

Example: Microsatellite data set

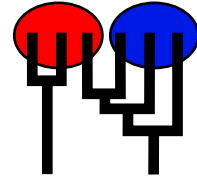
MIGRATION RATE AND POPULATION SIZE ESTIMATION

using the coalescent and maximum likelihood or Bayesian inference

Migrate-n version 3.0.6

Program started at Thu May 14 10:47:29 2009

Program finished at Thu May 14 10:51:44 2009



Options

Datatype: Microsatellite data [Brownian motion]
 Missing data: not included
 Random number seed: (with internal timer) 174874791
 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	2
1 population__num	*	*
2 population__num	*	*

Order of parameters:

1	Θ_1	<displayed>
2	Θ_2	<displayed>
3	$M_{2 \rightarrow 1}$	<displayed>
4	$M_{1 \rightarrow 2}$	<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])	100
Number of concurrent chains (replicates) [c]	2
Visited (sampled) parameter values [a*b*c]	32000
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)
2 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	0.220	0.167	0.193
19	0.040	0.071	0.056
18	0.060	0.119	0.090
15	0.220	0.024	0.122
21	0.020	0.167	0.093
23	0.020	0.119	0.070
17	0.280	0.095	0.188
22	0.060	0.119	0.090
25	0.060	0.024	0.042
24	0.020	0.000	0.010
26	0.000	0.024	0.012
27	0.000	0.048	0.024
29	0.000	0.024	0.012

Locus 2

Allele	Pop1	Pop2	All
16	0.520	0.571	0.546
19	0.040	0.000	0.020
18	0.220	0.119	0.170
17	0.160	0.167	0.163
15	0.020	0.000	0.010
21	0.020	0.071	0.046
20	0.020	0.024	0.022
22	0.000	0.048	0.024

Locus 3

Allele	Pop1	Pop2	All
19	0.240	0.262	0.251
20	0.280	0.476	0.378
18	0.080	0.095	0.088
21	0.280	0.119	0.200
22	0.120	0.048	0.084

Locus 4

Allele	Pop1	Pop2	All
--------	------	------	-----

Allele	Pop1	Pop2	All
16	0.080	0.071	0.076
24	0.180	0.024	0.102
15	0.020	0.048	0.034
25	0.160	0.167	0.163
14	0.020	0.048	0.034
19	0.100	0.143	0.121
12	0.060	0.000	0.030
20	0.080	0.190	0.135
23	0.060	0.119	0.090
28	0.020	0.000	0.010
22	0.060	0.024	0.042
21	0.160	0.119	0.140
13	0.000	0.024	0.012
26	0.000	0.024	0.012

Locus 5

Allele	Pop1	Pop2	All
20	0.400	0.524	0.462
21	0.420	0.357	0.389
19	0.180	0.119	0.150

Locus 6

Allele	Pop1	Pop2	All
19	0.060	0.000	0.030
20	0.100	0.024	0.062
18	0.300	0.214	0.257
22	0.200	0.119	0.160
21	0.120	0.476	0.298
16	0.060	0.000	0.030
24	0.160	0.048	0.104
17	0.000	0.119	0.060

Locus 7

Allele	Pop1	Pop2	All
23	0.040	0.238	0.139
20	0.660	0.143	0.401
22	0.180	0.190	0.185
21	0.100	0.333	0.217
19	0.020	0.095	0.058

Locus 8

Allele	Pop1	Pop2	All
19	0.520	0.524	0.522
17	0.040	0.048	0.044
18	0.100	0.071	0.086
20	0.140	0.190	0.165
16	0.080	0.000	0.040
22	0.100	0.048	0.074
15	0.020	0.048	0.034
23	0.000	0.071	0.036

Locus 9

Allele	Pop1	Pop2	All
24	0.080	0.024	0.052
19	0.300	0.429	0.364
20	0.300	0.167	0.233
23	0.180	0.143	0.161
22	0.080	0.024	0.052
18	0.020	0.071	0.046
21	0.040	0.095	0.068
25	0.000	0.048	0.024

