

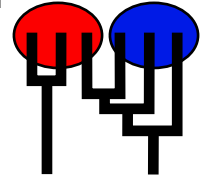
POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 4.4.4(git:) [June-1-2019]

Program started at Wed Oct 27 18:15:07 2021

Program finished at Thu Oct 28 15:43:58 2021 [Runtime:0000:21:28:51]



Options

Datatype: DNA sequence data

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]

Data set was subsampled: used a random sample of size:

10 and seed 198235

Random number seed: (with internal timer) 810236538

Start parameters:

Theta values were generated ERROR

M values were generated ERROR

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 POP1	*	0	0
2 POP2	*	*	0
3 POP3	*	0	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	$M_{1 \rightarrow 2}$	<displayed>
5	$M_{1 \rightarrow 3}$	<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	25.00	50.00	0.010	1500	0.10000
2	Theta **	Uniform	0.000000	25.00	50.00	0.010	1500	0.10000
3	Theta **	Uniform	0.000000	25.00	50.00	0.010	1500	0.10000
4	M **	Uniform	0.000000	500.0	1000.	1.000	1500	0.10000
5	M **	Uniform	0.000000	500.0	1000.	1.000	1500	0.10000

[* * means priors were set globally]

Markov chain settings:

Long chain

Number of chains

1

Recorded steps [a]

20000

Increment (record every x step [b])

300

Number of concurrent chains (replicates) [c]

3

Visited (sampled) parameter values [a*b*c]

18000000

Number of discard trees per chain (burn-in)

10000

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:	Migrate_all_locus2.txt
Haplotyping is turned on:	NO
Output file:	outfile_YYfinal_napo2yunnan_10i
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: Migrate_all_locus2.txt
 Datatype: Sequence data
 Number of loci: 2

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Felsenstein 84	[Bf:0.25 0.26 0.25 0.25, t/t ratio=2.000]
2	1	Felsenstein 84	[Bf:0.25 0.26 0.25 0.25, t/t ratio=2.000]

Data set was subsampled: used a random sample of size: 10

Sites per locus

Locus	Sites
1	13148
2	13148

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

Population	Locus	Gene copies
1 POP1	1	80
	2	80
2 POP2	1	94
	2	94
3 POP3	1	55
	2	55
Total of all populations	1	229
	2	229

Subsampled dataset

Data set was subsampled randomly per population: 10 samples taken

Locus	Population	Individuals	
1	POP1	VGNP32.cle VGNP110.cl VGNP39.cle VGNP1.clea VGNP27.cle VGNP105.cl VGNP1.clea VGNP2.clea VGNP13.cle VGNP103.cl	
	POP2	VXRM192.cl VXRM183.cl VXRM213.cl VXRM200.cl VXRM185.cl VXRM210.cl VXRM190.cl VXRM170.cl VXRM214.cl VXRM178.cl	
	POP3	VXRN136.cl VXRN04.cle VXRN139.cl VXRN126.cl VXRN130.cl VXRN107.cl VXRN137.cl VXRN142.cl VXRN128.cl VXRN119.cl	
	2	POP1	VGNP94.cle VGNP88.cle VGNP78.cle VGNP59.cle VGNP72.cle VGNP99.cle VGNP84.cle VGNP73.cle VGNP83.cle VGNP56.cle
		POP2	VXRM223.cl VXRM62.cle VXRM247.cl VXRM258.cl VXRM236.cl VXRM235.cl VXRM240.cl VXRM267.cl VXRM238.cl VXRM236.cl
		POP3	VXRN16.cle VXRN28.cle VXRN154.cl VXRN154.cl VXRN30.cle VXRN52.cle VXRN24.cle VXRN163.cl VXRN153.cl VXRN25.cle

