# *Example: Microsatellite data set*

MIGRATION RATE AND POPULATION SIZE ESTIMATION
using the coalescent and maximum likelihood or Bayesian inference
Migrate-n version 3.2.7 [1793]
  Compiled for a PARALLEL COMPUTER ARCHITECTURE
  One master and 2 compute nodes are available.
  Compiled for a SYMMETRIC MULTIPROCESSORS
Program started at   Wed Jan 12 16:51:18 2011
Program finished at   Wed Jan 12 16:51:55 2011
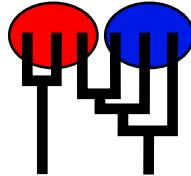
Options

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]   20
   Number of concurrent chains (replicates) [c]   2
   Visited (sampled) parameter values [a*b*c]   20000
   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	9.24000	10.44000	11.78000	13.56000	15.72000	13.50000	12.51716
2	Θ_1	2.80000	3.80000	5.98000	7.12000	13.24000	6.58000	7.33761
3	Θ_1	4.92000	7.04000	7.90000	9.20000	11.28000	8.14000	2.41386
4	Θ_1	0.72000	1.04000	2.02000	3.00000	3.32000	4.82000	1.78259
5	Θ_1	3.36000	4.32000	4.94000	5.84000	8.04000	5.34000	1.01963
6	Θ_1	2.60000	4.40000	5.18000	6.16000	7.56000	5.22000	0.95042
7	Θ_1	3.76000	5.24000	5.94000	6.92000	8.08000	6.02000	0.79825
8	Θ_1	3.00000	4.12000	4.74000	5.32000	6.56000	4.82000	0.64158
9	Θ_1	0.48000	0.68000	1.26000	1.80000	2.12000	5.26000	0.53929
10	Θ_1	4.00000	7.84000	10.22000	11.84000	13.16000	10.42000	1.14241
All	Θ_1	5.20000	5.80000	6.14000	6.40000	6.84000	6.14000	6.06364

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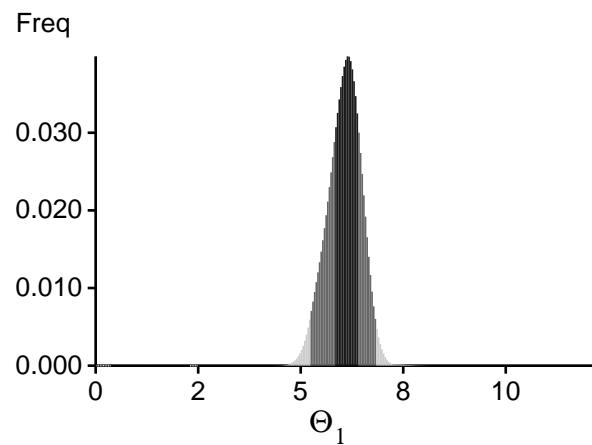
Log-Probability of the data given the model (marginal likelihood)

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	-12144.04	-2047.98	-95.26
2	-3324.29	-610.17	-72.93
3	-1935.16	-391.12	-55.87
4	-9647.05	-1658.33	-91.96
5	-1399.59	-285.80	-42.86
6	-5159.64	-902.95	-56.91
7	-3662.34	-665.00	-50.84
8	-2864.08	-535.03	-42.83
9	-4425.61	-774.89	-63.73
10	-5289.34	-940.22	-84.80
All	-49930.44	-8890.78	-737.30

(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 = -79.311089

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	1001210/1001210	1.00000
Genealogies	133670/501030	0.26679

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sampe Size
Θ_1	0.49501	10616.00
Ln[Prob(D|G)]	0.92430	1027.32