Example: Microsatellite data set

MIGRATION RATE AND POPULATION SIZE ESTIMATION
using the coalescent and maximum likelihood or Bayesian inference
Migrate-n version 3.2.4 [1772]
  Compiled for a PARALLEL COMPUTER ARCHITECTURE
  One master and 2 compute nodes are available.
  Compiled for a SYMMETRIC MULTIPROCESSORS
Program started at   Fri Nov 26 18:44:49 2010
Program finished at   Fri Nov 26 18:47:44 2010
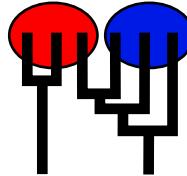
Options

Datatype: Microsatellite data [Brownian motion]
Missing data: not included

Inheritance scalers in use for Thetas: 1.00   1.00
1.00   1.00   1.00   1.00   1.00
1.00   1.00   1.00
[Each Theta uses the (true) ineritance scalar of the first locus as a reference]
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
 1 population__num   *   *
 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]	500
Increment (record every x step [b]	200
Number of concurrent chains (replicates) [c]	2
Visited (sampled) parameter values [a*b*c]	200000
Number of discard trees per chain (burn-in)	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.msat
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	(missing)
1 population__number___0	1	50	(0)
	2	50	(0)
	3	50	(0)
	4	50	(0)
	5	50	(0)
	6	50	(0)
	7	50	(0)
	8	50	(0)
	9	50	(0)
	10	50	(0)
1 population__number___1	1	42	(0)
	2	42	(0)
	3	42	(0)
	4	42	(0)
	5	42	(0)
	6	42	(0)
	7	42	(0)
	8	42	(0)
	9	42	(0)
	10	42	(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	Pop2	All
16	0.196	0.000	0.098
19	0.054	0.000	0.027
18	0.087	0.000	0.043
15	0.130	0.000	0.065
21	0.087	0.000	0.043
23	0.065	0.000	0.033
17	0.196	0.000	0.098
22	0.087	0.000	0.043
25	0.043	0.000	0.022
24	0.011	0.000	0.005
26	0.011	0.000	0.005
27	0.022	0.000	0.011
29	0.011	0.000	0.005

Locus 2

Allele	Pop1	Pop2	All
16	0.543	0.000	0.272
19	0.022	0.000	0.011
18	0.174	0.000	0.087
17	0.163	0.000	0.082
15	0.011	0.000	0.005
21	0.043	0.000	0.022
20	0.022	0.000	0.011
22	0.022	0.000	0.011

Locus 3

Allele	Pop1	Pop2	All
19	0.250	0.000	0.125
20	0.370	0.000	0.185
18	0.087	0.000	0.043
21	0.207	0.000	0.103
22	0.087	0.000	0.043

Locus 4

Allele	Pop1	Pop2	All

Allele	Pop1	Pop2	All
16	0.076	0.000	0.038
24	0.109	0.000	0.054
15	0.033	0.000	0.016
25	0.163	0.000	0.082
14	0.033	0.000	0.016
19	0.120	0.000	0.060
12	0.033	0.000	0.016
20	0.130	0.000	0.065
23	0.087	0.000	0.043
28	0.011	0.000	0.005
22	0.043	0.000	0.022
21	0.141	0.000	0.071
13	0.011	0.000	0.005
26	0.011	0.000	0.005

Locus 5

Allele	Pop1	Pop2	All
20	0.457	0.000	0.228
21	0.391	0.000	0.196
19	0.152	0.000	0.076

Locus 6

Allele	Pop1	Pop2	All
19	0.033	0.000	0.016
20	0.065	0.000	0.033
18	0.261	0.000	0.130
22	0.163	0.000	0.082
21	0.283	0.000	0.141
16	0.033	0.000	0.016
24	0.109	0.000	0.054
17	0.054	0.000	0.027

Locus 7

Allele	Pop1	Pop2	All
23	0.130	0.000	0.065
20	0.424	0.000	0.212
22	0.185	0.000	0.092
21	0.207	0.000	0.103
19	0.054	0.000	0.027

