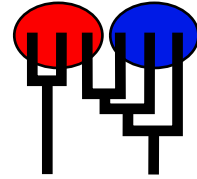


# Example: Microsatellite data set

MIGRATION RATE AND POPULATION SIZE ESTIMATION  
 using the coalescent and maximum likelihood or Bayesian inference  
 Migrate-n version 3.1.7 [1703]  
 Compiled for a SYMMETRIC MULTIPROCESSORS  
 Program started at Thu Aug 5 10:56:51 2010  
 Program finished at Thu Aug 5 11:13:24 2010



## 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	2
1 population__num	*	*
2 population__num	*	*

Order of parameters:

1	$\Theta_1$	<displayed>
2	$\Theta_2$	<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]						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				1000000.00	4 chains with temperatures	
					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

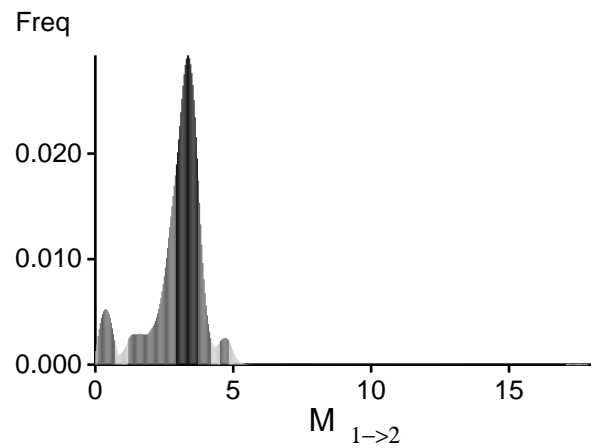
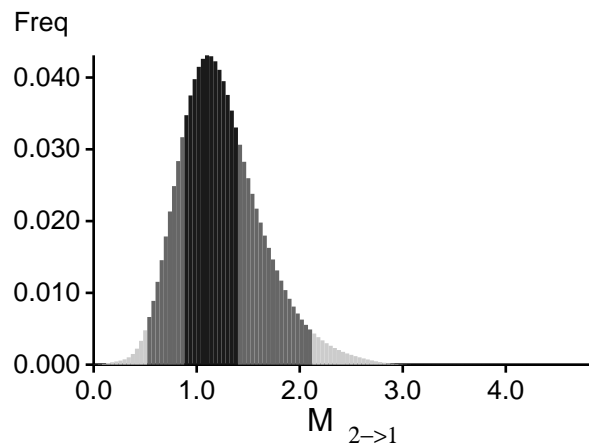
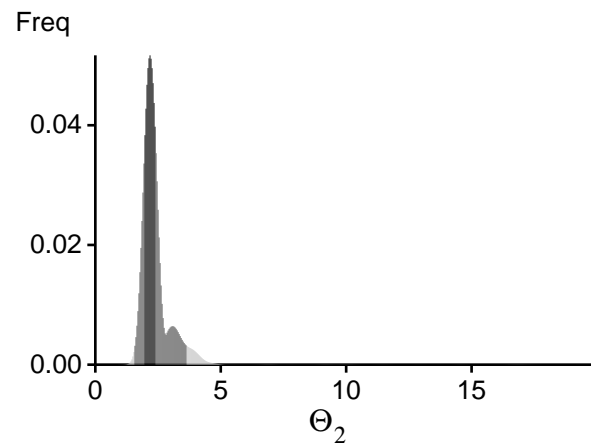
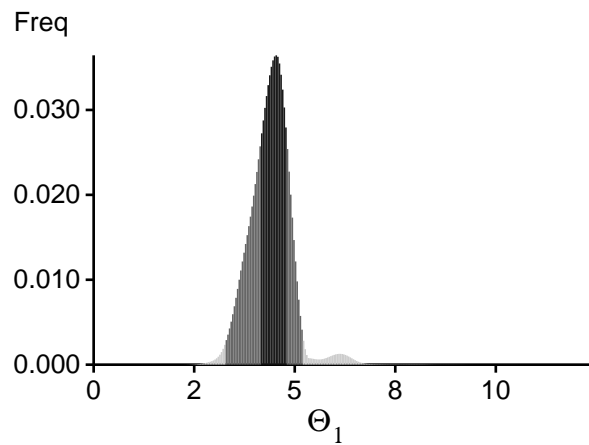
Allele	Pop1	Pop2	All
22	0.100	0.214	0.157
20	0.440	0.214	0.327
23	0.080	0.167	0.123
24	0.020	0.000	0.010
19	0.160	0.167	0.163
21	0.060	0.048	0.054
18	0.080	0.000	0.040
15	0.020	0.071	0.046
17	0.040	0.048	0.044
25	0.000	0.071	0.036

