

Chaotic genetic patchiness without sweepstakes reproduction in the shore crab *Hemigrapsus oregonensis*

Brendan H. Cornwell*, Jennifer L. Fisher, Steven G. Morgan, Joseph E. Neigel

*Corresponding author: bhcornwell@ucdavis.edu

Marine Ecology Progress Series 548: 139–152 (2016)

Supplement Ima2 Analysis

IMa2 uses Markov chain Monte Carlo simulations of gene genealogies to estimate posterior distributions of population genetic parameters (Hey & Nielsen 2007, Hey 2010b). Two M mode runs were used to generate a total of 137,278 gene genealogies for subsequent analysis. Priors for both runs were based on a value of 10 for θ ($2N\mu$), corresponding to the value of θ estimated by Waterson's (1975) method as implemented in Arlequin 3.5.1.2 (Excoffier & Lischer 2010). Each run used 5×10^5 burn-in steps, 20 chains, and a geometric chain heating model with parameters of 0.96 and 0.9. The marginal peak locations of the two independent runs were very similar, indicating convergence. For an initial L mode run a value of 10 was used for θ ($2N\mu$), as above for the M mode runs. Following recommendations in the IMA2 documentation (Hey 2010a), priors for the upper boundaries of population sizes, splitting time and migration rates were set to 5θ , 2θ , and $2/\theta$, respectively. However, with these priors the upper tails of the posterior probability distributions were truncated for some population size parameters, and so ten additional L mode runs were done with values of θ ranging from 10 to 100, with 5θ and 2θ again used as the upper limits on priors for population sizes and splitting time, and for migration rates an exponential prior was used with no upper bound and a mean of 0.01. IMA2 parameter estimates, which are scaled to mutation rate (μ), were converted to demographic estimates. This conversion was based on Petersen's (2007) approximation of two years for generation time in *H. oregonensis* and an estimate of 5% divergence per my for COI in species pairs of grapsoid crab in the genus *Sesarma* separated by the rise of the Isthmus of Panama (Schubart et al. 1998). The rate of divergence was converted to a rate of change for the entire sequence along one branch (9.38×10^{-6} y) by multiplying the value of 5% by the length of the analyzed region (375 bp) and dividing by 2.

The full isolation-with-migration model has 6 demographic parameters: population sizes for the 2 contemporary populations and their ancestral population, rates of gene flow in both directions between contemporary populations, and the time at which the 2 contemporary populations became isolated. We used likelihood ratios to test the full model against 25 nested models with fewer parameters. All the models with fewer parameters were rejected except for one, which differed from the full model (Model 1) only in having migration rates equal in both directions (Model 2).

Table S1. Joint peak locations and posterior probabilities from Ima2 L mode run with the following command line: `-icomined6agg.ima -oLmode_r3.txt -r0 -vRun1 -wmodels.txt -c2 -q150 -t60 -j7 -m0.01`. Model 1 is the full model, with separate parameters for ancestral and 2 daughter population sizes and separate rates of migration between populations. Model 2 has separate parameters for population sizes but a single bidirectional migration rate. Population 0 is COB, population 1 is DRA, 2 is ancestral population. These values were used to for likelihood ratio tests of alternative models rather than parameter estimates.

Model#	log(P)	#terms	df	2LLR	ESS	q0	q1	q2	m0>1	m1>0
1	-531.9	5	-	-	1.032	14.7399	68.0094	0.2468	0.2	0.098
2	-533.2	4	1	2.444	1.033	14.7387	68.008	0.2467	0.1747	[0.1747]
3	-544.5	4	1*	25.14	1	22.7114	65.8806	11.7286	[0.00000]	0
4	-544.5	4	1*	25.14	1	22.7115	65.8804	11.7286	0	[0.00000]
5	-544.5	3	2*	25.14	1	22.7115	65.8804	11.7286	[0.00000]	[0.00000]
6	-555.9	4	1	47.97	1.293	57.0925	[57.0925]	11.428	0	0.0798
7	-556.1	3	2	48.35	1	57.2467	[57.2467]	11.7286	0	[0.00000]
8	-555.9	3	2*	47.97	1.293	57.0925	[57.0925]	11.428	[0.00000]	0.0798
9	-556.1	3	2*	48.35	1	57.2467	[57.2467]	11.7286	0	[0.00000]
10	-556.1	2	3*	48.35	1	57.2467	[57.2467]	11.7286	[0.00000]	[0.00000]
11	-548.6	4	1	33.4	2.113	18.3843	71.7306	[18.3843]	0	0
12	-548.6	3	2	33.4	2.113	18.3843	71.7306	[18.3843]	0	[0.00000]
13	-548.6	3	2*	33.4	2.113	18.3843	71.7308	[18.3843]	[0.00000]	0
14	-548.6	3	2*	33.4	2.113	18.3843	71.7306	[18.3843]	0	[0.00000]
15	-548.6	2	3*	33.4	2.113	18.3843	71.7307	[18.3843]	[0.00000]	[0.00000]
16	-550.5	4	1	37.15	1.414	6.6836	54.4892	[54.4892]	0.2	0.0681
17	-551	3	2	38.11	1.417	6.6837	54.4893	[54.4893]	0.1143	[0.1143]
18	-557.4	3	2*	50.82	1.033	6.3817	56.9215	[56.9215]	[0.00000]	0
19	-557.4	3	2*	50.82	1.033	6.3817	56.9215	[56.9215]	0	[0.00000]
20	-557.4	2	3*	50.82	1.033	6.3817	56.9215	[56.9215]	[0.00000]	[0.00000]
21	-578.6	3	2	93.31	1.004	43.7034	[43.7034]	[43.7034]	0	0
22	-578.6	2	3	93.31	1.004	43.7034	[43.7034]	[43.7034]	0	[0.00000]
23	-578.6	2	3*	93.31	1.004	43.7034	[43.7034]	[43.7034]	[0.00000]	0
24	-578.6	2	3*	93.31	1.004	43.7034	[43.7034]	[43.7034]	0	[0.00000]
25	-578.6	1	4*	93.31	1.004	43.7034	[43.7034]	[43.7034]	[0.00000]	[0.00000]