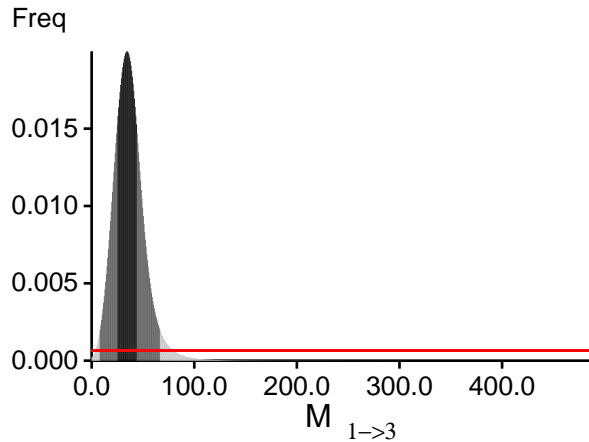
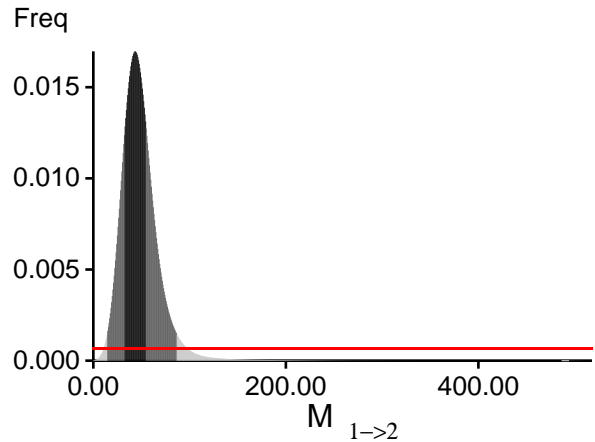
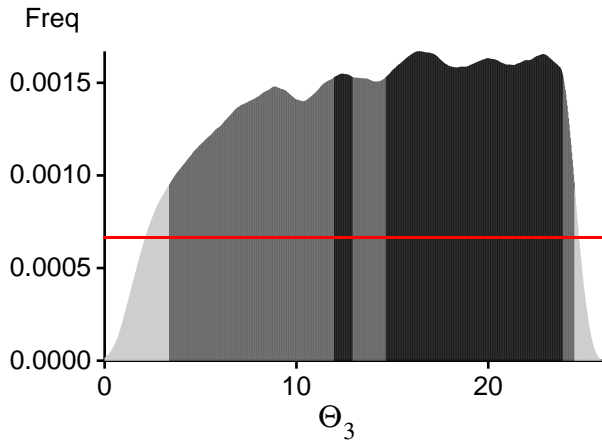
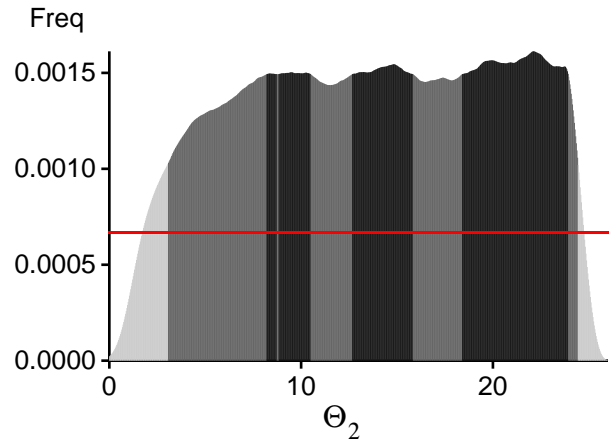
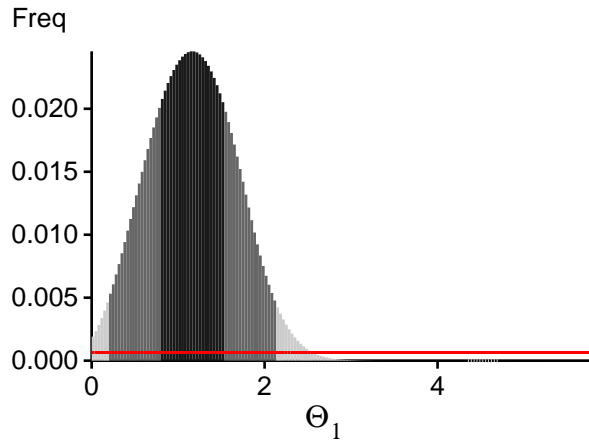
Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.73333	1.00000	1.25000	1.43333	1.96667	1.31667	1.30799
1	Θ_2	18.23333	18.36667	18.48333	18.50000	18.50000	26.58333	26.34028
1	Θ_3	34.43333	35.00000	35.18333	35.23333	35.80000	25.18333	25.28992
1	$M_{1 \rightarrow 2}$	14.0	23.3	40.3	64.0	122.7	51.0	60.2
1	$M_{1 \rightarrow 3}$	10.0	19.3	38.3	70.0	128.7	70.3	247.2
2	Θ_1	0.50000	0.53333	0.75000	1.00000	1.23333	0.95000	1.07920
2	Θ_2	21.20000	21.36667	21.41667	21.46667	22.03333	25.31667	25.39251
2	Θ_3	26.53333	28.06667	28.15000	28.23333	28.23333	27.58333	27.16772
2	$M_{1 \rightarrow 2}$	12.0	18.0	37.7	79.3	307.3	61.0	112.7
2	$M_{1 \rightarrow 3}$	7.3	20.7	24.3	28.0	123.3	31.7	46.3
All	Θ_1	0.16667	0.76667	1.15000	1.53333	2.13333	1.18333	1.17160
All	Θ_2	3.03333	18.36667	22.15000	23.93333	24.43333	13.81667	13.63807
All	Θ_3	3.33333	14.63333	16.41667	23.90000	24.50000	14.35000	13.99994
All	$M_{1 \rightarrow 2}$	14.0	32.0	43.7	54.7	86.7	47.0	49.1
All	$M_{1 \rightarrow 3}$	7.3	24.7	34.3	44.0	66.7	36.3	37.1

Citation suggestions:

- Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345
- Beerli P., 2007. Estimation of the population scaled mutation rate from microsatellite data, *Genetics*, 177:1967-1968.
- 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	-243518.66	-184245.70	-174522.66
2	-233528.84	-177650.22	-169600.01
All	-477049.68	-361898.10	-344124.85

(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 = -2.180542]

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	1784715/3601840	0.49550
Θ_2	3424295/3602026	0.95066
Θ_3	3394394/3598702	0.94323
$M_{1 \rightarrow 2}$	1584545/3599316	0.44024
$M_{1 \rightarrow 3}$	1850964/3599527	0.51422
Genealogies	43189/17998589	0.00240

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.77669	15942.40
Θ_2	0.01658	116268.37
Θ_3	0.01912	115746.70
$M_{1 \rightarrow 2}$	0.66051	26166.64
$M_{1 \rightarrow 3}$	0.82233	13712.25
Genealogies	0.98770	743.74

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