

D. aruanus dataset lumped by archipelago with 3 loci removed for HWE violation -- 1

POPULATION SIZE, MIGRATION, DIVERGENCE, ASSIGNMENT, HISTORY

Bayesian inference using the structured coalescent

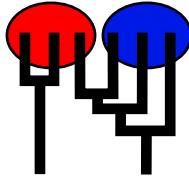
Migrate-n version 4.4.4(git:) [June-1-2019]

Compiled for PARALLEL computer architectures

One master and 31 compute nodes are available.

Program started at Thu Oct 6 04:24:49 2022

Program finished at Thu Oct 6 13:32:01 2022 [Runtime:0000:09:07:12]



Options

Datatype:

Microsatellite data [Brownian motion]

Missing data:

not included

Inheritance scalers in use for Thetas:

All loci use an inheritance scaler of 1.0

[The locus with a scaler of 1.0 used as reference]

Random number seed:

(with internal timer) 2783770042

Start parameters:

Theta values were generated

ERROR

xNm values were generated

Using a percent value of the prior

Connection matrix:

m = average (average over a group of Thetas or M,

s = symmetric migration M, S = symmetric 4Nm,

0 = zero, and not estimated,

* = migration free to vary, Thetas are on diagonal

d = row population split off column population, D = split and then migration

Population	1	2	3	4
1 NC	*	D	0	0
2 Fiji	0	*	D	0
3 Societies	0	0	*	D
4 Tuamotus	0	0	0	*

D. aruanus dataset lumped by archipelago with 3 loci removed for HWE violation -- 2

Order of parameters:

1	Θ_1			<displayed>
2	Θ_2			<displayed>
3	Θ_3			<displayed>
4	Θ_4			<displayed>
5	$xNm_{2 \rightarrow \bar{F}}$	Θ_1	$M_{2 \rightarrow 1}$	<displayed>
6	$xNm_{3 \rightarrow \bar{Z}}$	Θ_2	$M_{3 \rightarrow 2}$	<displayed>
7	$xNm_{4 \rightarrow \bar{S}}$	Θ_3	$M_{4 \rightarrow 3}$	<displayed>
8	$\Delta_{2 \rightarrow 1}$			<displayed>
9		$\sigma_{2 \rightarrow 1}$		<displayed>
10		$\Delta_{3 \rightarrow 2}$		<displayed>
11		$\sigma_{3 \rightarrow 2}$		<displayed>
12		$\Delta_{4 \rightarrow 3}$		<displayed>
13		$\sigma_{4 \rightarrow 3}$		<displayed>

Mutation rate among loci:

Mutation rate is constant for all loci

Analysis strategy:

Bayesian inference

-Population size estimation:

Exponential Distribution

-Geneflow estimation:

Exponential Distribution

-Divergence time estimation:

Normal Distribution Shortcut (mean and standard dev.)

Proposal distributions for parameter

Parameter	Proposal
Theta	Metropolis sampling
xNm	Metropolis sampling
Divergence	Metropolis sampling
Divergence Spread	Metropolis sampling
Genealogy	Metropolis-Hastings

Prior distribution for parameter

Parameter	Prior	Minimum	Mean	Maximum	Delta	Bins	UpdateFreq
1	Theta *Exp window	0.000000	0.100	1.000	0.100	1000	0.03846
2	Theta *Exp window	0.000000	0.100	1.000	0.100	1000	0.03846
3	Theta *Exp window	0.000000	0.100	1.000	0.100	1000	0.03846
4	Theta *Exp window	0.000000	0.100	1.000	0.100	1000	0.03846
5	xNm ** Uniform	0.000000	150.0	300.0	30.00	1000	0.03846
6	xNm ** Uniform	0.000000	150.0	300.0	30.00	1000	0.03846

D. aruanus dataset lumped by archipelago with 3 loci removed for HWE violation -- 3

7	xNm	** Uniform	0.000000	150.0	300.0	30.00	1000	0.03846
8	Splittime mean	*Exp window	0.000000	0.010	1.000	0.010	1500	0.03846
9	Splittime std	*Exp window	0.000000	0.010	1.000	0.010	1500	0.03846
10	Splittime mean	*Exp window	0.000000	0.010	1.000	0.010	1500	0.03846
11	Splittime std	*Exp window	0.000000	0.010	1.000	0.010	1500	0.03846
12	Splittime mean	*Exp window	0.000000	0.010	1.000	0.010	1500	0.03846
13	Splittime std	*Exp window	0.000000	0.010	1.000	0.010	1500	0.03846

[* * means priors were set globally]

Markov chain settings:

Number of chains

Long chain

1

Recorded steps [a]

10000

Increment (record every x step [b])

100

Number of concurrent chains (replicates) [c]

3

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

3000000

Number of discard trees per chain (burn-in)

2000

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:

..../Daruanus_All_8locus.mig

Haplotyping is turned on:

NO

Output file:

outfile.txt

Log file:

logfile.txt

Posterior distribution raw histogram file:

bayesfile

Raw data from the MCMC run:

bayesallfile

Print data:

No

Print genealogies [only some for some data type]:

None

Data summary

Data file:/Daruanus_All_8locus.mig

Datatype: Microsatellite data [Brownian]

[Fragment length is translated to repeats]

Number of loci: 8

Mutationmodel:

Locus	Sublocus	Mutationmodel	Mutationmodel parameters
-------	----------	---------------	--------------------------

1	1	Brownian Motion	[none]
2	1	Brownian Motion	[none]
3	1	Brownian Motion	[none]
4	1	Brownian Motion	[none]
5	1	Brownian Motion	[none]
6	1	Brownian Motion	[none]
7	1	Brownian Motion	[none]
8	1	Brownian Motion	[none]

Population	Locus	Gene copies	
		data	(missing)
1 NC	1	828	(0)
	2	828	(0)
	3	828	(0)
	4	828	(0)
	5	828	(0)
	6	828	(0)
	7	826	(2)
	8	828	(0)
2 Fiji	1	716	(0)
	2	716	(0)
	3	716	(0)
	4	714	(2)
	5	716	(0)
	6	716	(0)
	7	716	(0)
	8	716	(0)
3 Societies	1	882	(0)
	2	882	(0)
	3	882	(0)
	4	882	(0)
	5	882	(0)

D. aruanus dataset lumped by archipelago with 3 loci removed for HWE violation -- 5

	6	882	(0)
	7	882	(0)
	8	882	(0)
4 Tuamotus	1	146	(0)
	2	146	(0)
	3	146	(0)
	4	146	(0)
	5	144	(2)
	6	146	(0)
	7	146	(0)
	8	146	(0)
Total of all populations	1	2572	(0)
	2	2572	(0)
	3	2572	(0)
	4	2570	(2)
	5	2570	(2)
	6	2572	(0)
	7	2570	(2)
	8	2572	(0)

Bayesian Analysis: Posterior distribution table

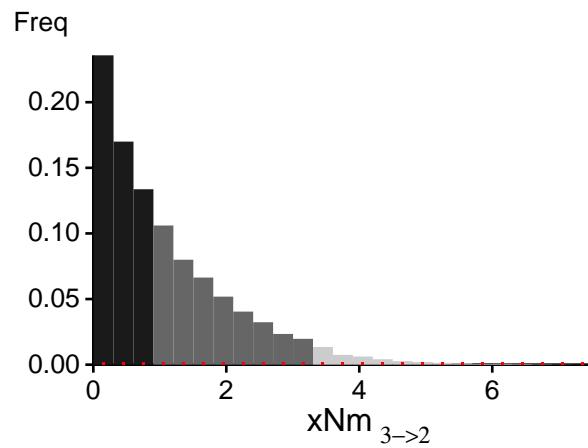
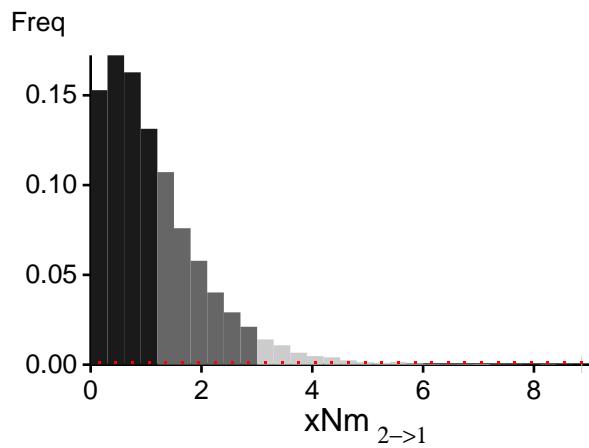
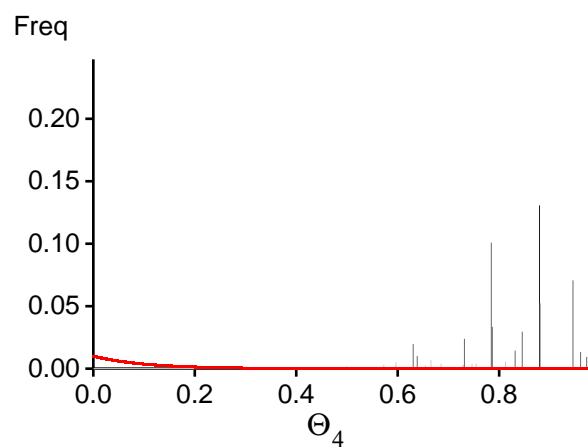
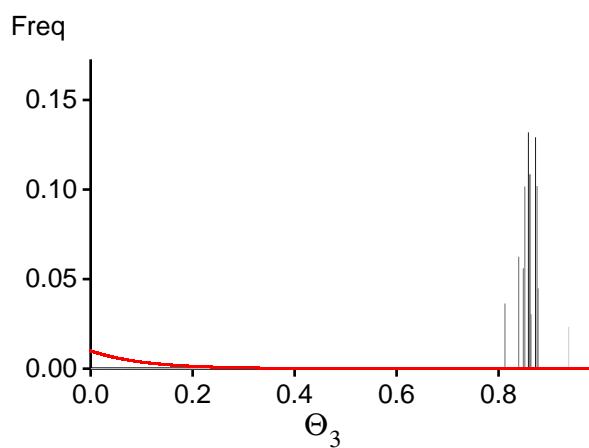
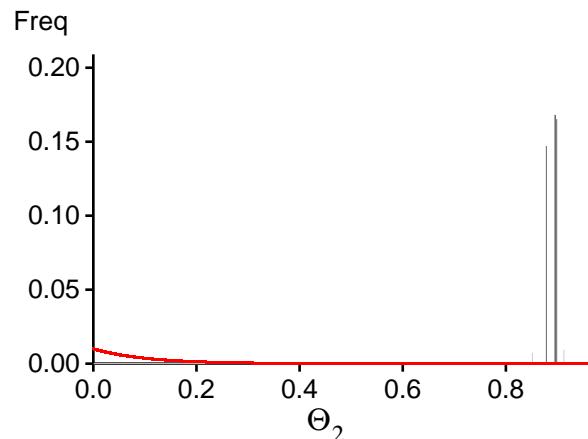
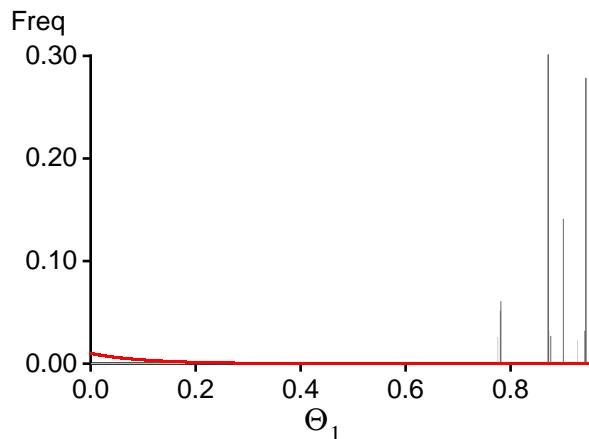
Locus	Parameter	2.5%	25.0%	Mode	75.0%	97.5%	Median	Mean
All	Θ_1	0.36000	0.42400	0.43850	0.45100	0.45700	0.43250	0.39546
All	Θ_2	0.33900	0.35200	0.36850	0.38300	0.47400	0.36950	0.33464
All	Θ_3	0.00000	0.00000	0.00050	0.01500	0.07200	0.28650	0.17455
All	Θ_4	0.46900	0.47800	0.49350	0.50600	0.51400	0.48450	0.45387
All	$xNm_{2>1}$	0.00000	0.00000	0.15000	0.60000	2.40000	0.75000	0.79953
All	$xNm_{3>2}$	0.00000	0.00000	0.15000	0.90000	2.70000	1.05000	0.86588
All	$xNm_{4>3}$	0.00000	0.00000	0.15000	0.90000	2.40000	1.05000	0.84887
All	$D_{2>1}$	0.01333	0.01867	0.02167	0.02400	0.02667	0.02167	0.02085
All	$S_{2>1}$	0.01133	0.01600	0.01967	0.02133	0.02467	0.01967	0.01844
All	$D_{3>2}$	0.01333	0.01867	0.02167	0.02400	0.02667	0.02167	0.02087
All	$S_{3>2}$	0.01133	0.01667	0.01967	0.02200	0.02467	0.01967	0.01852
All	$D_{4>3}$	0.01333	0.01867	0.02167	0.02400	0.02733	0.02167	0.02092
All	$S_{4>3}$	0.01133	0.01667	0.01967	0.02200	0.02467	0.01967	0.01882

