

two (fake) Swiss towns

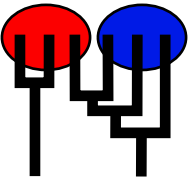
POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 5.0.7 [May-01-2025]

Program started at Sat May 3 12:22:58 2025

Program finished at Sat May 3 12:24:21 2025 [Runtime:0000:00:01:23]



Options

Inheritance multipliers in use for Thetas:

All loci use an inheritance multiplier of 1.0

Random number seed: (with internal timer) 127620206

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 Aadorf	*	*
2 Bern	*	*

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:

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	0.010	0.100	-	1500	0.10417
2	Theta	*	*	Exponential	0.000000	0.010	0.100	-	1500	0.10417
3	M	2	1	Uniform	0.000000	100.0	1000.	100.00000	2000	0.10417
4	M	1	2	Uniform	0.000000	100.0	1000.	100.00000	2000	0.10417

[* * 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])	50
Number of concurrent chains (replicates) [c]	1
Visited (sampled) parameter values [a*b*c]	500000
Number of discard trees per replicate (burn-in * b)	250000

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:

twoswisstowns
parmfile.twoswisstowns

Haplotyping is turned on:

NO

Output file:	outfile-twoswisstowns
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

Data summary

Data file:	twoswisstowns
Datatype:	Haplotype data
Number of loci:	3

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Tamura-Nei	[Bf:0.30 0.25 0.24 0.22, k1=1.300, k2=0.800]
1	2	Felsenstein 84	[Bf:0.24 0.28 0.22 0.27, t/t ratio=2.000]
2	1	Tamura-Nei	[Bf:0.27 0.23 0.24 0.26, k1=1.300, k2=2.000]
3	1	Jukes-Cantor	[Basefreq: =0.25]

Sites per locus

Locus	Sites
1	200 800
2	500
3	500

Site rate variation and probabilities:

Locus	Sublocus	Region type	Rate of change	Probability	Patch size
-------	----------	-------------	----------------	-------------	------------

1	1	1	1.000	1.000	1.000
1	2	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

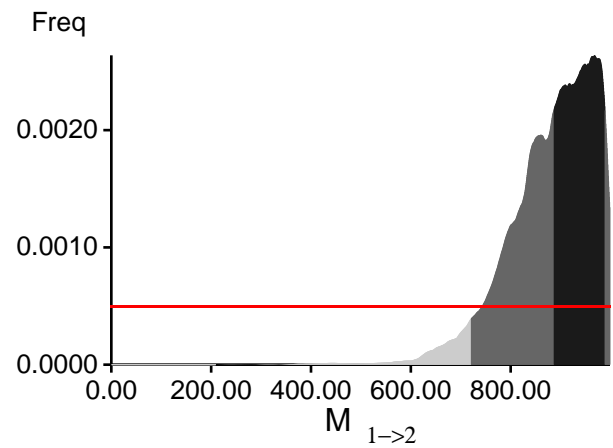
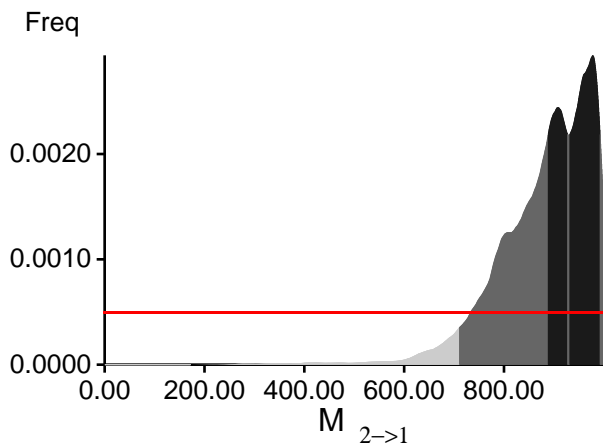
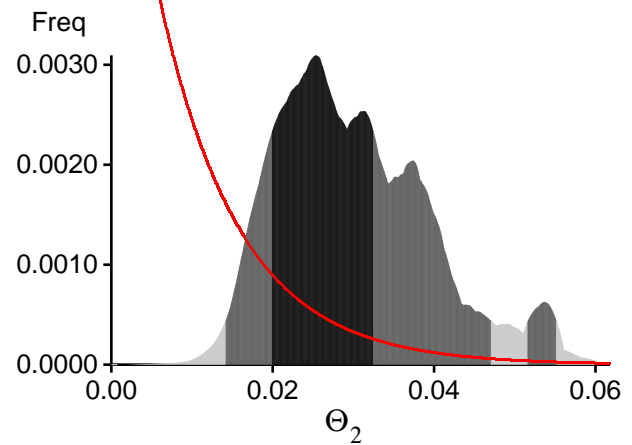
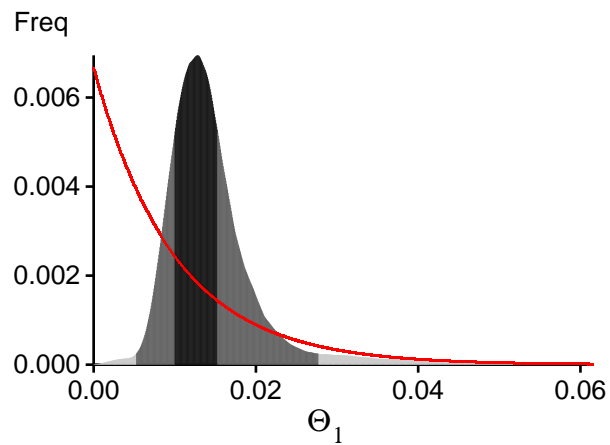
Population	Locus	Gene copies data	(missing)
1 Aadorf	1	10	
	2	10	
	3	10	
2 Bern	1	10	
	2	10	
	3	10	
Total of all populations	1	20	(0)
	2	20	(0)
	3	20	(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00360	0.00747	0.01030	0.01380	0.02433	0.01190	0.01292
1	Θ_2	0.00593	0.01187	0.01643	0.02180	0.03967	0.01883	0.02063
1	$M_{2 \rightarrow 1}$	576.000	842.000	979.250	993.500	999.500	848.750	826.422
1	$M_{1 \rightarrow 2}$	588.000	847.000	978.750	990.500	999.500	855.750	833.540
2	Θ_1	0.00187	0.00520	0.00770	0.01093	0.02173	0.00937	0.01045
2	Θ_2	0.00560	0.01033	0.01617	0.02227	0.03987	0.01950	0.02147
2	$M_{2 \rightarrow 1}$	490.000	829.500	970.750	991.000	999.500	805.750	780.837
2	$M_{1 \rightarrow 2}$	548.500	823.000	967.250	990.000	999.500	832.250	809.191
3	Θ_1	0.00560	0.01073	0.01450	0.02000	0.03653	0.01743	0.01898
3	Θ_2	0.00980	0.01753	0.02583	0.03180	0.05267	0.02790	0.02977
3	$M_{2 \rightarrow 1}$	582.500	832.000	960.250	988.500	999.500	842.250	822.657
3	$M_{1 \rightarrow 2}$	594.500	824.500	917.750	957.000	999.500	830.250	812.985
All	Θ_1	0.00520	0.00993	0.01283	0.01527	0.02773	0.01350	0.01464
All	Θ_2	0.01413	0.01987	0.02537	0.03247	0.04707	0.02890	0.03018
All	$M_{2 \rightarrow 1}$	710.500	931.500	977.750	992.500	999.500	899.250	882.102
All	$M_{1 \rightarrow 2}$	720.500	886.500	967.750	989.000	999.500	896.250	882.834

