

chr22-3pop.vcf.gz

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

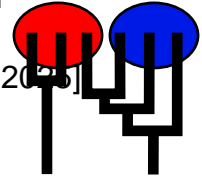
Migrate-n version 6.0.1 [Mittag (merged with main Oct 11 2025)] [October-11-2025]

Compiled for PARALLEL computer architectures

One master and 5 compute nodes are available.

Program started at Tue Jan 6 18:49:01 2026

Program finished at Tue Jan 6 18:49:38 2026 [Runtime:0000:00:00:37]



Options

Datatype:

DNA sequence data

Inheritance multipliers in use for Thetas:

All loci use an inheritance multiplier of 1.0

Random number seed:

(with internal timer)

3410082789

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 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	$M_{2 \rightarrow 3}$	<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
Genealogy	Metropolis-Hastings

Prior distribution for parameter

Parameter		Prior	Minimum	Mean*	Maximum	Delta	Bins	UpdateFreq
1 Theta	* *	Exponential	0.000000	0.100	0.200	-	1500	0.10000
2 Theta	* *	Exponential	0.000000	0.100	0.200	-	1500	0.10000
3 Theta	* *	Exponential	0.000000	0.100	0.200	-	1500	0.10000
4 M	* *	Exponential	0.000000	100.0	10000	-	1500	0.10000
5 M	* *	Exponential	0.000000	100.0	10000	-	1500	0.10000

[* * means priors were set globally]

Posterior distribution:

Parameter values were collected using MCMC, these values

were then used to generate the posterior histograms using KERNEL SMOOTHING (window=41)

and subsequent SAVITZKY-GOLAY SMOOTHING (window=41) for combination over loci

Markov chain settings:

Long chain

Number of chains	1
Recorded steps [a]	10000
Increment (record every x step [b])	100
Number of concurrent chains (replicates) [c]	1
Visited (sampled) parameter values [a*b*c]	1000000
Number of discard trees per replicate (burn-in * b)	100000

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

parmfile-x00xx00xx-5e

Haplotyping is turned on:

NO

Output file:

outfile-x00xx00xx-5e

Posterior distribution raw histogram file:

bayesfile

Raw data from the MCMC run:

bayesallfile.txt

Print data:

No

Print genealogies [only some for some data type]:

