

sea urchin (single population; split along a gap)

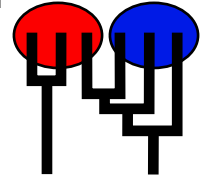
POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 5.0.2(git:v5.0.2-5-g36240df-dirty) [December-1-2021]

Program started at Wed Mar 30 07:41:24 2022

Program finished at Fri Apr 1 04:13:30 2022 [Runtime:0001:20:32:06]



Options

Inheritance multipliers in use for Thetas:

All loci use an inheritance multiplier of 1.0

Random number seed:

(with internal timer)

3349294830

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
1 urchins	*	*
2 nogap	*	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	$M_{2 \rightarrow 1}$	<displayed>
4	$M_{1 \rightarrow 2}$	<displayed>

Mutation rate among loci: Mutation rate is constant for all loci

Analysis strategy: Bayesian inference

-Population size estimation: Mittag-Leffler with alpha=0.90

-Geneflow estimation: Mittag-Leffler with alpha=0.90

Proposal distributions for parameter

Parameter	Proposal
Theta	Slice sampling
M	Slice sampling
Divergence	Slice sampling
Divergence Spread	Slice 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	0.12500
2	Theta	** Uniform	0.000000	0.050	0.100	0.010	1500	0.12500
3	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.12500
4	M	** Uniform	0.000000	2500.	5000.	500.0	1500	0.12500

[* * means priors were set globally]

Markov chain settings: Long chain

Number of chains	1
Recorded steps [a]	500
Increment (record every x step [b])	10
Number of concurrent chains (replicates) [c]	2
Visited (sampled) parameter values [a*b*c]	10000
Number of discard trees per chain (burn-in)	100

Multiple Markov chains:

Static heating scheme	4 chains with temperatures
	100000.00 3.00 1.50 1.00
	Swapping interval is 1

Print options:

Data file:	infile.gap
	parmfile_frac2.gap
Haplotyping is turned on:	NO
Output file:	outfile-fracgap2
Posterior distribution raw histogram file:	bayesfile
Raw data from the MCMC run:	bayesallfile-fracgap2
Print data:	No
Print genealogies [only some for some data type]:	None

Data summary

Data file: infile.gap
 Datatype: Haplotype data
 Number of loci: 5

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]

Sites per locus

Locus	Sites
1	252
2	921
3	425
4	459
5	713

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

Population	Locus	Gene copies data (missing)
1 urchins	1	24
	2	24
	3	24
	4	24
	5	24
2 nogap	1	34
	2	34
	3	34
	4	34

	5	34	
Total of all populations	1	58	(0)
	2	58	(0)
	3	58	(0)
	4	58	(0)
	5	58	(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00067	0.00300	0.00477	0.00700	0.01620	0.00597	0.00709
1	Θ_2	0.00000	0.00180	0.00350	0.00567	0.01527	0.00477	0.00615
1	$M_{2 \rightarrow 1}$	266.667	976.667	1335.000	1880.000	3513.333	1765.000	2009.768
1	$M_{1 \rightarrow 2}$	643.333	790.000	985.000	1143.333	3326.667	2175.000	2585.827
2	Θ_1	0.00547	0.00860	0.01150	0.01647	0.03513	0.01510	0.02001
2	Θ_2	0.00993	0.01493	0.01863	0.02793	0.06180	0.02590	0.03092
2	$M_{2 \rightarrow 1}$	213.333	313.333	441.667	680.000	2460.000	1051.667	1173.531
2	$M_{1 \rightarrow 2}$	173.333	223.333	381.667	773.333	1833.333	745.000	1237.320
3	Θ_1	0.01327	0.01793	0.02223	0.02793	0.05993	0.02690	0.03226
3	Θ_2	0.01160	0.01593	0.01997	0.02607	0.06487	0.03110	0.03478
3	$M_{2 \rightarrow 1}$	316.667	433.333	741.667	1093.333	1756.667	1031.667	1571.960
3	$M_{1 \rightarrow 2}$	920.000	2223.333	2695.000	2793.333	2930.000	1978.333	1981.538
4	Θ_1	0.00500	0.00680	0.01010	0.02533	0.07387	0.02423	0.03475
4	Θ_2	0.00207	0.00393	0.00563	0.00727	0.01640	0.00650	0.00757
4	$M_{2 \rightarrow 1}$	463.333	843.333	1145.000	1340.000	1786.667	1168.333	1293.660
4	$M_{1 \rightarrow 2}$	140.000	263.333	398.333	540.000	850.000	481.667	614.317
5	Θ_1	0.00887	0.01367	0.01690	0.02460	0.03920	0.02137	0.02244
5	Θ_2	0.00167	0.00567	0.00830	0.01067	0.02120	0.00903	0.01036
5	$M_{2 \rightarrow 1}$	363.333	443.333	538.333	673.333	986.667	621.667	645.622
5	$M_{1 \rightarrow 2}$	556.667	723.333	821.667	990.000	1240.000	891.667	896.948
All	Θ_1	0.00520	0.00793	0.01523	0.01653	0.02887	0.01530	0.01636
All	Θ_2	0.00173	0.00413	0.00570	0.00700	0.01907	0.00610	0.00762
All	$M_{2 \rightarrow 1}$	360.000	423.333	518.333	616.667	1360.000	655.000	745.077
All	$M_{1 \rightarrow 2}$	196.667	713.333	1028.333	1123.333	1846.667	908.333	1044.178

