θ_3	0.01093	0.01173	0.01183	0.01187	0.01267	0.01470	0.01682
5	$M_{1->2}$	76.667	76.667	77.667	78.000	78.000	343.667	380.363
5	$M_{2->3}$	5.333	59.333	61.000	64.000	132.667	76.333	92.934
6	Θ_1	0.00200	0.00353	0.00390	0.00467	0.00713	0.00490	0.00518
6	Θ_2	0.00133	0.00273	0.00297	0.00327	0.01167	0.00723	0.01114
6	Θ_3	0.02940	0.02987	0.02997	0.03007	0.03007	0.03357	0.03672
6	$M_{1->2}$	932.000	939.333	941.000	941.333	944.000	791.000	755.305
6	$M_{2->3}$	0.000	10.000	11.667	12.000	63.333	31.667	34.708

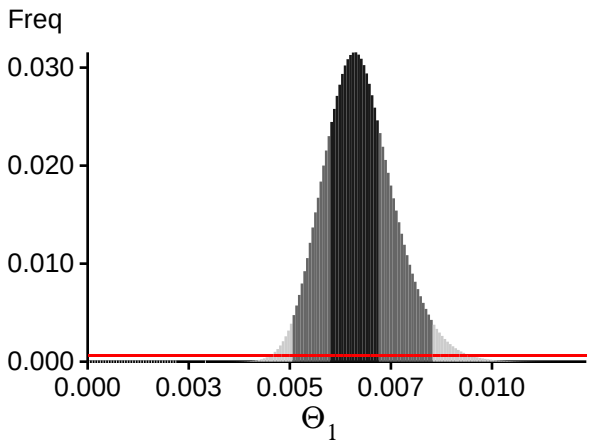
7	Θ_1	0.01767	0.01967	0.01977	0.01993	0.02433	0.01717	0.01752
7	Θ_2	0.00847	0.00847	0.00863	0.00867	0.00987	0.02443	0.03240
7	Θ_3	0.01087	0.01187	0.01197	0.01213	0.01213	0.01683	0.02036
7	$M_{1->2}$	983.333	986.000	987.000	987.333	991.333	722.333	700.932
7	$M_{2->3}$	13.333	16.667	18.333	18.667	26.667	270.333	290.564
8	Θ_1	0.00320	0.00500	0.00543	0.00580	0.00893	0.00683	0.00749
8	Θ_2	0.00240	0.00340	0.00357	0.00367	0.00913	0.00870	0.01097
8	Θ_3	0.00333	0.00867	0.00903	0.00920	0.01627	0.01117	0.01214
8	$M_{1->2}$	236.000	238.000	242.333	242.667	246.000	245.000	273.165
8	$M_{2->3}$	0.000	30.667	31.667	32.667	79.333	46.333	73.003
9	Θ_1	0.00600	0.00660	0.00703	0.00720	0.00753	0.00830	0.00877
9	Θ_2	0.00087	0.00187	0.00230	0.00233	0.00787	0.00370	0.00431
9	Θ_3	0.01160	0.01347	0.01357	0.01367	0.01887	0.01650	0.01856
9	$M_{1->2}$	314.667	322.000	323.000	323.333	331.333	348.333	386.274
9	$M_{2->3}$	20.000	20.667	21.667	22.000	71.333	61.000	79.776
10	Θ_1	0.00587	0.00593	0.00603	0.00633	0.00713	0.00883	0.00973
10	Θ_2	0.00427	0.00647	0.00657	0.00673	0.01160	0.00910	0.01216
10	Θ_3	0.01400	0.01400	0.01423	0.01427	0.01447	0.01677	0.01893
10	$M_{1->2}$	214.667	255.333	256.333	256.667	261.333	291.000	340.656
10	$M_{2->3}$	0.000	4.667	7.000	7.333	60.000	47.000	75.650
All	Θ_1	0.00500	0.00593	0.00663	0.00720	0.00853	0.00677	0.00673
All	Θ_2	0.00300	0.00393	0.00463	0.00533	0.00713	0.00490	0.00493
All	Θ_3	0.00627	0.00733	0.00857	0.00860	0.00860	0.00850	0.00962
All	$M_{1->2}$	156.000	178.667	193.667	194.000	194.000	270.333	252.722
All	$M_{2->3}$	12.000	25.333	31.667	40.667	70.667	36.333	36.839

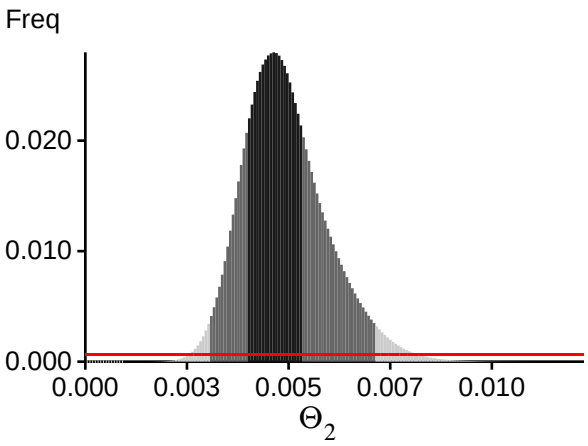
Citation suggestions:

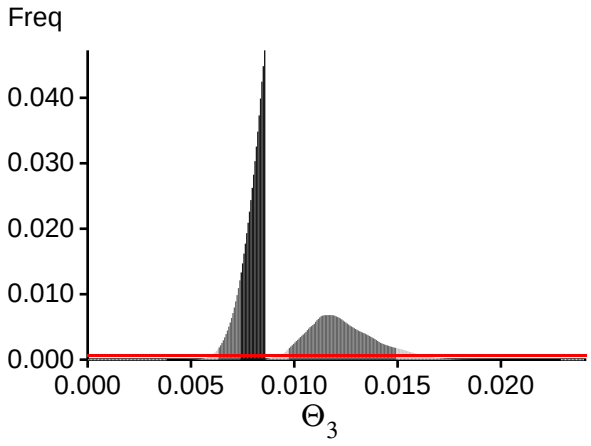
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341-345

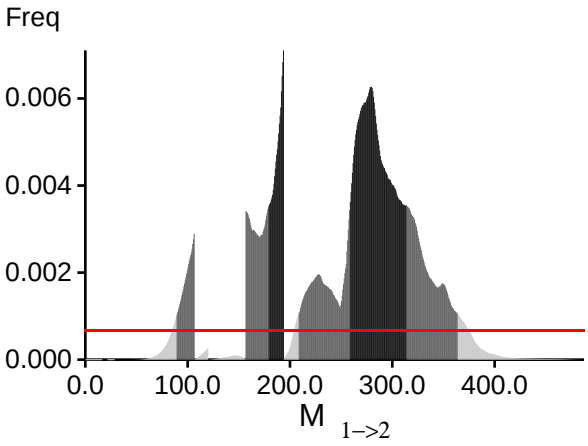
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp. 42-79.

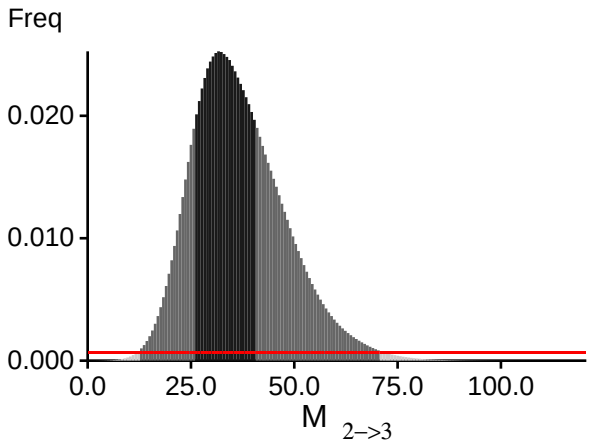
Bayesian Analysis: Posterior distribution over all loci











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	-2938.94	-2390.78	-2293.66
2	-2374.85	-2033.77	-1959.78
3	-2623.39	-2201.84	-2118.40
4	-2794.85	-2342.09	-2262.78
5	-2725.92	-2257.21	-2179.82
6	-2790.48	-2393.02	-2305.24
7	-2779.07	-2284.65	-2202.11
8	-2955.41	-2337.76	-2230.17
9	-2997.81	-2398.37	-2296.61
10	-2669.04	-2231.12	-2162.39
All	-27611.99	-22832.83	-21973.16

(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 = 37.779370]
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	22665/49795	0.45517
Θ_2	24055/49697	0.48403
Θ_3	21276/50053	0.42507
$M_{1->2}$	25443/49892	0.50996
$M_{2->3}$	24699/50129	0.49271
Genealogies	36092/250434	0.14412

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.86464	5415.62
Θ_2	0.86186	5196.11
Θ_3	0.80855	7230.76
$M_{1->2}$	0.75834	9667.60
$M_{2->3}$	0.81137	7319.79
Genealogies	0.82392	6987.23

Potential Problems

This section reports potential problems with your run, but such reporting is often not very accurate. Whith many parameters in a multilocus analysi s, it is very common that some parameters for some loci will not be very informative, triggering suggestions (for example to increase the prior ran ge) that are not sensible. This suggestion tool will improve with time, therefore do not blindly follow its suggestions. If some parameters are fla gged, inspect the tables carefully and judge wether an action is required. For example, if you run a Bayesian inference with sequence data, for mac roscopic 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 rou tes 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.

Param 6 (Locus 6): Upper prior boundary seems too low!
Param 6 (Locus 7): Upper prior boundary seems too low!