Citation suggestions:

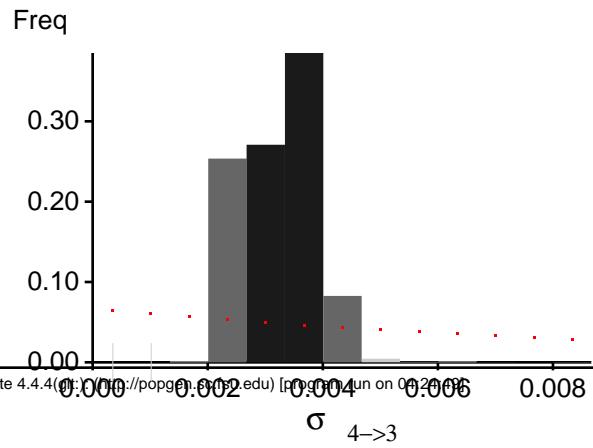
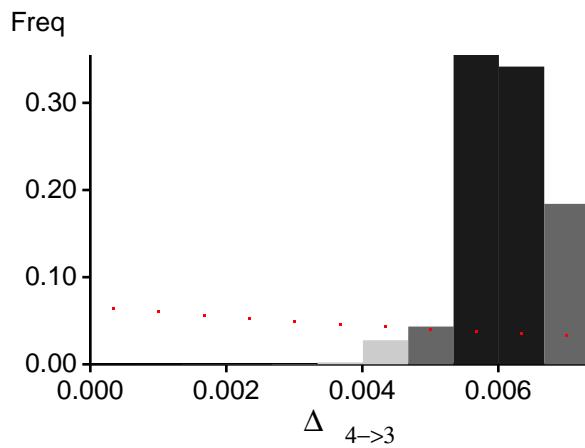
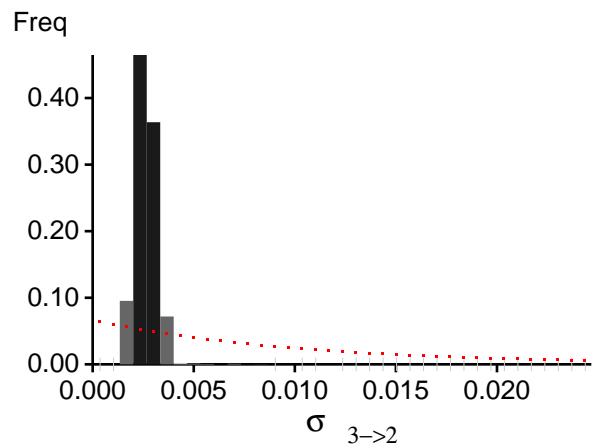
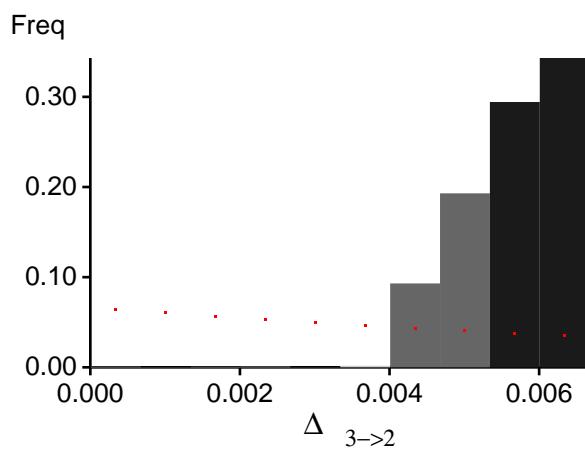
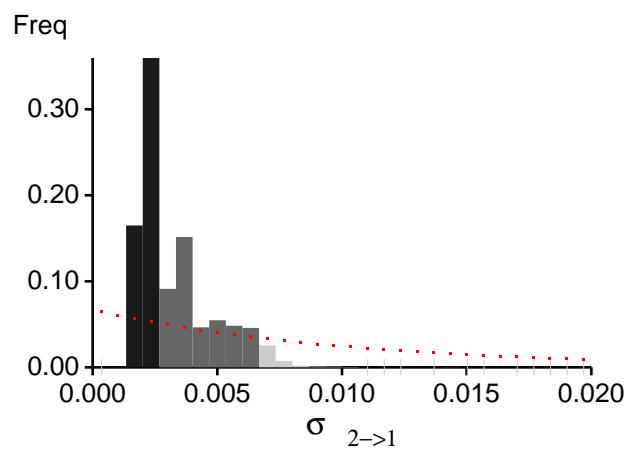
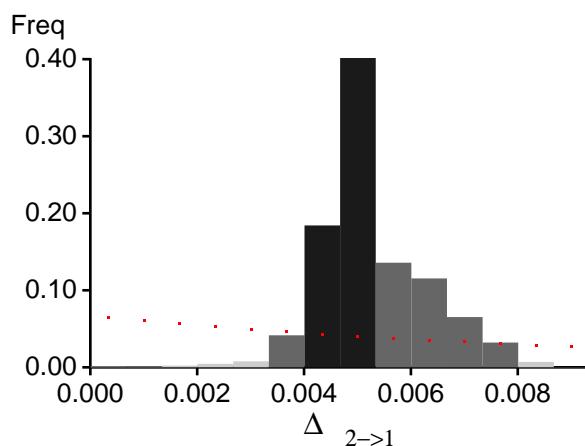
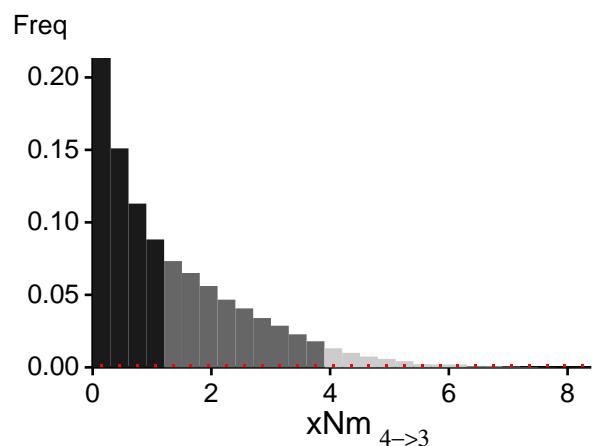
Beerli P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* 22:341-345

Beerli P., 2009. How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In Population Genetics for Animal Conservation, G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli, and C. Vernesi, eds., vol. 17 of Conservation Biology, Cambridge University Press, Cambridge UK, pp. 42-79.

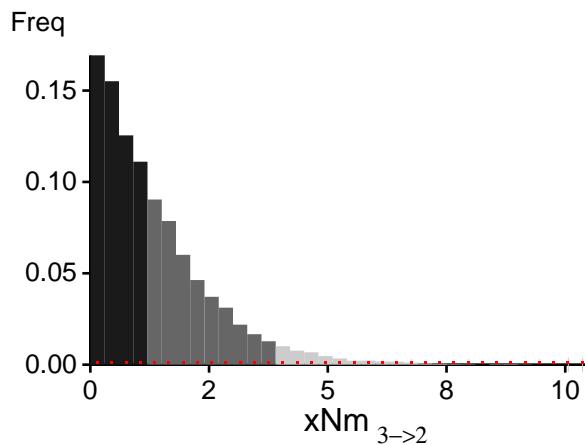
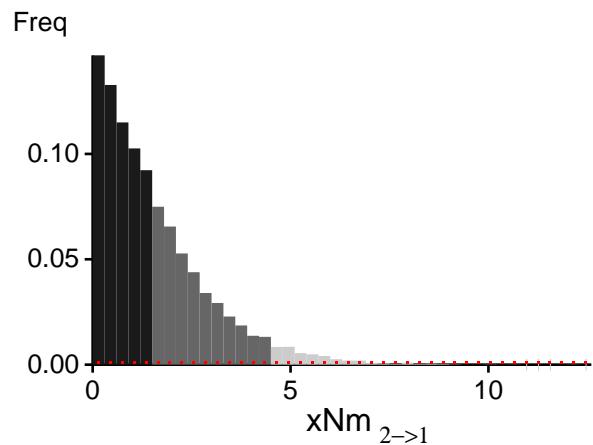
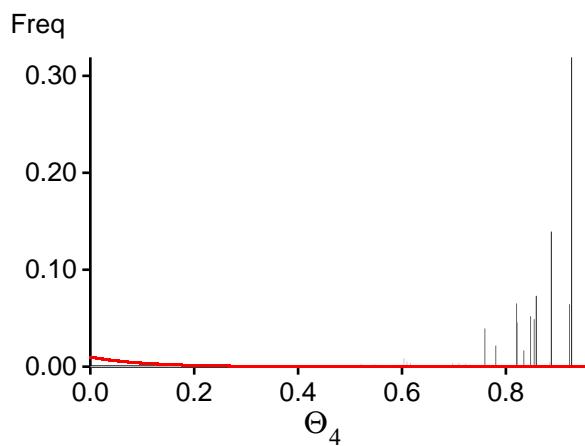
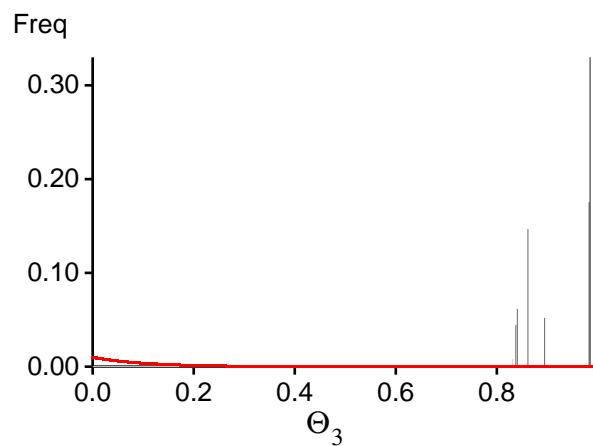
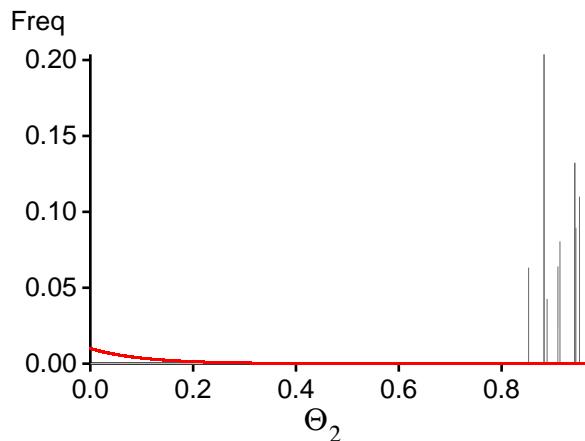
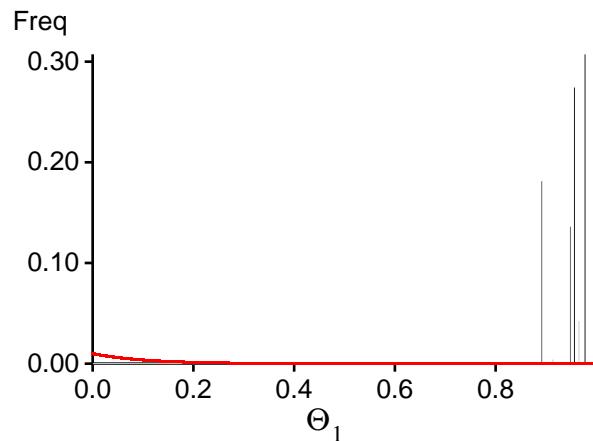
Bayesian Analysis: Posterior distribution for locus 1