Table S2 Parameter estimates from marginal peak locations of IM α 2 L mode analysis and corresponding demographic parameter estimates for the full isolation-with-migration model. Model parameters are scaled to θ for the mitochondrial genome in the female population ($2N\mu$). Demographic estimates are based on a mutation rate of 9.38×10^{-6} and a generation time of 2 y. Migration rates are shown as forward in time, rather than back in time towards the coalescent event. COB is Coos Bay, Oregon and DRA is Doran Park in Bodega Harbor, California. Estimates of the number of migrant individuals are the products of estimated migration rates and estimated population sizes.

Parameter	Parameter Estimate	Demographic Estimate
Effective population size (Northern - COB)	23.3	6.21×10^5 females
Effective population size (Bodega Harbor - DRA)	103.	2.75×10^6 females
Effective population size (Ancestral Population)	75.6	2.02×10^6 females
Migration rate (COB to DRA – forward time)	0.00	0.00
Upper 95% Confidence limit	0.029	5.44×10^{-7}
Migration rate (DRA to COB – forward time)	0.00	0.00
Upper 95% Confidence limit	0.023	4.31×10^{-7}
Number of Migrants (COB to DRA – forward time)	0.098	1.50 females
Number of Migrants (DRA to COB – forward time)	0.200	0.268 females

