Current protocols example dataset

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY
Bayesian inference using the structured coalescent
Migrate-n version 4.4.3(git:) [March-21-2019]
  Using Intel AVX (Advanced Vector Extensions)
  Compiled for PARALLEL computer architectures
  One master and 8 compute nodes are available.
  Compiled for a SYMMETRIC multiprocessors (Grandcentral)
Program started at   Mon Jun 3 09:46:14 2019
Program finished at   Mon Jun 3 10:02:12 2019 [Runtime:0000:00:15:58]

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)   967906974
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
1 Arbon_1	*	0	0
2 Berg_2	D	*	0
3 Chur_3	0	D	*

  Order of parameters:

1	Θ_1	\<displayed\>
2	Θ_2	\<displayed\>
3	Θ_3	\<displayed\>
4	$M_{1\rightarrow2}$	\<displayed\>
5	$M_{2\rightarrow3}$	\<displayed\>
6	$\Delta_{1\rightarrow2}$	\<displayed\>
7	$\sigma_{1\rightarrow2}$	\<displayed\>
8	$\Delta_{2\rightarrow3}$	\<displayed\>
9	$\sigma_{2\rightarrow3}$	\<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
 -Divergence time estimation: Normal Distribution Shortcut (mean and standard dev.)

 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	**	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
2	Theta	**	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
3	Theta	**	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
4	M	**	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
5	M	**	Uniform	0.000000	2500.	5000.	500.0	1500	0.05556
6	Splittime mean	**	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
7	Splittime std	**	Uniform	0.000000	0.050	0.100	0.100	1500	0.05556
8	Splittime mean	**	Uniform	0.000000	0.050	0.100	0.010	1500	0.05556
9	Splittime std	**	Uniform	0.000000	0.050	0.100	0.100	1500	0.05556

[* * 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)	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:	infile
Haplotyping is turned on:	NO
Output file:	outfile_model2
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: infile
Datatype: Haplotype data
Number of loci: 10

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]
10	1	Jukes-Cantor	[Basefreq: =0.25]

Sites per locus

Locus	Sites
1	1000
2	1000
3	1000
4	1000
5	1000
6	1000
7	1000
8	1000
9	1000
10	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
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
10	1	1	1.000	1.000	1.000

Population	Locus	Gene copies	
		data	(missing)
1 Arbon_1	1	10	
	2	10	
	3	10	
	4	10	
	5	10	
	6	10	
	7	10	
	8	10	
	9	10	
	10	10	
2 Berg_2	1	10	
	2	10	
	3	10	
	4	10	
	5	10	
	6	10	
	7	10	
	8	10	
	9	10	
	10	10	
3 Chur_3	1	10	
	2	10	
	3	10	
	4	10	
	5	10	
	6	10	
	7	10	
	8	10	
	9	10	
	10	10	
Total of all populations	1	30	(0)
	2	30	(0)
	3	30	(0)
	4	30	(0)
	5	30	(0)
	6	30	(0)
	7	30	(0)
	8	30	(0)
	9	30	(0)

10   30   (0)