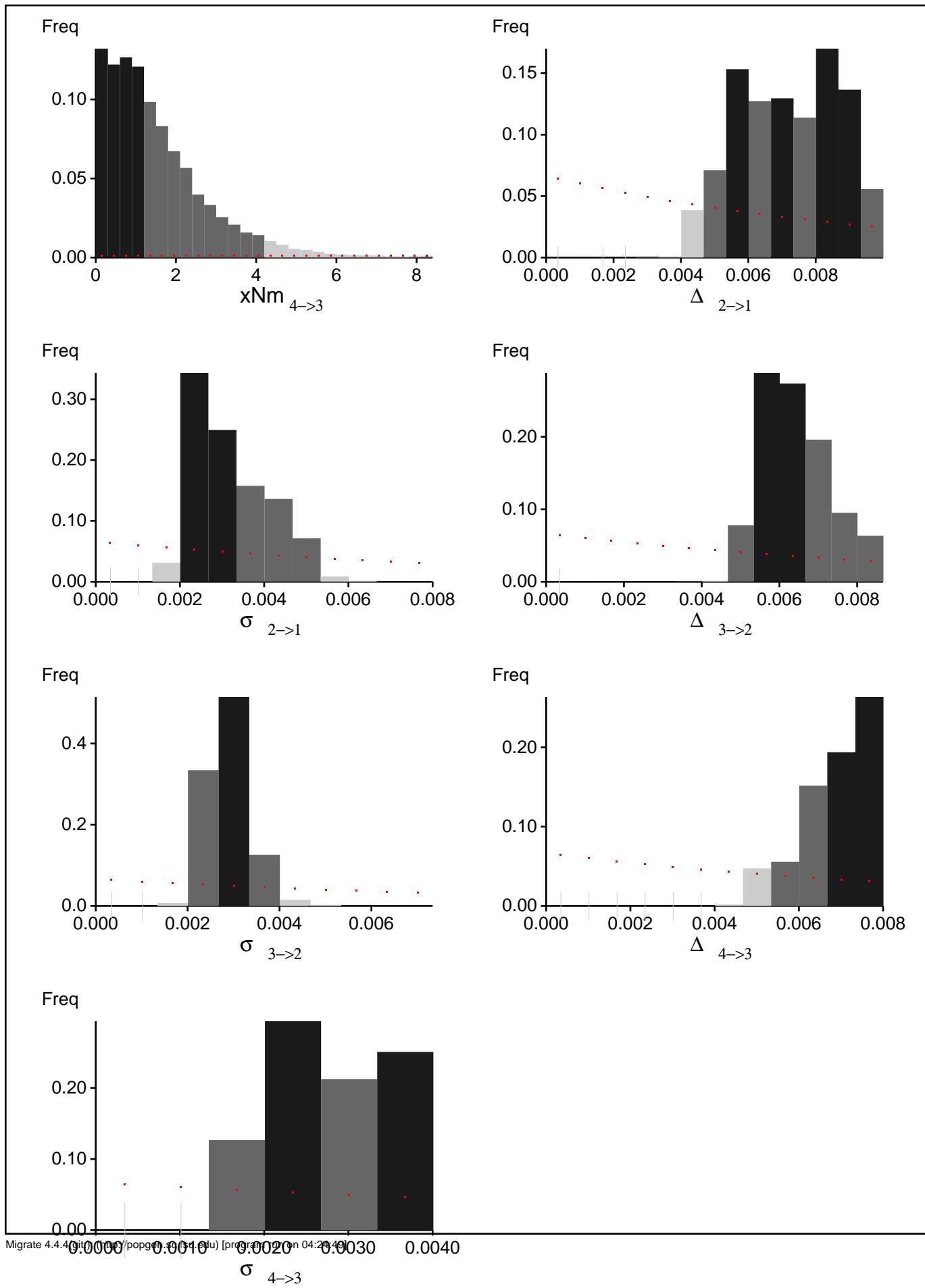


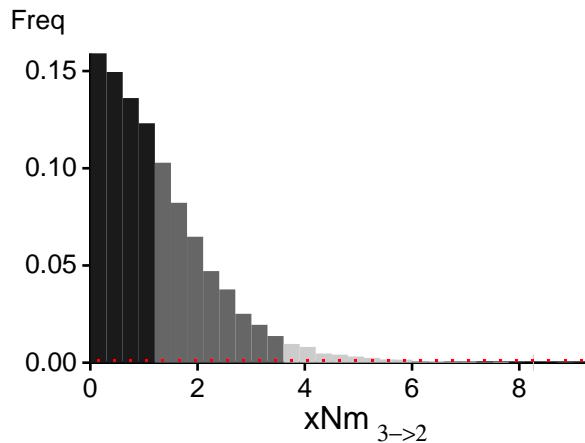
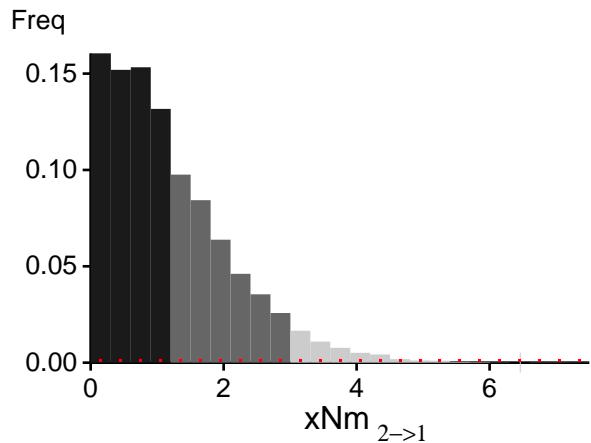
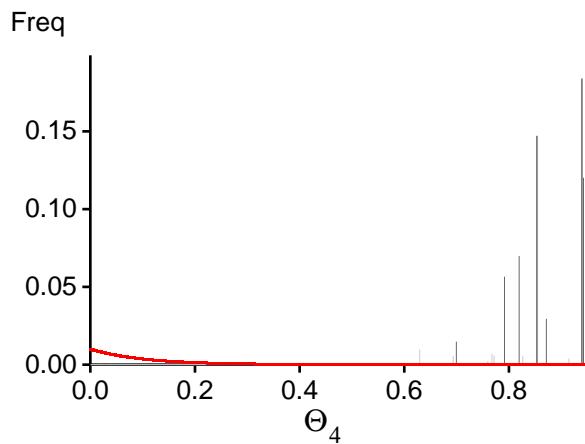
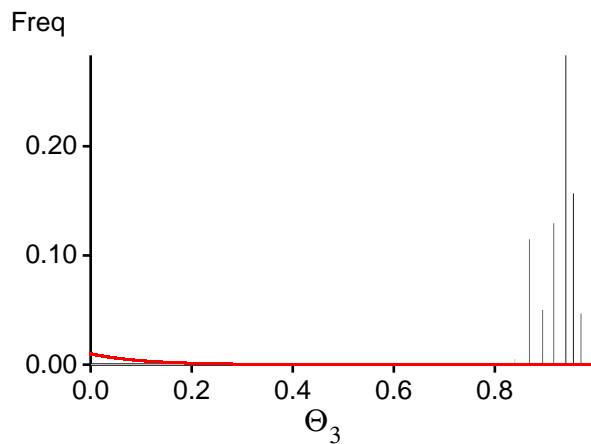
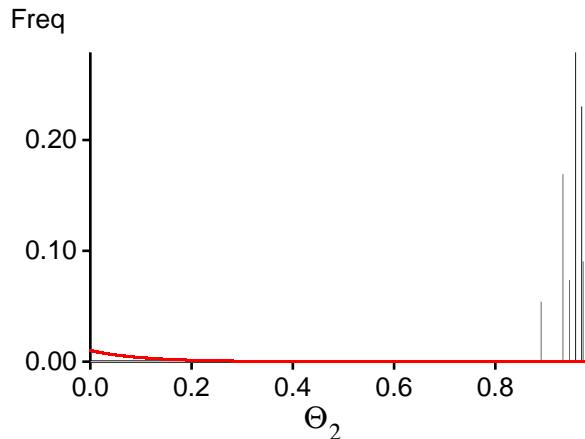
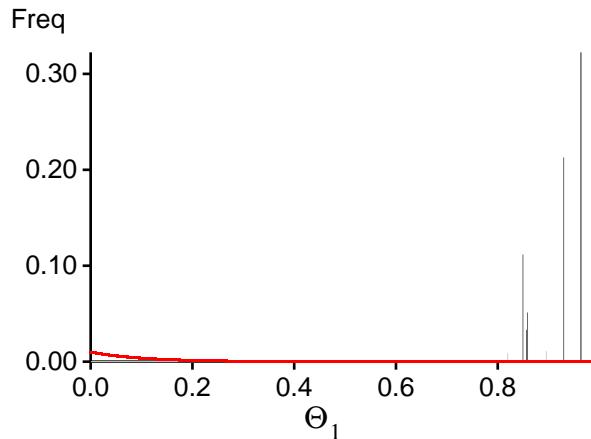
D. aruanus dataset lumped by archipelago with 3 loci removed for HWE violation -- 8