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., 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	-2617.36	-2310.13	-2474.44
2	-1423.27	-1257.04	-1299.25
3	-1579.71	-1350.93	-1447.27
All	-5633.87	-4931.62	-5234.48

(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 = -13.525702]

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	43460/155638	0.27924
Θ_2	34705/156366	0.22195
$M_{2 \rightarrow 1}$	77293/156245	0.49469
$M_{1 \rightarrow 2}$	71538/155933	0.45877
Genealogies	66140/875818	0.07552

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.56750	8322.11
Θ_2	0.60805	7417.09
$M_{2 \rightarrow 1}$	0.56266	8426.67
$M_{1 \rightarrow 2}$	0.52461	9384.51
Genealogies	0.56750	8322.11

Average temperatures during the run

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

1	1.00000
2	0.66667
3	0.33333
4	0.00000

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.

Param 3 (Locus 1): Upper prior boundary seems too low!
Param 4 (Locus 1): Upper prior boundary seems too low!
Param 3 (Locus 2): Upper prior boundary seems too low!
Param 4 (Locus 2): Upper prior boundary seems too low!
Param 3 (Locus 3): Upper prior boundary seems too low!
Param 4 (Locus 3): Upper prior boundary seems too low!
Param 3 (all loci): Upper prior boundary seems too low!
Param 4 (all loci): Upper prior boundary seems too low!

Summary Assignment of Individuals to Populations

Individual	Population	
	1	2
?BAH0	0.306	0.694
?BAF0	0.302	0.698
?BAG1	0.397	0.603
?BAJ1	0.451	0.549
?BAH1	0.354	0.646
?BAI1	0.390	0.610
?BAF1	0.342	0.658

Detailed Assignment of Individuals to Populations

Individual	Locus	Population	
		1	2
?BAH0	1	0.471	0.529
?BAH0	2	0.398	0.602
?BAH0	3	0.427	0.573
?BAH0	All	0.306	0.694
?BAF0	1	0.442	0.558
?BAF0	2	0.400	0.600
?BAF0	3	0.450	0.550
?BAF0	All	0.302	0.698
?BAG1	1	0.513	0.487
?BAG1	2	0.411	0.589
?BAG1	3	0.473	0.527
?BAG1	All	0.397	0.603
?BAJ1	1	0.511	0.489
?BAJ1	2	0.450	0.550
?BAJ1	3	0.490	0.510
?BAJ1	All	0.451	0.549
?BAH1	1	0.467	0.533
?BAH1	2	0.418	0.582
?BAH1	3	0.466	0.534
?BAH1	All	0.354	0.646
?BAI1	1	0.483	0.517
?BAI1	2	0.438	0.562
?BAI1	3	0.468	0.532
?BAI1	All	0.390	0.610
?BAF1	1	0.424	0.576
?BAF1	2	0.439	0.561
?BAF1	3	0.474	0.526
?BAF1	All	0.342	0.658