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.00440	0.00660	0.00763	0.00973	0.02153	0.01217	0.01538
1	Θ_2	0.00093	0.00193	0.00263	0.00447	0.01340	0.00550	0.00690
1	Θ_3	0.00440	0.01167	0.01177	0.01207	0.02580	0.01477	0.01662
1	$M_{1->2}$	0.000	10.000	45.000	110.000	756.667	275.000	393.661
1	$M_{2->3}$	0.000	0.000	5.000	30.000	140.000	31.667	46.552
1	$D_{1->2}$	0.00300	0.00300	0.00323	0.00327	0.00367	0.04330	0.04546
1	$S_{1->2}$	0.02980	0.02987	0.03017	0.03040	0.03040	0.05310	0.05292
1	$D_{2->3}$	0.00747	0.00860	0.00870	0.00880	0.00887	0.04297	0.04499
1	$S_{2->3}$	0.08380	0.08380	0.08397	0.08420	0.08460	0.05563	0.05465
2	Θ_1	0.00220	0.00673	0.00683	0.00727	0.01607	0.00890	0.00984
2	Θ_2	0.00087	0.00247	0.00297	0.00453	0.01007	0.00443	0.00520
2	Θ_3	0.00140	0.00287	0.00397	0.00520	0.01127	0.00550	0.00635
2	$M_{1->2}$	0.000	10.000	88.333	173.333	770.000	268.333	342.367
2	$M_{2->3}$	0.000	0.000	11.667	50.000	323.333	58.333	136.696
2	$D_{1->2}$	0.00273	0.00273	0.00290	0.00293	0.00440	0.03657	0.04100
2	$S_{1->2}$	0.00053	0.00053	0.00063	0.00080	0.00087	0.05383	0.05294
2	$D_{2->3}$	0.05020	0.05053	0.05063	0.05100	0.05140	0.04363	0.04616
2	$S_{2->3}$	0.03120	0.03127	0.03137	0.03140	0.03140	0.05537	0.05432
3	Θ_1	0.00233	0.00467	0.00530	0.00593	0.01387	0.00777	0.00843
3	Θ_2	0.00140	0.00473	0.00543	0.00793	0.02373	0.01143	0.01714
3	Θ_3	0.00520	0.01060	0.01210	0.01240	0.02327	0.01403	0.01646
3	$M_{1->2}$	0.000	30.000	38.333	136.667	706.667	298.333	387.846
3	$M_{2->3}$	0.000	0.000	1.667	76.667	233.333	35.000	97.577
3	$D_{1->2}$	0.00000	0.00000	0.00017	0.00300	0.02500	0.00917	0.01820
3	$S_{1->2}$	0.00020	0.00047	0.00057	0.00073	0.00133	0.02910	0.03806
3	$D_{2->3}$	0.00567	0.00593	0.00603	0.00607	0.00627	0.04630	0.04779
3	$S_{2->3}$	0.05207	0.05327	0.05357	0.05373	0.05373	0.05683	0.05527
4	Θ_1	0.00200	0.00347	0.00590	0.00907	0.01027	0.00930	0.01470
4	Θ_2	0.00120	0.00340	0.00490	0.00593	0.01860	0.00777	0.01036
4	Θ_3	0.00180	0.00553	0.00597	0.00620	0.01787	0.00930	0.01122
4	$M_{1->2}$	0.000	0.000	21.667	56.667	533.333	115.000	202.131
4	$M_{2->3}$	0.000	0.000	1.667	46.667	260.000	61.667	99.809
4	$D_{1->2}$	0.00000	0.00000	0.00017	0.00393	0.01553	0.00950	0.02008