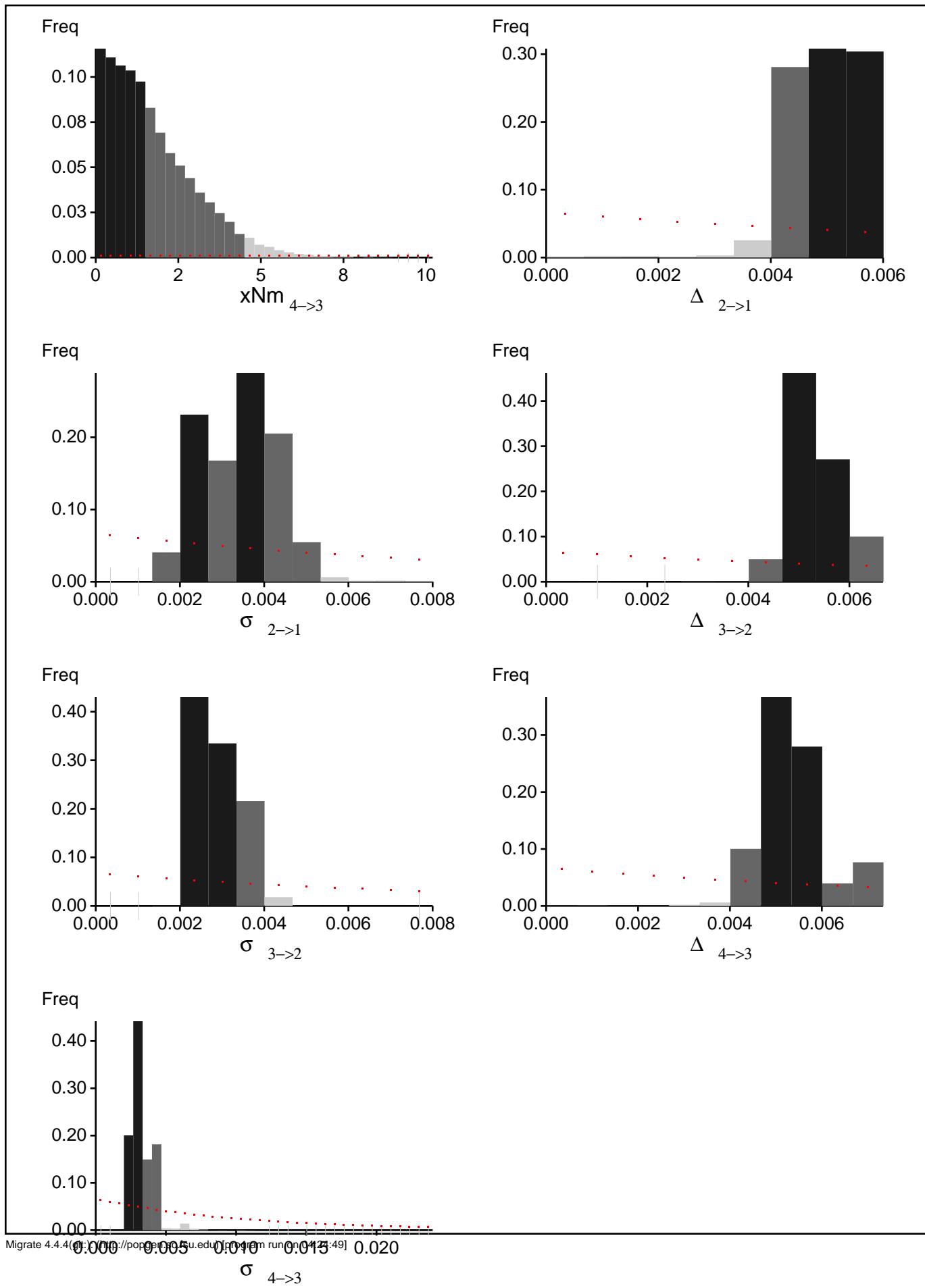


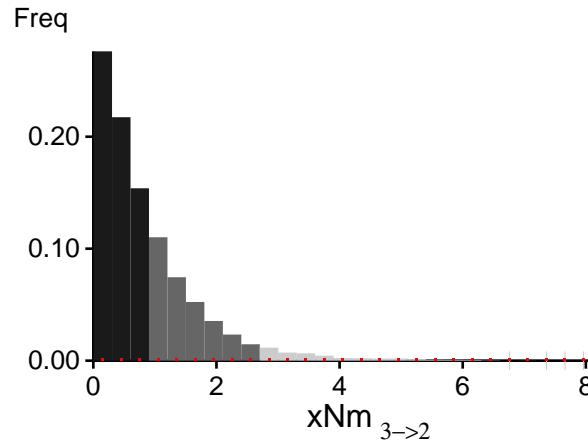
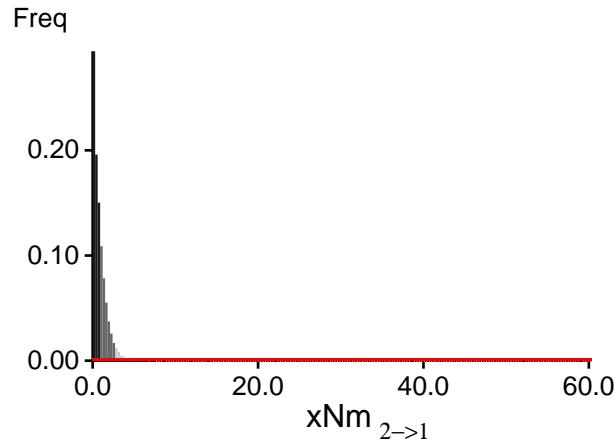
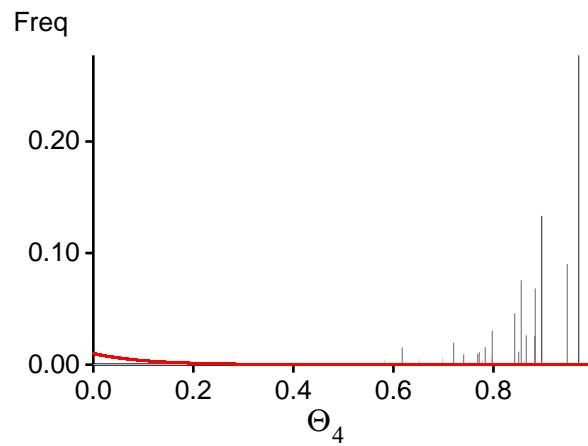
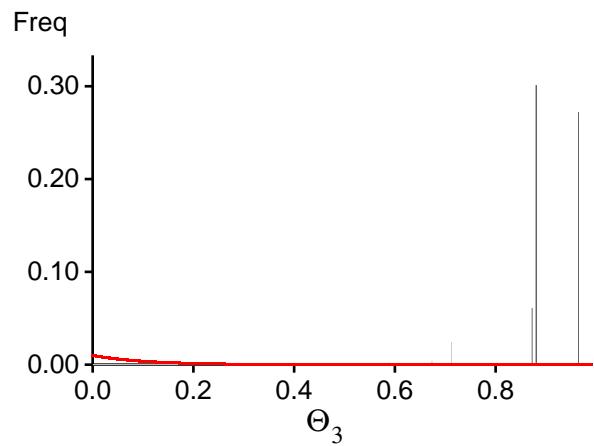
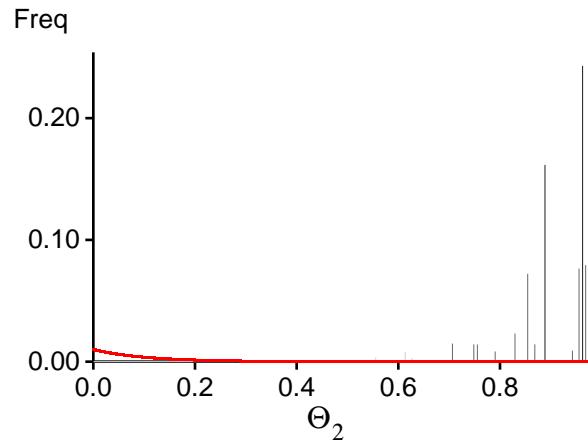
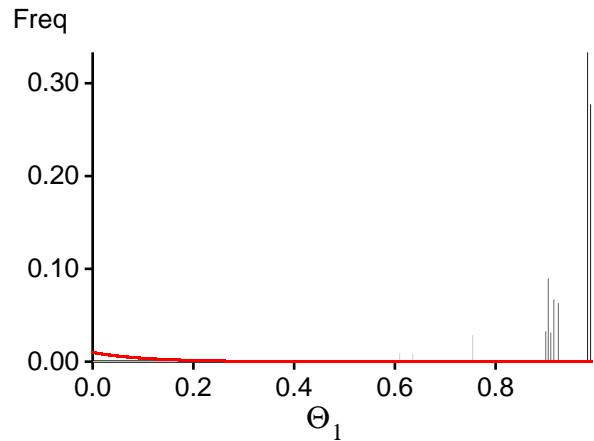
Bayesian Analysis: Posterior distribution for locus 2

