

AUTO

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

Migrate-n version 4.2.8 [June-24-2016]

Using Intel AVX (Advanced Vector Extensions)

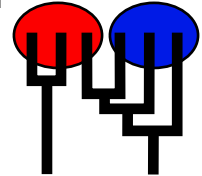
Compiled for PARALLEL computer architectures

One master and 4 compute nodes are available.

Compiled for a SYMMETRIC multiprocessors (Grandcentral)

Program started at Sun Aug 7 23:09:43 2016

Program finished at Mon Aug 8 00:10:34 2016 [Runtime:0000:01:00:51]



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) 1083285969

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	4
1 Romanshorn_0	*	*	0	t
2 Arbon_1	*	*	0	t
3 Kreuzlingen_2	0	0	*	0
4 ancestor	0	0	d	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	Θ_4	<displayed>
5	$M_{2 \rightarrow 1}$	<displayed>
6	$M_{1 \rightarrow 2}$	<displayed>
7	$\Delta_{4 \rightarrow 1}$	<displayed>
8	$\sigma_{4 \rightarrow 1}$	<displayed>
9	$\Delta_{4 \rightarrow 2}$	<displayed>
10	$\sigma_{4 \rightarrow 2}$	<displayed>
11	$\Delta_{3 \rightarrow 4}$	<displayed>
12	$\sigma_{3 \rightarrow 4}$	<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	Gamma	0.000000	0.020000	0.200000	0.020000	1500
Theta	Gamma	0.000000	0.020000	0.200000	0.020000	1500
Theta	Gamma	0.000000	0.020000	0.200000	0.020000	1500
Theta	Gamma	0.000000	0.020000	0.200000	0.020000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
M	Uniform	0.000000	500.000000	1000.000000	100.000000	1500
Splittime mean	Gamma	0.000000	0.100000	1.000000	0.100000	1500
Splittime std	Gamma	0.000000	1.000000	10.000000	1.000000	1500
Splittime mean	Gamma	0.000000	0.100000	1.000000	0.100000	1500

Splittime std	Gamma	0.000000	1.000000	10.000000	1.000000	1500
Splittime mean	Gamma	0.000000	0.100000	1.000000	0.100000	1500
Splittime std	Gamma	0.000000	1.000000	10.000000	1.000000	1500

Markov chain settings:

Number of chains	Long chain
Recorded steps [a]	1
Increment (record every x step [b])	10000
Number of concurrent chains (replicates) [c]	500
Visited (sampled) parameter values [a*b*c]	2
Number of discard trees per chain (burn-in)	10000000
	10000

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_2_0.25_10
Haplotyping is turned on:	NO
Output file:	outfile_2_0.25_10
Posterior distribution raw histogram file:	bayesfile
Raw data from the MCMC run:	bayesallfile_2_0.25_10
Print data:	No
Print genealogies [only some for some data type]:	None