None

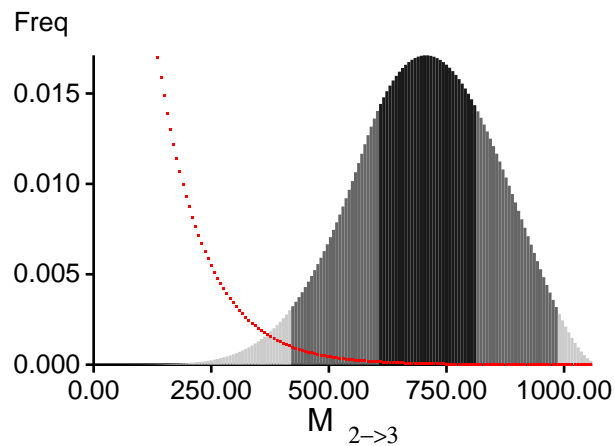
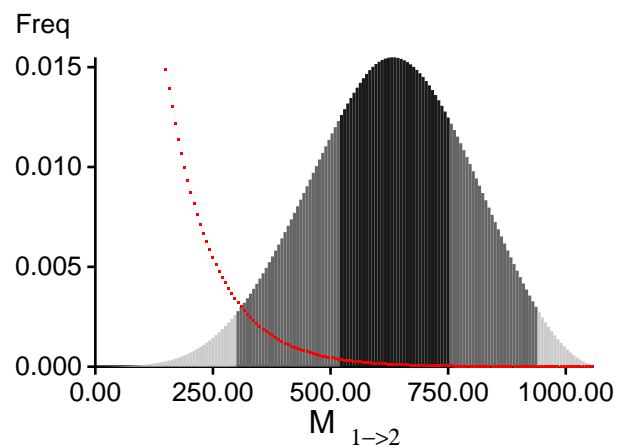
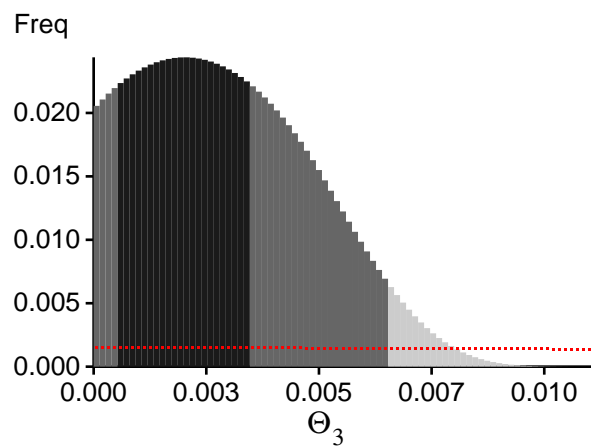
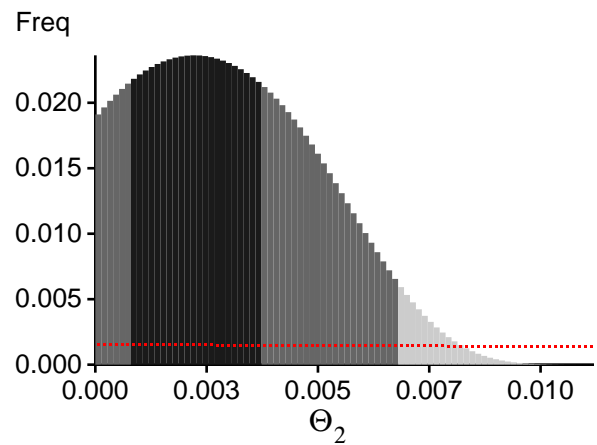
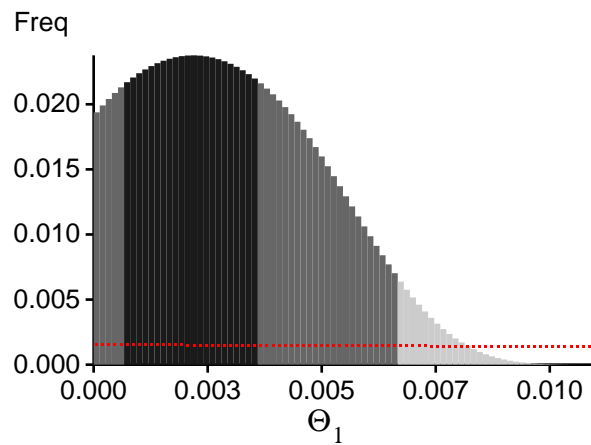
Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00000	0.00133	0.00340	0.00533	0.00960	0.00420	0.00299
1	Θ_2	0.00000	0.00107	0.00327	0.00520	0.01107	0.00420	0.00318
1	Θ_3	0.00000	0.00080	0.00300	0.00480	0.01187	0.00407	0.00283
1	$M_{1 \rightarrow 2}$	0.0	126.7	256.7	400.0	700.0	316.7	289.7
1	$M_{2 \rightarrow 3}$	0.0	180.0	330.0	486.7	800.0	376.7	367.2
2	Θ_1	0.00000	0.00093	0.00273	0.00453	0.00827	0.00367	0.00175
2	Θ_2	0.00000	0.00093	0.00287	0.00467	0.00893	0.00380	0.00209
2	Θ_3	0.00000	0.00080	0.00287	0.00480	0.01427	0.00407	0.00329
2	$M_{1 \rightarrow 2}$	0.0	86.7	210.0	326.7	600.0	256.7	210.8
2	$M_{2 \rightarrow 3}$	0.0	126.7	256.7	406.7	726.7	316.7	295.9
3	Θ_1	0.00000	0.00093	0.00287	0.00453	0.00853	0.00367	0.00184
3	Θ_2	0.00000	0.00080	0.00273	0.00440	0.00880	0.00367	0.00175
3	Θ_3	0.00000	0.00067	0.00260	0.00427	0.00880	0.00353	0.00170
3	$M_{1 \rightarrow 2}$	0.0	80.0	203.3	313.3	580.0	250.0	195.6
3	$M_{2 \rightarrow 3}$	0.0	80.0	203.3	320.0	600.0	256.7	201.5
4	Θ_1	0.00000	0.00080	0.00273	0.00427	0.00813	0.00353	0.00154
4	Θ_2	0.00000	0.00067	0.00260	0.00413	0.00827	0.00353	0.00141
4	Θ_3	0.00000	0.00080	0.00300	0.00507	0.02667	0.00433	0.00473
4	$M_{1 \rightarrow 2}$	0.0	113.3	243.3	380.0	686.7	303.3	270.7
4	$M_{2 \rightarrow 3}$	0.0	133.3	270.0	426.7	760.0	336.7	315.3
5	Θ_1	0.00000	0.00093	0.00300	0.00467	0.00880	0.00380	0.00209
5	Θ_2	0.00000	0.00120	0.00367	0.00587	0.01400	0.00487	0.00427
5	Θ_3	0.00000	0.00080	0.00287	0.00467	0.01080	0.00393	0.00245
5	$M_{1 \rightarrow 2}$	0.0	100.0	223.3	353.3	653.3	283.3	239.7
5	$M_{2 \rightarrow 3}$	0.0	126.7	250.0	400.0	713.3	316.7	289.3
All	Θ_1	0.00000	0.00053	0.00220	0.00360	0.00667	0.00313	0.00219
All	Θ_2	0.00000	0.00067	0.00220	0.00373	0.00680	0.00313	0.00222
All	Θ_3	0.00000	0.00040	0.00207	0.00347	0.00653	0.00300	0.00202
All	$M_{1 \rightarrow 2}$	293.3	513.3	630.0	753.3	940.0	636.7	624.1
All	$M_{2 \rightarrow 3}$	413.3	600.0	703.3	813.3	986.7	710.0	701.7