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
4	$S_{1->2}$	0.00233	0.00287	0.00317	0.00320	0.01613	0.03223	0.03902
4	$D_{2->3}$	0.00080	0.00080	0.00103	0.00107	0.00227	0.03523	0.04003
4	$S_{2->3}$	0.07553	0.07553	0.07577	0.07580	0.07987	0.05790	0.05629
5	Θ_1	0.00133	0.00247	0.00350	0.00593	0.01220	0.00557	0.00794
5	Θ_2	0.00100	0.00160	0.00283	0.00593	0.02113	0.00743	0.01299
5	Θ_3	0.00800	0.01787	0.01797	0.01807	0.03560	0.02230	0.02460
5	$M_{1->2}$	63.333	140.000	145.000	146.667	1296.667	868.333	1071.698
5	$M_{2->3}$	0.000	0.000	1.667	26.667	143.333	28.333	46.212
5	$D_{1->2}$	0.00000	0.00527	0.00537	0.00553	0.00553	0.03410	0.03984
5	$S_{1->2}$	0.00007	0.00020	0.00030	0.00033	0.00033	0.05017	0.04934
5	$D_{2->3}$	0.01547	0.01553	0.01563	0.01573	0.01587	0.05077	0.05099
5	$S_{2->3}$	0.06447	0.06453	0.06463	0.06467	0.06607	0.05437	0.05359
6	Θ_1	0.00200	0.00473	0.00523	0.00527	0.01020	0.00563	0.00610
6	Θ_2	0.00080	0.00320	0.00437	0.00727	0.01747	0.00830	0.01466
6	Θ_3	0.02240	0.02973	0.03010	0.03040	0.04667	0.03817	0.04140
6	$M_{1->2}$	66.667	543.333	551.667	553.333	1310.000	775.000	886.942
6	$M_{2->3}$	0.000	0.000	5.000	20.000	63.333	21.667	23.835
6	$D_{1->2}$	0.00000	0.00000	0.00110	0.00253	0.01840	0.02610	0.03391
6	$S_{1->2}$	0.00600	0.00687	0.00697	0.00700	0.00700	0.04943	0.05032
6	$D_{2->3}$	0.01707	0.01813	0.01823	0.01827	0.01860	0.05223	0.05251
6	$S_{2->3}$	0.00500	0.00507	0.00517	0.00520	0.00540	0.05063	0.05047
7	Θ_1	0.00413	0.00900	0.00930	0.00947	0.02013	0.01143	0.01281
7	Θ_2	0.00233	0.00640	0.00657	0.00727	0.02020	0.01297	0.01908
7	Θ_3	0.00580	0.01527	0.01603	0.01647	0.03160	0.01750	0.02160
7	$M_{1->2}$	0.000	0.000	35.000	120.000	376.667	168.333	207.315
7	$M_{2->3}$	0.000	0.000	5.000	43.333	280.000	88.333	268.394
7	$D_{1->2}$	0.00000	0.00000	0.00057	0.00393	0.01467	0.00717	0.01507
7	$S_{1->2}$	0.00053	0.00067	0.00083	0.00093	0.00107	0.02717	0.03618
7	$D_{2->3}$	0.00000	0.00280	0.00330	0.00360	0.00367	0.04177	0.04443
7	$S_{2->3}$	0.05400	0.05440	0.05450	0.05453	0.05513	0.05270	0.05194
8	Θ_1	0.00293	0.00507	0.00637	0.00640	0.01380	0.00757	0.00866
8	Θ_2	0.00233	0.00500	0.00570	0.00640	0.01867	0.01063	0.01385
8	Θ_3	0.00413	0.00747	0.00863	0.00907	0.02167	0.01230	0.01408
8	$M_{1->2}$	0.000	3.333	8.333	80.000	566.667	201.667	283.191
8	$M_{2->3}$	0.000	0.000	1.667	26.667	153.333	28.333	49.161
8	$D_{1->2}$	0.00000	0.00000	0.00003	0.00407	0.02640	0.01717	0.02555
8	$S_{1->2}$	0.00067	0.00067	0.00077	0.00080	0.00080	0.05070	0.05048

| Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$D_{2->3}$	0.00733	0.00780	0.00790	0.00793	0.00807	0.04750	0.04791
8	$S_{2->3}$	0.09793	0.09913	0.09923	0.09933	0.09993	0.05890	0.05723
9	Θ_1	0.00433	0.01133	0.01143	0.01173	0.02233	0.01270	0.01430
9	Θ_2	0.00093	0.00167	0.00297	0.00440	0.01307	0.00550	0.00672
9	Θ_3	0.00773	0.01853	0.01863	0.01993	0.03067	0.02543	0.03229
9	$M_{1->2}$	0.000	0.000	8.333	36.667	306.667	38.333	91.338
9	$M_{2->3}$	0.000	0.000	5.000	83.333	323.333	101.667	132.685
9	$D_{1->2}$	0.00000	0.00107	0.00143	0.00180	0.03113	0.01797	0.02558
9	$S_{1->2}$	0.00060	0.00060	0.00070	0.00073	0.00073	0.05563	0.05585
9	$D_{2->3}$	0.00000	0.00000	0.00210	0.00420	0.00813	0.02690	0.03446
9	$S_{2->3}$	0.00073	0.00280	0.00350	0.00520	0.00587	0.04497	0.04548
10	Θ_1	0.00247	0.00467	0.00543	0.00593	0.01207	0.00683	0.00834
10	Θ_2	0.00253	0.00593	0.00670	0.00747	0.02167	0.01117	0.01501
10	Θ_3	0.00693	0.01307	0.01323	0.01473	0.02973	0.01837	0.02085
10	$M_{1->2}$	0.000	0.000	25.000	110.000	696.667	195.000	338.985
10	$M_{2->3}$	0.000	0.000	1.667	30.000	220.000	51.667	122.416
10	$D_{1->2}$	0.00000	0.00000	0.00077	0.00427	0.01667	0.00677	0.01762
10	$S_{1->2}$	0.00020	0.00047	0.00077	0.00087	0.00200	0.02350	0.03366
10	$D_{2->3}$	0.00027	0.00027	0.00043	0.00047	0.00073	0.04297	0.04544
10	$S_{2->3}$	0.07307	0.07307	0.07330	0.07333	0.07500	0.05643	0.05525
All	Θ_1	0.00487	0.00587	0.00657	0.00727	0.00873	0.00677	0.00676
All	Θ_2	0.00313	0.00420	0.00490	0.00567	0.00740	0.00517	0.00519
All	Θ_3	0.00953	0.01187	0.01243	0.01367	0.01713	0.01323	0.01323
All	$M_{1->2}$	10.000	60.000	88.333	113.333	170.000	95.000	93.094
All	$M_{2->3}$	0.000	0.000	1.667	10.000	36.667	11.667	12.719
All	$D_{1->2}$	0.00000	0.00000	0.00003	0.00073	0.00273	0.00077	0.00096
All	$S_{1->2}$	0.00473	0.00707	0.00837	0.00960	0.02713	0.01110	0.01363
All	$D_{2->3}$	0.00000	0.00000	0.00150	0.01320	0.02127	0.01037	0.01483
All	$S_{2->3}$	0.02033	0.04087	0.04657	0.04833	0.04893	0.03990	0.03751

