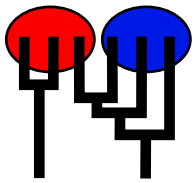
Example: Microsatellite data set

MIGRATION RATE AND POPULATION SIZE ESTIMATION
using the coalescent and maximum likelihood or Bayesian inference
Migrate-n version 3.0.7 [1547]
  Compiled for a PARALLEL COMPUTER ARCHITECTURE
  One master and 10 compute nodes are available.
Program started at Mon Jun 8 16:49:33 2009
Program finished at Mon Jun 8 16:49:35 2009



Options

Datatype:	Microsatellite data [Brownian motion]
Missing data:	not included
Random number seed:	(from parmfile) 310705631
Start parameters:	
Theta values were generated	RANDOM start value from U(min,msx)
M values were generated	from the FST-calculation

Connection type matrix:
  where m = average (average over a group of Thetas or M,
  s = symmetric M, S = symmetric 4Nm, 0 = zero, and not estimated,
  * = free to vary, Thetas are on diagonal

Population 1
  1 population__num *

  Order of parameters:
    1 Θ_1 <displayed>

Mutation rate among loci:	Mutation rate is constant for all loci
Analysis strategy:	Bayesian inference

Proposal distributions for parameter

Parameter	Proposal
Theta	Slice sampling
M	Slice sampling

Prior distribution for parameter

Parameter	Prior	Minimum	Mean*	Maximum	Delta	Bins
Theta	Uniform	0.000000	10.000000	20.000000	2.000000	500
M	Uniform	0.000000	10.000000	20.000000	2.000000	500

Markov chain settings: Long chain

Number of chains 1

Recorded steps [a] 160

Increment (record every x step [b] 2

Number of concurrent chains (replicates) [c] 2

Visited (sampled) parameter values [a*b*c] 640

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.msat1

Output file: outfile-bayes

Posterior distribution raw histogram file: bayesfile

Print data: No

Print genealogies [only some for some data type]: None

Data summary

Datatype: Microsatellite data
Number of loci: 10

Population	Locus	Gene copies data	Gene copies (missing)
1 population__number___0	1	92	(0)
	2	92	(0)
	3	92	(0)
	4	92	(0)
	5	92	(0)
	6	92	(0)
	7	92	(0)
	8	92	(0)
	9	92	(0)
	10	92	(0)
Total of all populations	1	92	(0)
	2	92	(0)
	3	92	(0)
	4	92	(0)
	5	92	(0)
	6	92	(0)
	7	92	(0)
	8	92	(0)
	9	92	(0)
	10	92	(0)

Allele frequency spectra

Locus 1

Allele	Pop1	All
16	0.196	0.196
19	0.054	0.054
18	0.087	0.087
15	0.130	0.130
21	0.087	0.087
23	0.065	0.065
17	0.196	0.196
22	0.087	0.087
25	0.043	0.043
24	0.011	0.011
26	0.011	0.011
27	0.022	0.022
29	0.011	0.011

Locus 2

Allele	Pop1	All
16	0.543	0.543
19	0.022	0.022
18	0.174	0.174
17	0.163	0.163
15	0.011	0.011
21	0.043	0.043
20	0.022	0.022
22	0.022	0.022

Locus 3

Allele	Pop1	All
19	0.250	0.250
20	0.370	0.370
18	0.087	0.087
21	0.207	0.207
22	0.087	0.087

Locus 4

Allele	Pop1	All

Allele	Pop1	All
16	0.076	0.076
24	0.109	0.109
15	0.033	0.033
25	0.163	0.163
14	0.033	0.033
19	0.120	0.120
12	0.033	0.033
20	0.130	0.130
23	0.087	0.087
28	0.011	0.011
22	0.043	0.043
21	0.141	0.141
13	0.011	0.011
26	0.011	0.011

Locus 5

Allele	Pop1	All
20	0.457	0.457
21	0.391	0.391
19	0.152	0.152

Locus 6

Allele	Pop1	All
19	0.033	0.033
20	0.065	0.065
18	0.261	0.261
22	0.163	0.163
21	0.283	0.283
16	0.033	0.033
24	0.109	0.109
17	0.054	0.054

