Open access • Journal Article • DOI:10.1111/MEC.12031

Coalescent and biophysical models of stepping-stone gene flow in neritid snails.

Eric D. Crandall, Eric A. Treml, Eric A. Treml, Paul H. Barber ...+1 more authors

Institutions: Boston University, University of Queensland, University of Melbourne, University of California, Los Angeles

Published on: 01 Nov 2012 - Molecular Ecology (Wiley)

Topics: Biological dispersal, Isolation by distance, Coalescent theory and Population

2012

Coalescent and Biophysical Models of Stepping-Stone Gene Flow in Neritid Snails

Eric D. Crandall
California State University, Monterey Bay, ecrandall@csumb.edu

Eric A. Treml

Paul H. Barber

Recommended Citation

Crandall, E. D., Treml, E. A. and Barber, P. H. (2012), Coalescent and biophysical models of stepping-stone gene flow in neritid snails. Mol Ecol, 21: 5579–5598. doi:10.1111/mec.12031

MOLECULAR ECOLOGY

Coalescent and Biophysical Models of Stepping-Stone Gene Flow in Neritid Snails

Journal:	*Molecular Ecology*
Manuscript ID:	MEC-12-0420.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Crandall, Eric; Southwest Fisheries Science Center, Fisheries Ecology Division Treml, Eric; University of Melbourne, Melbourne Barber, Paul; University of California, Los Angeles, Dept of Ecology and Evolutionary Biology
Keywords:	isolation-by-distance, habitat preference, model selection, amphidromy, Indo-Pacific, South Pacific

Coalescent and Biophysical Models of Stepping-Stone Gene Flow in Neritid Snails

Eric D. Crandall[1,2*]; Eric A. Treml[3,4]; Paul H. Barber[1,5]

1. Boston University Marine Program, 5 Cummington St., Boston, MA 02215, USA

2. Fisheries Ecology Division, Southwest Fisheries Science Center, Santa Cruz, CA 95060, USA

3. School of Biological Sciences, University of Queensland, St. Lucia, QLD, Australia

4. Department of Zoology, University of Melbourne, Victoria 3010, Australia

5. Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA

*Corresponding Author:

Eric D. Crandall
Southwest Fisheries Science Center
110 Shaffer Rd.
Santa Cruz, CA 95060
Email: eric.crandall@noaa.gov
Tel. +1 (831) 420-3959
Fax +1 (831) 420-3977

Running Title: Stepping-Stone Gene Flow in Neritid Snails

Keywords: Isolation-by-distance, habitat preference, model selection, amphidromy, Indo-Pacific, South Pacific

Abstract

Marine species in the Indo-Pacific have ranges that can span thousands of kilometers, yet studies increasingly suggest that mean larval dispersal distances are less than historically assumed. Gene flow across these ranges must therefore rely to some extent on larval dispersal among intermediate "stepping-stone" populations in combination with long-distance dispersal far beyond the mean of the dispersal kernel. We evaluate the strength of stepping-stone dynamics by employing a spatially explicit biophysical model of larval dispersal in the tropical Pacific to construct hypotheses for dispersal pathways. We evaluate these hypotheses with coalescent models of gene flow among high-island archipelagos in four neritid gastropod species. Two of the species live in the marine intertidal, while the other two are amphidromous, living in freshwater but retaining pelagic dispersal. Dispersal pathways predicted by the biophysical model were strongly favored in 16 of 18 tests against alternate hypotheses. In regions where connectivity among high-island archipelagos was predicted as direct, there was no difference in gene flow between marine and amphidromous species. In regions where connectivity was predicted through stepping-stone atolls only accessible to marine species, gene flow estimates between high-island archipelagos were significantly higher in marine species. Moreover, one of the marine species showed a significant pattern of isolation-by-distance consistent with stepping-stone dynamics. While our results support stepping-stone dynamics in Indo-Pacific species, we also see evidence for non-equilibrium processes such as range expansions or rare long-distance dispersal events. This study couples population genetic and biophysical models to help to shed light on larval dispersal pathways.

Introduction

It has long been believed that disjunct populations of broadly distributed marine species maintain genetic and demographic coherence through dispersal of planktonic larvae on ocean currents (Thorson 1950; Scheltema 1971). However, over the last decade a new paradigm has emerged in which the majority of larval dispersal[1] is thought to be much more spatially limited (Cowen *et al.* 2000; Swearer *et al.* 2002). There is mounting evidence that many species retain some proportion of their larvae within local populations (Jones *et al.* 1999; Swearer *et al.* 1999; Jones *et al.* 2005; Almany *et al.* 2007), and estimates of dispersal distance from direct and indirect methods suggest that larvae consistently disperse on smaller spatial scales than expected based on their pelagic larval duration (PLD; Barber *et al.* 2000; Palumbi 2003; Kinlan and Gaines 2003; Taylor & Hellberg 2003; Shanks 2009).

At the same time, many marine species have remarkably large ranges, demonstrating a clear potential for gene flow and biogeographic dispersal across enormous distances (Mora *et al.* 2012). In the Indo-Pacific, a biogeographic region that spans two oceans from East Africa to Easter Island (Ekman 1953; Spalding *et al.* 2007), many neritic marine species have distributions with maximum linear distances well over 10,000 km (Roberts *et al.* 2002; Lester & Ruttenberg 2005), with individual populations on islands or continental shelves separated by large expanses of open ocean. Nevertheless, many neritic Indo-Pacific species have little or no genetic structure and share mtDNA haplotypes across large portions of their ranges (Craig *et al.* 2007; Crandall *et al.* 2008; Horne *et al.* 2008; Reece *et al.* 2010; Eble *et al.* 2011).

[1] We follow Lowe & Allendorf (2010) in defining larval dispersal as movement and successful recruitment of larvae between spatially distinct and extant populations. We define migration as the population genetic consequence of such dispersal. We distinguish this type of dispersal from biogeographic dispersal, which results in the expansion of species ranges.

How then is genetic connectivity maintained across tens of thousands of kilometers in the Indo-Pacific if the scale of larval dispersal is two orders of magnitude less? One part of the solution to this apparent enigma is that between 1 and 10 effective migrants per generation can limit genetic differentiation between two populations (as measured by F_{ST}; Wright 1931; Lowe and Allendorf 2010). Therefore, even a few successful larvae may provide sufficient genetic connectivity across broad spatial scales over evolutionary time (Waples 1998; Hedgecock *et al.* 2007). A second part of the answer lies in the probabilistic nature of larval dispersal. The distances traveled by the larvae released from a given locality can be modeled as a probability distribution (the dispersal kernel). For timescales greater than ~40 years, the number of oceanographically independent releases of larvae will create a relatively smooth and anisotropic dispersal kernel, the mean and variance of which is primarily determined by the mean velocity of the current and its fluctuating components, (summarized as eddy diffusion), the pelagic larval duration (PLD) of the larvae, mortality in the plankton, and adult fecundity (Largier 2003; Siegel *et al.* 2003). Although average dispersal may be limited, a small proportion of larvae will always disperse far beyond the mean, potentially providing enough gene flow to maintain genetic cohesion even at large distances over evolutionary time scales (i.e. leptokurtic long-distance dispersal, Case II LDD; Kinlan *et al.* 2005). Moreover, it may be that stochastic events at various temporal and spatial scales can create dispersal events that go far beyond what can be modeled (Richmond 1990; Lessios & Robertson 2006; Siegel *et al.* 2008).

While dispersal of a small proportion of exceptional larvae may help promote genetic connectivity across broad ranges of the ocean, an additional important factor is the existence of intermediate "stepping-stones"— areas of available adult habitat that provide generational layovers between the end of one dispersal event and the beginning of another. The existence of a

potent biogeographic break at the "Eastern Pacific Barrier", a 5000 km wide region of the Pacific that lacks any sort of shallow-water habitat (Ekman 1953; Vermeij 1987) demonstrates that there are limits to the maximum larval dispersal distance of most species, and that the absence of stepping-stone populations can represent a significant barrier to dispersal and gene flow (but see Lessios & Robertson 2006 for species that cross even this barrier).

Evidence for stepping-stone facilitated dispersal in the marine environment comes from a number of species of fish, crustaceans, and echinoderms where populations spanning the Indo-Pacific region exhibit a pattern of isolation by distance (IBD) at various spatial scales (Nishida & Lucas 1988; Lavery *et al.* 1996; Palumbi *et al.* 1997; Williams & Benzie 1998; Planes & Fauvelot 2002; Thacker 2004; DeBoer *et al.* 2008; Pinsky *et al.* 2010), consistent with the predictions that alleles must pass through intermediate stepping-stone populations over several generations (Wright 1943; Kimura & Weiss 1964). However, the detection of IBD is relatively rare; no more than 200 cases have been found out of thousands of studies on marine population genetics (Kinlan & Gaines 2003; Weersing & Toonen 2009; Selkoe & Toonen 2011). Conversely, many other studies from the Indo-Pacific support non-equilibrium expansions or rare long-distance dispersal events that directly link distant populations, often causing genetic data to depart from expectations of stepping-stone gene flow (e.g. Benzie & Williams 1997; Bernardi *et al.* 2001; Lessios *et al.* 2003; Kirkendale & Meyer 2004; Craig *et al.* 2007; Eble *et al.* 2011). Thus, while empirical studies seemingly provide evidence for both processes, the relative importance of long distance dispersal and stepping-stones in maintaining connectivity in marine species remains unclear.

An ideal test of the importance of stepping-stones would be to compare gene flow across a common environment in a set of species that are similar in dispersal characteristics but differ

greatly in their adult habitats such that stepping-stones for some taxa are uninhabitable for others. The gastropod family neritidae provides such a comparison. Neritid snails occur abundantly throughout the tropics and include both marine and freshwater genera (Holthuis 1995). Interestingly, most of the freshwater lineages have retained a dispersive, planktotrophic marine larval stage (i.e. they are amphidromous; McDowall 2007). Weakly swimming veliger larvae from both marine and amphidromous lineages have a PLD that is estimated to be from 55 to more than 90 days (Scheltema 1971; Underwood 1975; Holthuis 1995; Kano 2006), and genetic studies confirm that larvae from amphidromous lineages are capable of pelagic dispersal across broad expanses of open ocean (Hodges & Allendorf 1998; Myers *et al.* 2000; Crandall *et al.* 2010). In the South Pacific, neritid larvae are much more likely to encounter a reef or atoll than a high island with freshwater streams, and as such there are more stepping stones to facilitate dispersal in marine species than amphidromous species. Therefore, if stepping-stones are important to genetic connectivity, gene flow should be greater between populations of marine neritids than between those of amphidromous neritids.

The first step in testing the effect of stepping-stones on gene flow is making clear predictions for regions where stepping-stones will be important, based on a dispersal kernel for the target species. A variety of biophysical models have been developed for this purpose by integrating physical ocean data (currents) and larval biology (Cowen *et al.* 2006; Treml *et al.* 2008; Mitarai *et al.* 2009). Some have been used, with varying degrees of success, to predict the often chaotic patterns of genetic structure in the sea (as measured by F_{ST}, Nei's D, or clustering algorithms; Galindo *et al.* 2006; Selkoe *et al.* 2010; White *et al.* 2010; Foster *et al.* 2012). However, genetic structure is sensitive to many other factors at the population and community levels (Hedrick 2005; Selkoe *et al.* 2010, Faurby and Barber 2012) and is often a poor proxy for

the parameter of interest, which is gene flow (Whitlock & McCauley 1999). Coalescent genealogy samplers provide a way to disentangle gene flow from other parameters by simulating an explicit population genetic model over a large sample of possible genealogies (Kuhner 2009). Because they are based in a probabilistic coalescent framework, these programs also offer the possibility of model selection and an appropriate assessment of error, which makes them ideal for testing predictions from biophysical models over long timescales.

In this study we use coalescent models of gene flow to investigate the relative importance of long-distance dispersal and stepping-stones in facilitating genetic cohesion among populations of widely distributed snail species in the Indo-Pacific. To this end, we employ a biophysical model of larval dispersal potential in the tropical Pacific to construct a network of most probable dispersal pathways among South Pacific high island archipelagos. We test this hypothetical matrix against alternatives, and then use Bayesian estimates of migration parameters to compare levels of gene flow among two marine and two amphidromous species from the neritidae. In regions where the biophysical model predicts connectivity between high islands (which contain freshwater habitat) through a single dispersal event, we expect to see high rates of gene flow in all four species. In regions where the model predicts that dispersal events must pass through one or more intermediate atoll stepping-stones that lie between sampled populations, we expect to see reduced gene flow in the marine species (due to isolation-by-distance), and negligible gene flow in the amphidromous species, which cannot use atolls as stepping-stones.

Materials and Methods

Study System

The marine snails *Nerita plicata* and *Nerita albicilla* and their amphidromous relatives *Neritina canalis* and *Neripteron dilatatum*[2] differ in adult habitat, but both retain marine pelagic larval dispersal. All species co-occur in the islands of the South Pacific (Figure 1), but *Nerita albicilla* does not occur to the east of Rarotonga, and *Neripteron dilatatum* does not occur to the east of the Society Islands. Although Rarotonga has freshwater streams, neither amphidromous species occurs there (D. Winter, G. McCormack, M. Frey, personal communication). The adults of both marine species are found at high densities (~ 10/m^2) only on rocky intertidal substrate (Vermeij 1971), while adults of the amphidromous species are found at high densities (> 20/m^2) in streams or estuaries (Liu & Resh 1997). Since they live almost exclusively on rock substrate, none of the species are subject to rafting events, meaning that pelagic larvae are likely their only means of dispersal.

Biophysical Model

We used a spatially explicit model of larval dispersal in the South Pacific (Treml *et al.* in press) to construct a hypothesis of potential dispersal pathways for the region highlighting where, and to what degree, atoll stepping-stones would facilitate gene flow among populations. The model simulates larval dispersal between all coral reef habitat patches throughout the Tropical Pacific (584 individual patches). Each dispersal simulation tracked a "cloud" of larvae (i.e. the dispersal kernel), with the equivalent of 1 million effective larvae released per square kilometer

[2] *Neripteron dilatatum* was incorrectly referred to as *Neripteron dilatatus* in Crandall *et al.* 2010.

158 of coral reef habitat. After release, the larval cloud was allowed to drift throughout the Tropical
159 Pacific on the ocean currents derived from the 12.5 km^2 Regional Ocean Modeling System
160 (Wang & Chao 2004). An advection transport algorithm (Smolarkiewicz & Margolin 1998) was
161 used to disperse the larval cloud through the ocean currents (see Treml *et al.* in press for model
162 framework). Both marine and freshwater neritid species have been observed to lay eggs at highest
163 densities during Austral Spring (Underwood 1975; Resh *et al.* 1992), so we modeled larval
164 dispersal from October through December. The larval cloud was allowed to drift with weak
165 swimming ability and no homing behavior for a maximum pelagic larval duration (PLD) of up to
166 90 days (Underwood 1975; Holthuis 1995; Kano 2006). To explore potential inter-annual
167 variability in ocean currents, we completed simulations using current velocities from an El Niño
168 (1997) and a La Niña (1999) year, as well as from a 'neutral' year (2001).
169 We used the results of these dispersal simulations to quantify the dispersal probabilities
170 from each locality to every other locality over time, where the dispersal probability is the
171 probability of a larva arriving at a downstream habitat patch after being released from a source
172 patch. Dispersal probabilities from all three years were combined and weighted to quantify the
173 maximum likelihood that larvae could pass between sample sites either directly or via
174 intermediate stepping-stones during any ENSO state (see methods in Treml *et al.* 2008). We
175 explored potential dispersal probabilities using a realistic larval mortality coefficient of 6% per
176 day (Rumrill 1990; Johnson & Shanks 2003; Nishikawa & Sakai 2005). We tracked dispersal
177 probabilities greater than 1×10^{-12} or 1 out of a trillion larvae released per generation from the
178 upstream site. This extremely low threshold reflects the sensitivity of genetic structure to even
179 small amounts of gene flow (0.1 to 10 female migrants per generation; Lowe and Allendorf

2010). We used the resultant maximum dispersal probability matrix to represent the potential larval connectivity of the South Pacific (Supplemental Tables S1-S3).

