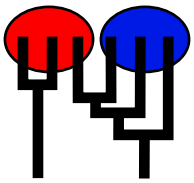
Japan

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 12:25:39 2025
Program finished at Sat May 3 12:30:10 2025 [Runtime:0000:00:04:31]



Options

Datatype: Single nucleotide polymorphism data(Hapmap formatting)

Inheritance multipliers in use for Thetas:
All loci use an inheritance multiplier of 1.0
Random number seed: (with internal timer) 79826771
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 Japan	*

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
Genealogy	Metropolis-Hastings

Prior distribution for parameter

Parameter			Prior	Minimum	Mean*	Maximum	Delta	Bins	UpdateFreq
1	Theta	* *	Uniform	0.000000	0.050	0.100	0.010000	1500	0.50000

[* * 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.hapmap
parmfile.hapmap
Haplotyping is turned on: NO
Output file: outfile_hapmap
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: infile.hapmap
Datatype: SNP data (Hapmap data)
Number of loci: 9

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Jukes-Cantor	[Basefreq: =0.25]
2	1	Jukes-Cantor	[Basefreq: =0.25]
3	1	Jukes-Cantor	[Basefreq: =0.25]
4	1	Jukes-Cantor	[Basefreq: =0.25]
5	1	Jukes-Cantor	[Basefreq: =0.25]
6	1	Jukes-Cantor	[Basefreq: =0.25]
7	1	Jukes-Cantor	[Basefreq: =0.25]
8	1	Jukes-Cantor	[Basefreq: =0.25]
9	1	Jukes-Cantor	[Basefreq: =0.25]

Sites per locus

Locus	Sites
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1

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
3	1	1	1.000	1.000	1.000
4	1	1	1.000	1.000	1.000
5	1	1	1.000	1.000	1.000
6	1	1	1.000	1.000	1.000
7	1	1	1.000	1.000	1.000
8	1	1	1.000	1.000	1.000

9	1	1	1.000	1.000	1.000

Population	Locus	Gene copies
1 Japan	1	72
	2	90
	3	88
	4	90
	5	90
	6	88
	7	90
	8	90
	9	90
Total of all populations	1	72
	2	90
	3	88
	4	90
	5	90
	6	88
	7	90
	8	90
	9	90

Allele frequency spectra

Locus 1

Allele	Pop1	All
G	0.542	0.542
C	0.458	0.458
	-	0.000
	-	0.000
Alleles	2	4
Samplesize	72	72
H_{exp}	0.497	0.497

Locus 2

Allele	Pop1	All
G	1.000	1.000
A	-	0.000
	-	0.000
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Locus 3

Allele	Pop1	All
C	1.000	1.000
A	-	0.000
	-	0.000
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Locus 4

Allele	Pop1	All
G	0.125	0.125
A	0.875	0.875
	-	0.000
	-	0.000

Allele	Pop1	All
Alleles	2	4
Samplesize	72	72
H_{exp}	0.219	0.219

Locus 5

Allele	Pop1	All
T	1.000	1.000
C	-	0.000
	-	0.000
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Locus 6

Allele	Pop1	All
C	1.000	1.000
T	-	0.000
	-	0.000
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Locus 7

Allele	Pop1	All
C	1.000	1.000
	-	0.000
	-	0.000
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Locus 8

Allele	Pop1	All
G	1.000	1.000
T	-	0.000
	-	0.000

Allele	Pop1	All
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Locus 9

Allele	Pop1	All
A	1.000	1.000
G	-	0.000
	-	0.000
	-	0.000
Alleles	1	4
Samplesize	72	72
H_{exp}	0.000	0.000

Average expected heterozygosity

	Pop1	All
H_{exp}	0.079	0.079

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.04427	0.06433	0.07523	0.08753	0.09993	0.07397	0.07296
2	Θ_1	0.03273	0.05293	0.06503	0.07827	0.09460	0.06337	0.06295
3	Θ_1	0.04900	0.08867	0.09610	0.09873	0.09993	0.08383	0.08096
4	Θ_1	0.04927	0.07367	0.08370	0.09247	0.09993	0.07810	0.07663
5	Θ_1	0.05213	0.07407	0.08263	0.08853	0.09887	0.07650	0.07491
6	Θ_1	0.05000	0.07193	0.07623	0.08753	0.09993	0.07723	0.07620
8	Θ_1	0.05933	0.09013	0.09530	0.09880	0.09993	0.07943	0.07904
9	Θ_1	0.05080	0.08200	0.09210	0.09627	0.09993	0.08050	0.07860
All	Θ_1	0.06773	0.07713	0.08257	0.08667	0.09700	0.08223	0.08219

Citation suggestions:

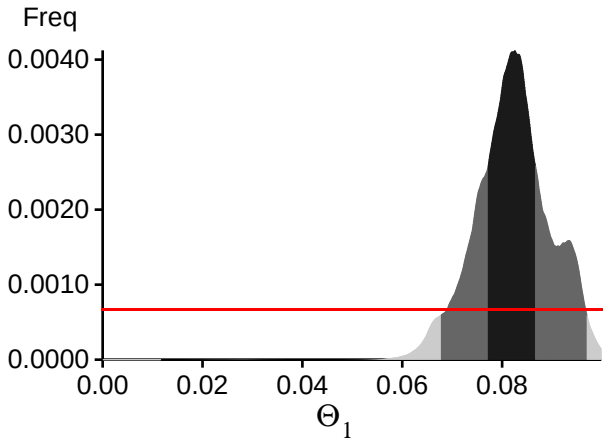
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341-345

Beerli, P., H. Ashki, S. Mashayekhi, and M. Palczewski, 2022. Population divergence time estimation using individual lineage label switching. G3 Genesâ Genomesâ 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 = 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	TI(1a)	BTI(1b)	HS(3)
1	-61.38	-39.68	-16.03
2	-8.20	-8.04	-8.18
3	-43.69	-36.15	-15.29
4	-34.27	-31.28	-16.07
5	-16.70	-16.29	-14.98
6	-16.67	-16.33	-14.97
8	-24.45	-23.01	-14.56
9	-28.86	-26.96	-15.67
All	-238.99	-202.50	-120.52

(1a) TI: Thermodynamic integration: log(Prob(D|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 = -4.763097]
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	2516714/3998514	0.62941
Genealogies	3342303/4001486	0.83527

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.78504	10388.93
Genealogies	0.78504	10388.93

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.

No warning was recorded during the run