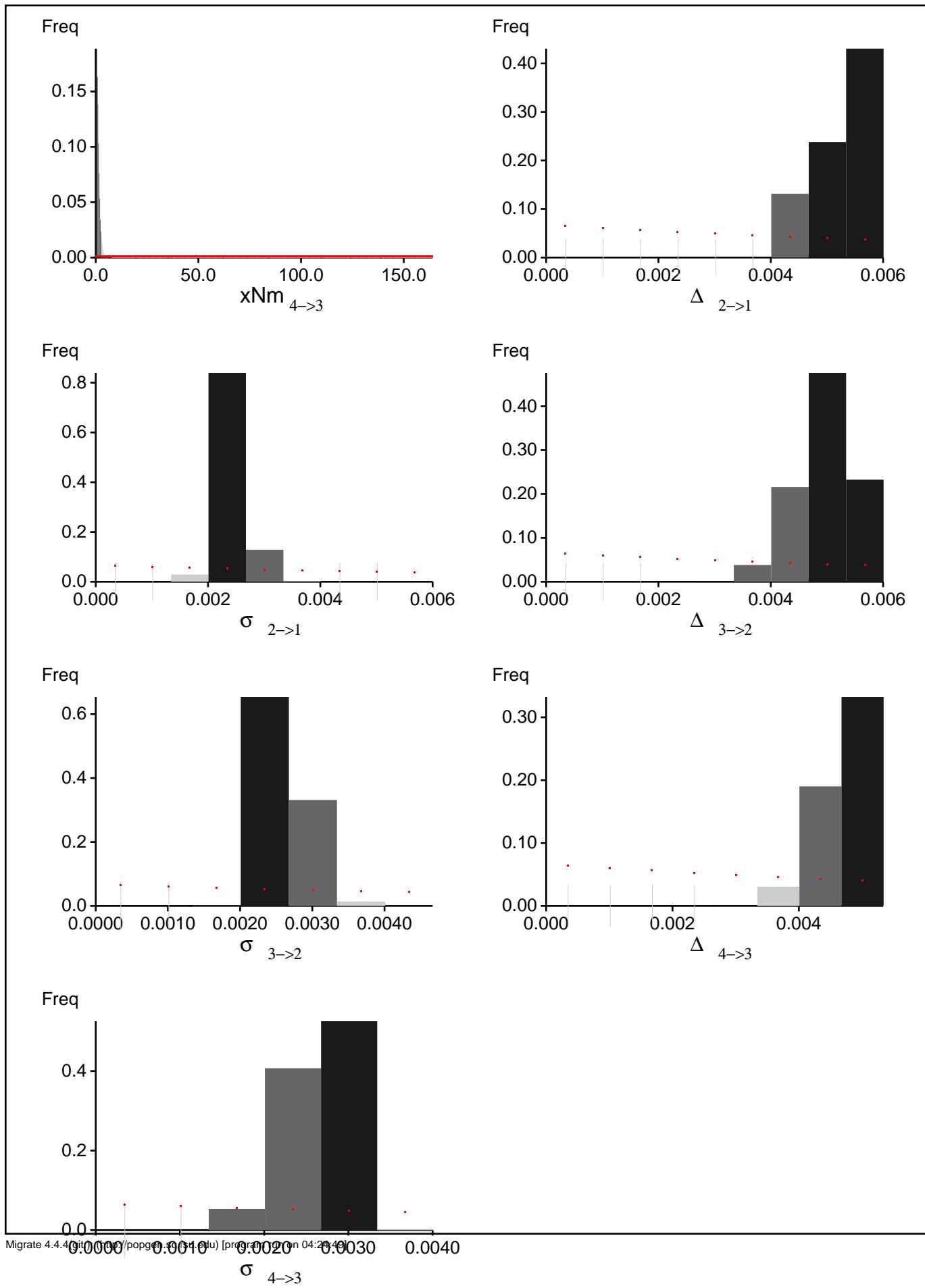


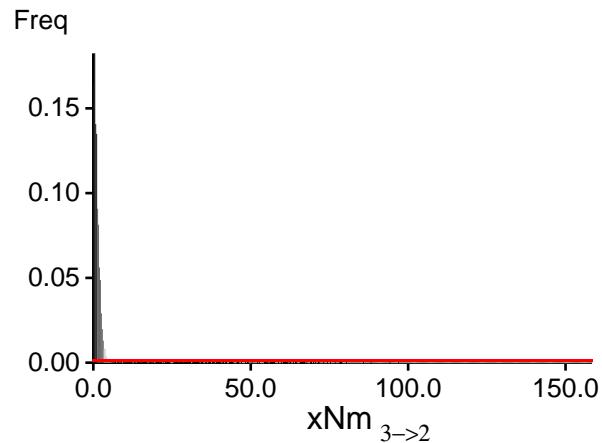
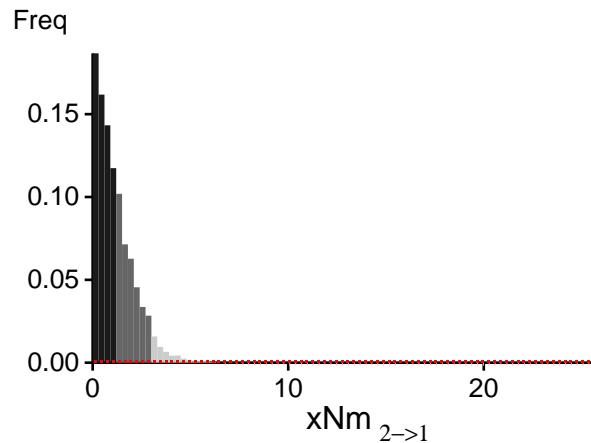
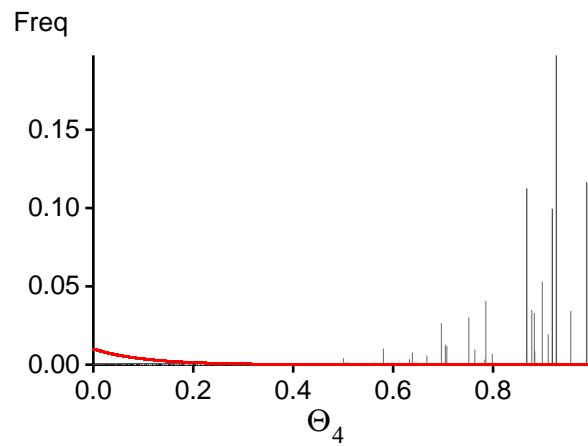
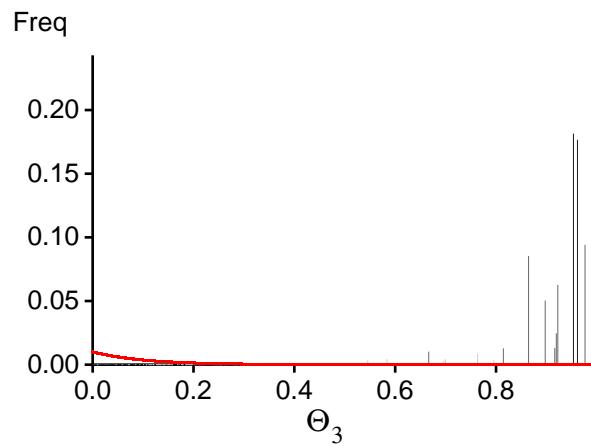
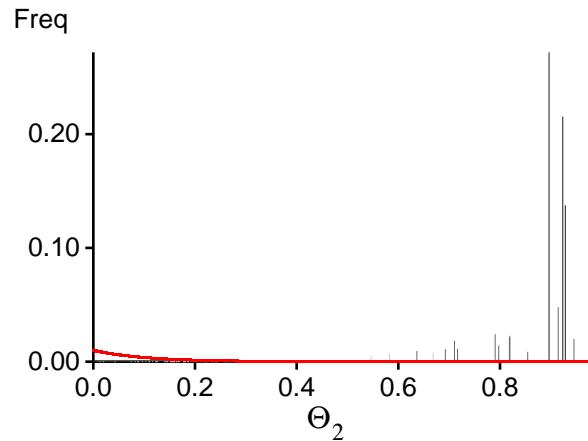
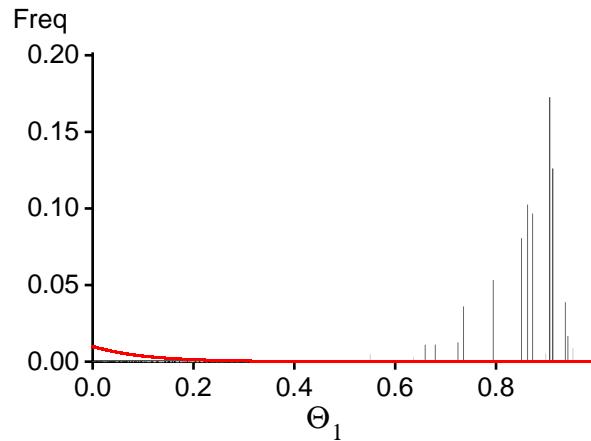


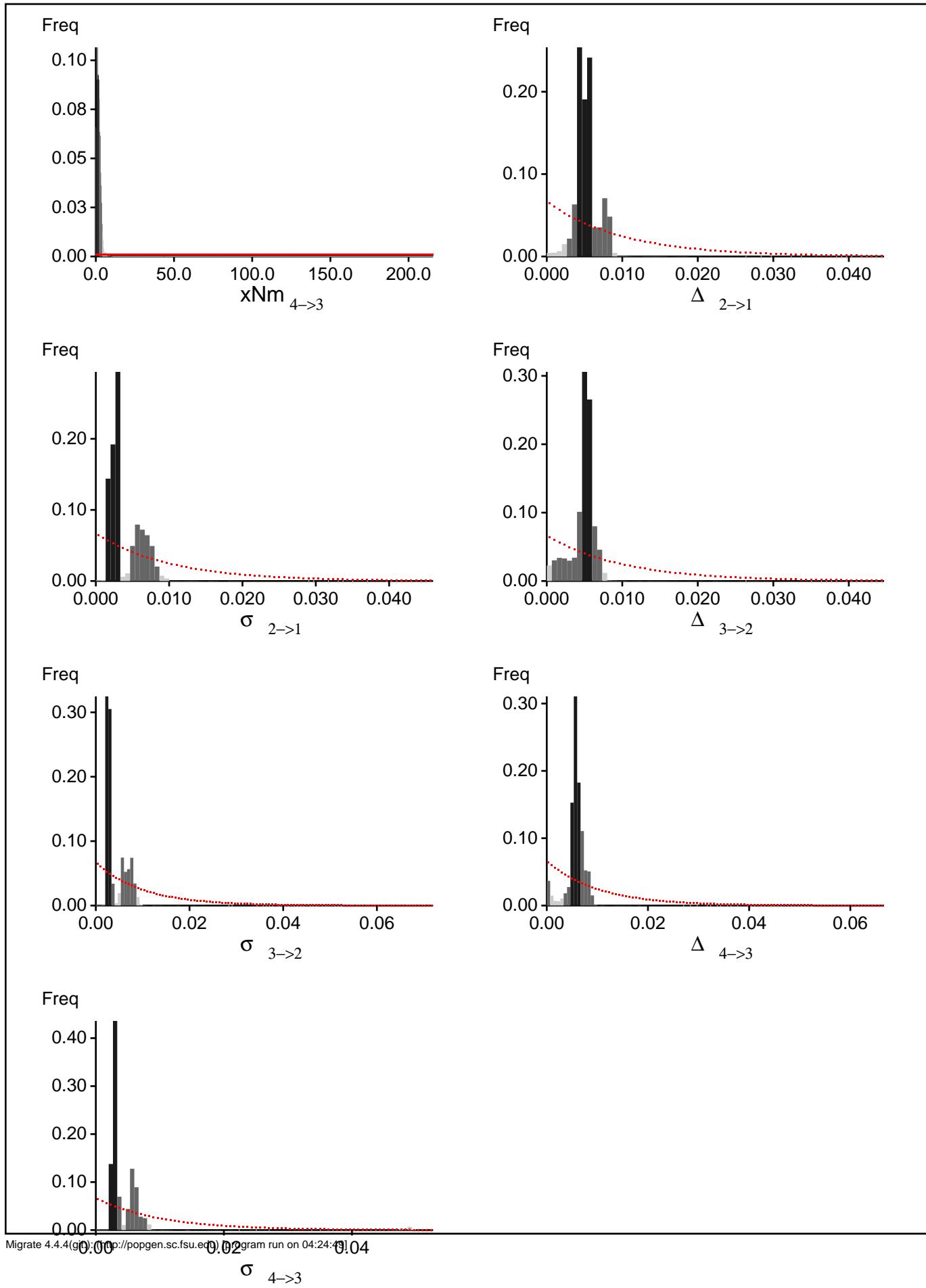
Bayesian Analysis: Posterior distribution for locus 3

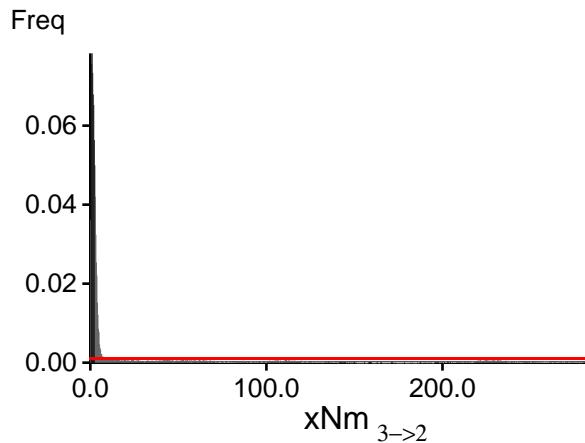
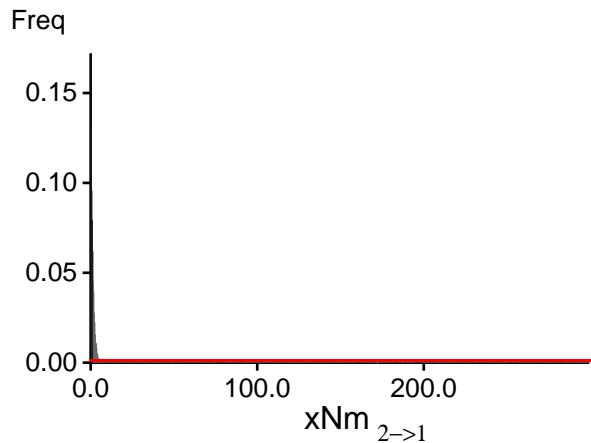
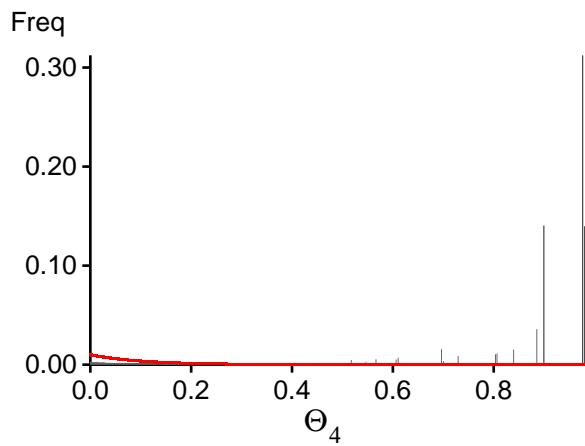
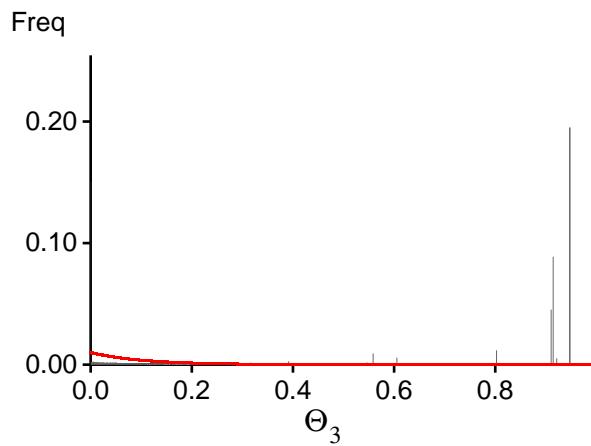
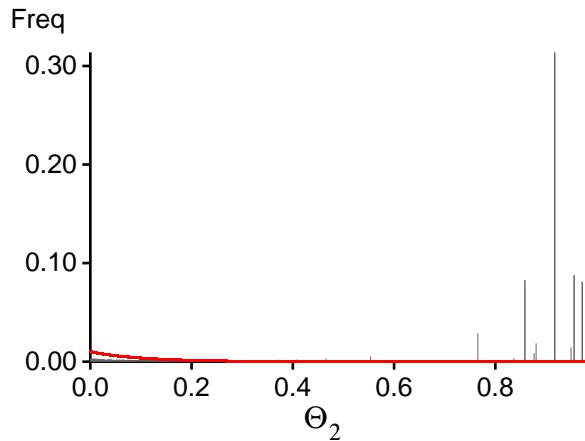
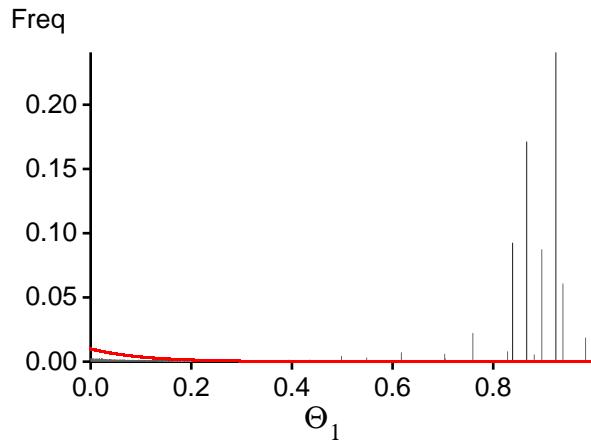