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
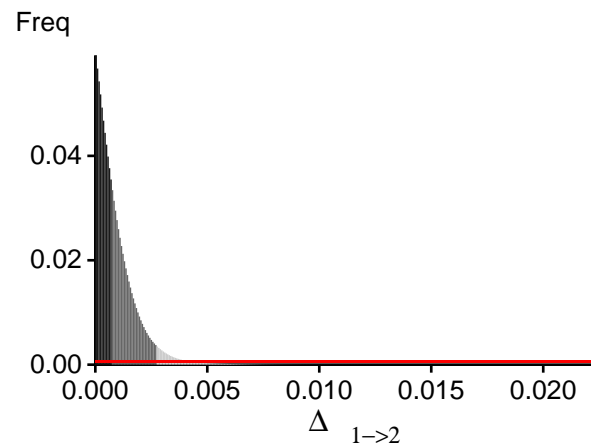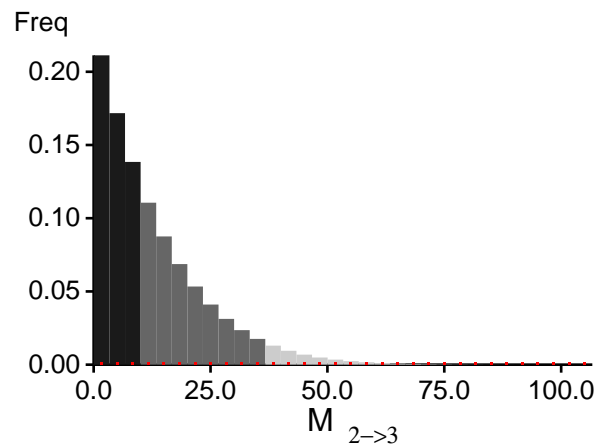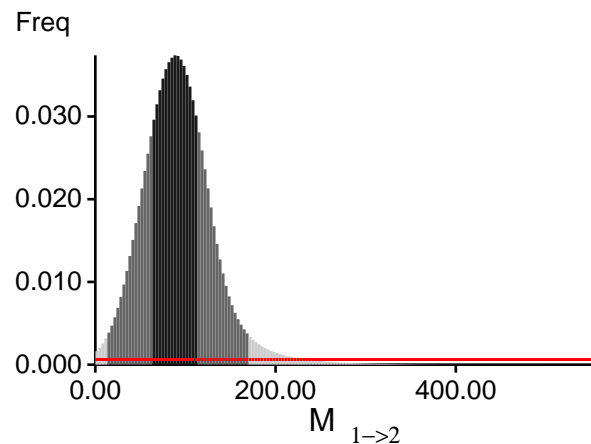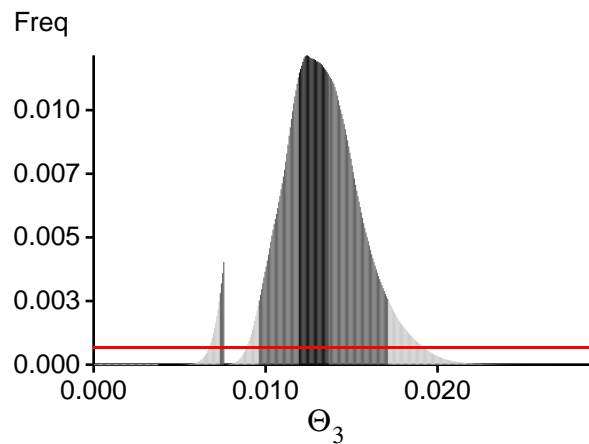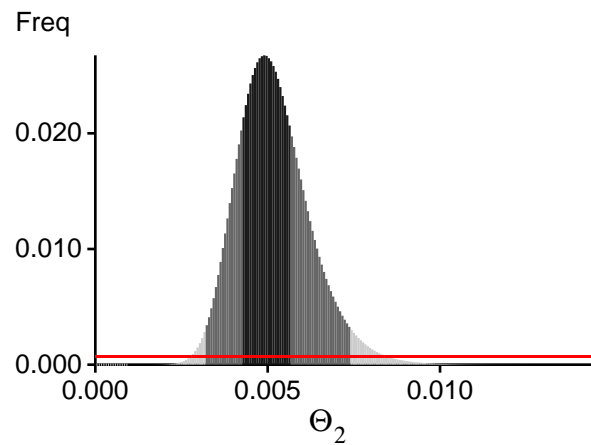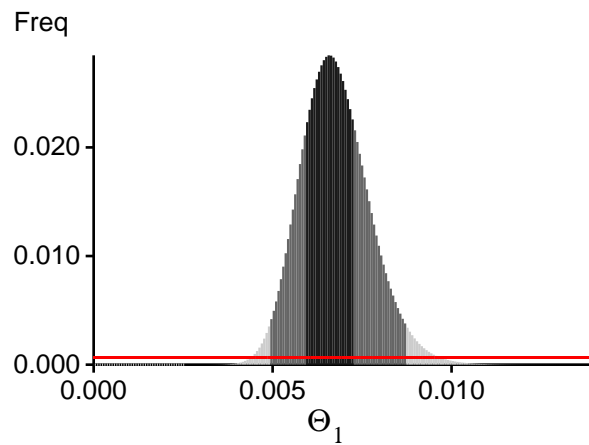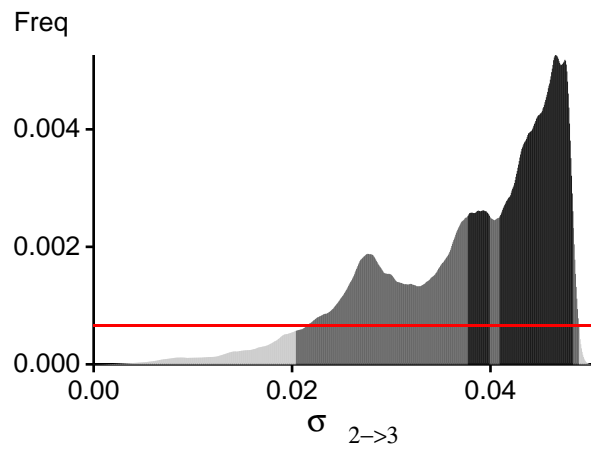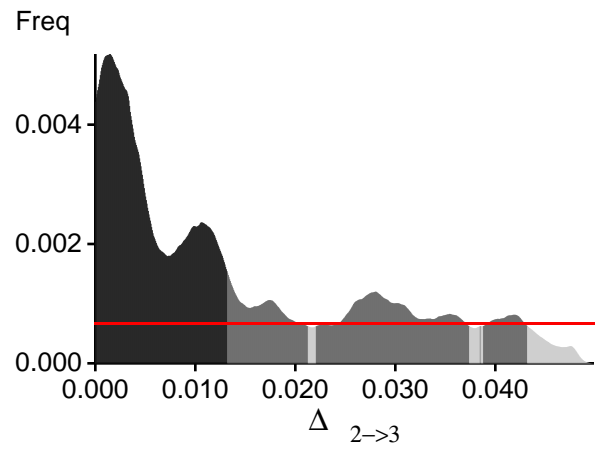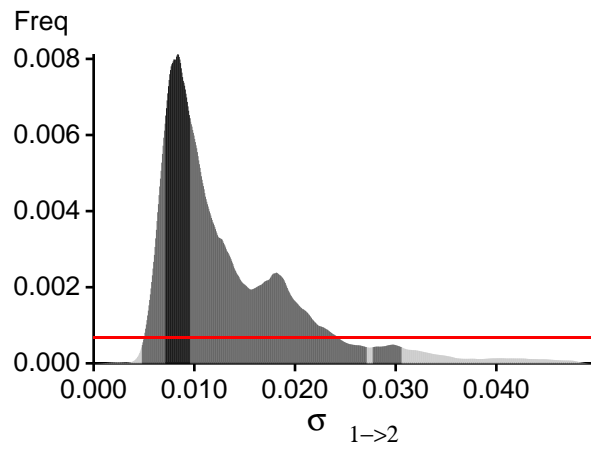
Log-Probability of the data given the model (marginal likelihood)