## *Bayesian Analysis: Posterior distribution table*

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
1	$\Theta_1$	0.76000	2.32000	3.34000	5.36000	12.64000	5.58000	6.59041
1	$\Theta_2$	6.08000	15.36000	18.54000	19.80000	20.00000	15.18000	14.22092
1	$M_{2 \rightarrow 1}$	3.640	4.360	6.060	6.680	11.920	6.780	7.194
1	$M_{1 \rightarrow 2}$	2.880	3.920	4.700	5.120	6.120	4.580	4.554
2	$\Theta_1$	2.12000	3.20000	5.02000	6.12000	10.92000	5.58000	3.02212
2	$\Theta_2$	1.32000	1.80000	2.46000	5.04000	12.96000	7.22000	4.23683
2	$M_{2 \rightarrow 1}$	0.000	0.080	0.500	0.960	2.840	0.900	0.526
2	$M_{1 \rightarrow 2}$	0.920	4.120	5.820	9.040	13.640	7.180	3.756
3	$\Theta_1$	4.40000	5.12000	7.46000	9.88000	17.32000	9.26000	3.47165
3	$\Theta_2$	0.08000	0.32000	1.38000	2.36000	6.36000	2.26000	1.56737
3	$M_{2 \rightarrow 1}$	0.760	1.200	2.100	2.920	5.520	2.620	0.933
3	$M_{1 \rightarrow 2}$	8.000	15.120	15.580	19.120	19.880	12.100	3.832
4	$\Theta_1$	10.48000	17.52000	18.18000	18.76000	20.00000	15.94000	3.84416
4	$\Theta_2$	9.04000	16.24000	19.34000	19.92000	20.00000	16.42000	3.87965
4	$M_{2 \rightarrow 1}$	0.160	0.280	0.740	1.200	2.920	1.340	0.434
4	$M_{1 \rightarrow 2}$	0.080	0.360	0.740	1.240	2.680	1.180	0.317
5	$\Theta_1$	0.60000	0.96000	2.26000	4.68000	6.48000	4.46000	1.40698
5	$\Theta_2$	1.08000	1.36000	2.22000	4.40000	9.12000	4.22000	1.42288
5	$M_{2 \rightarrow 1}$	0.000	0.000	0.340	1.920	6.160	5.260	1.572
5	$M_{1 \rightarrow 2}$	0.000	0.000	0.020	3.920	4.920	4.740	1.625
6	$\Theta_1$	2.32000	4.32000	4.74000	6.92000	10.56000	6.14000	1.07252
6	$\Theta_2$	0.12000	0.56000	0.98000	2.32000	7.08000	2.14000	0.45631
6	$M_{2 \rightarrow 1}$	0.000	0.360	0.740	1.200	2.160	1.060	0.213
6	$M_{1 \rightarrow 2}$	0.160	2.000	2.740	4.160	12.560	3.540	0.819
7	$\Theta_1$	0.04000	0.36000	0.86000	2.32000	6.00000	2.14000	0.36836
7	$\Theta_2$	1.12000	2.12000	3.58000	5.56000	13.92000	5.18000	0.94528
7	$M_{2 \rightarrow 1}$	0.080	1.000	1.580	2.600	9.080	4.180	0.623
7	$M_{1 \rightarrow 2}$	0.000	0.000	0.020	2.080	4.920	2.100	0.610
8	$\Theta_1$	2.40000	3.32000	4.26000	7.04000	15.64000	7.22000	1.02955

Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
8	$\Theta_2$	1.44000	1.84000	2.42000	6.00000	12.72000	6.58000	0.99931
8	$M_{2 \rightarrow 1}$	0.280	0.960	1.780	2.040	3.360	1.700	0.216
8	$M_{1 \rightarrow 2}$	1.560	3.320	5.020	5.640	13.360	7.020	0.920
9	$\Theta_1$	1.96000	3.68000	4.46000	6.68000	11.60000	5.86000	0.71100
9	$\Theta_2$	1.00000	1.68000	2.54000	5.24000	10.44000	6.10000	0.85300
9	$M_{2 \rightarrow 1}$	0.000	0.000	0.020	0.440	5.160	2.420	0.266
9	$M_{1 \rightarrow 2}$	1.520	1.880	3.740	6.960	15.920	6.660	0.887
10	$\Theta_1$	1.56000	2.60000	3.50000	7.60000	14.80000	7.86000	0.90640
10	$\Theta_2$	5.04000	6.08000	7.10000	8.16000	18.96000	11.26000	1.13829
10	$M_{2 \rightarrow 1}$	0.760	2.080	3.380	4.280	11.280	4.660	0.552
10	$M_{1 \rightarrow 2}$	0.000	0.880	1.260	1.800	3.480	1.500	0.170
All	$\Theta_1$	3.24000	4.12000	4.54000	4.80000	5.20000	4.46000	4.40411
All	$\Theta_2$	1.52000	1.92000	2.18000	2.40000	3.64000	2.30000	2.39049
All	$M_{2 \rightarrow 1}$	0.480	0.840	1.100	1.400	2.120	1.220	1.243
All	$M_{1 \rightarrow 2}$	1.160	2.880	3.380	3.720	4.200	3.220	2.952



*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]

Locus	Raw thermodynamic score(1a)	Bezier approximation score(1b)	Harmonic mean(2)
1	-7636.41	-1346.67	-115.66
2	-2913.00	-559.26	-78.37
3	-1204.14	-292.12	-74.13
4	-7792.32	-1378.82	-129.88
5	-883.89	-216.25	-55.86
6	-4265.14	-788.37	-77.29
7	-1593.18	-346.27	-76.84
8	-2813.17	-552.12	-91.10
9	-2858.41	-563.54	-77.50
10	-3804.78	-729.55	-99.33
All	-35837.26	-6845.79	-948.77

(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 = -72.817552

*Acceptance ratios for all parameters and the genealogies*

Parameter	Accepted changes	Ratio
$\Theta_1$	276107/276107	1.00000
$\Theta_2$	276097/276097	1.00000
$M_{2 \rightarrow 1}$	274642/274642	1.00000
$M_{1 \rightarrow 2}$	275913/275913	1.00000
Genealogies	310443/1001187	0.31007

## *MCMC-Autocorrelation and Effective MCMC Sample Size*

Parameter	Autocorrelation	Effective Sample Size
$\Theta_1$	0.79950	1190.18
$\Theta_2$	0.81841	1344.70
$M_{2 \rightarrow 1}$	0.95344	656.66
$M_{1 \rightarrow 2}$	0.92559	648.58
$\text{Ln}[\text{Prob}(D G)]$	0.92862	498.24