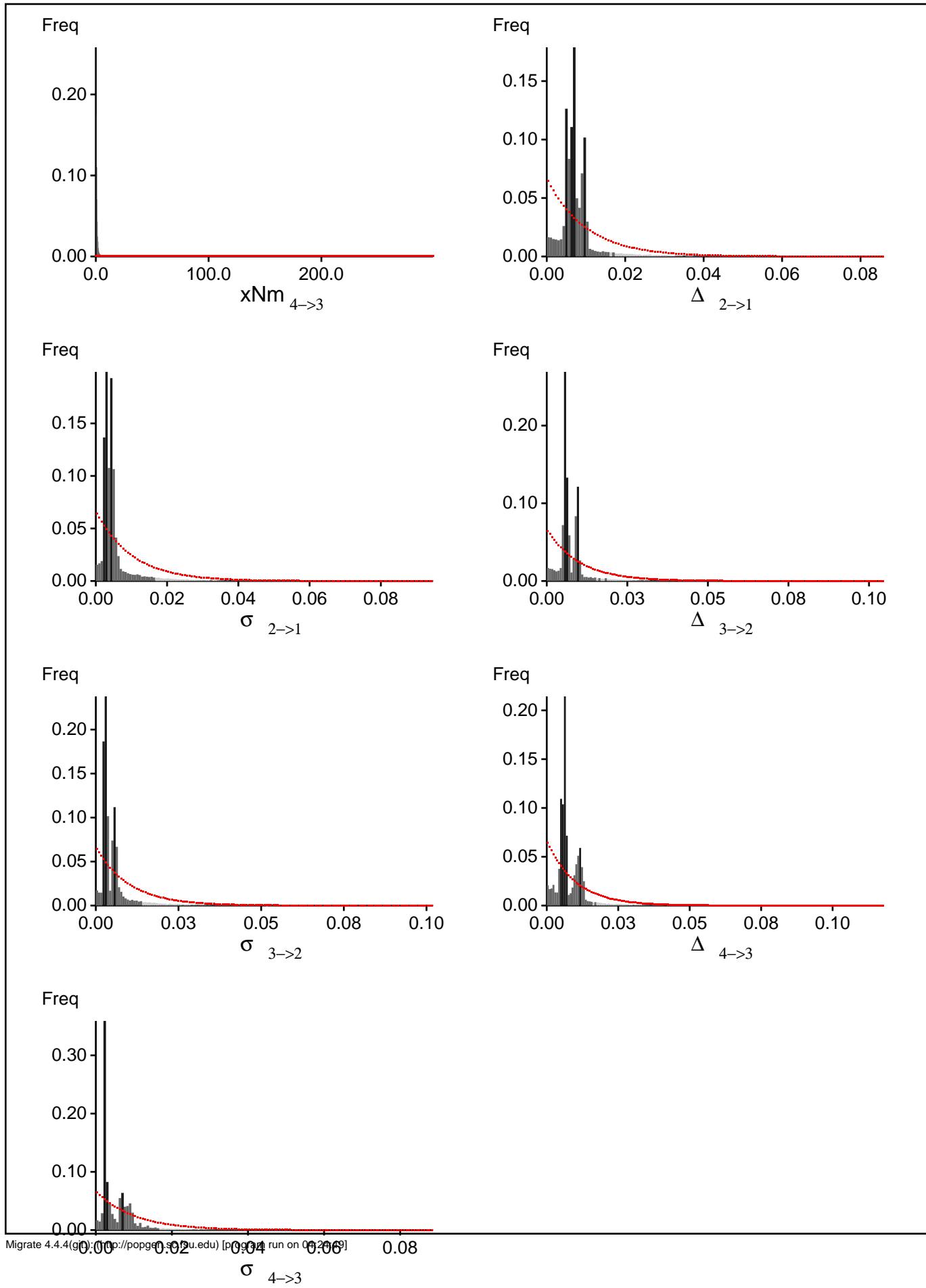
Bayesian Analysis: Posterior distribution for locus 4



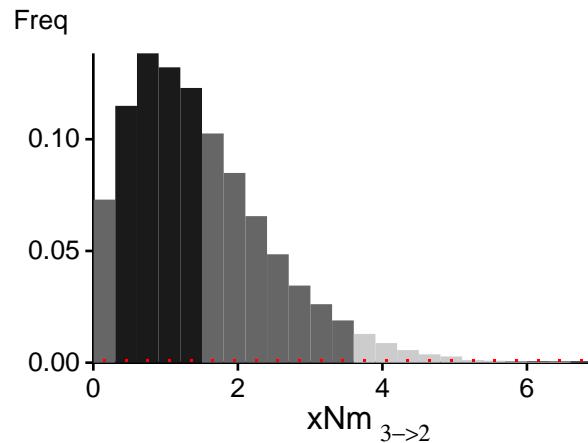
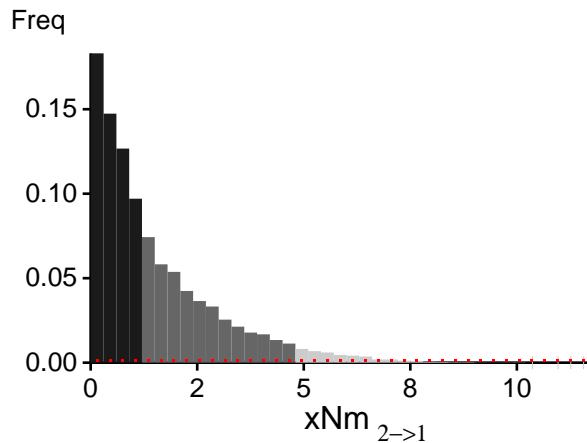
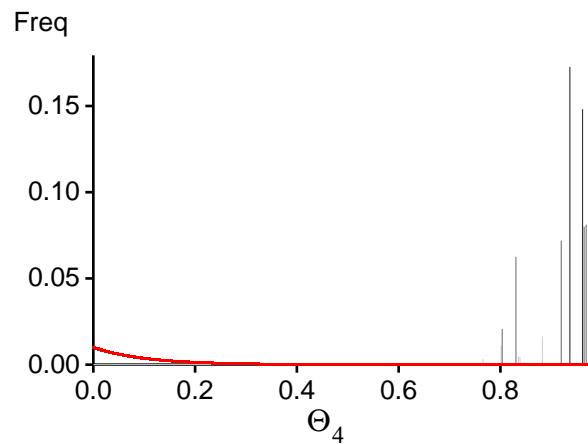
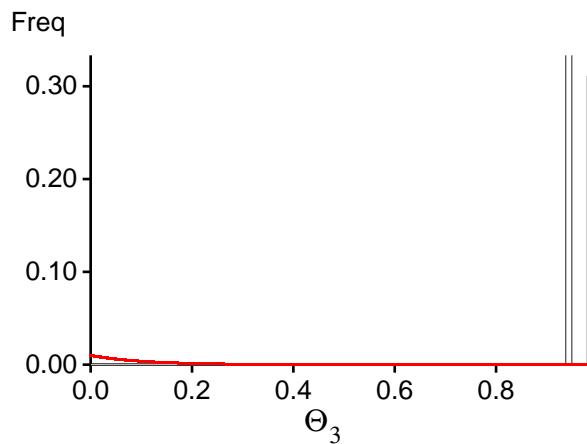
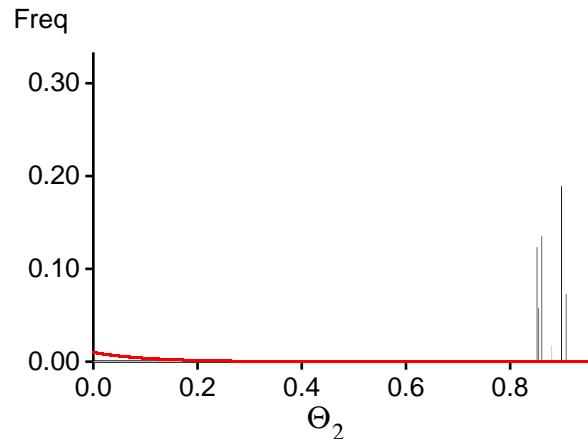
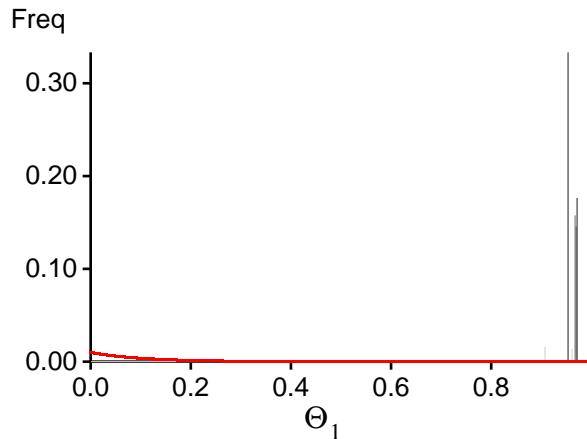
Bayesian Analysis: Posterior distribution for locus 5

