Supplement 1: Unified framework to evaluate panmixia and gene flow among multiple sampling locations using marginal likelihoods

Peter Beerli and Michal Palczewski

November 27, 2009

1 Comparison of simulated datasets: Expanded tables 2 and 3

Tables 2 and 3 in the article are abridged versions that do not highlight the strength of rejection of particular models. We present the full tables here and also give in Table 1S (Kass and Raftery 1995) the interpretation of the strength of support for the different values of the LBF. Table 2S and 3S give a more detailed answer than Table 2 and 3, but do not change the interpretation of the results. Unidirectional models models have high support even when the migration direction is incorrect when the number of parameters is small compared to the true model. Highest support among the incorrect models is given to the model with the correct migration direction and with constrained population sizes. It is worth noting that this support has a clear trend in the thermodynamic integration scheme; the harmonic mean estimator does not show such a trend, but shows a high variance.

Table 1S: Bayes factors and strength of acceptance of a model in comparison to a reference model (**?**). BF_{M_2,M_1} is the Bayes factor of model 2 versus model 1

$LBF_{M_2,M_1} = \log_e(\mathrm{BF}_{M_2,M_1})$	Evidence against Model 1
0 to 1	weak
1 to 3	positive
3 to 5	strong
>5	very strong

Table 2S: Comparison of the influence of the approximation on the power of LBF for simple models with different migration schemes. LBF compared a full model (Model $M_1 = [xxxx]$) with a panmictic population (Model $M_0 = [x]$). Models used to simulate the data were: (1a) a single population ($Nm \to \infty$), the sampled individuals were split randomly into two sets; (1b) two populations exchanging many migrants ($Nm = 1250$); (2a) two population exchanging a moderate number of migrants ($Nm = 0.25$); and (2b) two populations with very low migration rate ($Nm = 0.0025$). The marginal likelihoods used in the LBF were approximated with thermodynamic integration (TI) with 16 and 4 temperature bins and with the harmonic mean (HM_4). The reported counts are the number of replicates that fall into the categories outlined in Table **??**

Evidence (M_0: one population)		Counts [based on $LBF_{TI_{16}}$, LBF_{TI_4}, LBF_{HM}]											
	Model	(1a)			(1b)			(2a)			(2b)		
	Nm	∞			1250			0.25			0.0025		
	Method	16	4	H	16	4	H	16	4	H	16	4	H
against M_0	very strong	0	1	0	0	2	0	32	34	0	100	100	59
	strong	0	1	1	0	2	0	10	6	1	0	0	3
	positive	0	0	5	0	0	5	18	5	18	0	0	11
	weak	0	3	20	0	4	24	10	4	34	0	0	5
	Total	0	5	26	0	8	29	70	49	53	100	100	78
against M_1	weak	1	1	40	4	1	38	13	4	33	0	0	5
	positive	6	3	31	4	3	30	10	3	13	0	0	2
	strong	35	8	5	32	8	3	7	7	1	0	0	0
	very strong	57	82	0	60	80	0	0	37	0	0	0	15
	Total	100	94	73	100	92	71	30	51	47	0	0	22

Table 3S: Summary of support for specific models using LBF approximated with harmonic mean (HM) and thermodynamic integration (TI) using 16 chains with different temperatures. 100 single-locus data sets were analyzed, each with a total of 20 DNA sequences simulated using a 3-parameter model with 2 different population sizes, and unidirectional migration from population 2 to 1 (Model abbreviation is xx0x; see Methods for details). All other models 1 to 8 (M_i), such as the full model (xxxx) or the minimal model (mmmm) are compared with this 'true' model (xx0x) that represent the M_0 hypothesis. n_{param} accounts for the number of parameter estimated.

Evidence ($M_0 = xx0x$)		Counts [based on LBF_{TI} and LBF_{HM}]															
n_{param}		4		3		3		3		2		2		2		1	
Model		xxxx		x0xx		xmmx		mxxm		mx0m		m0xm		mmmm		x	
Approximation		TI	HM	TI	HM	TI	HM	TI	HM	TI	HM	TI	HM	TI	HM	TI	HM
against M_0	very str.	0	1	0	0	0	3	0	4	0	1	0	1	0	1	9	10
	strong	0	4	1	2	0	1	0	1	0	4	0	3	0	6	0	1
	positive	0	22	3	13	0	23	0	27	20	21	17	17	0	28	0	16
	weak	0	19	24	21	0	21	0	25	50	24	37	14	0	24	2	13
		0	46	28	36	0	48	0	57	70	50	54	35	0	59	11	40
against M_i	weak	0	26	38	18	0	24	0	17	22	16	24	24	0	19	0	19
	positive	2	21	31	31	2	21	1	23	7	25	20	26	1	18	18	23
	strong	66	5	3	6	63	4	46	3	0	5	1	10	44	3	18	4
	very str.	32	2	0	9	35	3	53	0	1	4	1	5	55	1	53	14
		100	54	72	64	100	52	100	43	30	50	46	65	100	41	89	60
Different data sets		76	76	97	97	89	89	88	88	99	99	99	99	98	98	96	96

2

2 Run conditions for Figure 1

Ten artificial two-population data sets were created with the programs migtree and migdata using the following settings:

Mutation model	F84-model with transition/transversion ratio=2.0	
Mutation rate	2×10^{-6}	
Sequence length	1000	
Population model	Population 1	Population 2
Population size $N_e^{(i)}$	625	1250
Immigration rate m_{ji}	0.0002	0.0
Sample size	10	10