Genetic Data and Analyses

To evaluate the predictions of the larval dispersal model, we collected data from mitochondrial cytochrome oxidase C (subunit 1) data from all four species. Existing data for marine (658 bp, Crandall *et al.* 2008) and amphidromous (520 bp; Crandall *et al.* 2010) species were augmented with additional data following previously published protocols (Crandall *et al.* 2008). Additional samples include *Nerita plicata* from the islands of Espiritu Santo and Tanna in Vanuatu, Taveuni in Fiji, Upolu in Samoa, Huahine in the Society Islands and Hiva Oa in the Marquesas and *Nerita albicilla* from Tanna, in Vanuatu. Samples from multiple islands within an archipelago (Figure 1, Table 1) ensure that we sampled as much intra-archipelagic variation as possible, and avoided potential Wahlund effects when analyzing gene flow among archipelagos.

We ran a preliminary analysis on each dataset using Arlequin 3.1 (Excoffier *et al.* 2005) to estimate minimum spanning trees and measure standard diversity indices and pairwise Φ_{ST} and F_{ST}, as well as Fu's F_s test of neutrality (Fu 1997). Significance of pairwise Φ_{ST} and F_{ST} values were tested with 10,000 random permutations of the data, with final p-value determined by a Bonferroni correction. There were no significant Φ_{ST} or F_{ST} values among islands within an archipelago for any of the four species, so we combined COI data from all localities within a given archipelago for use with coalescent estimates of gene flow, described below. Due to spurious Φ_{ST} values arising from deep divergences between clades (Bird *et al.* 2011), we only report F_{ST} values.

Coalescent Analyses

To evaluate empirical evidence for the dispersal pathways predicted by the biophysical model against alternative models in a Bayesian model-selection framework, we set up five to seven models of the structured coalescent for each species using the Bayesian implementation of Migrate 3.2.6 (Beerli & Felsenstein 2001). The stepping-stone migration matrix for the initial model followed predictions from the biophysical model (Figure 2a). A migration parameter (m/μ, where m is the fraction of migrants and μ is the mutation rate) between two sampled archipelagos was estimated if the biophysical model predicted that it would provide the smallest number of steps between them and each step had a dispersal probability greater than 1×10^{-12} (see above). Where the biophysical model indicated connectivity in both directions between archipelagos, we added migration parameters for both directions. Each sampled archipelago had an independent Θ ($=N_e\mu$) parameter. Because unsampled populations can have an unpredictable effect on parameter estimation (Beerli 2004; Slatkin 2005), we included a "ghost" population that contained no genetic data. This population had constant $\Theta = 1.0$, and exchanged migrants with Western Pacific populations at a constant rate of $m/\mu = 100$ (an expected value of 25 female migrants per generation).

We tested simpler and more complex alternatives to the hypothesis of stepping-stone gene flow predicted by the biophysical model by setting up alternate migration matrices in Migrate. For all four species we tested simpler models of panmixia (a single Θ parameter) and an island model (all populations share a single mean estimate of Θ and exchange genes with all other populations at the same mean rate). We also tested migration models that excluded all migration parameters running from west to east, against the prevailing flow of the South Equatorial Current, even though these connections were predicted to be possible under the biophysical model. For *N.*

plicata and *N. canalis*, we also tested the possibility of an additional connection (not predicted by the biophysical model, but suggested by previous analyses in both species; Crandall et al. 2008, 2010) between the Marquesas and Samoa that did not run through the Society Islands (Figure 2b, parameter L). Finally, we tested the most general model of migration under which we made independent estimates of Θ and pairwise gene flow among all sampled populations (diagrams of all models are given in Supp. Figure S1).

The Bayesian version of Migrate uses Metropolis-coupled Markov chain Monte Carlo methods (MC^3) to sample over all possible genealogies given a model and the data and returns posterior distributions for each parameter that reflect how often each parameter value was visited (Beerli 2006). We ran Migrate analyses under an F84 mutational model, the parameters of which were determined for each species on a neighbor-joining tree by PAUP* 4.b10 (Supp. Table S2, Swofford 2002). After several exploratory runs, we chose a windowed exponential prior for Θ and m/μ, the bounds of which are given in Table S2. We conducted MC^3 searches of parameter space using four chains with relative temperatures of 1.0, 1.5, 3.0, and 10000.0. The marginal likelihood of each model (i.e. $L(Model_i) = P(Data|Model_i)$ was calculated using the thermodynamic integration method implemented in Migrate (Beerli & Palczewski 2010) which takes advantage of the large area of parameter space searched by the four chains. We calculated Log Bayes Factors (LBF) for each model i as $2(\ln(L_i) - (\ln(L_a))$, where L_a was the highest marginal likelihood and interpreted them following Kass and Raftery (1995).

For each species and model we ran two replicates using Markov chains of 20 to 200 million steps, which sampled 1 out of every 200 iterations. The outputs from these replicate runs were checked for convergence in Tracer 1.5 (Rambaut & Drummond 2007) after removing the

first 10 to 50 million steps from each run as burn-in depending on where the marginal likelihood values reached a plateau. Models that had the highest marginal likelihoods for each species were run a third time, and the trimmed output files were merged using LogCombiner 1.6.1, part of the BEAST software package (Drummond et al. 2012). Within each logfile, we created the parameter N_em as the product of the values for $\Theta = N_e\mu$ and m/μ for each sampled step. Modified logfiles were then analyzed by using Tracer to estimate 95% highest posterior density (HPD) intervals for each parameter and estimate effective sample sizes (ESS) to determine whether the chains had mixed sufficiently.

Comparison of Gene Flow Across Stepping-Stones in Marine and Amphidromous Species

To test the hypothesis that marine species will have higher rates of gene flow than amphidromous species across regions requiring atoll stepping-stones we considered what fraction of the proportion of migration (m/μ) posteriors were greater in the marine species for these regions. We used the sample function in R 2.11.1 (2010) to take 10,000 random samples of the m/μ parameters from the posterior distributions generated for each species by Migrate for the biophysical-based model (we did not use N_em here, to avoid correlations arising from incorporation of N_e). We divided each random sample of m/μ values into 6 groups: 1) the proportion of migrants across stepping-stone atolls in marine species (n=5), 2) the proportion of migrants across stepping-stone atolls in amphidromous species (n=3), 3) the proportion of migrants directly exchanged between high-island archipelagos in marine species (n=13), 4) the proportion of migrants directly exchanged between high-island archipelagos in amphidromous species (n=6), 5) a random selection from m/μ values in all 4 species (n=5), 6) another random

selection from all 4 species (n=5). We made all possible combinations of values from a given sampling operation and tested the above stated hypothesis by simply taking the difference of values between group 1 (marine) and group 2 (amphidromous). The fraction of differences for which marine gene flow is higher than amphidromous gene flow can be taken as the probability that gene flow in marine species is higher across stepping-stone regions. We further tested the hypothesis that direct gene flow between high island archipelagos is greater in marine species than it is in amphidromous species by taking the difference of group 3 (marine) and group 4 (amphidromous). We represented the "null" hypothesis that there is no difference between marine and amphidromous groups by taking the difference between group 5 (random) over group 6 (random), and tested for significant differences from this null distribution using a one-tailed Kolmogorov-Smirnov test in R.

Stepping-Stone Gene Flow in Nerita plicata *and* Nerita albicilla

To determine whether gene flow among Pacific Ocean populations of both marine species conforms to a stepping-stone model, we tested for the expected correlation with distance using posterior distributions of the N_em parameter from the biophysical-based coalescent model in Migrate. We did not perform this test for the amphidromous species because they cannot use intermediate atoll stepping-stones. We tried three different distances for the independent variable. First, we used the geographic distance between the two closest points in each archipelago. We also used the minimum stepping-stone distance between archipelagos, defined as the number of larval dispersal events required to connect two archipelagos with highest probability, as measured from dispersal probability networks derived from the biophysical model (Figure 2a, Supp. Table S3). For example, although gene flow between the Marquesas and the Societies is possible in a

292 single event with a probability of 9.36×10^{-11}, it is more probable that it will happen in two
293 dispersal events with joint probability of 3.04×10^{-6}. Finally, we used the joint probability of
294 dispersal among archipelagos calculated as the sum of alternate routes between archipelagoes,
295 each route being the product of dispersal probabilities across all intermediate stepping-stones
296 (Supp. Tables S2 and S3).

297 Continuing in a Bayesian framework, we took 10,000 random samples of N_em parameters
298 from the posterior of both marine species as described above and paired them with the
299 appropriate distances. We then set up 10,000 OLS regressions for the linear equation $\log_{10}(N_em)$
300 $= a + b(\log_{10}(\text{distance})$ using R's lm function. These values were log-transformed due to clear
301 heteroscedasticity in the posteriors for N_em, (Supp. Figure S2) and to bring them in line with
302 theoretical expectations (Slatkin 1993). Before log-transforming, we added 0.0001 to all matrix
303 members that were equal to zero. We used the output of these models to construct posterior
304 distributions for each regression parameter, as well as for the log-likelihood of the model as
305 measured by the logLik.lm function in R.

306 Due to the non-independence of pairwise comparisons of distances, we evaluated the
307 strength and significance of the linear relationship with 10,000 more regressions in which the
308 distances were randomly permuted among N_em parameters, representing a null hypothesis of no
309 relationship of gene flow with distance. This is essentially a Bayesian implementation of a
310 Mantel test. We evaluated the probability that the slope of the linear model is less than zero as the
311 number of instances out of 10,000 for which it was more negative than the slope for the null
312 model.

313

314 **Results**

315 *Gene diversity and genetic structure*

316 We analyzed 658 bp of CO1 sequence data for *Nerita plicata* (342 total sequences), *N.*
317 *albicilla* (152 total sequences), *Neritina canalis* (198 total sequences) and *Neripteron dilatatum*
318 (150 total sequences). There were no non-synonymous changes in the new data for either species.
319 Minimum spanning trees of COI haplotypes had similar topologies for all four species,
320 with multiple star-like polytomies surrounding high-frequency haplotypes (Supp. Figure S2).
321 *Nerita plicata* was exceptional in having two deeply divergent clades, A and B (3.4% average
322 divergence) that occur in a cline across the Pacific, with relatively high frequencies of clade B in
323 the Central Pacific that decline to the west (see results and discussion in Crandall *et al.* 2008).
324 Only *Nerita plicata* and *Neritina canalis* showed significant pairwise F_{ST} values, and only
325 *Neritina canalis* showed F_{ST} values greater than 0.01 between archipelagos (Supp. Table S5).
326 With the exception of two *N. dilatatum* demes, all demes had significantly negative values of
327 Fu's F_s, indicating departures from the Wright-Fisher neutral model (Table 1).
328
329 *Biophysical Model*
330 The dispersal probability matrix showed two well-connected regions, one in the Central
331 Pacific and one in the Western Pacific (Figure 2A, Supp. Table S1). The most probable dispersal
332 events occur from east to west, following the South Equatorial Current. However, particularly in
333 the Western Pacific region, there is a lesser probability of west-to-east dispersal as well. There
334 were also two east-to-west connections for larval dispersal between the Central and Western
335 Pacific regions, with one dispersal route running through Suwarrow Atoll, and the other running
336 through Niue and the Tongan archipelago. Dispersal networks from individual years show that
337 these connections between the Central and Western Pacific only occur during ocean conditions

338 associated with La Niña events. We translated this dispersal probability matrix into two sets of
339 connectivity predictions for the Tropical Pacific, one for the marine neritids and one for the
340 amphidromous neritids (Figure 2B, Supp. Tables S2 and S3). The predictions differed on routes
341 where atoll stepping-stones were required for connectivity. For marine species we calculated the
342 joint probability of dispersal through these stepping-stones as the product of each independent
343 dispersal event, plus the probability of direct dispersal if it was higher than our threshold. For
344 amphidromous species we predicted zero probability of gene flow along these routes due to lack
345 of freshwater habitat on atolls.
346
347 *Coalescent Analyses*
348 Final model runs for all species took approximately 6528 CPU hours on two dual core
349 desktop iMacs. Replicate runs of each model were very similar for each species, usually
350 converging to within one unit of log-likelihood. In cases where the difference between runs was
351 larger than one log-likelihood unit (mostly for the general model), we reported the higher value.
352 After three replicates were combined for the best model in each species, ESS values for each
353 parameter in each species were well above 200, as suggested by the authors of Tracer. Posterior
354 distributions for all parameters can be viewed in Supplemental Figure S3.
355 Results from coalescent models were unanimous in their support of the dispersal
356 pathways predicted by the biophysical model over alterative hypotheses and gave unequivocally
357 strong support in 16 out of 18 cases (Table 2; LBFs can be interpreted on the same scale as
358 likelihood ratio tests; Kass & Raftery 1995). Panmixia was the worst model for all four species,
359 with the biophysical model favored by LBF >> 100 (odds >> 10^{21}:1 against panmixia). The
360 classic island model of equal Θ and migration rates among all population pairs was also strongly

rejected for three species by LBF > 25 (odds > 200,000:1 against the island model). Evidence against the island model was weaker in *N. albicilla*, with LBF = 5.7 which equates to a relative probability for the island model of just over 5%. The migration model delineated by the biophysical model of larval dispersal was also strongly supported over simpler east-to-west-only models (LBF > 6 or favored by odds of > 20 to 1). For the two species with ranges that reach the Marquesas, the addition of a gene flow parameter between the Marquesas and Samoa was very strongly rejected for *N. plicata* but could not be completely rejected for *N. canalis* (LBF = 3.53 or relative model probability of about 15%). Finally, a generalized model of migration (independently estimated parameters for all Θ and m parameters) was very strongly rejected for all species (LBF > 50 or odds of 10^{10} to 1), although these models did not converge well due to a high number of parameters.

Gene flow generally differed between marine and amphidromous species in regions where the biophysical model predicted gene flow through at least one atoll stepping-stone. Between the Marquesas and Society Islands, marine *Nerita plicata* had modal gene flow of about 140 female migrants per generation (95% Highest Posterior Density – HPD was 90.8 to 206.1), while the amphidromous *Neritina canalis* had significantly lower gene flow close to 2 female migrants per generation (95% HPD 0.7 to 4.6). Between the Societies and Samoa, *Nerita plicata* had modal gene flow of about 122 female migrants per generation (95% HPD 53.1 to 1058.0). The two amphidromous species differed significantly in this region with *Neritina canalis* having ~2 effective female migrants and *Neripteron dilatatum* having 67 (95% HPD 28.2 to 151.7). The biophysical model predicted no genetic connectivity for any species between the Marquesas and Samoa, and models with this gene flow parameter were rejected: strongly for *Nerita plicata* (model probability 2.02×10^{-11}) but inconclusively for *N. canalis* (model probability 0.15).

Comparison of Gene Flow Across Stepping-Stones in Marine and Amphidromous Species

A comparison of m/μ across atoll stepping-stones showed that the two marine species had an 85% probability of having a higher proportion of migrants crossing these regions than did the amphidromous species (Figure 3). The modal difference in proportion of migrants between marine and amphidromous species across stepping-stone regions was m/μ = 1027. In regions where direct dispersal was possible between high-island archipelagos, marine species had only a 41% probability of having a greater proportion of migrants, which was lower than the 52% probability found between two random vectors of m/μ values. The modal difference in proportion of migrants between marine and amphidromous species across high island archipelago regions was m/μ = -1540. A Kolmogorov-Smirnov test showed that the difference between marine and amphidromous species was significantly greater than random across stepping-stone regions ($p < 2.2 \times 10^{-16}$) but not significantly greater than random across areas where direct dispersal was possible ($p = 0.15$).

Stepping-Stone Gene Flow in Nerita plicata *and* Nerita albicilla

The mean regression slopes for log-transformed gene flow in *Nerita plicata* were close to 1 for both the stepping-stone and geographic distances, as predicted for equilibrium gene flow under a one-dimensional stepping-stone model (Figure 4a-d; see figure 9b in Slatkin 1993). For both distances, there was a 95% probability that the negative relationship was real, (note that these are Bayesian posterior probabilities, rather than frequentist p-values). The posterior distribution of r^2 for the linear models based on stepping-stone distance had a mean value of 0.37 and an HPD that ranged between 0.04 and 0.6. The r^2 posterior distribution for models based on

geographic distance had a slightly smaller mean of 0.35 and a similar HPD interval. The mean slope of the IBD relationship with inverse dispersal probability was much smaller ($\overline{b}$ = 0.44, Figure 4e,f), with an 86% probability of being larger than it would be under the null hypothesis. The IBD relationship with the three distance metrics was much weaker in *N. albicilla* (Figures 5a-f), with none of the 3 distances significantly different than the null hypothesis.