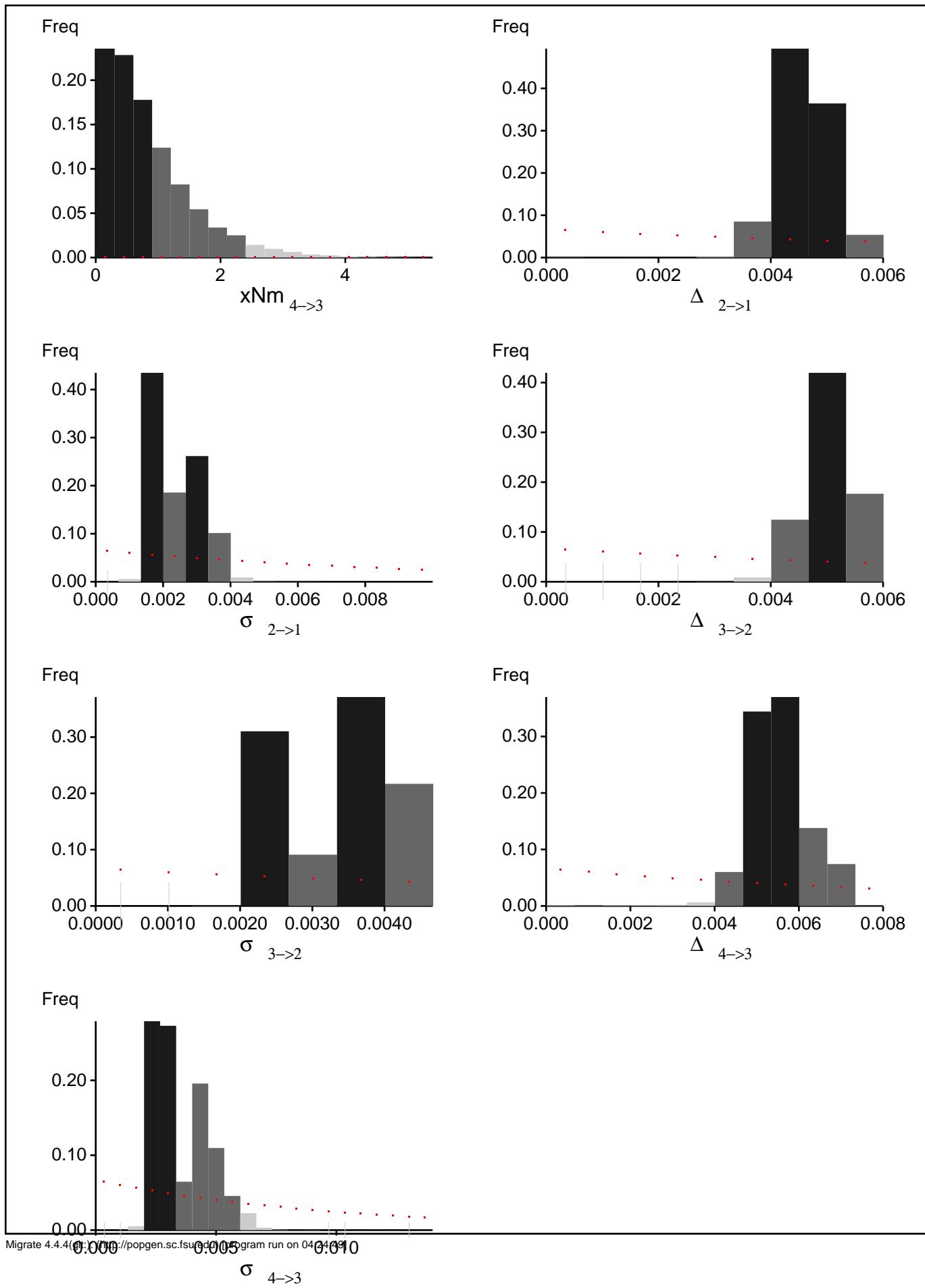


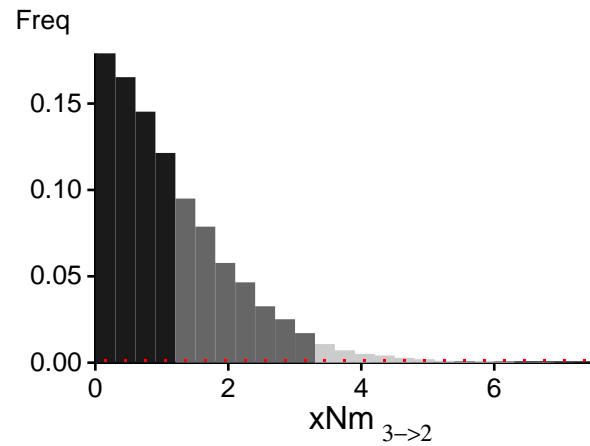
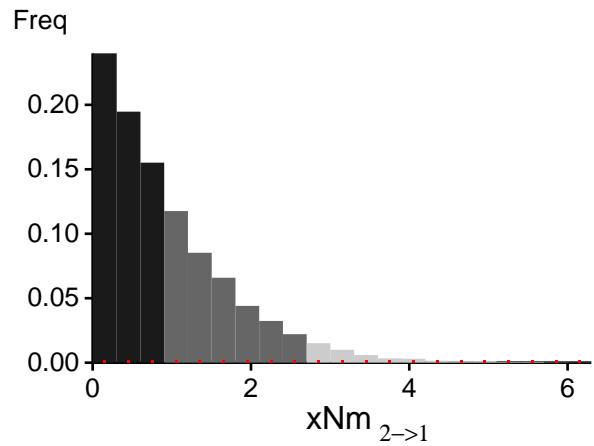
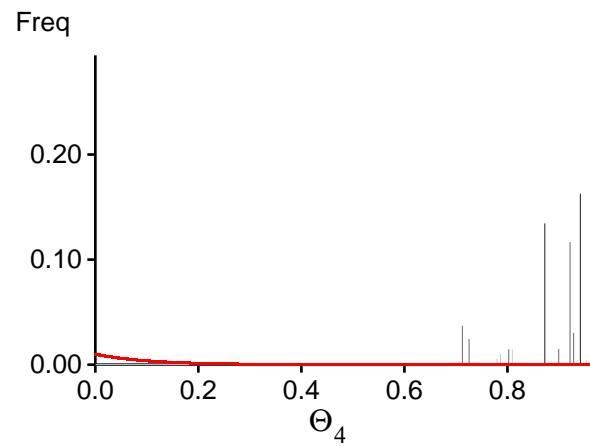
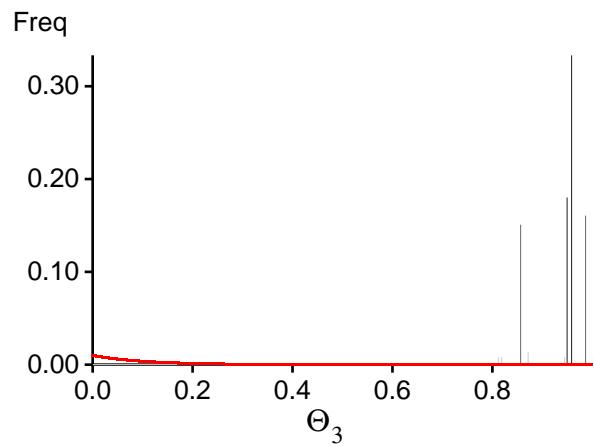
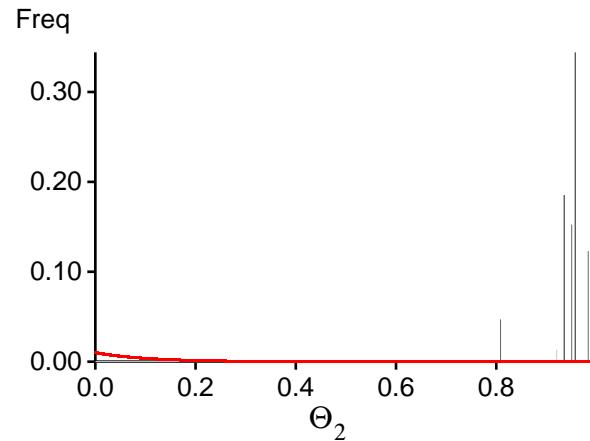
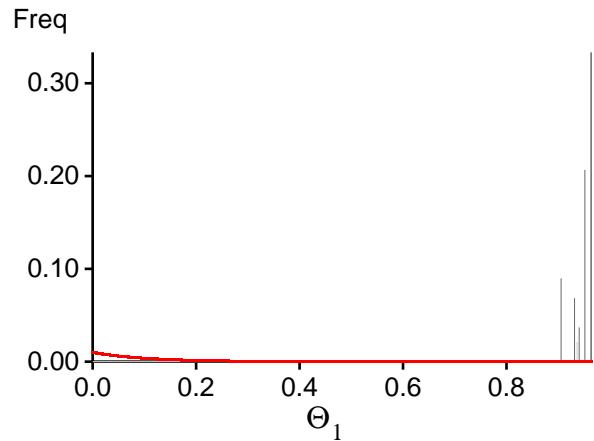
Bayesian Analysis: Posterior distribution for locus 6