Data summary

Data file: infile_2_0.25_10
 Datatype: Haplotype data
 Number of loci: 2

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Felsenstein 84	[Bf:0.25 0.25 0.25 0.26, t/t ratio=2.000]
2	1	Felsenstein 84	[Bf:0.23 0.26 0.25 0.26, 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
					data (missing)
1 Romanshorn_0			1	20	
			2	20	
2 Arbon_1			1	20	
			2	20	
3 Kreuzlingen_2			1	1	
			2	1	
4 ancestor			1	0	
			2	0	
Total of all populations			1	41	(0)
			2	41	(0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00253	0.00667	0.00940	0.01200	0.01813	0.00993	0.00940
1	Θ_2	0.00080	0.00453	0.00700	0.00933	0.01413	0.00740	0.00690
1	Θ_3	0.00120	0.00133	0.00553	0.01493	0.01587	0.01380	0.01620
1	Θ_4	0.00000	0.00027	0.00540	0.01587	0.05480	0.01553	0.01852
1	$M_{2 \rightarrow 1}$	0.000	12.667	71.667	188.000	284.000	139.667	169.476
1	$M_{1 \rightarrow 2}$	22.000	118.000	199.000	277.333	482.667	227.000	255.924
1	$D_{4 \rightarrow 1}$	0.00533	0.01333	0.03433	0.07933	0.10867	0.08900	0.09749
1	$S_{4 \rightarrow 1}$	0.00000	0.00000	0.19667	0.42000	1.82000	0.41667	0.51950
1	$D_{4 \rightarrow 2}$	0.00933	0.00933	0.03433	0.09333	0.09333	0.08900	0.09749
1	$S_{4 \rightarrow 2}$	0.00000	0.00000	0.19667	0.42000	1.82000	0.41667	0.51950
1	$D_{3 \rightarrow 4}$	0.00867	0.00867	0.02967	0.08267	0.08267	0.09633	0.10472
1	$S_{3 \rightarrow 4}$	0.00000	0.00667	0.20333	0.42000	1.78000	0.41000	0.52038
2	Θ_1	0.00120	0.00507	0.00753	0.00987	0.01480	0.00793	0.00800
2	Θ_2	0.00080	0.00467	0.00713	0.00933	0.01400	0.00740	0.00753
2	Θ_3	0.00000	0.00120	0.00553	0.01453	0.04800	0.01340	0.01710
2	Θ_4	0.00133	0.00133	0.00540	0.01307	0.01307	0.01540	0.01983
2	$M_{2 \rightarrow 1}$	9.333	15.333	58.333	133.333	156.000	135.000	174.233
2	$M_{1 \rightarrow 2}$	1.333	53.333	106.333	248.667	477.333	197.000	243.684
2	$D_{4 \rightarrow 1}$	0.00000	0.00800	0.03833	0.09600	0.12733	0.08900	0.10492
2	$S_{4 \rightarrow 1}$	0.00000	0.00667	0.19667	0.42000	1.79333	0.41000	0.55408
2	$D_{4 \rightarrow 2}$	0.00000	0.00800	0.03833	0.09600	0.26667	0.08900	0.10492
2	$S_{4 \rightarrow 2}$	0.00000	0.00667	0.19667	0.42000	1.79333	0.41000	0.55408
2	$D_{3 \rightarrow 4}$	0.01067	0.01067	0.02967	0.08000	0.08000	0.10033	0.11717
2	$S_{3 \rightarrow 4}$	0.00000	0.01333	0.21000	0.43333	1.79333	0.41667	0.56829
All	Θ_1	0.00253	0.00600	0.00833	0.01040	0.01427	0.00847	0.00848
All	Θ_2	0.00133	0.00467	0.00687	0.00880	0.01227	0.00700	0.00692
All	Θ_3	0.00000	0.00173	0.00567	0.01373	0.04400	0.01233	0.01578
All	Θ_4	0.00000	0.00040	0.00540	0.01600	0.05613	0.01553	0.02019
All	$M_{2 \rightarrow 1}$	18.667	28.667	61.667	114.667	141.333	101.667	123.742
All	$M_{1 \rightarrow 2}$	36.667	94.000	193.000	256.667	364.667	195.000	218.659
All	$D_{4 \rightarrow 1}$	0.01133	0.01133	0.06367	0.14933	0.14933	0.10100	0.11449
All	$S_{4 \rightarrow 1}$	0.00000	0.02000	0.16333	0.28667	1.17333	0.26333	0.34322
All	$D_{4 \rightarrow 2}$	0.01133	0.01133	0.06367	0.14867	0.14867	0.10100	0.11441
All	$S_{4 \rightarrow 2}$	0.00000	0.02000	0.16333	0.28667	1.17333	0.26333	0.34318