Discussion

Our results provide three distinct lines of evidence that speak to the importance of intermediate stepping-stones in maintaining genetic connectivity across large species ranges in the Indo-Pacific. First, the biophysical model shows that even dispersal events at the very tail of the dispersal kernel are generally only able to span the distance between neighboring archipelagos, and no further (Figure 2a). The model's predicted dispersal pathways were strongly upheld by the genetic data, which supported these pathways against simpler models of panmixia and island model migration, as well as against more complex models. Second, while all four species had similarly high rates of gene flow in regions where neritid larvae can disperse between high island archipelagos in a single generation, we found that the amphidromous species have significantly lower rates of gene flow than marine species in regions where atoll stepping-stones are required by the biophysical model. Finally, one of the marine species showed a significant decrease in gene flow across stepping-stone regions, as predicted under a model of isolation by distance. We will examine each of these results in turn.

Assessing the Biophysical Model with Genetic Data

Results from the biophysical model of larval dispersal indicate that species with high dispersal potential are able to maintain genetic cohesion between neighboring high island archipelagos through long-distance larval dispersal across much of the Western Pacific where habitat for both marine and freshwater neritid snails is relatively common and closely spaced. In contrast, maintaining genetic connectivity into and among high islands in the Central Pacific required dispersal through intermediate atoll stepping-stones even when considering the extreme tails of the dispersal kernel. More specifically, dispersal across regions of sparse atoll stepping-stones occurs less frequently and relies on variable ocean currents associated with ENSO events, a result that is consistent with the pattern of genetic structure between the Western and Central Pacific that is often found in marine species distributed across this region (e.g. Palumbi *et al.* 1997; Bernardi *et al.* 2001; Lessios *et al.* 2001; Thacker 2004). The results are also consistent with output from a similar biophysical model, which shows multiple independent voyages required for biogeographic dispersal (colonization) across the Pacific (Mora *et al.* 2012).

Coalescent models of gene flow in four neritid snail species confirmed the importance of the dispersal pathways identified by the biophysical model, conforming consistently to its predictions, and significantly so in 16 out 18 tests against alternative hypotheses (Table 2). Models of panmixia were unambiguously rejected, even for the two species that had no significant F_{ST} values. Simple and more general versions of the island model, in which genes are exchanged between all population pairs were also strongly rejected. Interestingly, even simpler models of unidirectional gene flow moving only with the South Equatorial Current from east to west were significantly less probable than the bi-directional set of pathways identified by the biophysical model. Finally, coalescent gene flow estimates were correlated with modeled

451 dispersal probabilities, with an average r^2 of 0.16 in a Bayesian Mantel test. All of these results
452 support predictions of the biophysical model, at least at large spatial and temporal scales.
453
454 *Comparing gene flow in marine and amphidromous neritids*
455 Although use of a single locus yields large uncertainties in the inference of gene flow
456 (Table 3) there was enough information in the data to distinguish different levels of gene flow
457 among the species. In regions where the biophysical model predicts that larvae can potentially
458 move between high island archipelagos in a single dispersal event, results show no significant
459 difference in the proportions of migrants across four species (Figure 3; KS Test $p = 0.15$). In
460 contrast, when the biophysical model predicted that connectivity between high island
461 archipelagos would require dispersal through atoll stepping-stones, the marine species had
462 significantly higher rates of gene flow than amphidromous species (with a modal difference ~
463 1000 times greater than the mutation rate; KS Test $p < 2.2 \times 10^{-16}$). In terms of female migrants
464 per generation, the modal values across stepping-stone regions in the marine species was between
465 90 and 208, while modal N_em fell between 1 and 67 in the amphidromous species (Figure 2C).
466 These results are generally consistent with our predictions that amphidromous species will have
467 lower gene flow where atoll stepping-stones are required because of the absence of suitable
468 freshwater habitat.
469 However, there were also significant departures from biophysical model predictions that
470 require further examination: (1) both amphidromous species have non-zero gene flow between
471 the Society Islands and Samoa, where the biophysical model predicted the need for atoll
472 stepping-stones, and (2) when we added a migration parameter between the Marquesas and
473 Samoa, where atoll stepping-stones are also predicted to be necessary, it could not be statistically

474 rejected by the *N. canalis* dataset. These results are somewhat puzzling given that the
475 amphidromous species are not able to use atoll stepping-stones as the intertidal species do. It is
476 certainly possible that the observed gene flow is the residue of incomplete lineage sorting
477 following a colonization event. However, analysis under the Isolation with Migration model (Hey
478 & Nielsen 2004), which explicitly considers this scenario, was not able to reject migration
479 following the time of population splitting for either species (Crandall *et al.* 2010), and modal
480 estimates of east to west gene flow in both species that were similar in both programs (0.5 vs. 1.5
481 migrants per generation in *N. canalis* and 104 vs 67 migrants per generation for *N. dilatatum*, for
482 IM and Migrate, respectively). We must therefore consider the possibility that the larvae of these
483 amphidromous species are occasionally able to cross this region in a single dispersal event.
484 Given the multiple instances where amphidromous neritid species had higher rates of
485 gene flow than predicted by the model, an explanation that is common to all of them would be
486 most parsimonious. One potential explanation is that neritid larvae may be able to delay
487 metamorphosis far beyond the 90-day period used in the model. Delayed metamorphosis is a
488 well-documented phenomenon in the planktotrophic larvae of invertebrates (Pechenik 1990;
489 Miller & Hadfield 1994). Long-distance dispersal through delayed metamorphosis is likely to be
490 somewhat rare because of significant deferred costs to successful recruitment (Highsmith &
491 Emlet 1986; Pechenik 2006; Burgess et al. 2012). However, given that gene flow estimates are
492 averaged across millennia, even rare events can significantly impact patterns of genetic exchange.
493 Another potential explanation may lie in variability of currents. ENSO events are highly variable
494 (Quinn *et al.* 1998), and current velocities may occasionally exceed those from the three years
495 that were used to drive the physical oceanographic model, decreasing the time required for larvae
496 to cross the area between Central and Western Pacific archipelagos. The combination of these

two factors may have allowed a few extremely long-distance dispersal events (on the scale of one per century) to occur beyond the tails of the dispersal distribution suggested by the biophysical model. Finally, it is important to consider that ocean currents may have been stronger during glacial periods that occurred during the timescale sampled by CO1 mutations (Benzie and Williams 1997).

Stepping-Stone Gene Flow in Nerita plicata*, but not* N. albicilla

Gene flow estimates in *Nerita plicata* were negatively correlated with both geographic and stepping-stone distance suggesting that gene flow in *N. plicata* can be clearly described by a stepping-stone model (Figure 4; Kimura & Weiss 1964; Slatkin 1993). The slopes for both relationships were close to -1.0, as classically predicted for a one-dimensional stepping-stone model that is at a rough equilibrium between gene flow and genetic drift (Slatkin 1993; Hellberg 1995; Hutchison & Templeton 1999). This isolation-by-distance relationship is relatively uncommon in studies of marine population genetics, where gene flow is usually approximated from pairwise F-statistics, and geographic distance is generally used as a proxy for stepping-stone distance (see Selkoe & Toonen 2011 for a discussion of why IBD may sometimes escape detection). However, our use of a parametric estimate of gene flow, and a stepping-stone distance predicted by the biophysical model comes much closer to theory (See Figure 9b in Slatkin 1993), and shows that a stepping-stone model of gene flow can apply even when genetic structure is too small to be measured by traditional means.

The relationship of gene flow with distance is just as clearly absent in *N. albicilla*, indicating that it is further out of gene flow/drift equilibrium than *N. plicata* (Figure 5). This inference is supported by the failure to statistically reject the island model of gene flow for this

520 species. This lack of equilibrium suggests that it has expanded its range more recently than *N.*
521 *plicata* (see Figure 7 in Slatkin 1993), which supports the idea that *N. albicilla* is more sensitive
522 to the environmental fluctuations that occurred during the Last Glacial Maximum. We have
523 suggested this previously as a reason for why *N. plicata* has apparently maintained gene flow
524 across the Coral Triangle, while *N. albicilla* has not (Crandall *et al*. 2008).

525 The parameters of the regression model for *N. plicata* provide useful insight into the
526 dispersal kernel averaged over long timescales. From the intercept of the regression on stepping-
527 stone distance (Figure 4b), we can see that there will be an average of 600 effective female
528 migrants/generation between two populations that can be linked by a single larval dispersal event
529 (with confidence intervals between 400 and 1000). From the slope (Figure 4d), we can project
530 that demes separated by ten stepping-stones will have an effective rate of gene flow of about ten
531 female migrants per generation, which is completely consistent with the absence of measurable
532 genetic structure in *Nerita plicata* from the Marquesas to Africa (Crandall *et al.* 2008). The
533 intercept of the regression on geographic distance (~1 million effective female migrants per
534 generation with confidence intervals of about 2 orders of magnitude on either side) supports our
535 choice of 1 million larvae/km^2 in the biophysical model, while the slope confirms massive
536 reductions in dispersal probability for every extra kilometer dispersed (Buston *et al.* 2011), while
537 still allowing for 25 effective female migrants to cross 10,000 km. That the coalescent model in
538 Migrate is able to estimate gene flow higher than ten migrants per generation and reject panmixia
539 when F_{ST} estimators cannot is worthy of further discussion below.

540 *Coalescent Estimates of Gene Flow in Marine Species*

541 High rates of larval dispersal among Western Pacific archipelagos, as suggested by our
542 biophysical model, can drive down genetic structure, leading to difficulties in estimating gene

543 flow (Waples 1998). Empirical estimates of N_em have traditionally been based on the nonlinear
544 relationship between migrants per generation and some version of Wright's F_{st} ($N_em = [1 -$
545 $F_{ST}]/4F_{ST}$; Wright, 1931). However, the island model underlying this conversion makes a number
546 of assumptions, most notably that both populations have the same effective sizes (N_e) and
547 exchange the same proportion of migrants (m) and that F_{ST} is measured without error (Whitlock
548 & McCauley 1999). Furthermore, high levels of gene flow in marine species can lead to sizable
549 genetic neighborhoods with large effective population sizes that can harbor high levels of genetic
550 diversity (DeWoody & Avise 2000; Palumbi 2004). The resultant high diversity at a sampled
551 locus can lower the maximum value of F_{ST} (Hedrick 2005, Bird *et al.* 2012), making it even more
552 difficult to detect population structure without extremely large sample sizes, even though the
553 population is not strictly panmictic.
554 Unlike the traditional method of converting genetic structure into estimates of gene flow,
555 coalescent simulation methods make full use of the genealogical information in sequence data
556 (Slatkin & Maddison 1989), and can provide an accounting of the uncertainty in the model given
557 the data. Existing models of the structured coalescent (e.g. Beerli & Felsenstein 2001; Hey &
558 Nielsen 2004; Kuhner *et al.* 2005) estimate the two components of N_em separately (scaled by the
559 mutation rate: $\Theta = N_e\mu$ and proportion of migrants = m/μ). It is therefore possible, under the
560 structured coalescent, to estimate relatively high values of N_em, as long as $N_e >> m >> \mu$ (the
561 diffusion limit, or when many sub-populations contribute to Ne, the "many-demes limit";
562 Wakeley 2004; Wakeley & Takahashi 2004). This limit seems to apply for the four species of
563 neritid gastropods in this study, which have a minimum N_e of several hundred thousand and a
564 maximum m of 10^{-4}. These parameters result in estimates of N_em that are much higher than what

565 would be possible using F_{ST}, while still having reasonable confidence intervals in many cases
566 (Table 3). Thus, while estimates of gene flow from F_{ST} values are poorly suited to data from
567 marine species because of their large effective population sizes, coalescent methods may be able
568 to measure high levels of gene flow ($N_em > 10$) with greater precision in marine species than in
569 terrestrial species for exactly the same reason.

570 Nevertheless, coalescent models of gene flow are not immune to the effects of non-
571 equilibrium gene frequencies resulting from range-expansion. Large 95% HPD intervals for most
572 gene flow parameters (Table 3) in *Nerita albicilla* and *Neripteron dilatatum* suggest that both of
573 these species have expanded their ranges more recently than the other two species. This leads to
574 significant uncertainty in gene flow estimates because recent range expansions will mimic high
575 rates of gene flow between distant populations (Slatkin 1993), and growth tends to push
576 coalescence times towards the root of the genealogical tree (Slatkin & Hudson 1991). In addition,
577 Kuhner *et al*. (1998) found a strong correlation between growth rate and Θ, meaning that
578 estimates of Θ in these species are also probably biased upward. Preliminary analysis in
579 LAMARC 2.1.3 (Kuhner 2006) confirms that these two species have significantly larger growth
580 rates than *Nerita plicata* or *Neritina canalis* (results not shown). The inference of non-
581 equilibrium processes is underscored by the absence of a significant relationship of gene flow and
582 any distance metric in *N. albicilla*, and the failure to reject the island model for this species. Data
583 from additional loci will be required before inferences about growth rate can be made from a
584 structured model of the coalescent.

585 *Conclusions*

586 Although the importance of intermediate stepping-stones is in many ways intuitively
587 obvious for marine environments (since islands are, after all, the archetype for the population

genetic theories discussed herein), this work provides the first *a priori* test of this idea. By comparing predictions from a biophysical model about the geographic availability of stepping-stones to coalescent gene flow models for species with distinct habitat requirements, we have shown that the genetic coherence of neritid species across the Indo-Pacific relies on intermediate stepping-stones in combination with dispersal in the tails of the dispersal kernel ($P_{dispersal} \geq 1 \times 10^{-12}$). However, there also seems to be an important role for range expansions and extremely rare long-distance dispersal events (occurring only a few times per millenium) beyond the scope of the biophysical model.

Conclusions for the current coalescent and biophysical models can only be drawn at evolutionary timescales. At ecological timescales (< 40 years), larval dispersal likely occurs at smaller spatial scales, but with even higher variance than depicted here (Siegel *et al.* 2003, 2008). However, the approach outlined herein provides a promising new method to empirically evaluate biophysical models of larval dispersal (Werner *et al.* 2007). By setting model predictions of dispersal pathways as explicit hypotheses to be evaluated with empirical genetic data in a probabilistic model, we can identify specific areas of disagreement that will lead to better tuning of one or both of the models. Moreover, the coalescent sampler used here can detect restricted gene flow at levels that would wash out a traditional approach based on F_{ST}, and provide confidence intervals that are appropriate to the data. As next-generation sequencing technology makes available genetic data from hundreds of loci, there is an exciting prospect for increasingly precise estimation of parameters in both the coalescent and biophysical models that will eventually be of direct utility to managers of marine reserve networks.

Acknowledgements

This dissertation project had its genesis during E.D.C.'s undergraduate research in the 1998 Biology and Geomorphology of Tropical Islands course, taught at the UC Berkeley Gump Station in French Polynesia. We gratefully acknowledge the people and governments of American Samoa, Fiji, French Polynesia, Samoa and Vanuatu for permitting our sample collections and for assistance in collections. In particular we thank E. Bardi, S. Crandall, N. Davies, K. Feussner, S. Ravuna, S. Siolo, R. Utzurrum, D. Vargo, L. Wade, and the Turner Family for assistance with field collections and permits. M. Frey graciously loaned additional samples from the Cook Islands, New Caledonia, Fiji and American Samoa, and J. Taffel assisted with molecular data collection. Grants to E.D.C. from the Lerner-Grey Fund (American Museum of Natural History) and the Walter Sage Memorial Fund (Conchologists of America) and the NSF Doctoral Dissertation Improvement Grant — DEB-0508788 to E.D.C. funded sample collections. NSF grants DEB-0338566 and OCE-0349177 to P.H.B funded sequencing and part of E.D.C.'s graduate stipend. E.A.T.'s modeling work was funded in part by the Australian Research Council (DP0878306), the Kathryn Fuller Post-Doctoral Fellowship (World Wildlife Fund), the NSF Coral Triangle Partnerships for International Research and Education, OISE-0730256 to P.H.B., K. Carpenter, and P. Halpin. E. Anderson, T. Eichhorst, D. Drown, J. Finnerty, L. Kaufman, L. Mullineaux, J. Philibotte, C. Riginos, E. Sbrocco, C. Schneider, L. Smith, S. Thorrold, R. Toonen and two anonymous reviewers provided very useful comments and discussion. The Crandall Family provided computational resources for Migrate analyses as well as moral support. E.D.C. thanks S. Crandall for her continued patience and support.