Citation suggestions:

- Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345
- Beerli, P., H. Ashki, S. Mashayekhi, and M. Palczewski, 2022. Population divergence time estimation using individual lineage label switching. *G3 Genes & Genomes & Genetics*, 12(4), 02 2022.
- 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.
- Beerli, P., S. Mashayekhi, M. Sadeghi, M. Khodaei, and K. Shaw, 2019. Population genetic inference with migrate. *Current Protocols in Bioinformatics*, 68(1):e87.

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	-1598.84	-1469.69	-1444.54
2	-1564.18	-1435.28	-1415.90
3	-1554.06	-1429.43	-1413.82
4	-1558.05	-1433.36	-1416.90
5	-1575.51	-1453.51	-1438.21
All	-7822.47	-7193.10	-7101.20

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

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

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

[Scaling factor = 28.166743]

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.

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	12339/500036	0.02468
Θ_2	16659/499762	0.03333
Θ_3	22080/498505	0.04429
$M_{1 \rightarrow 2}$	186270/499843	0.37266
$M_{2 \rightarrow 3}$	139084/499991	0.27817
Genealogies	1254711/499991	2.50947

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.50668	16409.94
Θ_2	0.33263	25237.61
Θ_3	0.39266	21870.57
$M_{1 \rightarrow 2}$	0.39385	22348.39
$M_{2 \rightarrow 3}$	0.50162	17487.64
Genealogies	0.50668	16409.94

Static temperatures during the run using standard heating

Chain	Temperatures	log(marginal likelihood)	log(mL_steppingstone)
1	1.00000	-1413.52449	-1186.04577
2	1.00000	-1420.72931	-787.93795
3	1.00000	-1451.98512	-400.77501
4	1.00000	-2261.81955	8.26252

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 common that some parameters for some loci will not be informative. These parameters then trigger suggestions to increase the prior range that are not sensible. Do not blindly follow the suggestions given. If some parameters are flagged, inspect the tables carefully and judge whether an action is required. Suppose you run a Bayesian inference with sequence data for macroscopic species. In that case, there is rarely the need to increase the prior for Theta beyond 0.1. If you use microsatellites data, 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 prior range will not help in such situations, but reducing the number of parameters may help.

No warning was recorded during the run