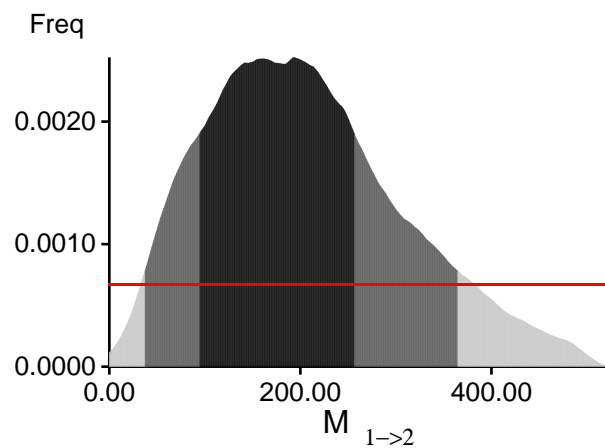
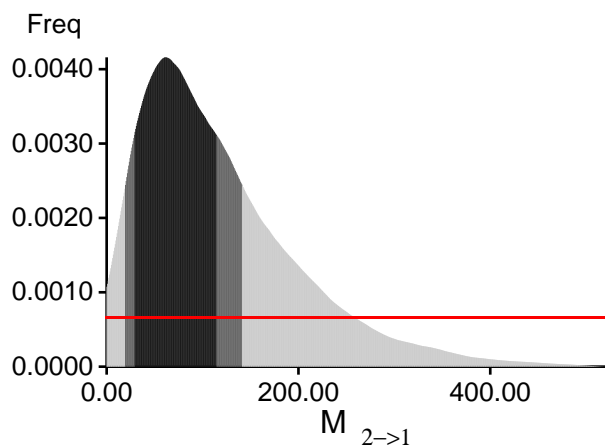
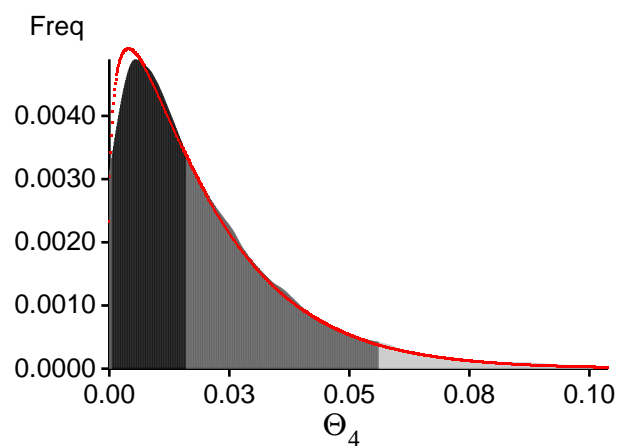
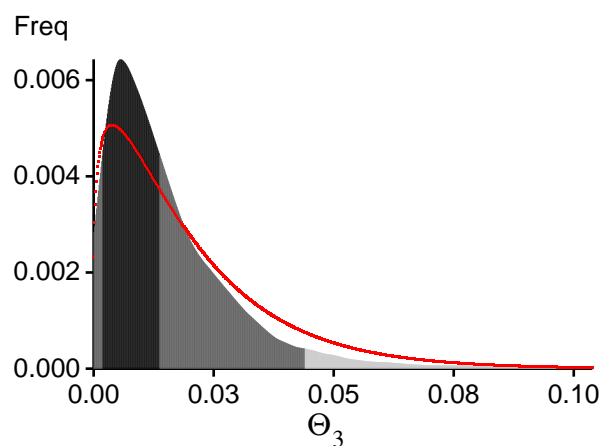
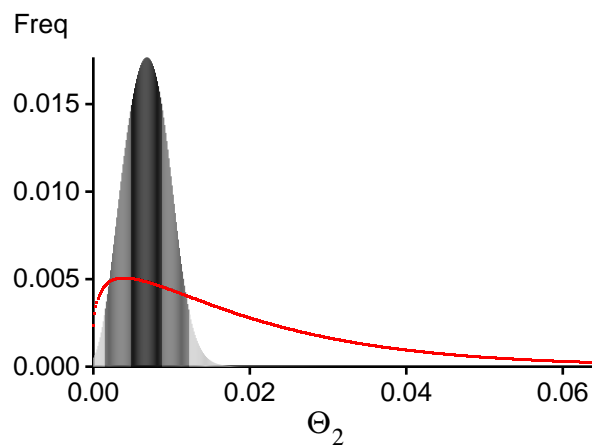
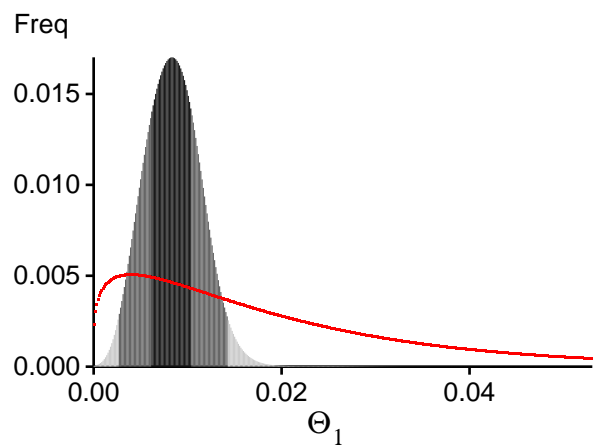
Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
All	D _{3->4}	0.01333	0.10867	0.15367	0.17667	0.18800	0.12433	0.13046
All	S _{3->4}	0.00000	0.04000	0.18333	0.31333	1.16667	0.27667	0.35927