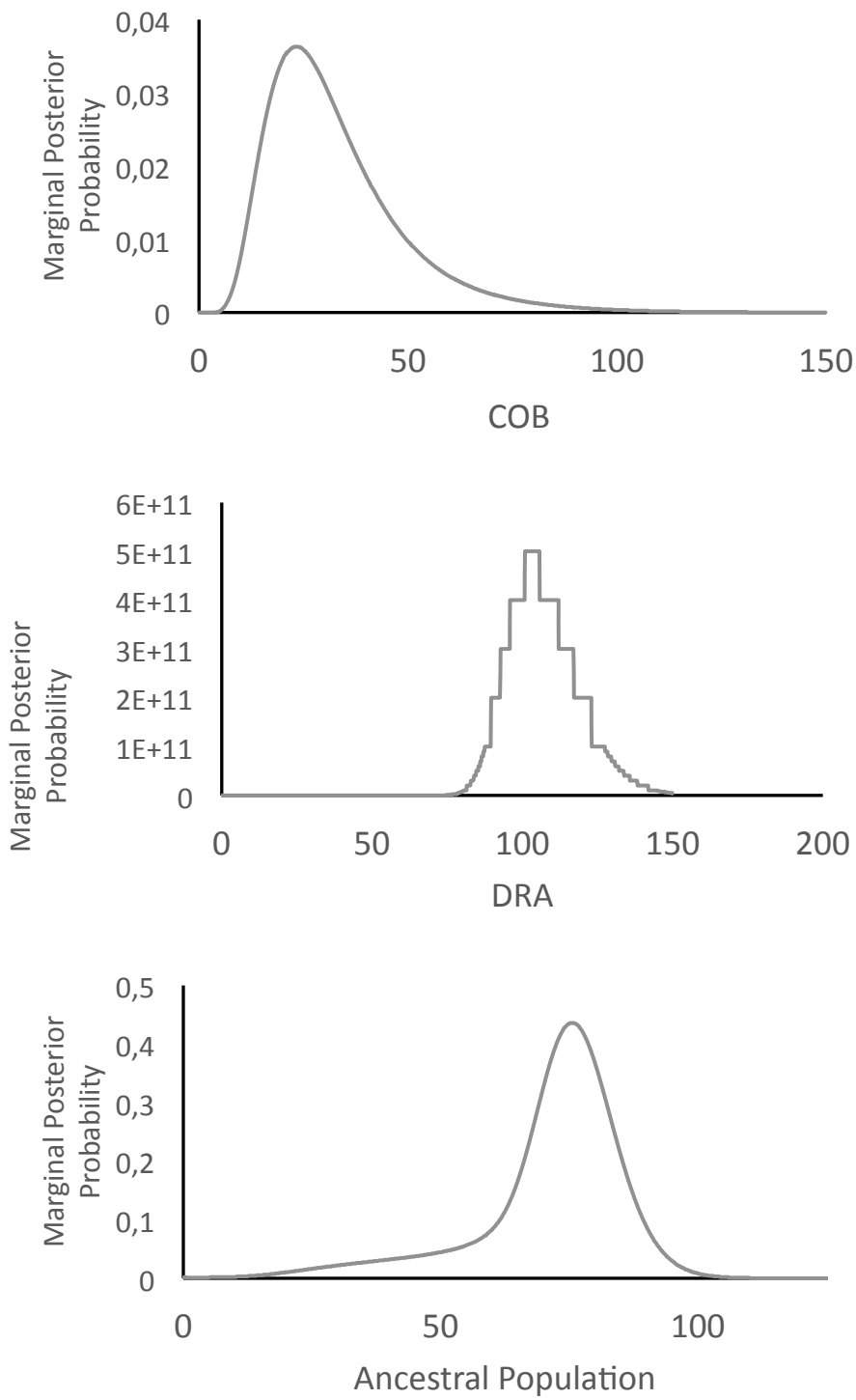


Figure S1. Estimated marginal posterior probabilities for population size parameters from MCMC chains in Ima2 L mode run.

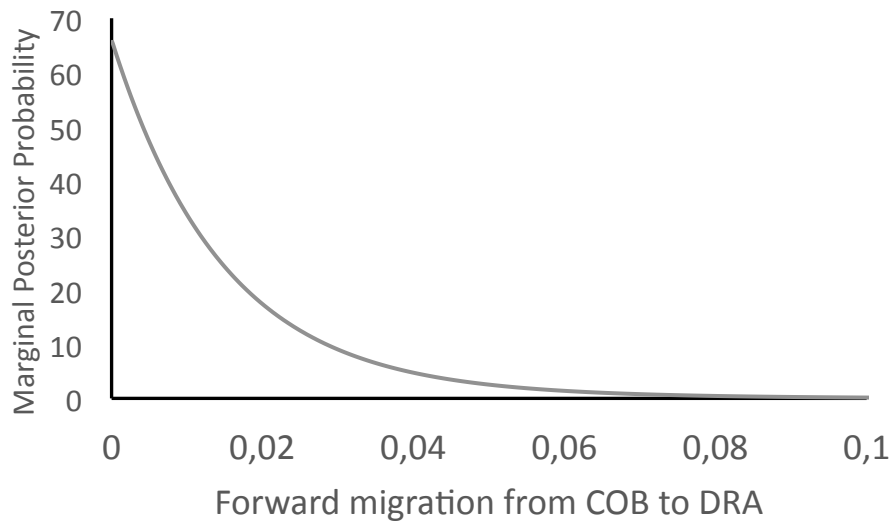
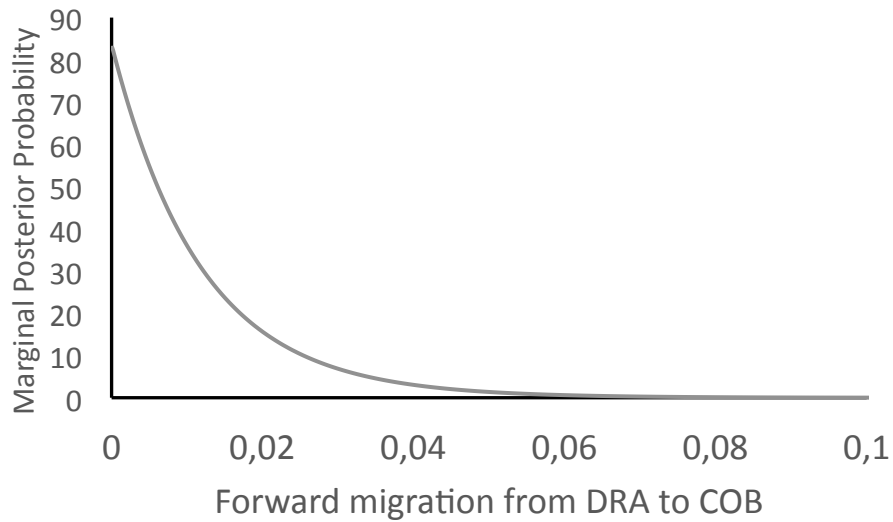


Figure S2 Estimated marginal posterior probabilities for migration rate parameters from MCMC chains in Ima2 L mode run.

LITERATURE CITED

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetic analyses under Linux and Windows *Molecular Ecology Resources* 10(3):564-567

Hey J (2010a) Documentation for IMA2, p 60

Hey J (2010b) Isolation with migration models for more than two populations. *Molecular Biology and Evolution* 27:905-920

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences USA* 104:2785-2790

Petersen C (2007) Historical demography and contemporary spatial genetic structure of an estuarine crab in the northeast Pacific (*Hemigrapsus oregonensis*). *Mar Biol* 150:1289-1300
doi:10.1007/s00227-006-0442-x

Schubart CD, Diesel R, Hedges SB (1998) Rapid evolution to terrestrial life in Jamaican crabs. *Nature* 393:363-365

Waterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 7(2):256-276