

# *two (fake) Swiss towns*

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version debug 4.2.7 [April-1-2016]

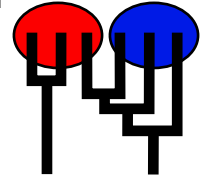
Using Intel AVX (Advanced Vector Extensions)

Compiled for PARALLEL computer architectures

One master and 2 compute nodes are available.

Program started at Sun Apr 3 22:16:32 2016

Program finished at Sun Apr 3 22:30:35 2016



## *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)

882827702

Start parameters:

Theta values were generated

Using a percent value of the prior

M values were generated

from guessed values

M-matrix:

- 0.0,  
0.0, -

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	*	c
2 Bern	*	*

## Order of parameters:

1	$\Theta_1$	<displayed>
2	$\Theta_2$	<displayed>
3	$M_{1 \rightarrow 2}$	<displayed>

## Mutation rate among loci:

Mutation rate is constant for all loci

## Analysis strategy:

Bayesian inference

## Proposal distributions for parameter

Parameter	Proposal
Theta	Metropolis sampling
M	Metropolis sampling

## Prior distribution for parameter

Parameter	Prior	Minimum	Mean*	Maximum	Delta	Bins
Theta	Uniform	0.000000	0.050000	0.100000	0.010000	1500
Theta	Uniform	0.000000	0.050000	0.100000	0.010000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500

## Markov chain settings:

Long chain

Number of chains	1
Recorded steps [a]	5000
Increment (record every x step [b])	50
Number of concurrent chains (replicates) [c]	4
Visited (sampled) parameter values [a*b*c]	1000000
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:	twoswisstowns
Haplotyping is turned on:	NO
Output file:	outfile-twoswisstowns_c
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: twoswisstowns  
 Datatype: Haplotype data  
 Number of loci: 3

### Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Felsenstein 84	[Bf:0.26 0.27 0.22 0.25, t/t ratio=2.000]
1	2	Felsenstein 84	[Bf:0.24 0.27 0.22 0.27, t/t ratio=2.000]
2	1	Felsenstein 84	[Bf:0.26 0.23 0.25 0.25, t/t ratio=2.000]
3	1	Felsenstein 84	[Bf:0.26 0.24 0.24 0.26, t/t ratio=2.000]