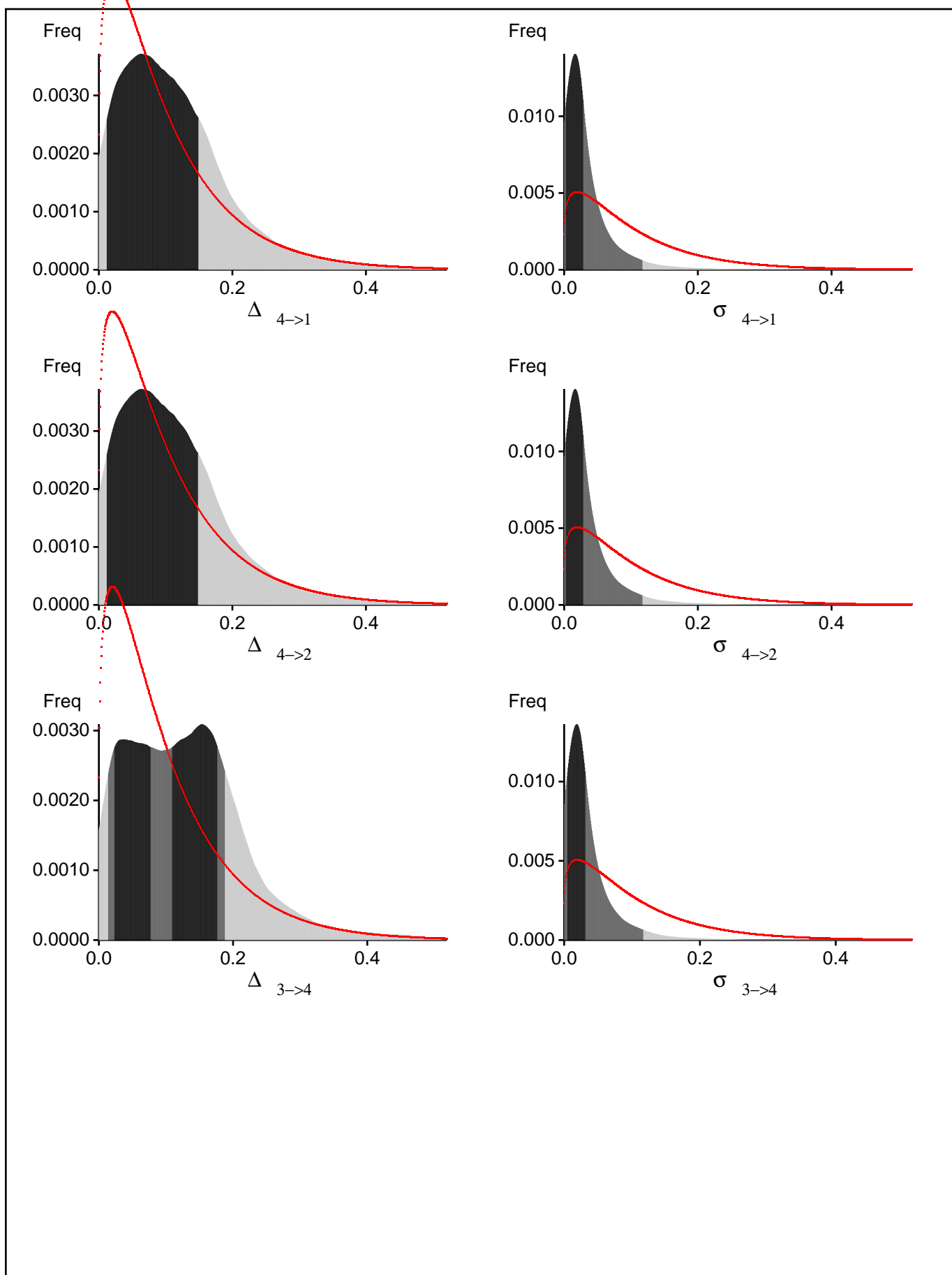
Citation suggestions:

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	-3300.51	-2910.47	-2862.14
2	-3101.37	-2813.49	-2779.33
All	-6434.53	-5756.60	-5674.11

(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 = -32.644864]

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	180165/833536	0.21615
Θ_2	164654/833206	0.19761
Θ_3	554556/832957	0.66577
Θ_4	830738/833221	0.99702
$M_{2 \rightarrow 1}$	463197/833453	0.55576
$M_{1 \rightarrow 2}$	416166/833579	0.49925
$\Delta_{4 \rightarrow 1}$	719221/833554	0.86284
$\sigma_{4 \rightarrow 1}$	498451/831504	0.59946
$\Delta_{4 \rightarrow 2}$	717022/831815	0.86200
$\sigma_{4 \rightarrow 2}$	498824/832890	0.59891
$\Delta_{3 \rightarrow 4}$	710547/832549	0.85346
$\sigma_{3 \rightarrow 4}$	503549/832657	0.60475
Genealogies	1225027/10005079	0.12244

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.28742	21805.07
Θ_2	0.29819	21713.63
Θ_3	0.12099	31529.10
Θ_4	-0.00572	39498.53
$M_{2 \rightarrow 1}$	0.84158	4675.64
$M_{1 \rightarrow 2}$	0.67118	6394.01
$\Delta_{4 \rightarrow 1}$	0.01766	38629.23
$\sigma_{4 \rightarrow 1}$	-0.00188	39536.97
$\Delta_{4 \rightarrow 2}$	0.01766	38629.23
$\sigma_{4 \rightarrow 2}$	-0.00188	39536.97
$\Delta_{3 \rightarrow 4}$	0.00282	39717.32
$\sigma_{3 \rightarrow 4}$	0.00825	39552.65
$\text{Ln}[\text{Prob}(\text{D} \text{G})]$	0.44167	15154.23

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.

No warning was recorded during the run