Locus 7

Allele	Pop1	All
23	0.130	0.130
20	0.424	0.424
22	0.185	0.185
21	0.207	0.207
19	0.054	0.054

Locus 8

Allele	Pop1	All
19	0.522	0.522
17	0.043	0.043
18	0.087	0.087
20	0.163	0.163
16	0.043	0.043
22	0.076	0.076
15	0.033	0.033
23	0.033	0.033

Locus 9

Allele	Pop1	All
24	0.054	0.054
19	0.359	0.359
20	0.239	0.239
23	0.163	0.163
22	0.054	0.054
18	0.043	0.043
21	0.065	0.065
25	0.022	0.022

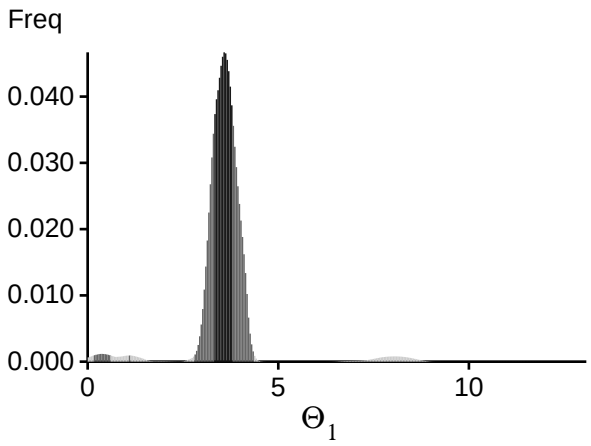
Locus 10

Allele	Pop1	All
22	0.152	0.152
20	0.337	0.337
23	0.120	0.120
24	0.011	0.011
19	0.163	0.163
21	0.054	0.054
18	0.043	0.043
15	0.043	0.043
17	0.043	0.043
25	0.033	0.033

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	1.60000	1.88000	2.34000	2.84000	3.28000	6.30000	5.09656
2	Θ_1	0.44000	0.52000	0.94000	1.28000	1.40000	5.90000	2.93101
3	Θ_1	0.00000	0.00000	0.18000	0.52000	0.56000	6.38000	0.49365
4	Θ_1	2.52000	2.92000	3.38000	4.00000	4.52000	4.90000	3.89325
5	Θ_1	2.88000	3.24000	3.82000	4.36000	4.96000	5.30000	1.53574
6	Θ_1	2.60000	2.80000	3.30000	3.88000	4.36000	5.82000	3.53977
7	Θ_1	2.56000	2.92000	3.46000	3.84000	4.48000	4.74000	1.49064
8	Θ_1	0.84000	0.92000	1.38000	1.80000	1.92000	5.42000	0.84496
9	Θ_1	0.12000	0.24000	0.54000	0.80000	0.92000	1.10000	0.35526
10	Θ_1	3.12000	3.52000	4.42000	4.92000	5.28000	7.02000	1.89188
All	Θ_1	2.76000	3.28000	3.58000	3.80000	4.36000	3.62000	3.60198

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	-5429.28	-959.93	-49.73
2	-4056.26	-810.97	-39.44
3	-600.00	-268.83	-41.04
4	-3057.13	-672.36	-49.84
5	-240.44	-194.02	-40.17
6	-1914.30	-481.16	-30.96
7	-1115.46	-341.52	-27.28
8	-1603.10	-422.11	-29.19
9	-883.24	-310.57	-37.03
10	-1520.19	-417.56	-36.57
All	-20419.40	-4879.03	-381.27

(1a, 1b and 2) is an approximation to the marginal likelihood, make sure the program run long enough!
(1a, 1b) and (2) should give a similar result, (2) is considered more
crude than (1), but (1) needs heating with several well-spaced chains,
(1b) is using a Bezier-curve to get better approximations for runs with low number
of heated chains

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	3172/3172	1.00000
Genealogies	767/3268	0.23470

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.29338	1712.15
Ln[Prob(D|G)]	0.94975	101.58