### Sites per locus

Locus	Sites
1	500
2	300
3	700

### 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*

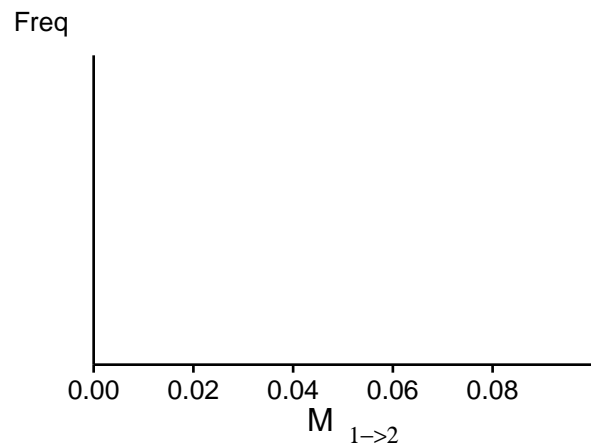
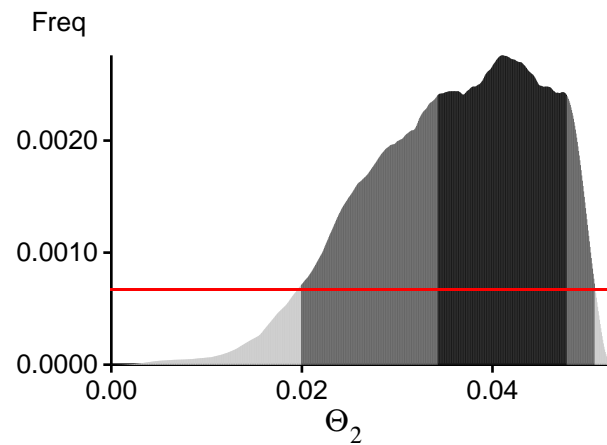
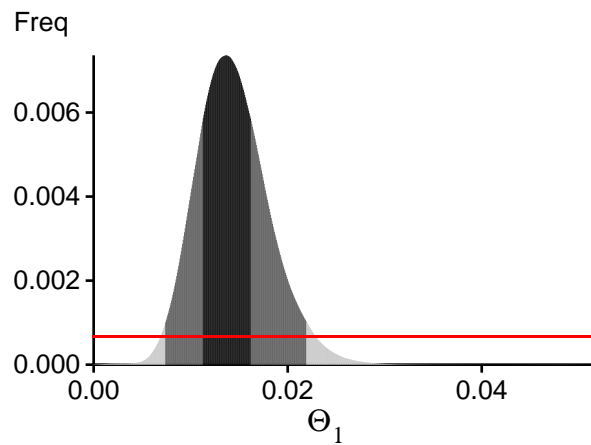
Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	$\Theta_1$	0.00160	0.00807	0.01077	0.01287	0.02633	0.01210	0.01297
1	$\Theta_2$	0.00960	0.01307	0.02543	0.03893	0.04860	0.02717	0.03920
1	$M_{1 \rightarrow 2}$	74.667	140.667	215.000	294.667	473.333	248.333	299.350
2	$\Theta_1$	0.00627	0.01280	0.01403	0.01520	0.02760	0.01623	0.01756
2	$\Theta_2$	0.01760	0.03700	0.04757	0.04893	0.05087	0.03717	0.06275
2	$M_{1 \rightarrow 2}$	110.667	196.000	231.667	293.333	466.000	261.000	273.536
3	$\Theta_1$	0.00673	0.01253	0.01597	0.01907	0.03140	0.01763	0.01878
3	$\Theta_2$	0.01407	0.03367	0.04290	0.04893	0.05080	0.03523	0.06169
3	$M_{1 \rightarrow 2}$	143.333	230.667	265.667	310.667	466.000	291.000	309.743
All	$\Theta_1$	0.00733	0.01120	0.01363	0.01620	0.02193	0.01417	0.01442
All	$\Theta_2$	0.01987	0.03420	0.04110	0.04780	0.05073	0.03703	0.03600
All	$M_{1 \rightarrow 2}$	0.000	0.000	0.333	1000.000	1000.000	0.333	nan

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	Raw thermodynamic score(1a)	Bezier approximation score(1b)	Harmonic mean(2)
1	-2561.88	-2285.05	-2255.96
2	-870.21	-777.79	-755.99
3	-2005.91	-1764.46	-1731.58

2346638528859811704183484516925440.00

(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 = -340282346638528859811704183484516925440.000000]

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
$\Theta_1$	181347/417545	0.43432
$\Theta_2$	285712/416759	0.68556
$M_{1 \rightarrow 2}$	187557/416548	0.45027
Genealogies	160586/1749148	0.09181

## *MCMC-Autocorrelation and Effective MCMC Sample Size*

Parameter	Autocorrelation	Effective Sample Size
$\Theta_1$	0.55717	29173.19
$\Theta_2$	0.24996	60797.59
$M_{1 \rightarrow 2}$	0.54963	31347.93
$\text{Ln}[\text{Prob}(D G)]$	0.70710	16834.16



## *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.

Param 4 (all loci): Upper prior boundary seems too low!

## *Summary Assignment of Individuals to Populations*

Individual	Population	
	1	2
?BAG	0.394	0.606
?BAJ	0.000	1.000
?BAH	0.802	0.198
?BAI	1.000	0.000
?BAF	0.091	0.909

## *Detailed Assignment of Individuals to Populations*

Individual	Locus	Population	
		1	2
?BAG	1	0.530	0.470
?BAG	2	0.000	1.000
?BAG	3	1.000	0.000
?BAG	All	0.394	0.606
?BAJ	1	0.412	0.588
?BAJ	2	0.000	1.000
?BAJ	3	0.831	0.169
?BAJ	All	0.000	1.000
?BAH	1	0.187	0.813
?BAH	2	0.781	0.219
?BAH	3	0.831	0.169
?BAH	All	0.802	0.198
?BAI	1	0.600	0.400
?BAI	2	0.646	0.354
?BAI	3	1.000	0.000
?BAI	All	1.000	0.000
?BAF	1	0.187	0.813
?BAF	2	0.081	0.919
?BAF	3	0.831	0.169
?BAF	All	0.091	0.909