Use this value for Bayes factor calculations:
BF = 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	-3178.28	-2425.17	-2294.45
2	-2648.52	-2065.69	-1960.43
3	-2822.60	-2217.62	-2117.16
4	-3315.23	-2399.92	-2236.83
5	-2941.33	-2291.64	-2179.13
6	-2922.09	-2389.35	-2300.13
7	-2916.79	-2304.35	-2201.36
8	-3209.99	-2374.57	-2227.80
9	-3249.05	-2418.66	-2275.20
10	-2879.86	-2264.32	-2156.97
All	-30014.64	-23082.19	-21880.35

(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 = 69.107640]
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	117322/278167	0.42177
Θ_2	128291/277311	0.46262
Θ_3	121408/278449	0.43602
$M_{1\rightarrow2}$	116659/278338	0.41913
$M_{2\rightarrow3}$	157708/277833	0.56764
$\Delta_{1\rightarrow2}$	186528/277605	0.67192
$\sigma_{1\rightarrow2}$	195962/277719	0.70561
$\Delta_{2\rightarrow3}$	238240/278673	0.85491
$\sigma_{2\rightarrow3}$	236393/277693	0.85127
Genealogies	380771/2498212	0.15242

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.67166	15811.72
Θ_2	0.62880	17499.99
Θ_3	0.46738	24439.30
$M_{1 \to 2}$	0.67686	13804.62
$M_{2 \to 3}$	0.81895	7171.94
$\Delta_{1 \to 2}$	0.45450	26765.87
$\sigma_{1 \to 2}$	0.46794	26310.06
$\Delta_{2 \to 3}$	0.26755	40660.38
$\sigma_{2 \to 3}$	0.24313	42653.26
Genealogies	0.24313	42653.26

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.

Param 13 (Locus 8): Upper prior boundary seems too low!