633 **Author Contributions**

634 E.D.C. conceived the project and E.D.C and P.H.B. designed the genetic approach. E.A.T.
635 conceived, designed and implemented the biophysical model and provided model output. E.D.C.
636 collected the samples, sequenced them, analyzed the data and wrote the manuscript. P.H.B.
637 provided monetary and logistical support for travel and laboratory work and advised and oversaw
638 the work. E.D.C., E.A.T. and P.H.B. reviewed and edited the manuscript.
639

Literature Cited

Almany GR, Berumen ML, Thorrold SR, Planes S, Jones GP (2007) Local replenishment of coral reef fish populations in a marine reserve. *Science* **316**, 742-744.

Barber PH, Palumbi SR, Erdmann MV, Moosa MK (2000) A marine Wallace's line? *Nature* **406**, 692-693.

Bird CE, Karl SA, Smouse PE, Toonen RJ (2011). Detecting and measuring genetic differentiation. In: *Phylogeography and Population Genetics in Crustacea* (Ed. C Held, S Koenemann, C Schubart). Crustacean Issues **19**, pp. 31–73). CRC Press.

Beerli P (2004) Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology* **13**, 827-836.

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

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4563-4568.

Beerli P, Palczewski M (2010) Unified Framework to Evaluate Panmixia and Migration Direction Among Multiple Sampling Locations. *Genetics* **185**, 313-U463.

Benzie JAH, Williams ST (1997) Genetic structure of giant clam (*Tridacna maxima*) populations in the west Pacific is not consistent with dispersal by present-day ocean currents. *Evolution* **51**, 768-783.

Bernardi G, Holbrook SJ, Schmitt RJ (2001) Gene flow at three spatial scales in a coral reef fish, the three-spot dascyllus, *Dasyllus trimaculatus*. *Marine Biology* **138**, 457-465.

Burgess SC, Treml EA, Marshall DJ (2012) How do dispersal costs and habitat selection influence realized population connectivity? *Ecology* **93**, 1378–1387.

Buston PM, Jones GP, Planes S, Thorrold SR (2011) Probability of successful larval dispersal declines fivefold over 1 km in a coral reef fish. *Proceedings Of The Royal Society B-Biological Sciences*.

Cowen RK, Lwiza KMM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of Marine Populations: Open or Closed? *Science* **287**, 857-859.

Cowen RK, Paris CB, Srinivasan A (2006) Scaling of connectivity in marine populations. *Science* **311**, 522-527.

Craig MT, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series* **334**, 245-254.

Crandall ED, Frey MA, Grosberg RK, Barber PH (2008) Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. *Molecular Ecology* **17**, 611-626.

Crandall ED, Taffel JR, Barber PH (2010) High gene flow due to pelagic larval dispersal among South Pacific archipelagos in two amphidromous gastropods (Neritomorpha: Neritidae). *Heredity* **104**.

DeBoer TS, Subia MD, Ambariyanto*, et al.* (2008) Phylogeography and Limited Genetic Connectivity in the Endangered Boring Giant Clam across the Coral Triangle. *Conservation Biology* **22**, 1255-1266.
DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology* **56**, 461-473.
Drummond AJ, Suchard MA, Xie, D and Rambaut, A. (in press) Bayesian phylogenetics with BEAUTi and the BEAST 1.7. *Molecular Biology and Evolution*. doi: 10.1093/molbev/mss075
Eble JA, Rocha LA, Craig MT, Bowen BW (2011) Not All Larvae Stay Close to Home: Insights into Marine Population Connectivity with a Focus on the Brown Surgeonfish (Acanthurus nigrofuscus). *Journal of Marine Biology* **2011**, 1-12.
Ekman S (1953) *Zoogeography of the Sea* Sidgwick and Jackson, London.
Excoffier L, Laval LG, Schneider S (2005) Arlequin v.3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.
Faurby S, Barber PH (2012) Theoretical limits to the correlation between pelagic larval duration and population genetic structure. *Molecular Ecology **14, 3419-3432.***
Foster NL, Paris CB, Kool JT*, et al.* (2012) Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. *Molecular Ecology* **21**, 1143-1157.
Fu Y-X (1997) Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics* **147**, 915-925.
Galindo HM, Olson DB, Palumbi SR (2006) Seascape genetics: A coupled oceanographic-genetic model predicts population structure of Caribbean corals. *Current Biology* **16**, 1622-1626.
Hedgecock D, Barber PH, Edmands S (2007) Genetic Approaches to Measuring Connectivity. *Oceanography* **20**, 70-79.
Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59**, 1633-1638.
Hellberg ME (1995) Stepping-Stone Gene Flow in the Solitary Coral *Balanophyllia elegans* - Equilibrium and Nonequilibrium at Different Spatial Scales. **123**, 573-581.
Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747-760.
Highsmith RC, Emlet RB (1986) Delayed metamorphosis: effect on growth and survival on juvenile sand dollars (Echinoidea: Clypeasteroidea). *Bulletin of Marine Science* **39**, 347-361.
Hodges MH, Allendorf FW (1998) Population genetics and patterns of larval dispersal of the endemic Hawaiian freshwater amphidromous gastropod *Neritina granosa*. *Pacific Science* **52**, 237-249.
Holthuis B (1995) *Evolution between marine and freshwater habitats: a case study of the gastropod suborder Neritopsina* Doctoral Thesis, University of Washington.

Horne JB, van Herwerden L, Choat JH, Robertson DR (2008) High population connectivity across the Indo-Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics & Evolution* **49**, 629-638.
Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**, 1898-1914.
Johnson KB, Shanks AL (2003) Low rates of predation on planktonic marine invertebrate larvae. *Marine Ecology-Progress Series* **248**, 125-139.
Jones GP, Millicich MJ, Emslie MJ, Lunow C (1999) Self-recruitment in a coral reef fish population. *Nature* **402**, 802-804.
Jones GP, Planes S, Thorrold SR (2005) Coral reef fish larvae settle close to home. *Current Biology* **15**, 1314-1318.
Kano Y (2006) Usefulness of the Opercular Nucleus for Inferring Early Development in Neritomorph Gastropods. *Journal of Morphology* **267**, 1120-1136.
Kass RE, Raftery AE (1995) Bayes Factors. *Journal of the American Statistic Association* **90**, 773-795.
Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* **49**, 561-576.
Kinlan BP, Gaines SD (2003) Propagule dispersal in marine and terrestrial environments: a community perspective. *Ecology* **84**, 2007-2020.
Kinlan BP, Gaines SD, Lester SE (2005) Propagule dispersal and the scales of marine community process. *Diversity and Distributions* **11**, 139-148.
Kirkendale LA, Meyer CP (2004) Phylogeography of the *Patelloida profunda* group (Gastropoda: Lottidae): diversification in a dispersal-driven marine system. *Molecular Ecology* **13**, 2749-2762.
Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. *Bioinformatics* **22**, 768-770.
Kuhner MK (2009) Coalescent genealogy samplers: windows into population history. *Trends In Ecology & Evolution* **24**, 86-93.
Kuhner MK, Yamato J, Beerli P*, et al.* (2005) LAMARC v2.0.2. University of Washington.
Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149**, 429-434.
Largier JL (2003) Considerations in estimating larval dispersal distances from oceanographic data. *Ecological Applications* **13**, S71-S89.
Lavery S, Moritz C, Fielder DR (1996) Indo-Pacific population structure and evolutionary history of the coconut crab *Birgus latro*. *Molecular Ecology* **5**, 557-570.
Lessios HA, Kane J, Robertson DR (2003) Phylogeography of the pantropical sea urchin *Tripneustes*: Contrasting patterns of population structure between oceans. *Evolution* **57**, 2026-2036.
Lessios HA, Kessing BD, Pearse JS (2001) Population structure and speciation in tropical seas: Global phylogeography of the sea urchin *Diadema*. *Evolution* **55**, 955-975.

Lessios HA, Robertson DR (2006) Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B-Biological Sciences* **273**, 2201-2208.
Lester SE, Ruttenberg BI (2005) The relationship between pelagic larval duration and range size in tropical reef fishes: a synthetic analysis. *Proceedings of the Royal Society B-Biological Sciences* **272**, 585-591.
Liu H-TT, Resh VH (1997) Abundance and microdistribution of freshwater gastropods in three streams in Moorea, French Polynesia. *Annales De Limnologie-International* Journal of Limnology **33**, 235-244.
Lowe WH, Allendorf FW (2010). What can genetics tell us about population connectivity? *Molecular Ecology* **19**, 3038–3051.
McDowall RM (2007) On amphidromy, a distinct form of diadromy in aquatic organisms. *Fish and Fisheries* **8**, 1-13.
Miller SE, Hadfield MG (1994) Developmental arrrest during larval life and life-span extension in a marine mollusc. *Science* **248**, 356-358.
Mitarai S, Siegel DA, Watson JR, Dong C, McWilliams JC (2009) Quantifying connectivity in the coastal ocean with application to the Southern California Bight. Journal of Geophysical Research-Oceans **114**, 1-21.
Mora C, Treml E, Roberts J, Crosby K, Roy D (2012). High connectivity among habitats precludes the relationship between dispersal and range size in tropical reef fishes. *Ecography* **35**, 89-96.
Myers MJ, Meyer CP, Resh VH (2000) Neritid and thiarid gastropods from French Polynesian streams: how reproduction (sexual, parthenogenetic) and dispersal (active, passive) affect population structure. *Freshwater Biology* **44**, 535-545.
Nishida M, Lucas JS (1988) Genetic-Differences between Geographic Populations of the Crown-of-Thorns Starfish Throughout the Pacific Region. *Marine Biology* **98**, 359-368.
Nishikawa A, Sakai K (2005) Genetic connectivity of the scleractinian coral Goniastrea aspera around the Okinawa Islands. *Coral Reefs* **24**, 318-323.
Palumbi S (2004) Marine reserves and ocean neighborhoods: The spatial scale of marine populations and their management. *Annual Review of Environmental Resources* **29**, 31-68.
Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications* **13**, S146-S158.
Palumbi SR, Grabowsky G, Duda T, Geyer L, Tachino N (1997) Speciation and population genetic structure in tropical Pacific Sea urchins. *Evolution* **51**, 1506-1517.
Pechenik JA (1990) Delayed Metamorphosis by Larvae of Benthic Marine-Invertebrates - Does It Occur? Is There a Price to Pay? *Ophelia* **32**, 63-94.
Pechenik JA (2006) Larval experience and latent effects - metamorphosis is not a new beginning. *Integrative and Comparative Biology* **46**, 323-333.
Pinsky ML, Montes Jr HR, Palumbi SR (2010) Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution* **64**, 2688-2700.

Planes S, Fauvelot C (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. *Evolution* **56**, 378-399.
Quinn TM, Crowley TJ, Taylor FW*, et al.* (1998) A multicentury stable isotope record from a New Caledonia coral: Interannual and decadal sea surface temperature variability in the southwest Pacific since 1657 AD. *Paleoceanography* **13**, 412-426.
Rambaut A, Drummond A (2007) Tracer v1.4. Available from http://beast.bio.ed.ac.uk/.
Reece JS, Bowen BW, Joshi K, Goz V, Larson A (2010) Phylogeography of Two Moray Eels Indicates High Dispersal Throughout the Indo-Pacific. *J Hered* **101**, 391-402.
Resh VH, Barnes JR, Benissteger B, Craig DA (1992) Life-history features of some macroinvertebrates in French-Polynesian streams. *Studies on Neotropical Fauna and Environment* **27**, 145-153.
Richmond RH (1990) The effects of the El Nino/Southern Oscillation on the dispersal of corals and other marine organisms. *Elsevier Oceanography. Ser.* **52**, 127-140.
Roberts CM, McClean CJ, Veron JEN*, et al.* (2002) Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* **295**, 1280-1284.
Rumrill (1990) Natural mortality of marine invertebrate larvae. *Ophelia* **32**, 163-198.
Scheltema RS (1971) Larval dispersal as a means of genetic exchange between geographically seperated populations of shallow-water benthic marine gastropods. *Biological Bulletin* **140**, 284-322.
Selkoe KA, Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. Marine Ecology Progress Series **436**, 291–305.
Selkoe KA, Watson JR, White C*, et al.* (2010) Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology* **19**, 3708-3726.
Shanks, AL (2009). Pelagic larval duration and dispersal distance revisited. *Biological Bulletin* **216**, 373–385.
Siegel DA, Kinlan BP, Gaylord B, Gaines S (2003) Lagrangian descriptions of marine larval dispersion. *Marine Ecology Progress Series* **260**, 83-96.
Siegel DA, Mitarai S, Costello CJ*, et al.* (2008) The stochastic nature of larval connectivity among nearshore marine populations. *Proceedings Of The National Academy Of Sciences Of The United States Of America* **105**, 8974-8979.
Slatkin M (1993) Isolation by Distance in Equilibrium and Nonequilibrium Populations. *Evolution* **47**, 264-279.
Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology* **14**, 67-73.
Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **123**, 603-613.
Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. *Genetics* **123**, 603-613.
Smolarkiewicz PK, Margolin LG (1998) MPDATA: A Finite-Difference Solver for Geophysical Flows. *Journal of Computational Physics* **140**, 459-480.

Spalding MD, Fox HE, Halpern BS*, et al.* (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *Bioscience* **57**, 573-583.
Swearer SE, Caselle JE, Lea DW, Warner RR (1999) Larval retention and recruitment in an island population of a coral-reef fish. *Nature* **402**, 799-802.
Swearer SE, Shima JS, Hellberg ME*, et al.* (2002) Evidence of self-recruitment in demersal marine populations. *Bulletin of Marine Science* **70**, 251-271.
Swofford DL (2002) PAUP*: Phylogenetic Analysis Using Parsimony *and Other Methods. Sinauer Associates, Sunderland, MA.
Taylor MS, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science* **299**, 107-109.
Team RDC (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Thacker CE (2004) Population structure in two species of the reef goby *Gnatholepis* (Teleostei: Perciformes) among four South Pacific island groups. *Coral Reefs* **23**, 357-366.
Thorson G (1950) Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews of the Cambridge Philosophical Society* **25**, 1-45.
Treml EA, Halpin PN, Urban DL, Pratson LF (2008) Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology* **23**, 19-36.
Treml EA, Roberts JJ, Chao Y, Halpin PN, Possingham HP, Riginos C (in press) Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. *Integrative and Comparative Biology*.
Underwood AJ (1975) Comparative studies on the biology of *Nerita atramentosa* Reeve, *Bembicium nanum* (Lamarck) and *Cellana tramoserica* (Sowerby) (Gastropods: Prosobranchia) in S.E. Australia. *Journal of Experimental Marine Biology and Ecology* **18**, 153-172.
Vermeij GJ (1971) Substratum relationships of some tropical Pacific intertidal gastropods. *Marine Biology* **10**, 315-320.
Vermeij GJ (1987) The dispersal barrier in the tropical Pacific: implications for molluscan speciation and extinction. *Evolution* **41**, 1046-1058.
Wakeley J (2004) Recent trends in population genetics: More data! More math! Simple models? *Journal of Heredity* **95**, 397-405.
Wakeley J, Takahashi T (2004) The many-demes limit for selection and drift in a subdivided population. *Theoretical Population Biology* **66**, 83-91.
Wang X, Chao Y (2004) Simulated Sea Surface Salinity variability in the tropical Pacific. *Geophysical Research Letters* **31**, -.
Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity* **89**, 438-450.
Werner F, Cowen R, Paris C (2007) Coupled Biological and Physical Models: Present Capabilities and Necessary Developments for Future Studies of Population Connectivity. *Oceanography* **20**, 54-69.
Weersing K, Toonen RJ (2009). Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* **393**, 1–12.

White C, Selkoe KA, Watson J*, et al.* (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences* **277**, 1685-1694.
Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: F-ST not equal 1/(4Nm+1). *Heredity* **82**, 117-125.
Williams ST, Benzie JAH (1998) Evidence of a biogeographic break between populations of a high dispersal starfish: Congruent regions within the Indo-West Pacific defined by color morphs, mtDNA, and allozyme data. *Evolution* **52**, 87-99.
Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.
Wright S (1943) Isolation by distance. *Genetics* **28**, 114-137.