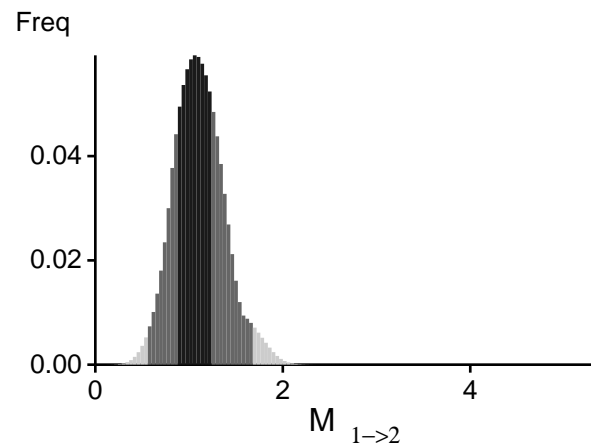
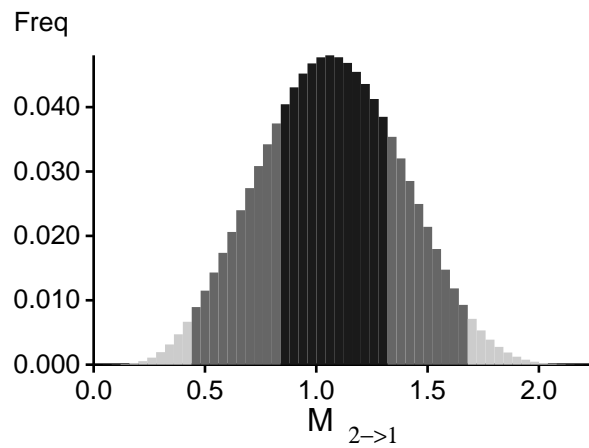
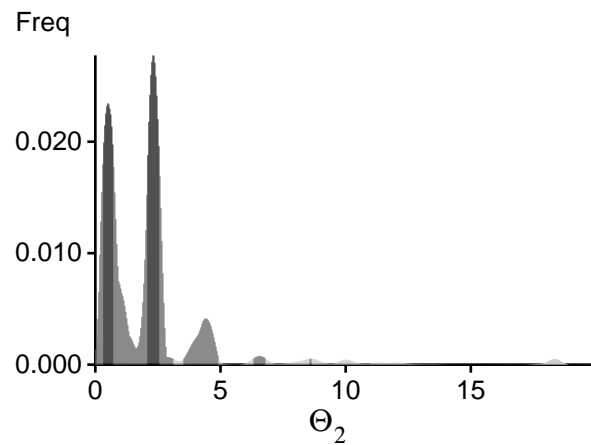
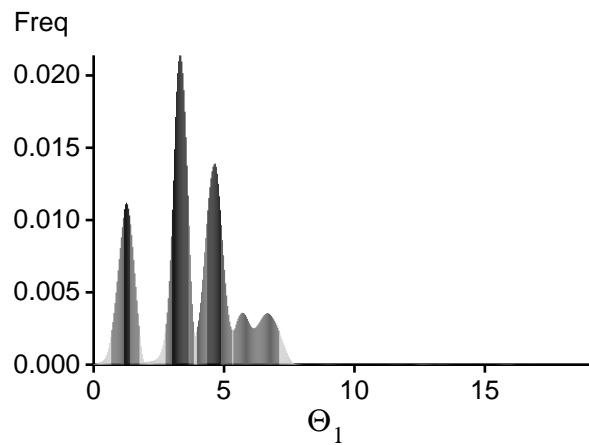
Locus 10

Allele	Pop1	Pop2	All
22	0.100	0.214	0.157
20	0.440	0.214	0.327
23	0.080	0.167	0.123
24	0.020	0.000	0.010
19	0.160	0.167	0.163
21	0.060	0.048	0.054
18	0.080	0.000	0.040
15	0.020	0.071	0.046
17	0.040	0.048	0.044
25	0.000	0.071	0.036