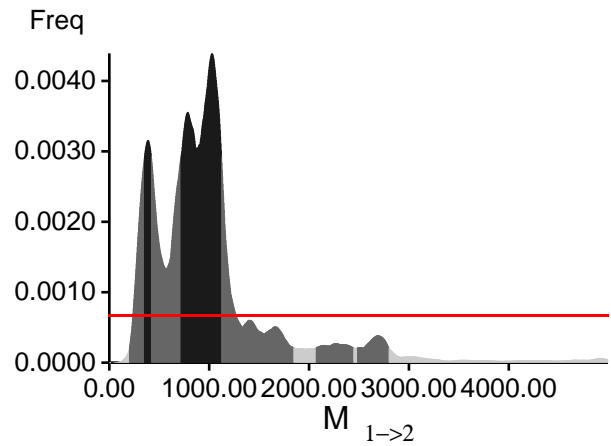
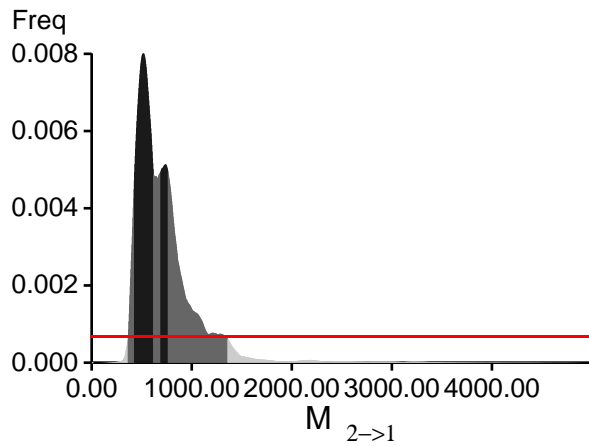
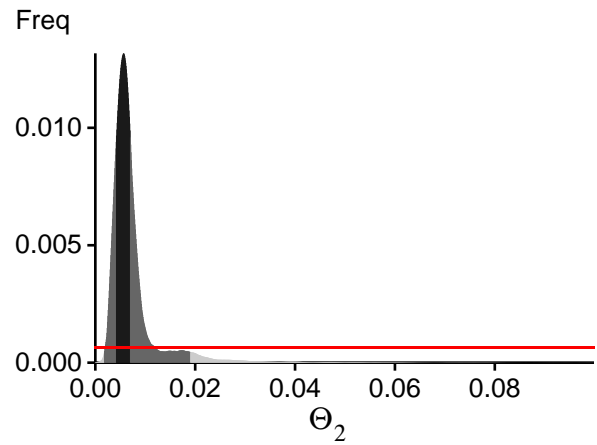
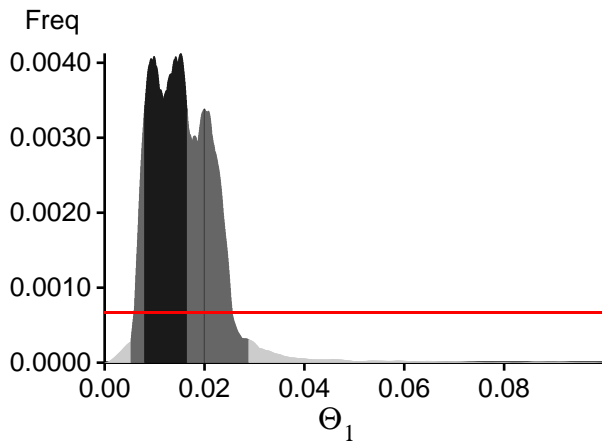
Citation suggestions:

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	TI(1a)	BTI(1b)	HS(3)
1	-579.72	-534.96	-556.04
2	-2362.48	-2171.48	-2353.60
3	-1581.85	-1391.57	-1522.26
4	-1183.80	-1121.24	-1105.12
5	-4656.63	-4008.45	-3692.05
All	-10365.93	-9229.15	-9230.53

(1a) TI: Thermodynamic integration: $\log(\text{Prob}(D|\text{Model}))$: Good approximation with many temperatures

(1b) BTI: Bezier-approximated Thermodynamic integration: when using few temperatures USE THIS!

(2) SS: Steppingstone Sampling (Xie et al 2011)

(3) HS: Harmonic mean approximation: Overestimates the marginal likelihood, poor variance

[Scaling factor = -1.453341]

Citation suggestions:

Beerli P. and M. Palczewski, 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations, *Genetics*, 185: 313-326.

Palczewski M. and P. Beerli, 2014. Population model comparison using multi-locus datasets.

In M.-H. Chen, L. Kuo, and P. O. Lewis, editors, *Bayesian Phylogenetics: Methods, Algorithms, and Applications*, pages 187-200. CRC Press, 2014.

Xie W., P. O. Lewis, Y. Fan, L. Kuo, and M.-H. Chen. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology*, 60(2):150â 160, 2011.

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	6378/6378	1.00000
Θ_2	6220/6220	1.00000
$M_{2 \rightarrow 1}$	6281/6281	1.00000
$M_{1 \rightarrow 2}$	6186/6186	1.00000
Genealogies	5571/24935	0.22342

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.67566	979.72
Θ_2	0.73462	794.94
$M_{2 \rightarrow 1}$	0.77209	682.40
$M_{1 \rightarrow 2}$	0.79184	622.79
Genealogies	0.67566	979.72

Average temperatures during the run

Chain	Temperatures
-------	--------------

1	1.00000
2	1.50000
3	3.00000
4	100000.00000

Adaptive heating often fails, if the average temperatures are very close together
try to rerun using static heating! If you want to compare models using marginal
likelihoods then you **MUST** use static heating

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