Data Accessibility

- Full dataset in Migrate format: doi:10.5061/dryad.vh21c
- Migrate parameter files are available from the corresponding author
- Biophysical Model Output: See Supplemental Information.

Figure Legends

Figure 1. Map of the South Pacific, showing island localities sampled in eight archipelagoes.

Figure 2. a) Map illustrating dispersal probabilities calculated by the biophysical model. Arrow sizes are scaled to dispersal probability among archipelagos. Archipelagos surrounded by a dotted line do not host populations of the amphidromous neritids studied in this paper, and are therefore expected to be stepping-stones only for marine neritids. For clarity, not all intermediate atolls are shown, just the ones that provide the most probable pathways for stepping-stone dispersal. b) Schematic of the biophysical model predictions translated into a migration model for Migrate. Lettered arrows depict

migration parameters and numbered, colored circles show sampled populations for which Θ was calculated. Arrow thickness is scaled to the joint probability of dispersal between sampled archipelagoes, red for marine neritids and blue for amphidromous neritids (also note differently shaped arrowheads). Where unsampled atoll stepping-stones are necessary, the joint probability for marine neritids is the product of dispersal probabilities across all intermediate populations, while the joint probability for amphidromous neritids is 0. Populations in the Western Pacific exchanged genes at a constant rate with a "ghost" population (Beerli 2004). c) Map illustrating modal estimates of N_em from the stepping-stone model of the structured coalescent in Migrate for all four species. N_em is the product of the migration rate m/μ and θ for the recipient population.

Figure 3. Density plot depicting differences in proportion of migrants between marine species and amphidromous species across regions requiring and not requiring atoll stepping-stones. We used m/μ to avoid spurious correlations arising from incorporation of N_e. For comparison we also depict the difference between random sets of values picked from the posteriors of all four species as the null hypothesis that there is no difference in gene flow between the two species types.

Figure 4. Isolation-by-distance in *Nerita plicata*. a,c,e) OLS regression model for $\log_{10}(N_em) = a + b(\log_{10}(\text{distance})$ for 10,000 random samples from the posterior distributions of each of 11 lettered migration parameters (the modes of which are shown as filled circles) depicted in figure 3a. Distances are a) great circle distance in km, c)

number of steps between sampled populations from the biophysical model, e) the inverse raw probability of larval dispersal among populations from the biophysical model. The mean value and posterior distributions for the slope are given in b,d, and f. Because stepping-stone distances are not independent, the slope histogram is overlaid on one in light grey for which distance values were permuted randomly among the migration parameters, representing a hypothesis of no relationship with distance (similar to a Mantel test).

Figure 5. Isolation-by-distance in *Nerita albicilla*. a,c,e) OLS regression model for $\log_{10}(N_em) = a + b(\log_{10}(\text{distance})$ for 10,000 random samples from the posterior distributions of each of 11 lettered migration parameters (the modes of which are shown as filled circles) depicted in figure 3a. Distances are a) great circle distance in km, c) number of steps between sampled populations from the biophysical model, e) the raw probability of larval dispersal among populations from the biophysical model. The mean value and posterior distributions for the slope are given in b,d, and f. Because stepping-stone distances are not independent, the slope histogram is overlaid on one for which distance values were permuted randomly among the migration parameters, representing a hypothesis of no relationship with distance (similar to a Mantel test).

Table 1. Summary statistics and neutrality test statistics for each Western Pacific island deme shown in Figure 1. Haplotype diversity (h), nucleotide diversity (π) and F_s (Fu, 1997) calculated in Arlequin 3.1 (Excoffier *et al*. 2005).

Western Pacific		*Nerita plicata*				*Nerita albicilla*				*Neritina canalis*				*Neritipteron dilatatum*			
Archipelago	Island	n	h	π	F_s	n	h	π	F_s	n	h	π	F_s	n	h	π	F_s
New Caledonia	1. New Caledonia	40	0.99	0.012	**-33.51**	30	0.99	0.012	**-12.27**								
Vanuatu	2. Espiritu Santo													19	0.87	0.009	-3.13
	3. Efate	24	1.00	0.021	**-10.45**									23	0.94	0.007	**-12.76**
	4. Tanna	21	0.99	0.013	**-9.37**	17	1.00	0.010	**-12.78**	5	1.00	0.004	**-2.86**				
Fiji	5. Viti Levu	40	1.00	0.016	**-24.62**	40	0.98	0.009	**-24.90**	17	0.99	0.008	**-9.95**	2	1.00	0.012	n/a
	6. Taveuni	12	1.00	0.018	**-4.39**					11	0.96	0.005	**-5.05**	23	0.98	0.013	**-11.60**
Samoa	7. Upolu	28	1.00	0.013	**-24.01**					22	0.94	0.005	**-14.85**	23	0.94	0.007	**-11.01**
	8. Tutuila	38	0.99	0.014	**-17.40**	36	0.99	0.009	**-24.97**	23	0.97	0.005	**-19.76**	19	0.89	0.008	-2.52

Table 1. (con't) Summary statistics and neutrality test statistics for each Central Pacific island deme shown in Figure 1. Haplotype diversity (h), nucleotide diversity (π) and F_s (Fu, 1997) calculated in Arlequin 3.1 (Excoffier *et al*. 2005).

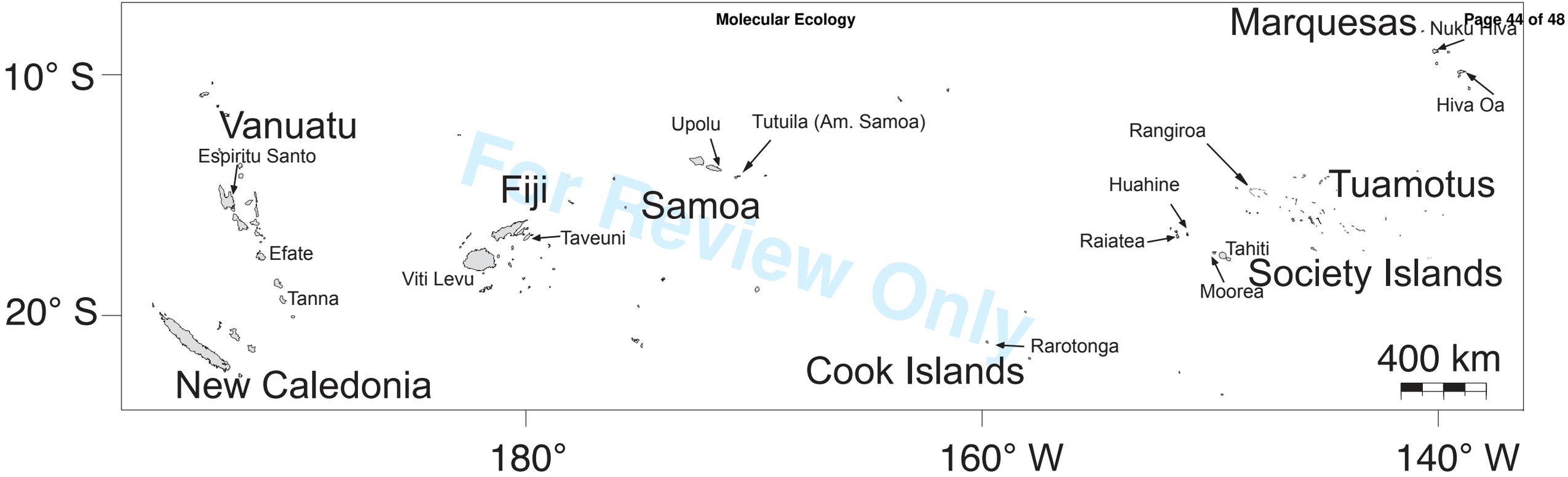
Central Pacific		*Nerita plicata*				*Nerita albicilla*				*Neritina canalis*				*Neritipteron dilatatum*			
Archipelago	Island	n	h	π	F_s	n	h	π	F_s	n	h	π	F_s	n	h	π	F_s
Cook Islands	9. Rarotonga	40	1.00	0.022	**-18.85**	29	0.98	0.009	**-12.55**								
Society Islands	10. Raiatea									17	0.91	0.005	**-6.11**	24	0.96	0.008	**-7.16**
	11. Huahine	12	1.00	0.023	**-3.46**					18	0.97	0.007	**-8.73**				
	12. Moorea	38	0.98	0.022	**-12.55**					25	0.97	0.005	**-21.20**	17	0.98	0.011	**-7.70**
	13. Tahiti									18	0.96	0.006	**-11.98**				
Tuamotus	14. Rangiroa	18	0.99	0.022	**-5.38**												
Marquesas	15. Nuku Hiva	20	1.00	0.017	**-8.61**					23	0.92	0.006	**-5.07**				
	16. Hiva Oa	11	1.00	0.018	**-3.70**					19	0.92	0.006	**-5.52**				

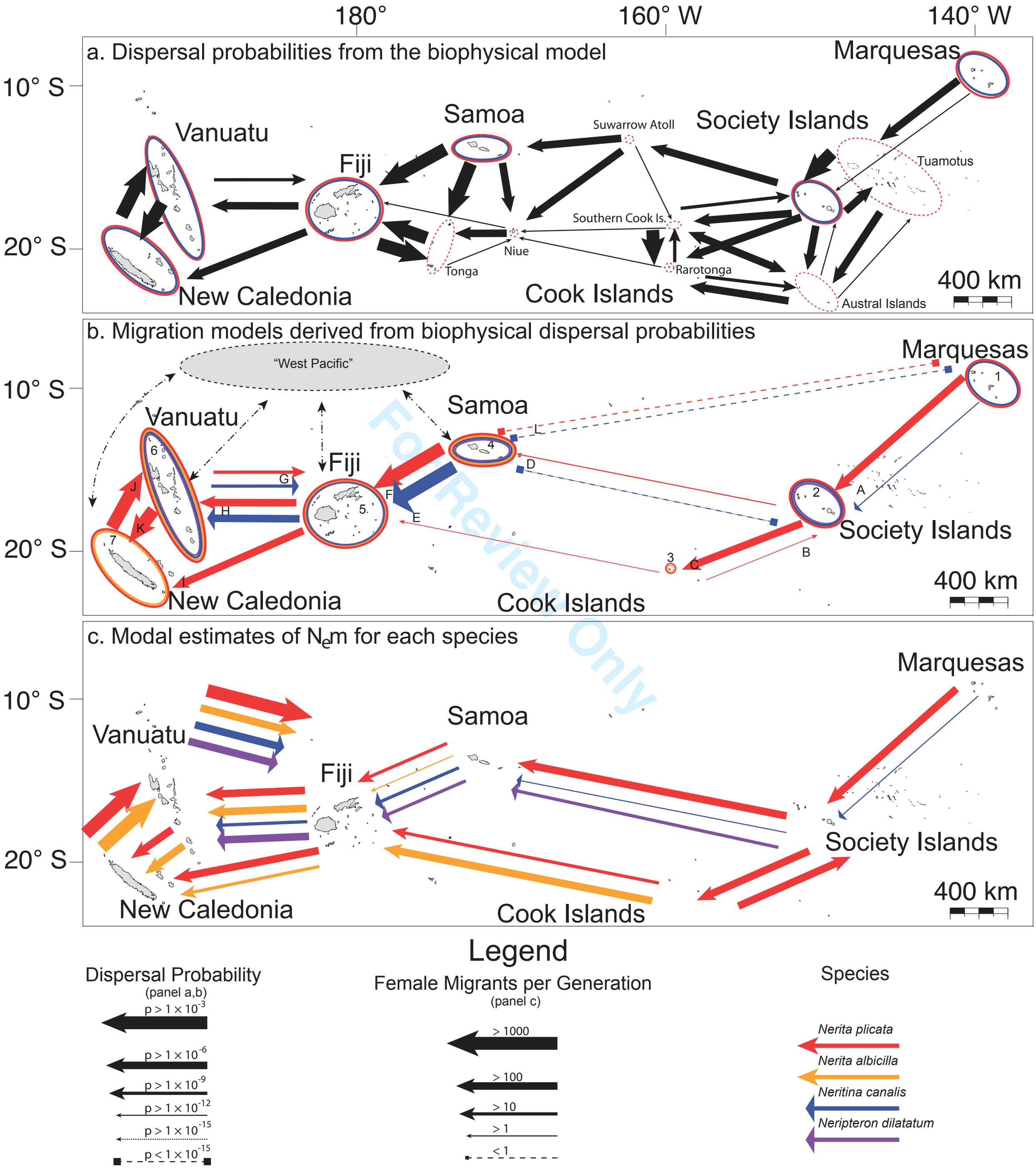
Table 2. Model comparison using 2ln Bayes Factors (LBF), which can be interpreted on the same scale as likelihood ratio tests.

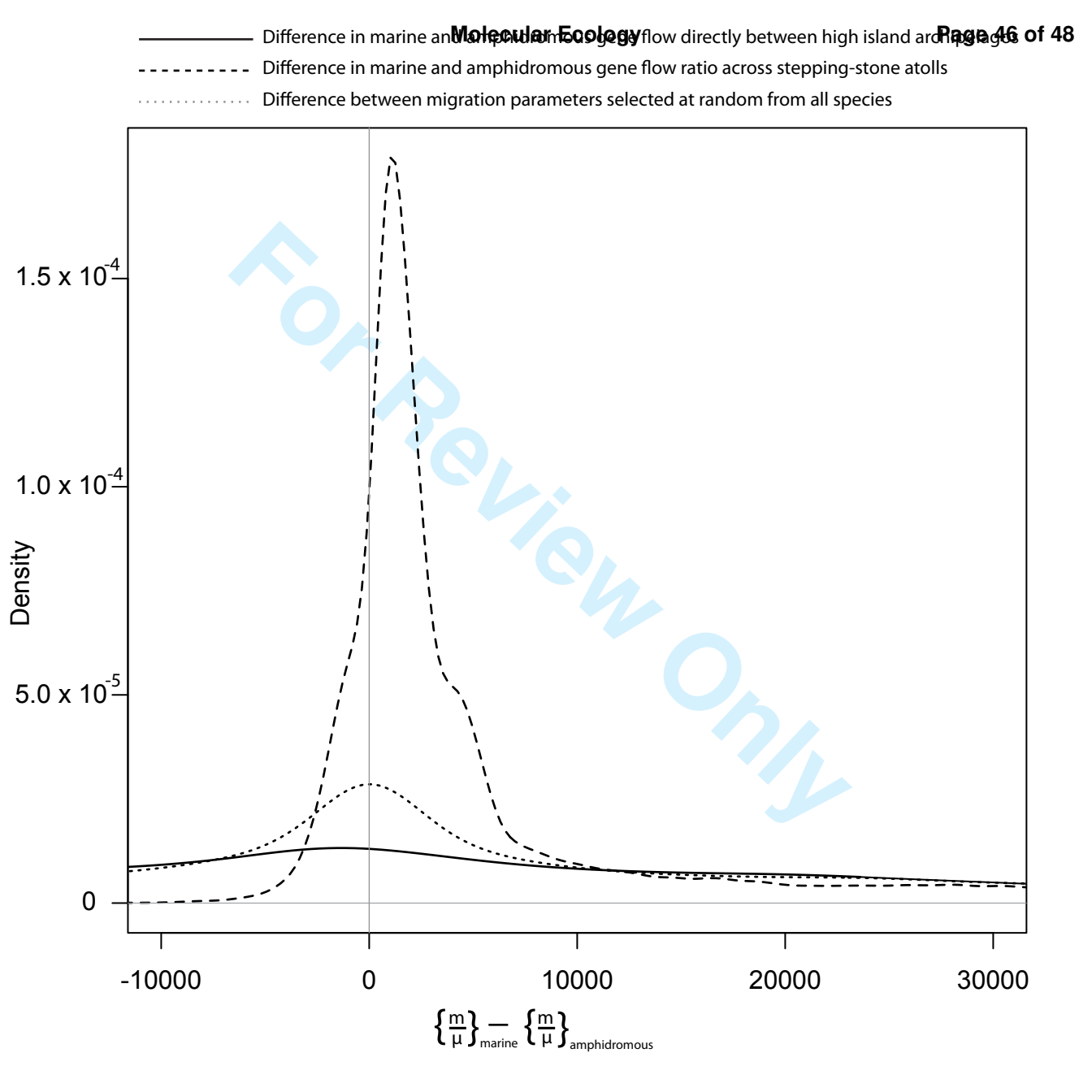
Species	Model	k[1]	Parameters Included[2]	Marginal LnL	LBF	Relative Probability	Rank
Nerita plicata	Panmixia	1	Θ (mean across all pops)	-9001.49	-1198.39	5.91×10^{-261}	6
	Island Model	2	Θ, m (mean across all pops)	-8535.75	-266.92	1.10×10^{-58}	4
	Biophysical 1-way	15	Θ = {1,2,3,4,5,6,7} m ={A,C,D,E,F,H,I,K}	-8450.59	-96.60	1.05×10^{-21}	3
	Biophysical Stepping-Stone	**18**	**Θ = {1,2,3,4,5,6,7} m = {A,B,C,D,E,F,G,H,I,J,K}**	**-8402.29**	**0.00**	**~1**	**1**
	Biophysical +Marquesas→Samoa	19	Θ = {1,2,3,4,5,6,7} m = {A,B,C,D,E,F,G,H,I,J,K,L}	-8426.92	-49.25	2.02×10^{-11}	2
	Full Model	49	All possible Θ and m	-8622.39	-440.19	2.59×10^{-96}	5
Nerita albicilla	Panmixia	1	Θ (mean across all pops)	-3123.26	-205.67	1.71×10^{-45}	5
	Island Model	2	Θ, m (mean across all pops)	-3023.10	-5.7	5.42×10^{-2}	2
	Biophysical 1-way	10	Θ = {3,4,5,6,7} m = {E,F,H,I,K}	-3025.17	-9.48	8.03×10^{-3}	3
	Biophysical Stepping-Stone	**12**	**Θ = {3,4,5,6,7} m = {E,F,G,H,I,J,K}**	**-3020.43**	**0.00**	9.38×10^{-1}	**1**
	Full Model	25	All possible Θ and m	-3047.71	-54.93	1.11×10^{-12}	4
Neritina canalis	Panmixia	1	Θ (mean across all pops)	-3399.87	-345.16	9.54×10^{-76}	6
	Island Model	2	Θ, m (mean across all pops)	-3285.10	-115.62	6.68×10^{-26}	4
	Biophysical 1-way	9	Θ = {1,2,4,5,6} m ={A,D,F,H}	-3235.61	-16.65	2.07×10^{-4}	3
	Biophysical	**10**	**Θ = {1,2,4,5,6} m ={A,D,F,G,H}**	**-3227.29**	**0.00**	$\mathbf{8.54\times10^{-1}}$	**1**
	Biophysical +Marquesas→Samoa	11	Θ = {1,2,4,5,6} m ={A,D,F,G,H,L}	-3229.06	-3.53	1.46×10^{-1}	2
	Full Model	25	All possible Θ and m	-3301.09	-147.60	7.60×10^{-33}	5
Neripteron dilatatum	Panmixia	1	Θ (mean across all pops)	-2417.84	-158.36	3.95×10^{-35}	5
	Island Model	2	Θ, m (mean across all pops)	-2351.64	-25.96	2.22×10^{-6}	3
	1-way	7	Θ = {2,4,5,6} m = {D,F,H}	-2342.02	-6.72	3.35×10^{-2}	2
	Biophysical	**8**	**Θ = {2,4,5,6} m ={D,F,G,H}**	**-2338.66**	**0.00**	$\mathbf{9.66\times10^{-1}}$	**1**
	Full Model	16	All possible Θ and m	-2365.14	-52.97	3.04×10^{-12}	4