Bayesian Analysis: Posterior distribution table

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	Θ_1	2.92000	5.12000	6.54000	7.08000	9.60000	6.22000	6.37143
1	Θ_2	7.64000	12.20000	12.82000	13.28000	20.00000	13.34000	12.74655
1	$M_{2 \rightarrow 1}$	0.200	0.680	1.220	1.400	1.880	1.100	1.064
1	$M_{1 \rightarrow 2}$	0.000	0.040	0.340	0.680	2.320	0.940	0.989
2	Θ_1	2.00000	4.28000	5.90000	6.32000	7.80000	5.14000	2.53746
2	Θ_2	1.40000	1.84000	2.54000	4.12000	5.76000	3.94000	3.35140
2	$M_{2 \rightarrow 1}$	0.320	0.680	0.980	1.280	2.040	1.140	0.579
2	$M_{1 \rightarrow 2}$	1.560	2.120	2.740	3.200	6.000	3.340	1.762
3	Θ_1	4.04000	5.00000	7.10000	7.60000	14.84000	9.10000	3.29339
3	Θ_2	0.64000	1.12000	1.58000	2.00000	10.56000	5.50000	1.85571
3	$M_{2 \rightarrow 1}$	0.120	1.600	2.140	2.840	3.280	1.980	0.612
3	$M_{1 \rightarrow 2}$	0.360	1.440	2.820	3.200	5.680	2.700	0.986
4	Θ_1	8.92000	10.12000	11.10000	11.64000	17.32000	13.22000	3.42245
4	Θ_2	6.96000	8.72000	9.14000	9.48000	12.32000	12.14000	3.22863
4	$M_{2 \rightarrow 1}$	0.000	0.120	0.380	0.600	1.720	0.860	0.212
4	$M_{1 \rightarrow 2}$	0.480	1.000	1.340	1.640	2.320	1.420	0.354
5	Θ_1	0.56000	0.80000	1.70000	3.24000	9.32000	3.66000	0.98051
5	Θ_2	0.24000	0.36000	0.94000	1.60000	1.80000	1.66000	1.42074
5	$M_{2 \rightarrow 1}$	0.000	1.480	2.700	3.520	4.480	7.140	1.444
5	$M_{1 \rightarrow 2}$	2.520	3.560	4.300	5.320	10.360	4.780	1.024
6	Θ_1	2.52000	3.08000	4.34000	5.24000	8.28000	10.34000	1.69012
6	Θ_2	2.28000	6.00000	6.46000	7.60000	12.68000	7.22000	1.24396
6	$M_{2 \rightarrow 1}$	0.280	0.720	1.180	1.520	2.280	1.260	0.206
6	$M_{1 \rightarrow 2}$	0.400	0.800	1.260	1.600	3.000	1.420	0.249
7	Θ_1	0.40000	0.60000	1.82000	2.36000	7.04000	2.94000	0.48604
7	Θ_2	3.12000	4.68000	6.26000	7.00000	11.04000	6.66000	1.04660
7	$M_{2 \rightarrow 1}$	0.000	0.280	0.620	1.040	2.920	0.900	0.163
7	$M_{1 \rightarrow 2}$	0.200	0.440	1.020	1.800	2.520	1.700	0.414
8	Θ_1	1.84000	2.04000	3.34000	4.28000	13.04000	5.10000	0.87159

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	Θ_2	5.64000	9.44000	10.06000	11.04000	18.96000	11.42000	1.39802
8	$M_{2 \rightarrow 1}$	0.560	0.960	1.220	1.760	2.640	1.540	0.194
8	$M_{1 \rightarrow 2}$	0.360	0.600	1.460	2.040	3.600	4.300	0.569
9	Θ_1	2.04000	2.40000	3.54000	4.00000	11.32000	6.74000	0.85737
9	Θ_2	1.08000	1.48000	3.30000	4.92000	6.96000	5.42000	0.84154
9	$M_{2 \rightarrow 1}$	0.280	0.920	1.500	2.120	4.840	2.020	0.270
9	$M_{1 \rightarrow 2}$	0.000	0.160	0.460	0.840	4.320	2.140	0.239
10	Θ_1	6.52000	14.92000	15.34000	15.64000	16.12000	12.34000	1.22867
10	Θ_2	6.72000	7.92000	8.54000	9.08000	12.28000	13.10000	1.28600
10	$M_{2 \rightarrow 1}$	0.040	0.520	0.900	1.160	1.720	0.940	0.090
10	$M_{1 \rightarrow 2}$	0.080	0.360	0.660	1.000	1.720	0.860	0.086
All	Θ_1	2.72000	2.96000	3.30000	3.64000	3.84000	3.58000	3.73338
All	Θ_2	0.00000	2.04000	2.30000	2.56000	3.12000	2.14000	2.38092
All	$M_{2 \rightarrow 1}$	0.400	0.800	1.060	1.320	1.680	1.100	1.068
All	$M_{1 \rightarrow 2}$	0.520	0.840	1.060	1.240	1.680	1.140	1.110

Bayesian Analysis: Posterior distribution over all loci

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

Use this value for Bayes factor calculations:

$BF = \text{Exp}[\ln(\text{Prob}(D \mid \text{thisModel})) - \ln(\text{Prob}(D \mid \text{otherModel}))]$

or as $LBF = 2 (\ln(\text{Prob}(D \mid \text{thisModel})) - \ln(\text{Prob}(D \mid \text{otherModel})))$

shows the support for thisModel]

Method	$\ln(\text{Prob}(D \mid \text{Model}))$	Notes
Thermodynamic integration	-34730.331872	(1a)
	-6570.385819	(1b)
Harmonic mean	-898.559432	(2)

(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	65304/65304	1.00000
Θ_2	65199/65199	1.00000
$M_{2 \rightarrow 1}$	65103/65103	1.00000
$M_{1 \rightarrow 2}$	64998/64998	1.00000
Genealogies	51107/161057	0.31732

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.74314	628.47
Θ_2	0.58919	638.96
$M_{2 \rightarrow 1}$	0.72010	621.32
$M_{1 \rightarrow 2}$	0.74088	514.13
$\text{Ln}[\text{Prob}(D G)]$	0.87499	161.95