Each data set was run under 3 different heating schemes with the following temperature settings:

Chains	Temperature settings $T_i = 1/t_i$. Ordering is $T_1...T_n$
4	1.0, 1.5, 3.0, 1000000.0
16	1.0, 1.071, 1.154, 1.25, 1.364, 1.5, 1.667, 1.875, 2.143, 2.5, 3.0, 3.75, 5.0, 7.5, 15.0, 1000000.0
32	1.0, 1.03, 1.069, 1.107, 1.148, 1.19, 1.24, 1.29, 1.35, 1.41, 1.48, 1.55, 1.63, 1.72, 1.82, 1.94, 2.07, 2.21, 2.38, 2.58, 2.82, 3.10, 3.44, 3.875, 4.429, 5.167, 6.2, 7.75, 10.33, 15.5, 31.0, 1000000.0

All other settings were at the default values except the following:

Increment (sampling every x state)	1,000
Sampled states	20,000
Discarded states	1,000,000

3 Run conditions for Figure 3

One random dataset from the artificial data sets used in Figure 1 was used. Same temperatures as for Figure 1, but run parameters where changed to

Relative run-length	Increment (sampling every x state)	Sampled states	Discarded states
1	100	200	10,000
2	100	400	20,000
4	100	800	40,000
8	100	1,600	80,000
16	100	3,200	160,000
32	100	6,400	320,000
64	100	12,800	640,000
128	100	25,600	1,280,000
256	100	51,200	2,560,000

4   Run conditions for Figure 4

Run conditions were identical to figure 1.

5   Run conditions for Table 2

100 artificial data sets for the Model *xx0x* were generated:

Population model	Parameters	Population 1	Population 2
Two populations (Model xx0x)	Population size $N_e^{(i)}$	625	1250
	mutation rate	2×10^{-6}	
	Immigration rate m_{ji}	0.0002	0.0
	Sample size	10	10

The mutation model was F84 with a mutation rate of 0.000002. Each sequence was 1000 base pairs long. All run parameters were identical to Figure 1, but the runs used different population models as indicated in Table 2. The runs were executed on the High-performance cluster at Florida State University using the commonly available backfill queue. This queue allows runs maximally 4 hours long, which resulted in some table cells with fewer than 100 runs. A total of 900 runs were executed for Table 2.

6   Run conditions for Table 3

100 artificial data sets for each of the following population models were generated:

Population model	Parameters	Population 1	Population 2
Single population (Model 1a)	Population size $N_e^{(i)}$	1250	-
	Mutation rate	2×10^{-6}	
	Immigration rate m_{ji}	-	-
	Sample size	20	
Two populations (Model 1b)	Population size $N_e^{(i)}$	625	625
	Mutation rate	2×10^{-6}	2×10^{-6}
	Immigration rate m_{ji}	1.0	1.0
	Sample size	10	10
Two populations (Model 2a)	Population size $N_e^{(i)}$	625	625
	Mutation rate	2×10^{-6}	2×10^{-6}
	Immigration rate m_{ji}	0.0002	0.0002
	Sample size	10	10
Two populations (Model 2b)	Population size $N_e^{(i)}$	625	625
	Mutation rate	2×10^{-6}	2×10^{-6}
	Immigration rate m_{ji}	0.000002	0.000002
	Sample size	10	10

The mutation model was F84 with a mutation rate of 0.000002. Each sequence was 1000 base pairs long. All run parameters were identical to Figure 1. Each data set was run twice for each of the approximation methods (TI$_4$, TI$_{16}$), with the single population model *x* and with the unrestricted two-population model *xxxx*.

7 Run conditions for Table 4: Effect of number of loci on Bayes factors

All parameter settings were default, except	
Prior distribution for mutation-scaled population size	Uniform with range 0.0 to 0.1
Prior distribution for mutation-scaled migration rates	Uniform with range 0.0 to 1000
Increment between samples	100
Samples per replicate	1,000
Burn-ins per replicate	100,000
Replicates	10
Heating	static with temperatures 1, 1.5, 3, 10^6

8 Run conditions for Table 5: Effect of prior distribution on Bayes factors

All parameter settings were default, except		
Type	Priors for	
	Mutation-scaled population size	Mutation-scaled migration rates
	Minimum – Mean – Maximum	Minimum – Mean – Maximum
Uniform narrow	0 – 0.05 – 0.1	0.0 – 2500 – 5000
Uniform wide	0 – 0.25 –0.1	0.0 – 25,000 – 50,000
Exponential narrow	0 –0.01 – 0.1	0.0 – 100 – 5,000
Exponential wide	0 – 0.1– 0.5	0.0 – 2,000 – 50,000

9 Run conditions for Table 6: Humpback whale example

Mutation model	F84-model
Transition/transversion ratio	11.400000
Site rate modifier (3 groups)	0.416751 2.274676 6.216591
Probabilities of site rates	0.708460 0.280989 0.010551
Prior distribution for mutation-scaled population size	Uniform with range 0.0 to 0.1
Prior distribution for mutation-scaled migration rates	Uniform with range 0.0 to 5000
Increment between samples	200
Samples per replicate	5,000
Burn-ins per replicate	100,000
Replicates	50

Proposal distribution for parameters was Slice-sampling, whereas the genealogy proposals were using Metropolis-Hastings.