Table 3. Modal parameter estimates and 95% highest posterior density estimates made by Tracer 1.4 for a stepping-stone model of gene flow run in Migrate 3.2.6

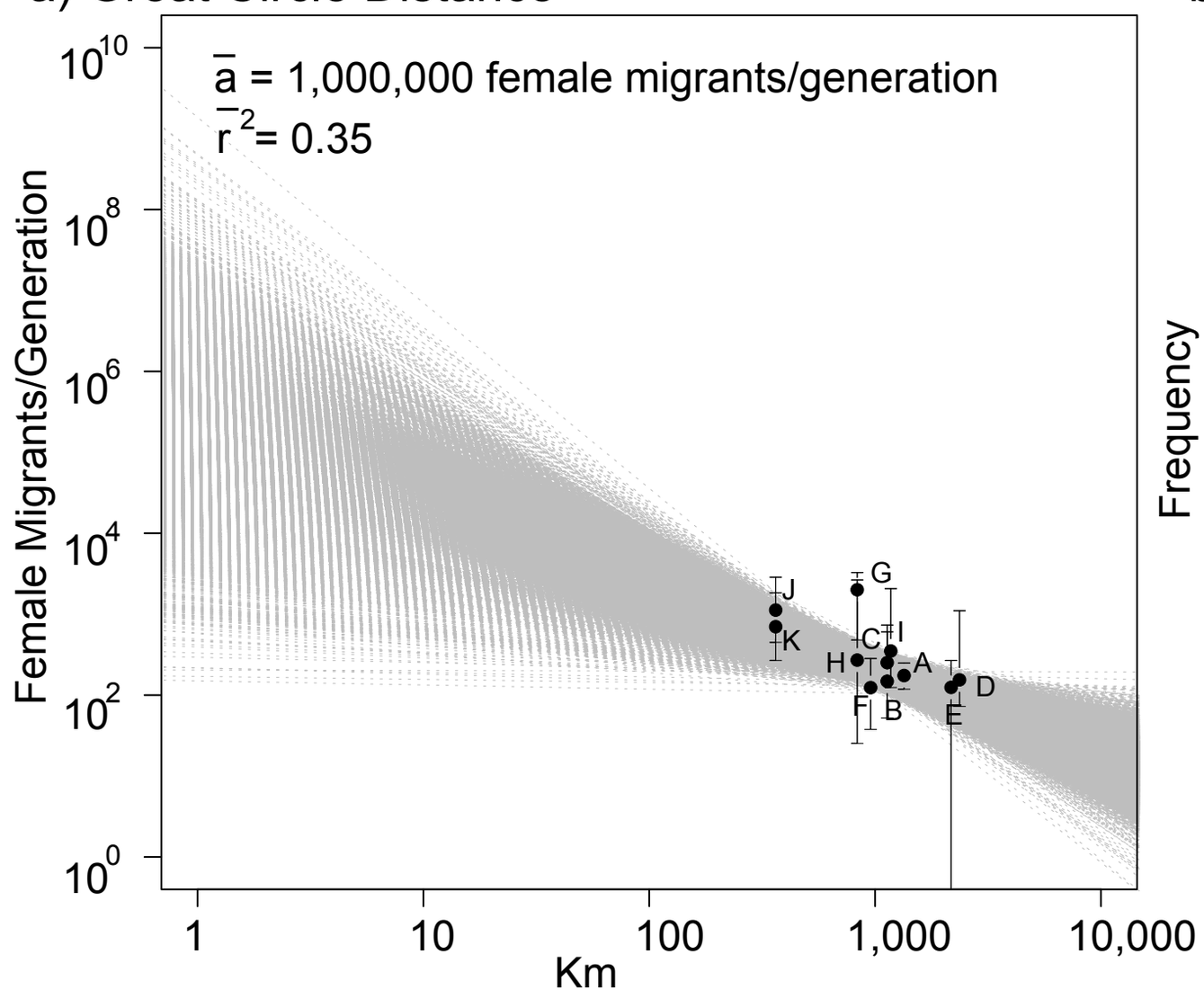
		Nerita plicata Marine			*Nerita albicilla* Marine			*Neritina canalis* Amphidromous			*Neripteron dilatatum* Amphidromous		
			95% HPD			95% HPD			95% HPD			95% HPD	
Parameter #	Name	Mode	Lower	Upper	Mode	Lower	Upper	Mode	Lower	Upper	Mode	Lower	Upper
1	Θ Marquesas	**0.069**	0.050	0.095	-	-	-	**0.015**	0.010	0.026	-	-	-
2	Θ Societies	**0.079**	0.045	0.198	-	-	-	**0.038**	0.026	0.052	**0.034**	0.023	0.049
3	Θ Rarotonga	**0.036**	0.014	0.105	**0.038**	0.025	0.057	-	-	-	-	-	-
4	Θ Samoa	**0.013**	0.005	0.087	**0.034**	0.018	0.068	**0.042**	0.020	0.051	**0.03**	0.012	0.063
5	Θ Fiji	**0.028**	0.010	0.093	**0.01**	0.002	0.032	**0.009**	0.001	0.027	**0.011**	0.002	0.035
6	Θ Vanuatu	**0.069**	0.039	0.107	**0.015**	0.004	0.044	**0.01**	0.001	0.033	**0.02**	0.007	0.046
7	Θ NewCal	**0.011**	0.004	0.081	**0.014**	0.003	0.043	-	-	-	-	-	-
A	Nm Marq→Soc	**139.3**	90.8	206.1	-	-	-	**2.0**	0.7	4.6	-	-	-
B	Nm Rar →Soc	**116.4**	36.9	546.3	-	-	-	-	-	-	-	-	-
C	Nm Soc →Rar	**208.7**	120.0	676.1	-	-	-	-	-	-	-	-	-
D	Nm Soc→Sam	**121.7**	53.1	1058.0	-	-	-	**1.5**	0.4	4.0	**67.0**	28.2	151.7
E	Nm Rar →Fiji	**96.1**	0.3	223.2	**142.8**	46.1	1194.9	-	-	-	-	-	-
F	Nm Sam→ Fiji	**95.4**	25.8	237.4	**9.2**	0.0	533.1	**50.9**	12.4	787.1	**83.4**	18.8	647.6
G	Nm Van→Fiji	**2041.4**	423.6	3492.5	**677.0**	0.0	2418.1	**221.9**	0.0	1658.1	**818.4**	64.7	2472.9
H	Nm Fiji→Van	**225.8**	16.7	2776.2	**109.8**	0.0	1446.2	**88.8**	5.0	1631.0	**514.5**	59.1	2449.5
I	Nm Fiji→NC	**300.3**	95.7	2114.6	**44.4**	0.0	1583.1	-	-	-	-	-	-
J	Nm NC→Van	**1080.2**	393.0	3021.3	**1179.6**	3.9	2875.6	-	-	-	-	-	-
K	Nm Van→NC	**642.3**	223.4	1852.5	**602.6**	0.0	2407.2	-	-	-	-	-	-
L	Nm Marq→Sam	-	-	-	-	-	-	-	-	-	-	-	-



