Zika virus

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY
Bayesian inference using the structured coalescent
Migrate-n version 4.2.8 [June-24-2016]
  Compiled for PARALLEL computer architectures
  One master and 4 compute nodes are available.
  Compiled for a SYMMETRIC multiprocessors (Grandcentral)
Program started at   Sat Jul 16 21:41:26 2016
Program finished at   Sat Jul 16 21:51:17 2016 [runtime: Runtime: 0: 0: 9:51]

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) 3306427103
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	3	3
1 Africa	*	0	0	0	0
2 China	d	*	0	0	0
3 Brazil	0	d	*	*	*
3 Mexico_and_neig	0	d	*	*	*
3 Puerto_Rico	0	d	*	*	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	Θ_3	<displayed>
4	$\Delta_{1->2}$	<displayed>
5	$\sigma_{1->2}$	<displayed>
6	$\Delta_{2->3}$	<displayed>
7	$\sigma_{2->3}$	<displayed>

Mutation rate among loci: Mutation rate is constant

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
Theta	Uniform	0.000000	0.050000	0.100000	0.010000	1500
M	Uniform	0.000000	2500.000000	5000.000000	500.000000	1500
M	Uniform	0.000000	2500.000000	5000.000000	500.000000	1500
M	Uniform	0.000000	2500.000000	5000.000000	500.000000	1500
M	Uniform	0.000000	2500.000000	5000.000000	500.000000	1500
M	Uniform	0.000000	2500.000000	5000.000000	500.000000	1500
M	Uniform	0.000000	2500.000000	5000.000000	500.000000	1500
Splittime mean	Uniform	0.000000	0.200000	0.400000	0.010000	1500
Splittime std	Uniform	0.000000	0.200000	0.400000	0.010000	1500
Splittime mean	Uniform	0.000000	0.200000	0.400000	0.010000	1500
Splittime std	Uniform	0.000000	0.200000	0.400000	0.010000	1500

Markov chain settings:	Long chain
Number of chains	1
Recorded steps [a]	9000
Increment (record every x step [b]	100
Number of concurrent chains (replicates) [c]	2
Visited (sampled) parameter values [a*b*c]	1800000
Number of discard trees per chain (burn-in)	1000

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_eastsplam
Posterior distribution raw histogram file:	bayesfile
Raw data from the MCMC run:	bayesallfile_eastsplam
Print data:	No
Print genealogies [only some for some data type]:	None

Data summary

Data file:	infile
Datatype:	Haplotype data
Number of loci:	1

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
1	1	Felsenstein 84	[Bf:0.27 0.22 0.29 0.21, t/t ratio=2.000]

Sites per locus

Locus	Sites
1	10269

Site rate variation and probabilities:

Locus	Sublocus	Region type	Rate of change	Probability	Patch size
1	1	1	0.428	0.680	1.000
1	1	2	2.081	0.308	1.000
1	1	3	5.491	0.012	1.000

Population	Locus	Gene copies	
		data	(missing)
1 Africa	1	5	
2 China	1	12	
3 Brazil	1	13	
3 Mexico_and_neighbors	1	8	
3 Puerto_Rico	1	2	
Total of all populations	1	40	(0)