Locus 8

Allele	Pop1	Pop2	All
19	0.522	0.000	0.261
17	0.043	0.000	0.022
18	0.087	0.000	0.043
20	0.163	0.000	0.082
16	0.043	0.000	0.022
22	0.076	0.000	0.038
15	0.033	0.000	0.016
23	0.033	0.000	0.016

Locus 9

Allele	Pop1	Pop2	All
24	0.054	0.000	0.027
19	0.359	0.000	0.179
20	0.239	0.000	0.120
23	0.163	0.000	0.082
22	0.054	0.000	0.027
18	0.043	0.000	0.022
21	0.065	0.000	0.033
25	0.022	0.000	0.011

Locus 10

Allele	Pop1	Pop2	All
22	0.152	0.000	0.076
20	0.337	0.000	0.168
23	0.120	0.000	0.060
24	0.011	0.000	0.005
19	0.163	0.000	0.082
21	0.054	0.000	0.027
18	0.043	0.000	0.022
15	0.043	0.000	0.022
17	0.043	0.000	0.022
25	0.033	0.000	0.016

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	11.76000	18.08000	19.62000	19.96000	20.00000	17.54000	16.32212
2	Θ_1	3.68000	4.84000	6.22000	8.08000	14.80000	7.90000	8.78449
3	Θ_1	1.08000	3.88000	5.98000	6.56000	11.32000	5.82000	1.91158
4	Θ_1	10.64000	17.24000	19.70000	20.00000	20.00000	17.34000	4.69047
5	Θ_1	0.36000	0.56000	0.98000	1.84000	6.96000	2.90000	0.60244
6	Θ_1	4.24000	6.44000	9.22000	10.68000	16.72000	9.18000	1.70238
7	Θ_1	1.72000	4.36000	6.06000	7.80000	12.28000	6.70000	0.95899
8	Θ_1	3.64000	4.56000	5.98000	7.32000	16.04000	10.26000	1.38630
9	Θ_1	3.40000	7.56000	9.98000	11.04000	14.00000	9.78000	1.07352
10	Θ_1	7.68000	16.92000	19.22000	19.68000	20.00000	13.46000	1.37594
All	Θ_1	3.72000	6.80000	7.50000	7.92000	10.72000	7.58000	7.67416

Bayesian Analysis: Posterior distribution over all loci
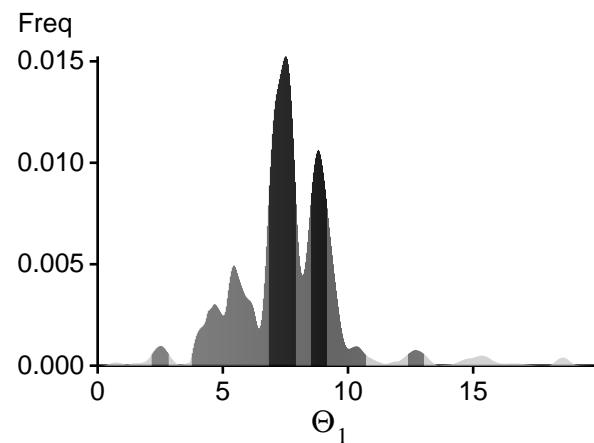
Log-Probability of the data given the model (marginal likelihood)

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	-16055.74	-2685.50	-101.48
2	-3621.55	-662.99	-72.10
3	-2501.45	-488.78	-81.28
4	-17573.15	-2927.30	-97.81
5	-1102.64	-242.21	-47.92
6	-7697.22	-1322.38	-80.55
7	-3159.38	-591.23	-75.77
8	-3582.20	-664.09	-83.72
9	-5475.79	-970.84	-84.49
10	-6609.30	-1162.11	-101.68
All	-67464.92	-11803.94	-913.32

(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
[Scaling factor = -86.512263

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	5505505/5505505	1.00000
Genealogies	1425865/5015290	0.28430

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.88946	3832.67
Ln[Prob(D|G)]	0.93313	3832.67

Warnings

You should most likely rerun your analysis after improving run parameters.

Param 1 (Locus 1): Upper prior boundary seems too low!
Param 1 (Locus 4): Upper prior boundary seems too low!
Param 1 (Locus 10): Upper prior boundary seems too low!