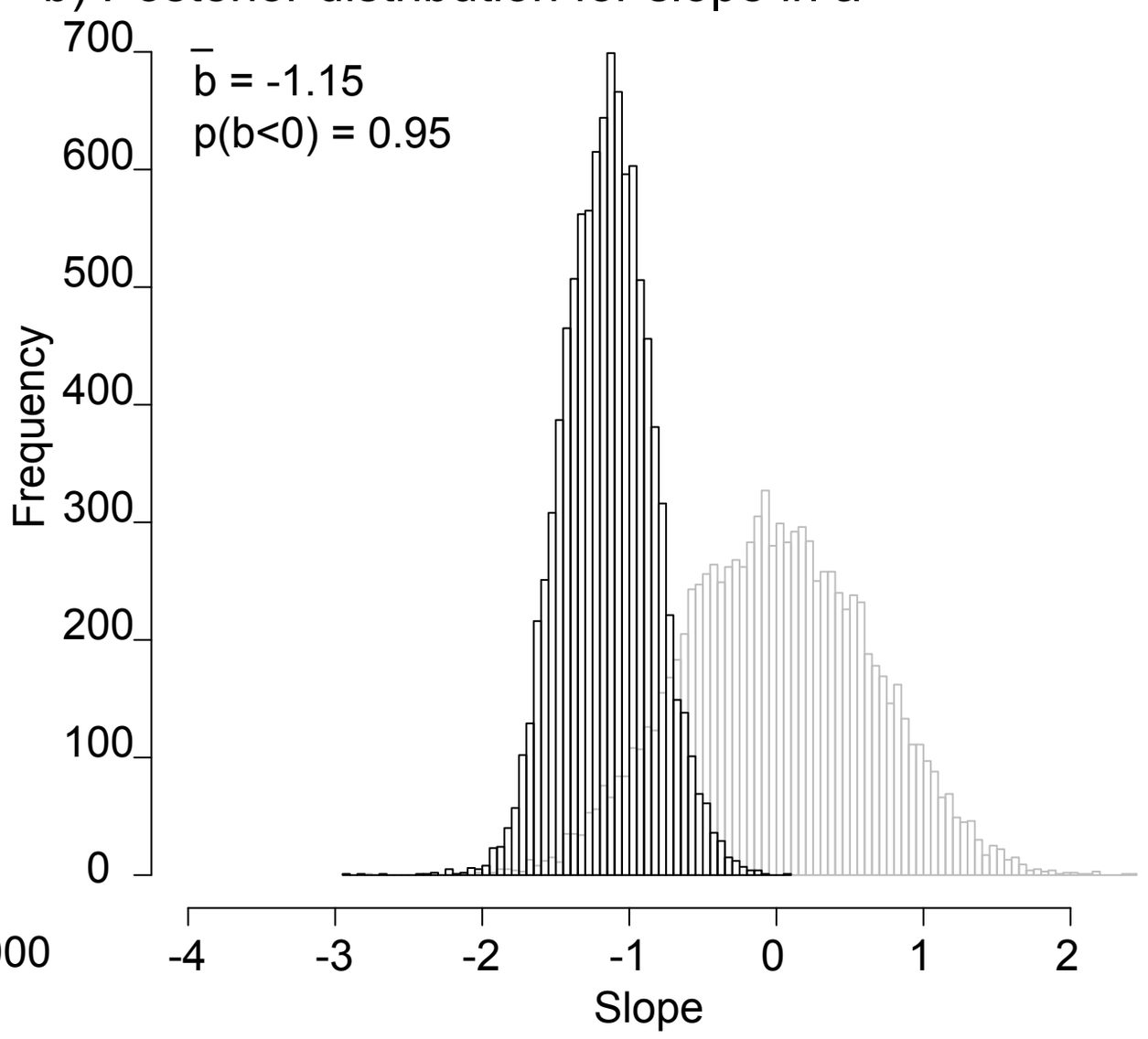




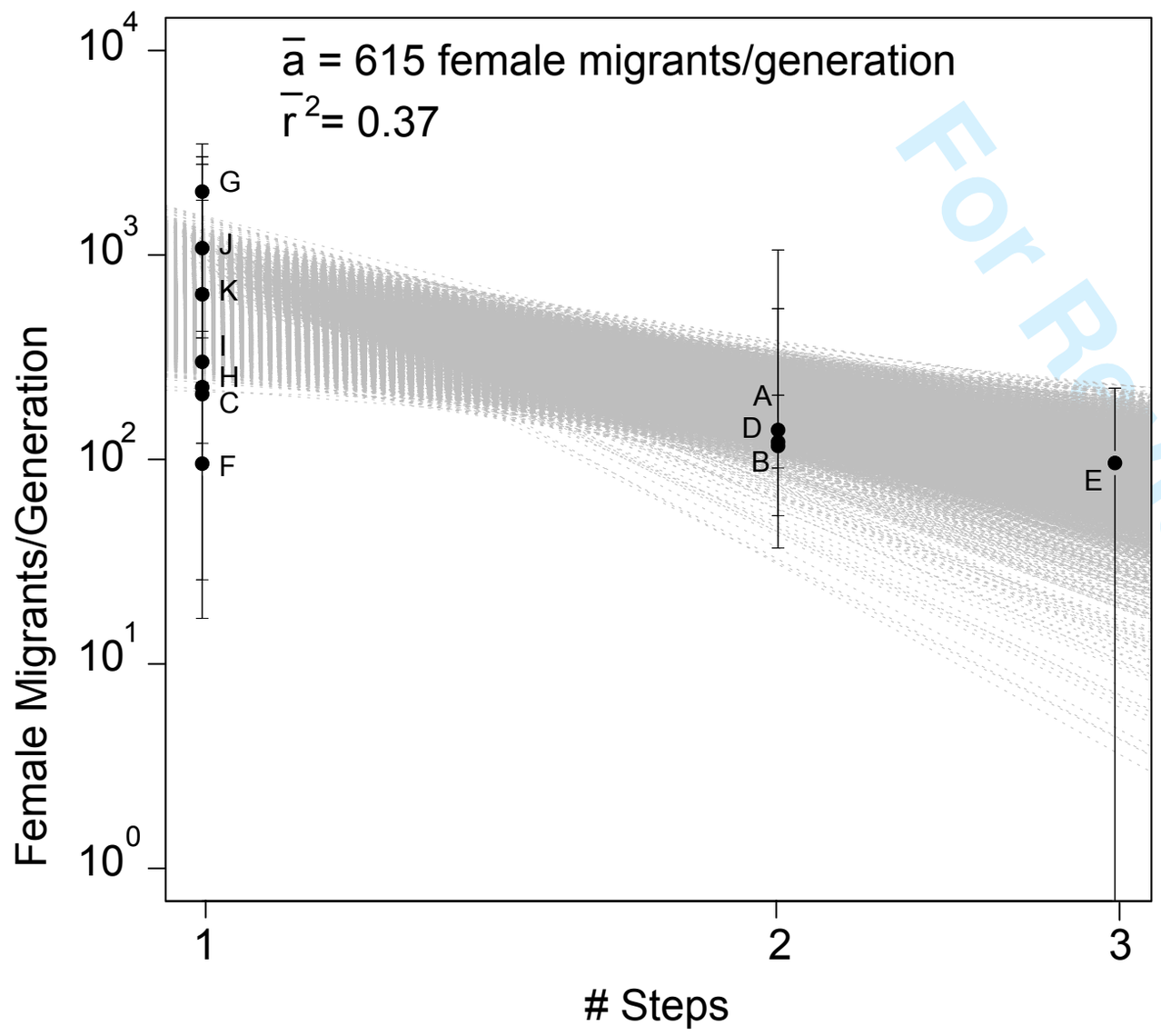
a) Great Circle Distance



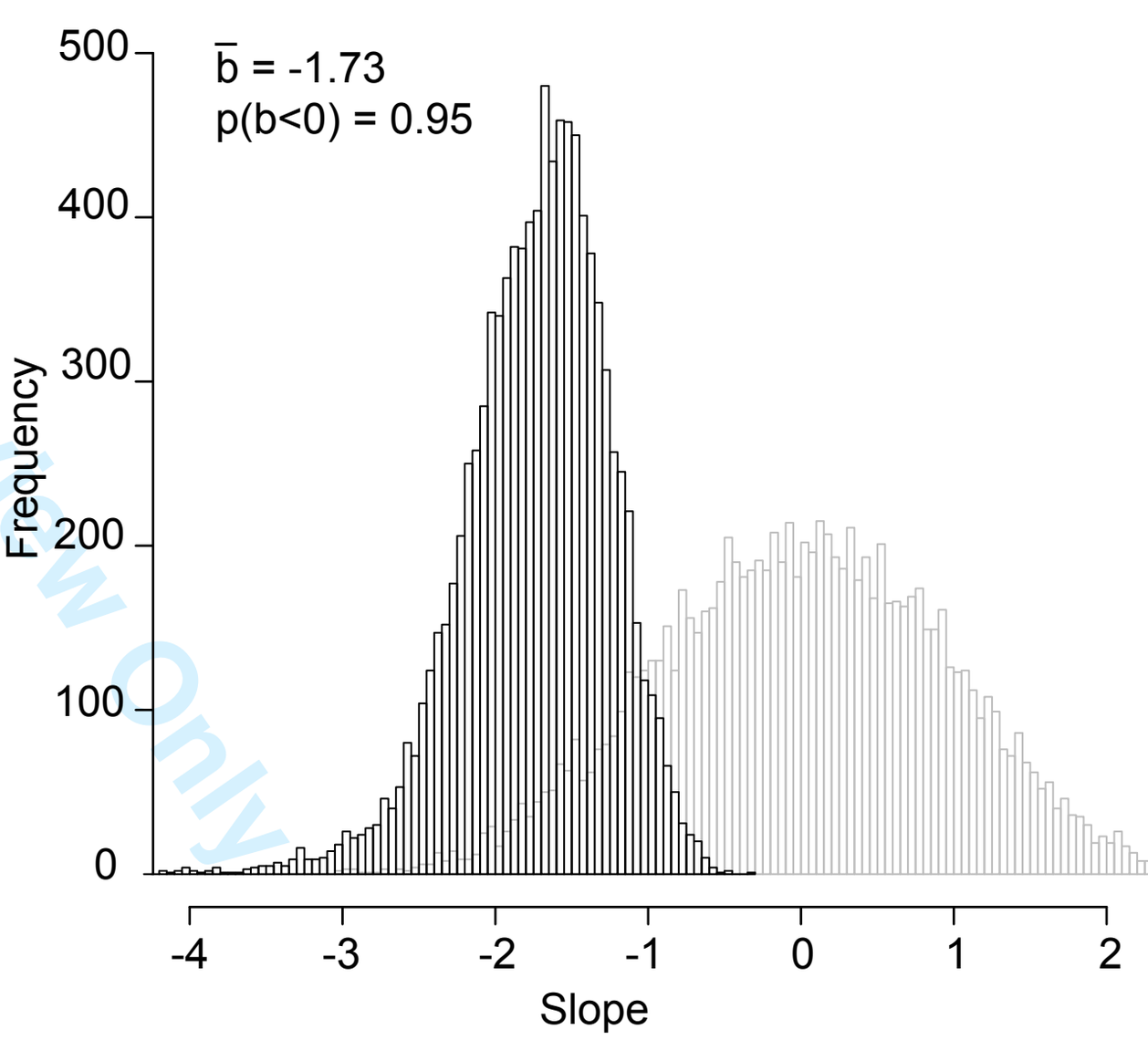
b) Posterior distribution for slope in a



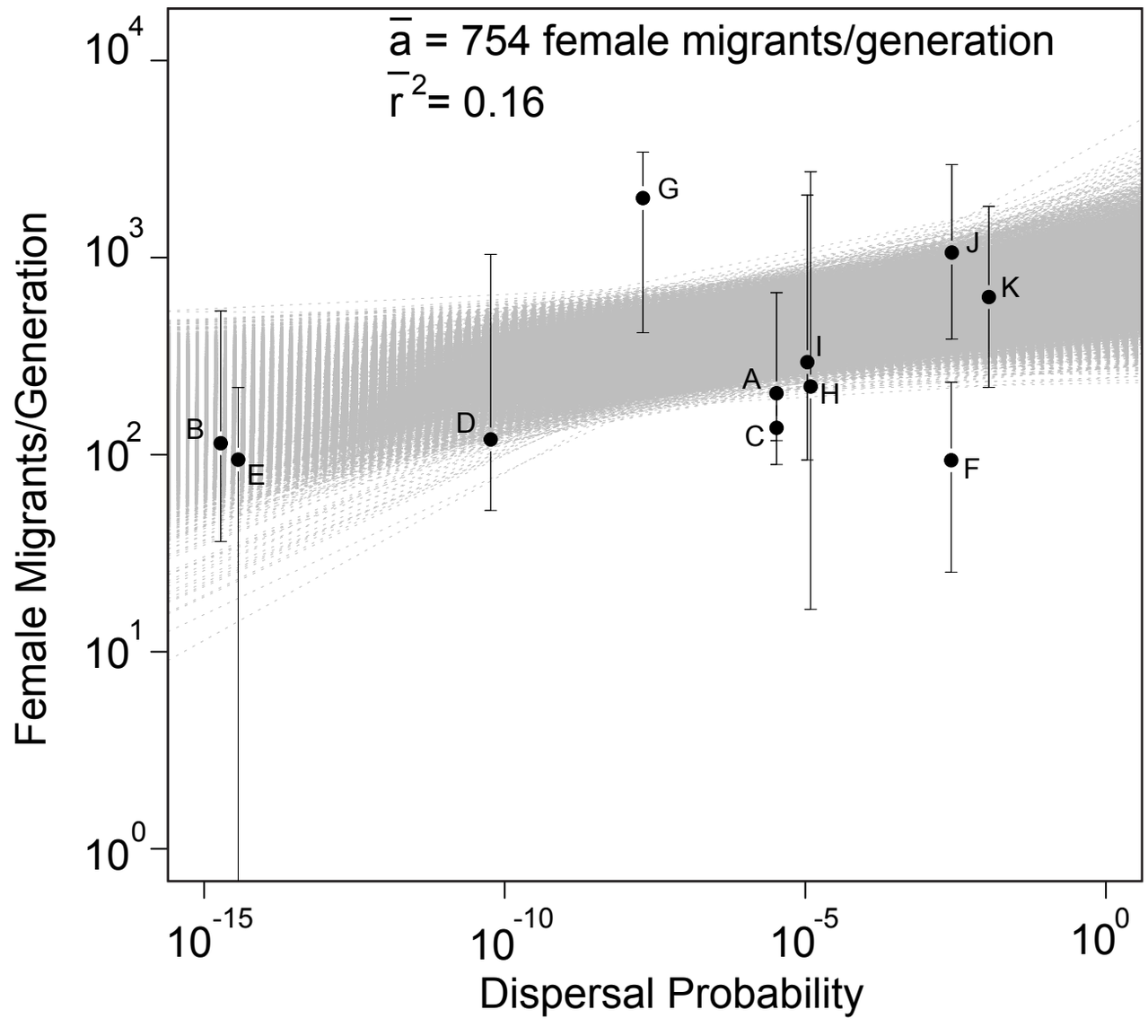
c) Predicted Steps from Biophysical Model



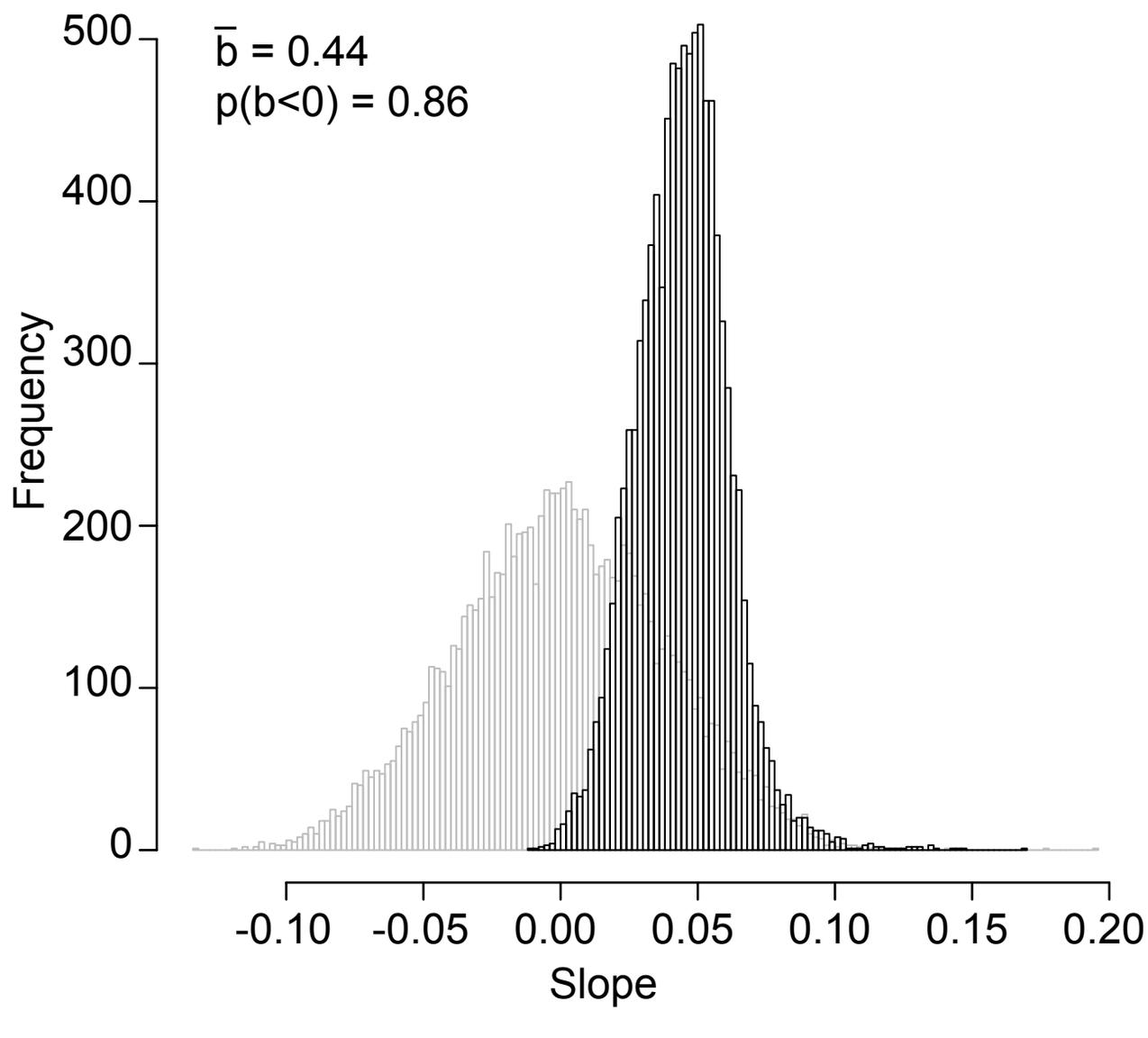
d) Posterior distribution for slope in c



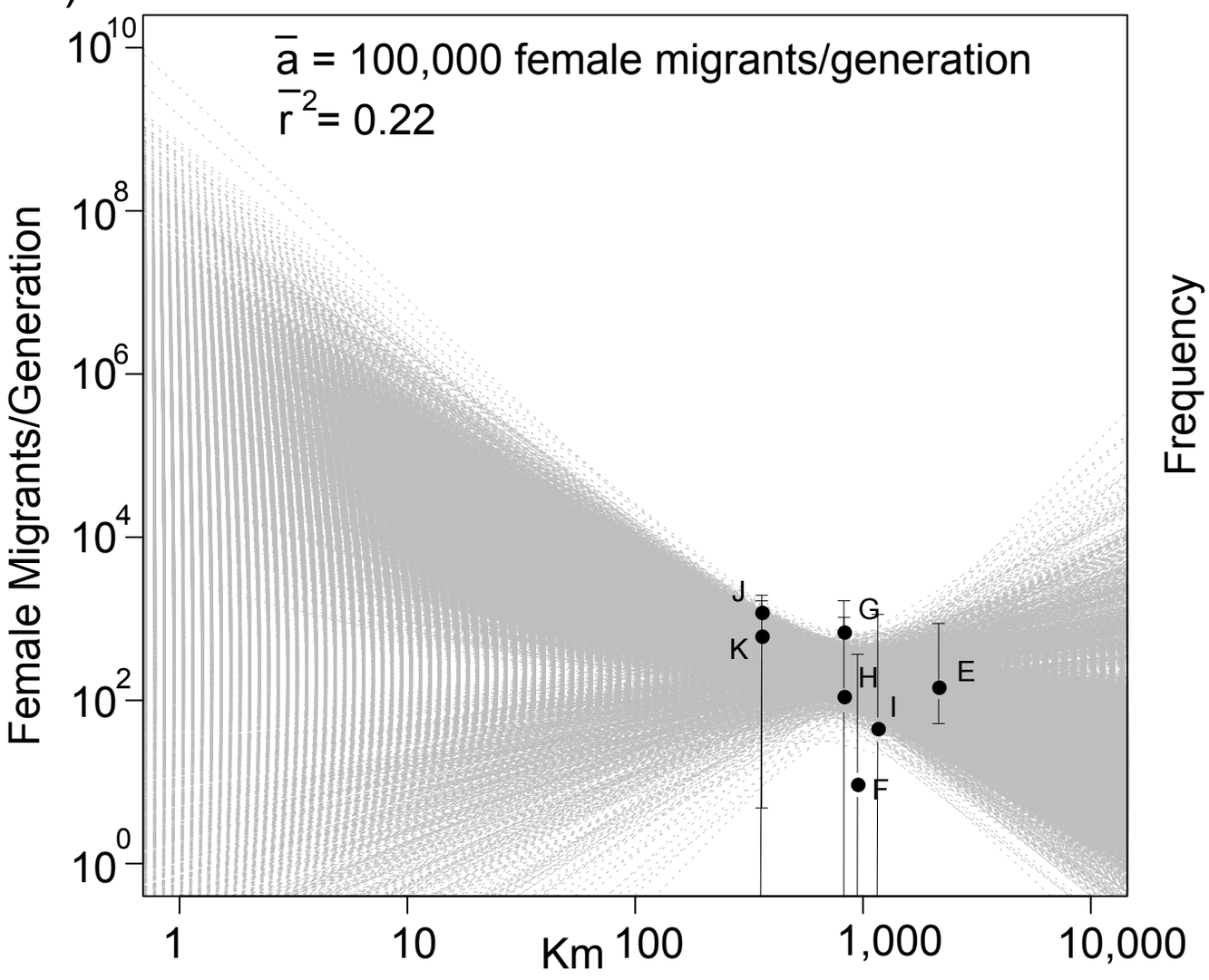
e) Dispersal Probability from Biophysical Model



