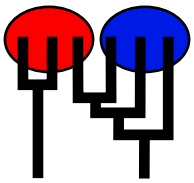
POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY
Bayesian inference using the structured coalescent
Migrate-n version 4.2.14 [April-16-2017]
Using Intel AVX (Advanced Vector Extensions)
Program started at Thu Apr 26 17:17:24 2018
Program finished at Thu Apr 26 17:23:46 2018 [Runtime:0000:00:06:22]



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]
Random number seed: (with internal timer) 1289200999
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	1	1	1	1
1 Romanshorn_0	*	*	*	*	*
1 Arbon_1	*	*	*	*	*
1 Kreuzlingen_2	*	*	*	*	*
1 Frauenfeld_3	*	*	*	*	*
1 Guendelhart_4	*	*	*	*	*

Order of parameters:

1 Θ_1 <displayed>

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

Analysis strategy: Bayesian inference
-Population size 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	00	Uniform	0.000000	0.010	0.100	0.010	1500	0.50000

[-1 -1 means priors were set globally]

Markov chain settings: Long chain
Number of chains 1
Recorded steps [a] 5000
Increment (record every x step [b] 100
Number of concurrent chains (replicates) [c] 1
Visited (sampled) parameter values [a*b*c] 500000
Number of discard trees per chain (burn-in) 500

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: infile1
Haplotyping is turned on: NO
Output file: outfile_0
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
Histogram of the frequency of migration events mighistfile_0

Data summary

Data file: infile1
Datatype: Sequence data
Number of loci: 2

Mutationmodel:

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

Sites per locus

Locus	Sites
1	1000
2	1000

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 Romanshorn_0	1	5
	2	5
1 Arbon_1	1	5
	2	5
1 Kreuzlingen_2	1	5
	2	5
1 Frauenfeld_3	1	5
	2	5
1 Guendelhart_4	1	5
	2	5
Total of all populations	1	25
	2	25

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00000	0.00000	0.00003	0.09993	0.09993	0.00003	0.09848
2	Θ_1	0.00000	0.00000	0.00003	0.09993	0.09993	0.00003	0.09639
All	Θ_1	0.00000	0.00000	0.00003	0.09993	0.09993	0.00003	0.09883

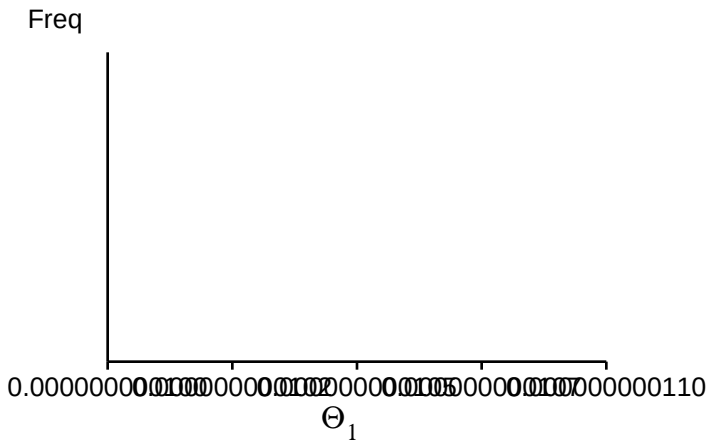
Citation suggestions:

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.