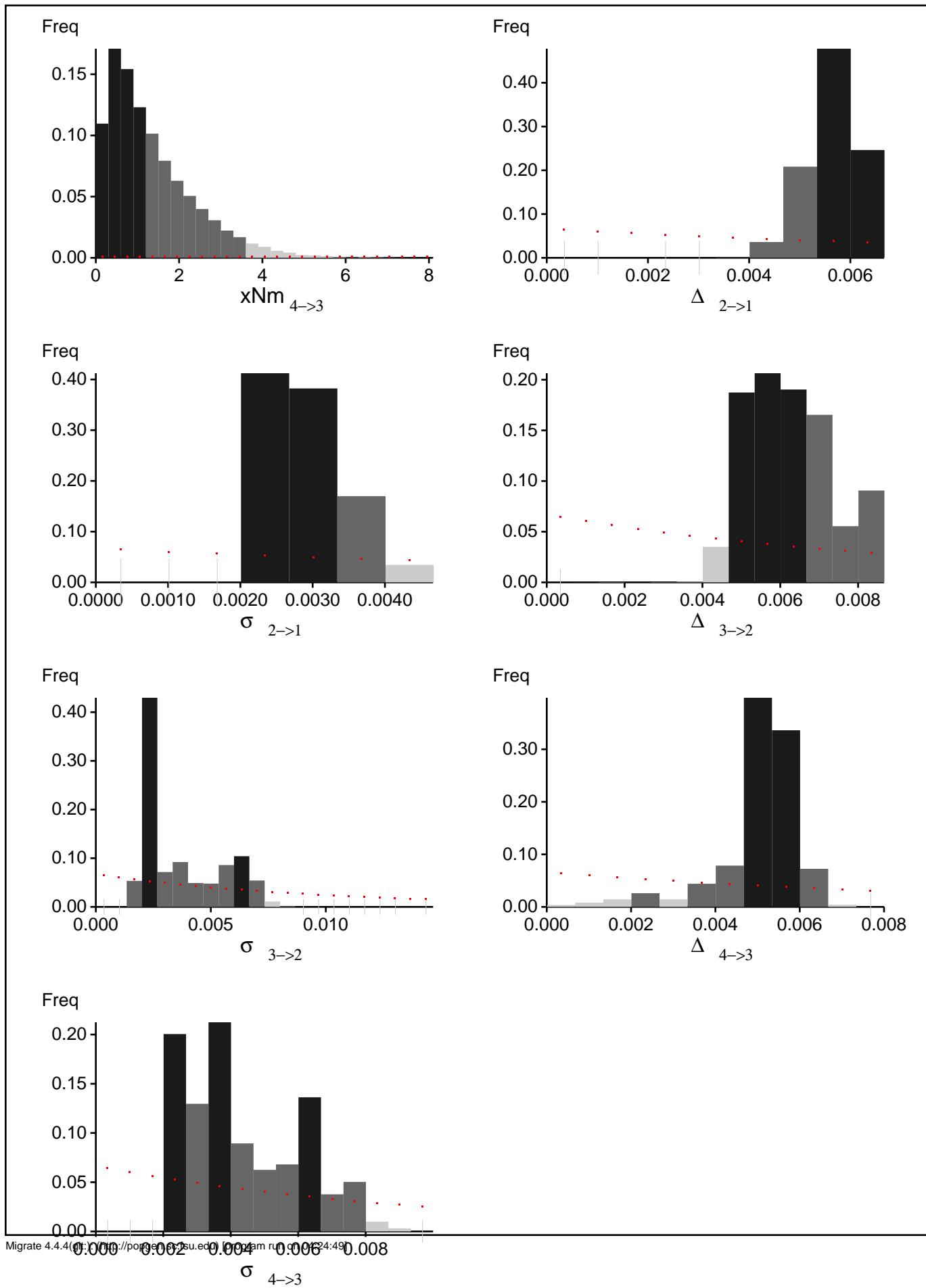


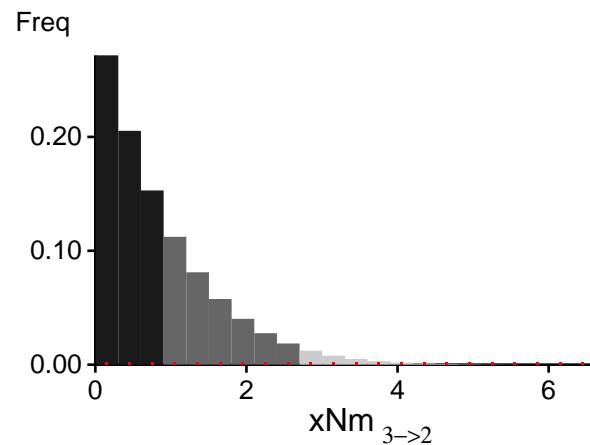
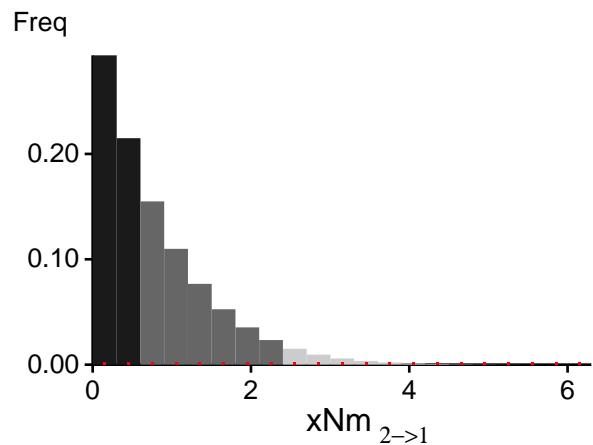
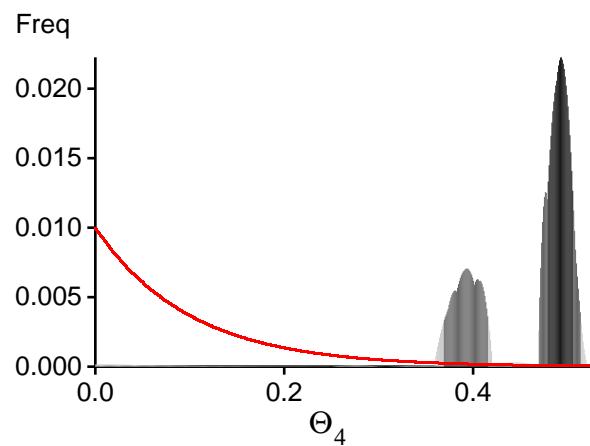
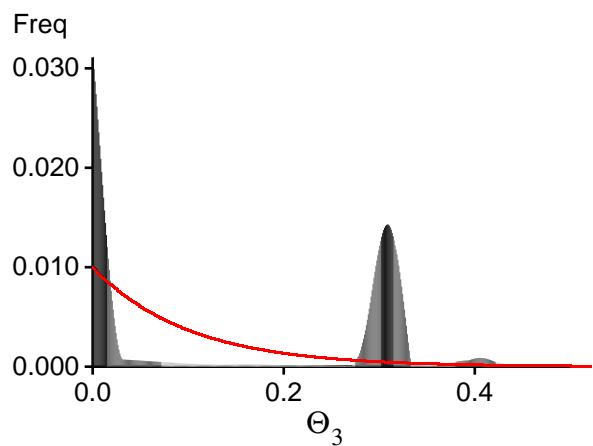
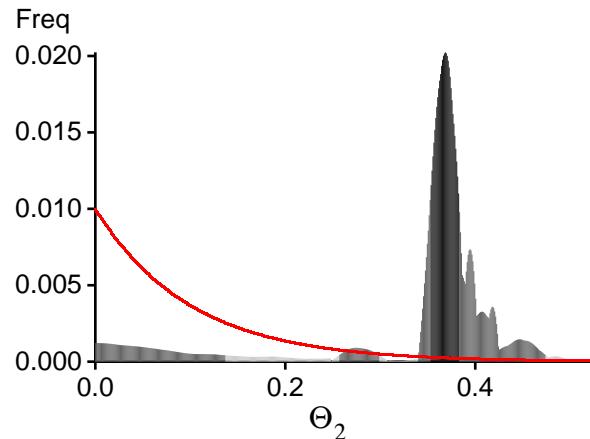
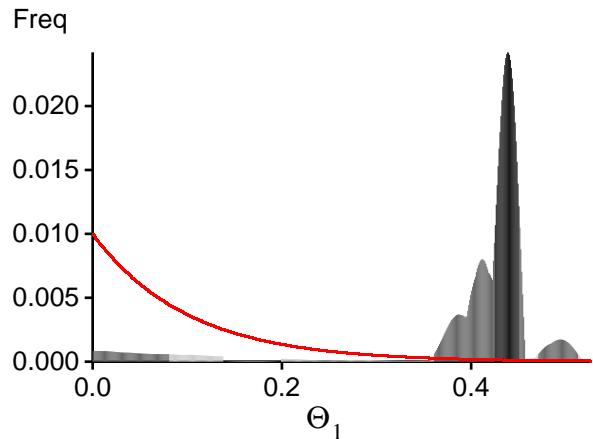
Bayesian Analysis: Posterior distribution for locus 7

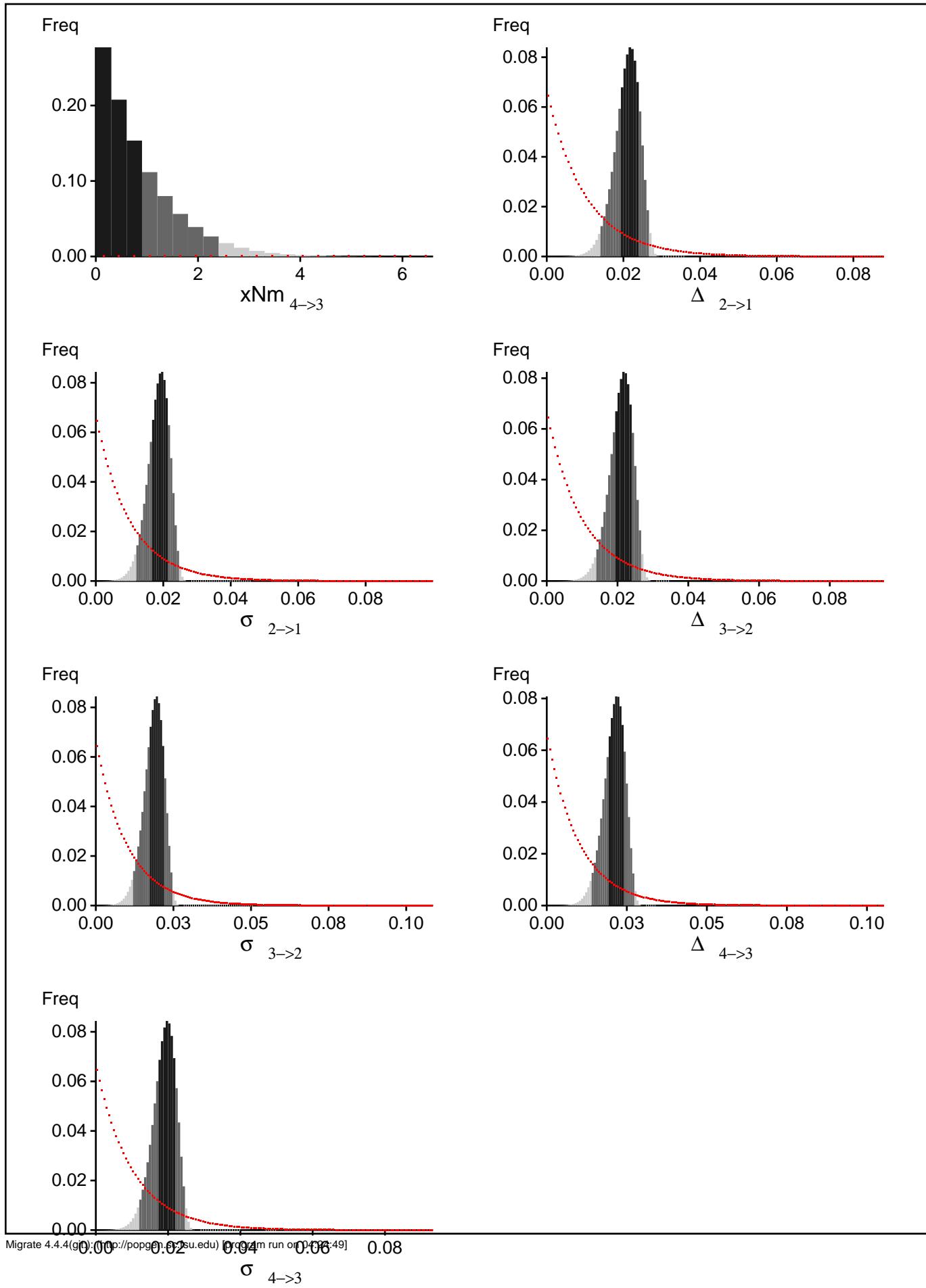




Bayesian Analysis: Posterior distribution for locus 8



Bayesian Analysis: Posterior distribution over all loci



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

Use this value for Bayes factor calculations:

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

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

shows the support for thisModel]

Locus	Raw thermodynamic score(1a)	Bezier approximation score(1b)	Harmonic mean(2)
1	-18731019.59	-3005913.63	-51504.25
2	-4371481.52	-700375.98	-3181.96
3	-51504213.99	-8393101.87	-537834.59
4	-31828909.44	-5116441.23	-141920.90
5	-15447277.42	-2487284.29	-61217.43
6	-14092015.79	-2261949.46	-39353.20
7	-64088387.10	-10388836.08	-662451.89
8	-38560865.77	-6217942.48	-313553.40
All	-238624111.07	-38571785.45	-1810958.06

(1a, 1b and 2) are approximations to the marginal likelihood, make sure that the program run long enough!
 (1a, 1b) and (2) should give similar results, in principle.

But (2) is overestimating the likelihood, it is presented for historical reasons and should not be used

(1a, 1b) needs heating with chains that span a temperature range of 1.0 to at least 100,000.

(1b) is using a Bezier-curve to get better approximations for runs with low number of heated chains

[Scaling factor = 59.561686]

Citation suggestions:

Beerli P. and M. Palczewski, 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations, Genetics, 185: 313-326.

Acceptance ratios for all parameters and the genealogies

Parameter	Accepted changes	Ratio
Θ_1	32232/923746	0.03489
Θ_2	31919/922618	0.03460
Θ_3	31943/923900	0.03457
Θ_4	32012/923530	0.03466
$xNm_{2 \rightarrow 1}$	327062/924477	0.35378
$xNm_{3 \rightarrow 2}$	397765/921965	0.43143
$xNm_{4 \rightarrow 3}$	392646/922698	0.42554
$\Delta_{2 \rightarrow 1}$	64065/923906	0.06934
$\sigma_{2 \rightarrow 1}$	66396/922527	0.07197
$\Delta_{3 \rightarrow 2}$	62562/923166	0.06777
$\sigma_{3 \rightarrow 2}$	59452/922317	0.06446
$\Delta_{4 \rightarrow 3}$	62749/921166	0.06812
$\sigma_{4 \rightarrow 3}$	56337/923230	0.06102
Genealogies	1487080/12000754	0.12392

MCMC-Autocorrelation and Effective MCMC Sample Size

Parameter	Autocorrelation	Effective Sample Size
Θ_1	0.99223	670.18
Θ_2	0.99314	584.52
Θ_3	0.99185	719.73
Θ_4	0.98979	912.00
$xNm_{2 \rightarrow 1}$	0.73525	44921.26
$xNm_{3 \rightarrow 2}$	0.72832	43420.34
$xNm_{4 \rightarrow 3}$	0.71259	43994.09
$\Delta_{2 \rightarrow 1}$	0.86549	18424.96
$\sigma_{2 \rightarrow 1}$	0.85439	18570.23
$\Delta_{3 \rightarrow 2}$	0.87587	16812.45
$\sigma_{3 \rightarrow 2}$	0.85990	18406.72
$\Delta_{4 \rightarrow 3}$	0.86198	20377.40
$\sigma_{4 \rightarrow 3}$	0.86814	16978.86
Genealogies	0.86814	16978.86

Potential Problems

This section reports potential problems with your run, but such reporting is often not very accurate. With many parameters in a multilocus analysis, it is very common that some parameters for some loci will not be very informative, triggering suggestions (for example to increase the prior range) that are not sensible. This suggestion tool will improve with time, therefore do not blindly follow its suggestions. If some parameters are flagged, inspect the tables carefully and judge whether an action is required. For example, if you run a Bayesian inference with sequence data, for macroscopic species there is rarely the need to increase the prior for Theta beyond 0.1; but if you use microsatellites it is rather common that your prior distribution for Theta should have a range from 0.0 to 100 or more. With many populations (>3) it is also very common that some migration routes are estimated poorly because the data contains little or no information for that route. Increasing the range will not help in such situations, reducing number of parameters may help in such situations.

No warning was recorded during the run