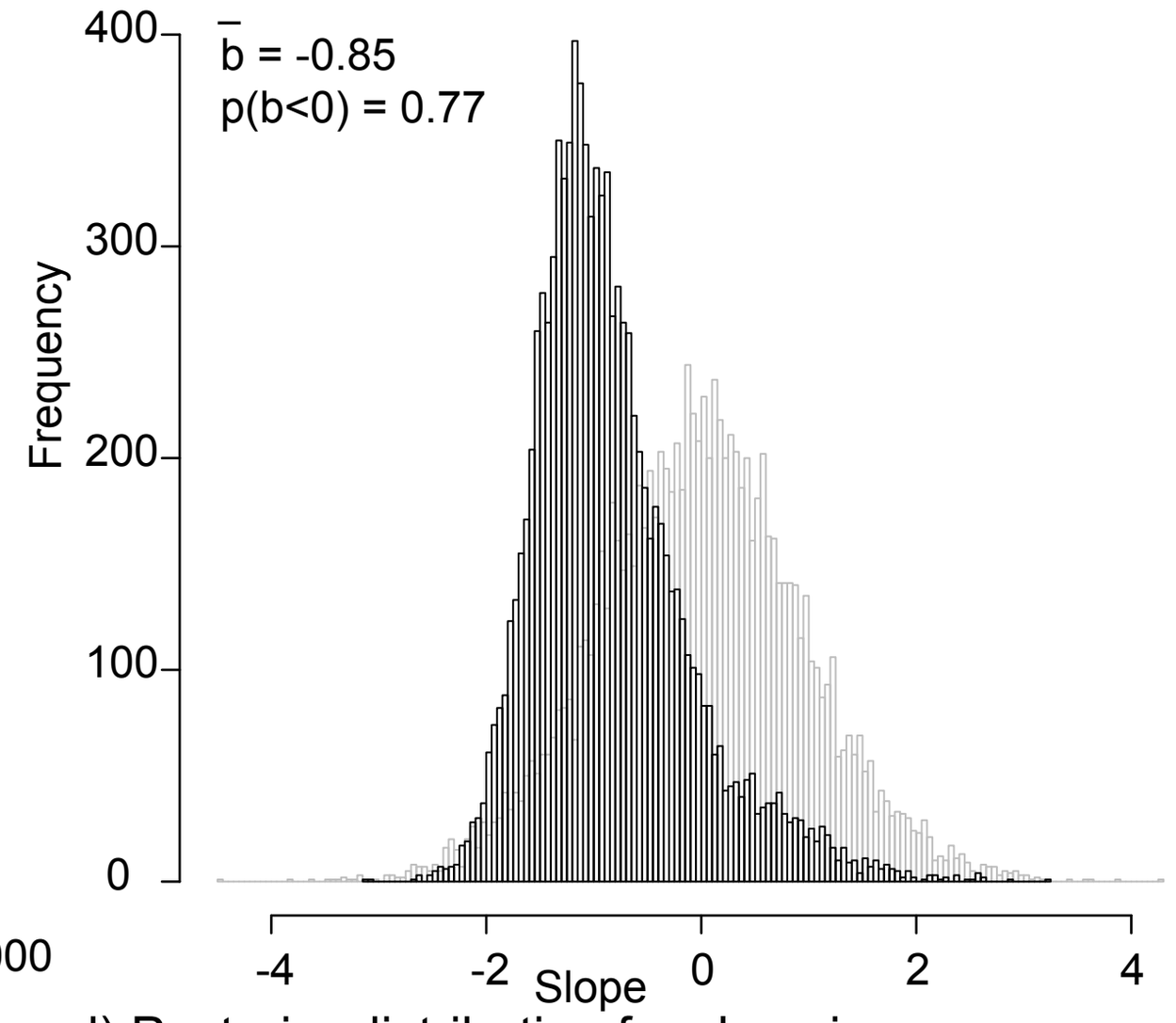
f) Posterior distribution for slope in e



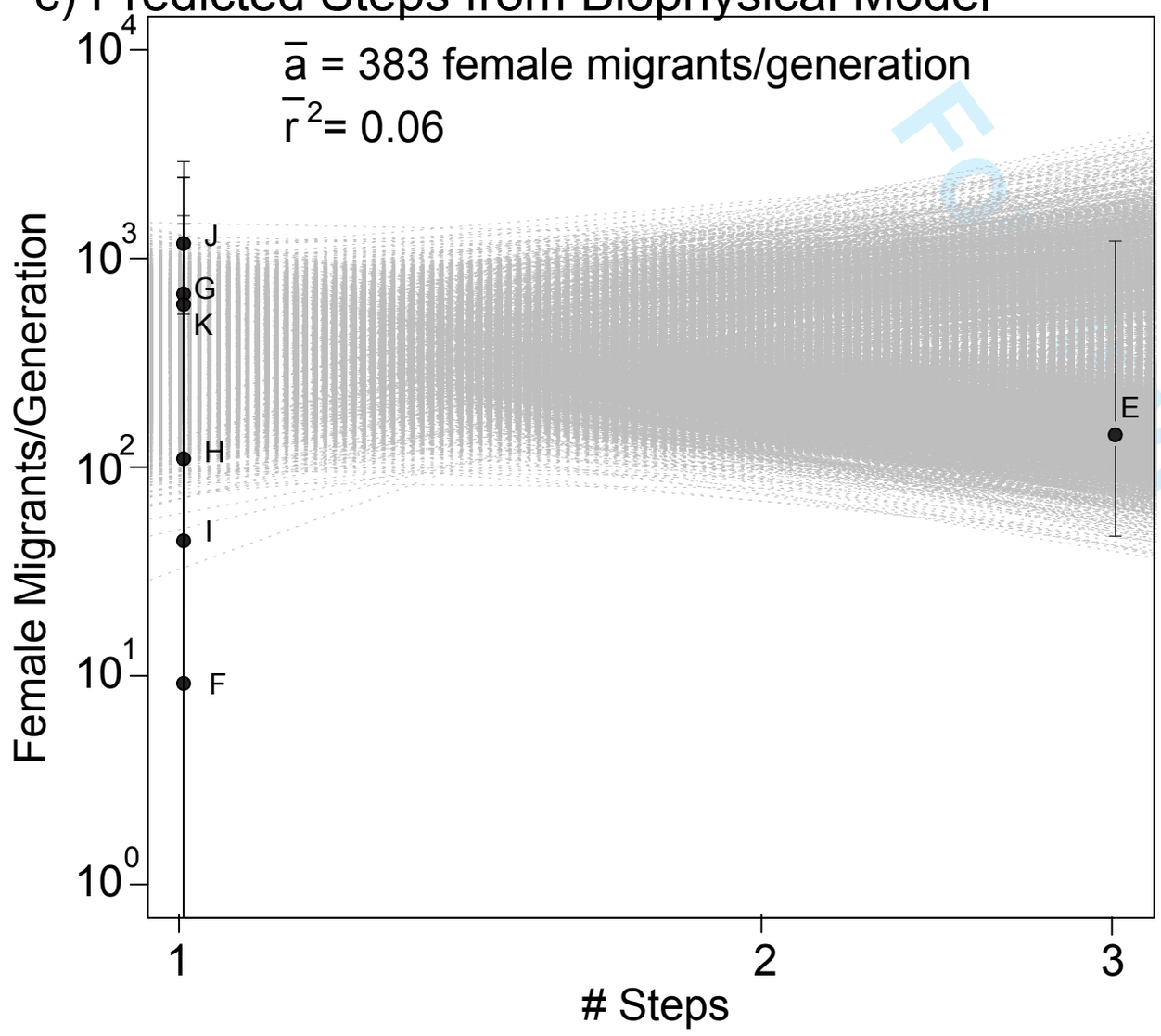
a) Great Circle Distance



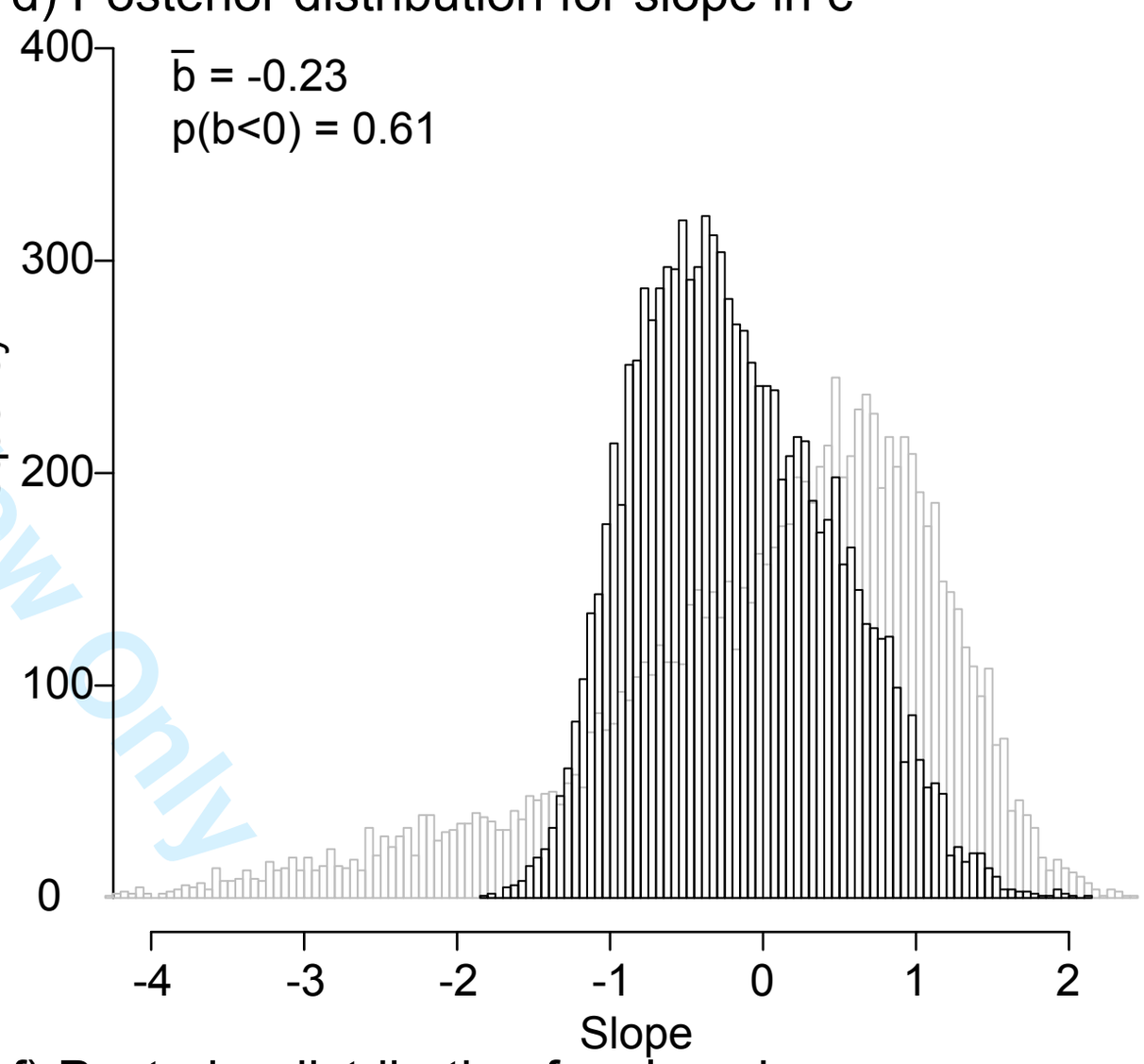
b) Posterior distribution for slope in a



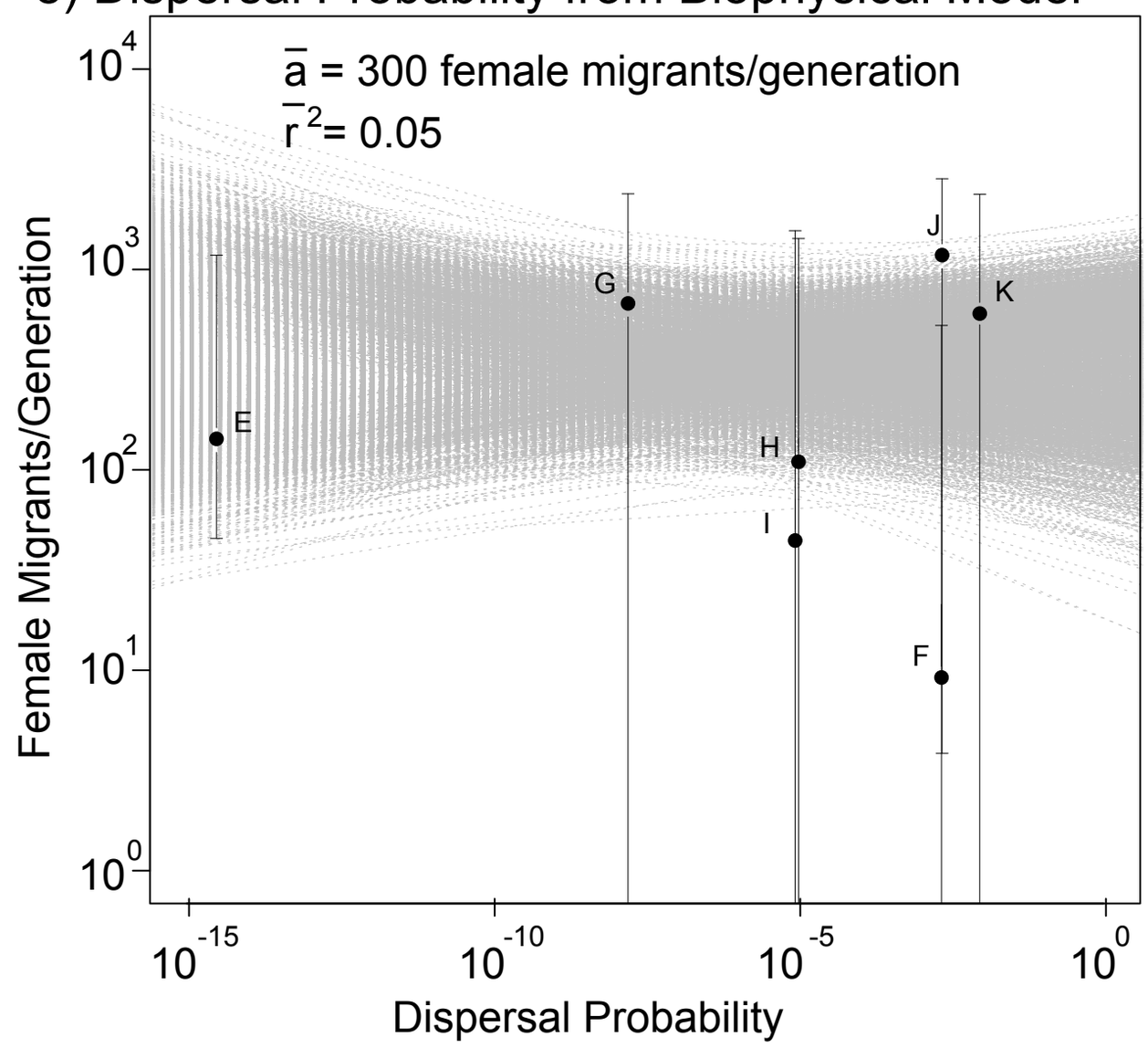
c) Predicted Steps from Biophysical Model



d) Posterior distribution for slope in c



e) Dispersal Probability from Biophysical Model



f) Posterior distribution for slope in e