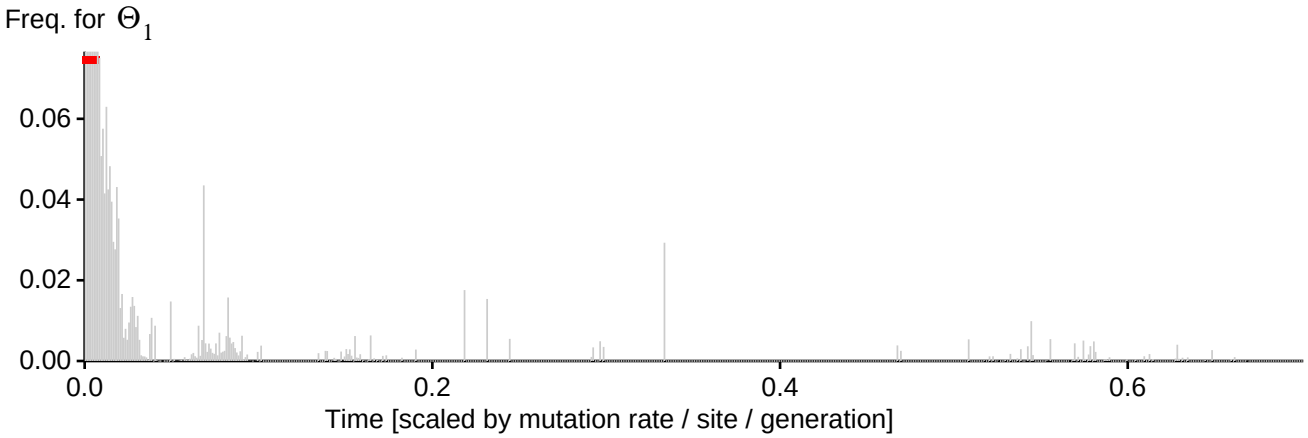
Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp. 42-79.

Bayesian Analysis: Posterior distribution over all loci



nan

Summary of events through time over all loci



Summary statistics of events through time

Locus 1

Population		Time			Frequency
From	To	Average	Median	Std	
1	1	0.050527	0.013500	0.111197	1.000000

Locus 2

Population		Time			Frequency
From	To	Average	Median	Std	
1	1	0.054230	0.007500	0.132683	1.000000

All loci

Population		Time			Frequency
From	To	Average	Median	Std	
1	1	0.052378	0.021000	0.122412	1.000000

Time and probability of location of most recent common ancestor

Locus 1

Population	Time			Frequency
	Average	Median	Std	
1	0.533276	0.544500	0.033322	1.000000

Locus 2

Population	Time			Frequency
	Average	Median	Std	
1	0.595362	0.581500	0.040041	1.000000

All loci

Population	Time			Frequency
	Average	Median	Std	
1	0.564319	1.126000	0.036835	1.000000

Legend for Skyline and Event plots

Skyline plots:

Skyline plots visualize the changes of population sizes and migration rates through time (today is on the left side and time is measured into the past. The time scale is in units of expected mutations per generation. To calculate the absolute time scale you must supply an mutation rate per year and the duration of a generation in years in the data option. You can calculate the absolute time by multiplying the scale by generation time times mutation rate per year (per site for DNA; per locus for all other datatypes).

With estimated mutation rate only the combined rate modifier is plotted. [this will change to mutation rate plot].

The gray bars cover 1.96 * approximate standard error (std in file skylinefile/number of observations in the bin) up and down from the expected value. The bar with different shades of gray on top of each plot indicates the number of values that were used to to calculate the expected value, white means there were very few and black means that there were many thousands of samples per bin.

On some plots one can see red squares below the grayscale bar, these suggest that either the upper quantile and/or the main value was higher than the visible part of the axis.

Event histograms:

All accepted events (migration events, coalescent events) are recorded and their frequency are shown as histograms over time with recent time on the left side. The frequency plots of populations with constant size and constant immigration rates show histograms that are similar to exponential distribution, if the populations come from a divergence model without migration then the frequency of migration events can show a peak in the past.

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	-17443.04	-9885.77	-8451.65
2	-17391.76	-9263.65	-7543.13
All	-34841.51	-19156.12	-16001.49

(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 = -6.710124]
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	306497/499899	0.61312
Genealogies	7124/500101	0.01425

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.49008	5174.75
Genealogies	0.99727	13.68

Potential Problems

This section reports potential problems with your run, but such reporting is often not very accurate. Whith many parameters in a multilocus analysi s, it is very common that some parameters for some loci will not be very informative, triggering suggestions (for example to increase the prior ran ge) that are not sensible. This suggestion tool will improve with time, therefore do not blindly follow its suggestions. If some parameters are fla gged, inspect the tables carefully and judge wether an action is required. For example, if you run a Bayesian inference with sequence data, for mac roscopic 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 rou tes 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.

Genealogies 2: Effective sample size of run seems too short!