Bayesian Analysis: Posterior distribution table

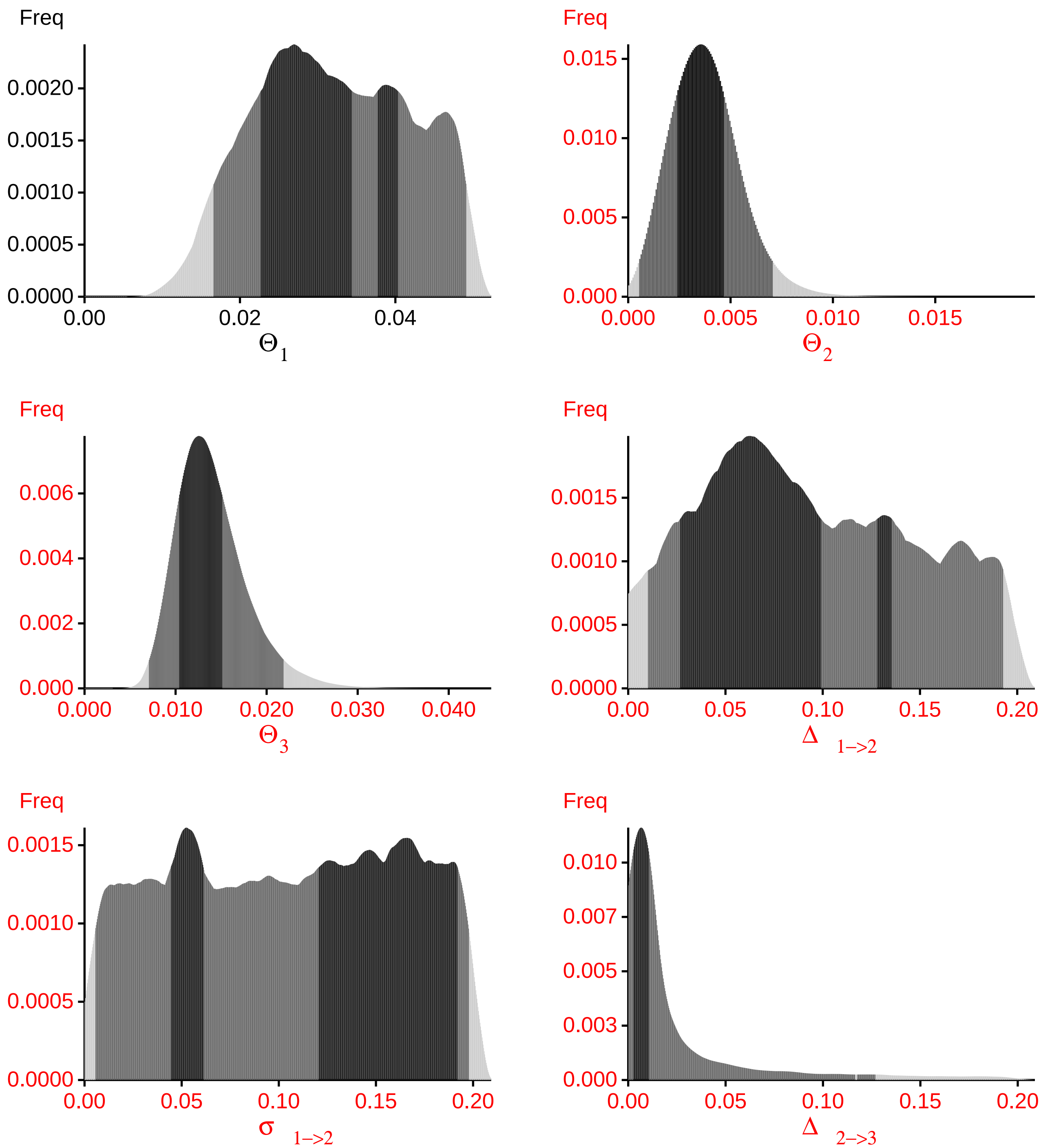
Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	0.01653	0.02260	0.02697	0.03440	0.04913	0.03183	0.04295
1	Θ_2	0.00047	0.00233	0.00357	0.00467	0.00707	0.00370	0.00376
1	Θ_3	0.00700	0.01033	0.01257	0.01513	0.02187	0.01350	0.01398
1	$D_{1->2}$	0.00987	0.02640	0.06227	0.09920	0.19280	0.08813	0.16230
1	$S_{1->2}$	0.00533	0.04427	0.05240	0.06133	0.19787	0.10547	0.19271
1	$D_{2->3}$	0.00000	0.00240	0.00680	0.01067	0.11653	0.01267	0.04588
1	$S_{2->3}$	0.00000	0.00800	0.01107	0.01387	0.13867	0.01667	0.06551

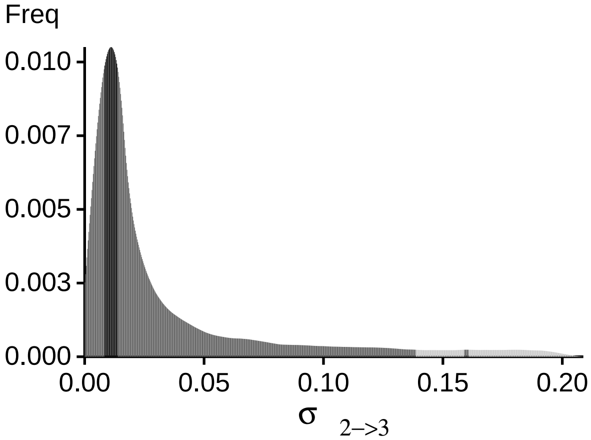
Citation suggestions:

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 for locus 1




Freq
0.010
0.007
0.005
0.003
0.000
0.00
0.05
0.10
0.15
0.20
σ 2->3

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]

Method	ln(Prob(D|Model))	Notes
Thermodynamic integration	-28151.903579	(1a)
	-25971.506703	(1b)
Harmonic mean	-25685.279502	(2)

(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

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	68518/127989	0.53534
Θ_2	45010/128324	0.35075
Θ_3	31301/128114	0.24432
$\Delta_{1->2}$	100126/128851	0.77707
$\sigma_{1->2}$	92200/128707	0.71636
$\Delta_{2->3}$	36750/128525	0.28594
$\sigma_{2->3}$	34246/128464	0.26658
Genealogies	38963/901026	0.04324

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.22360	12526.05
Θ_2	0.42810	9165.04
Θ_3	0.32972	10445.21
$\Delta_{1->2}$	0.28479	10845.47
$\sigma_{1->2}$	0.29155	10209.57
$\Delta_{2->3}$	0.65873	4192.40
$\sigma_{2->3}$	0.77594	2711.21
Ln[Prob(D|G)]	0.99711	35.54

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.

No warning was recorded during the run