Marine Biology (1995) 123: 573–581

M. E. Hellberg

Stepping-stone gene flow in the solitary coral *Balanophyllia elegans*: equilibrium and nonequilibrium at different spatial scales

Received: 24 April 1995/Accepted: 27 April 1995

Abstract Limited dispersal should result in genetic differences between populations proportional to geographic distances of separation. This association between gene flow and distance can be disrupted by (1) continuing genetic exchange among distant populations, (2) historical changes in gene flow, and (3) physical barriers or corridors to dispersal. The movements of larvae are thought to determine dispersal capability in benthic marine invertebrates. The solitary scleractinian *Balanophyllia elegans* Verrill possesses crawling larvae capable of only limited dispersal. Paradoxically, however, inferred levels of gene flow between pairs of localities spread over much of the 4000 km range of *B. elegans* exhibited a weaker relationship with geographical separation than that expected for a linear array of populations in which all genetic exchange takes place between adjacent populations. In this paper, I examined the pattern of gene flow (inferred from the frequencies of eight polymorphic allozyme loci) in *B. elegans* at a smaller (1 to 50 km) spatial scale to determine (1) whether gene flow at this spatial scale conformed to the expectations of the stepping-stone model, and (2) whether continuing long-distance gene flow or historical changes in gene flow were responsible for the weak relationship between gene flow and distance observed previously at the rangewide spatial scale. Between May and August 1992, I collected 75 adults from each of 18 localities along the coast of Sonoma County, California, USA. These populations of *B. elegans* were significantly subdivided both among localities separated by 1 to 50 km ($F_{LT} = 0.053$, SE = 0.0075) and among patches separated by 4 to 8 m ($F_{PL} = 0.026$, SE = 0.0023). The observed slope and correlation ($r^2 = 0.54$) between inferred levels of gene flow and the geographic distance at the 1 to 50 km spatial scale conformed to equilibrium expectations (obtained by simulation) for a linear stepping-stone model, although those from the rangewide spatial scale did not. This implies that the mechanisms conferring patterns of inferred genetic differentiation between localities in *B. elegans* differ fundamentally with spatial scale. At a scale of 1 to 50 km, continuing gene flow and drift have equilibrated and the process of isolation-by-distance may facilitate local adaptive change. At a broader spatial scale, historical changes in gene flow, perhaps affected by late Pleistocene climatic fluctuations, disrupt the equilibration of gene flow and genetic drift, so that genetic differentiation may not increase continuously with separation between populations.

Communicated by M.F. Strathmann, Friday Harbor

M.E. Hellberg
Center for Population Biology and Department of Zoology
University of California at Davis, Davis, California 95616, USA

Present address:
Center for Marine Biotechnology and Biomedicine
Scripps Institution of Oceanography
University of California at San Diego
La Jolla, California 92093-0202, USA

Introduction

At equilibrium, the degree of genetic homogenization resulting from gene flow depends on both the number of migrants moving among populations and the relative proximity of source and recipient populations. For a particular average level of gene flow, free exchange of genes between all populations, regardless of their geographical proximity (the "island model" of Wright 1951), diminishes divergence relative to the opposite extreme in which gene flow is restricted to immediately adjacent populations [as under "isolation-by-distance" Wright (1943) or the "stepping-stone model" (Kimura and Weiss 1964)]. Determining the qualitative pattern of gene flow between geographically isolated populations is therefore critical to understanding

microevolutionary processes of differentiation. Inferences about present-day patterns of gene flow, however, may be confounded by the vestiges of historical associations among populations (Larson et al. 1984; Ayre 1991). A full understanding of the consequences of geographic separation on gene flow requires: (1) estimates of the magnitude of gene flow, (2) inferences concerning the geographical relationships of exchanging populations, and (3) determinations of whether inferred patterns of exchange result from past or present processes.

To identify patterns of gene flow using genetic patterns of differentiation, Slatkin and Maddison (1990) and Slatkin (1993) advocate regressing pairwise estimates of gene flow ($\hat{M}$) on the distance separating populations. If gene flow passes largely between immediately adjacent populations then, at equilibrium, the slope of the log-log regression of $\hat{M}$ on the number of steps between populations assumes characteristic values (determined by the spatial distribution of populations) (Slatkin and Maddison 1990; Slatkin 1993). Following a rapid radiation into vacant areas, transient disequilibrium between gene flow and drift disrupts this relationship (Slatkin 1993). The rate at which gene flow and drift between populations returns to an equilibrium correlates directly with the migration rate (Crow and Aoki 1984). Therefore, subsequent to a radiation, gene flow and drift will equilibrate first between neighboring populations, because under a stepping-stone model rates of migration are highest between adjacent populations. More distant populations will carry the mark of historical change longer, and thus exhibit a weaker relationship between gene flow and distance. A recent range expansion may thus produce a pattern indistinguishable from that resulting from continuing dispersal between distant populations at some spatial scales.

In benthic marine invertebrates, the dispersal of larvae constitutes the primary source of continuing gene flow (Scheltema 1977; Knowlton and Jackson 1993). Species with limited larval dispersal capabilities have been associated with higher levels of genetic subdivision (Scheltema 1977; Knowlton and Jackson 1993), greater rates of speciation and extinction (Hansen 1978; Jablonski 1986), and more restricted geographical ranges (Shuto 1974) than species with more broadly dispersing planktonic larvae.

The solitary coral *Balanophyllia elegans* ranges over 4000 km of shoreline, from Baja California north into Alaska. However, *B. elegans* has no planktonic larval stage – its larvae can only crawl and usually settle within 40 cm of their birthplace (Gerrodette 1981). At a broad spatial scale ranging from 40 to 3000 km, distance explains little of the variance in gene flow between populations ($r^2 = 0.086$), and the relationship between inferred levels of gene flow and the distance between populations is weaker than expected for the stepping-stone model (Hellberg 1994). This suggests that either continuing long-distance gene flow or recent historical changes in gene flow act to determine rangewide patterns of gene flow in *B. elegans*.

In this paper, I examine the relationship between inferred levels of gene flow and geographical distance in *Balanophyllia elegans* at a smaller (1 to 50 km) spatial scale in order to distinguish between these alternative mechanisms. I infer levels of genetic subdivision and relationships between gene flow and geographical distance using the same seven polymorphic allozyme loci as at the larger spatial scale (plus one additional locus, *MPI*). I then use simulations to compute expected slopes and correlations between gene flow and geographical distance for a stepping-stone model at equilibrium in which levels of gene flow match those observed in *B. elegans* and are inferred using the same number of polymorphic genetic markers.

Materials and methods

Natural history of *Balanophyllia elegans*

Balanophyllia elegans Verrill lives on rocky substrata down to 500 m depth (Durham and Barnard 1952). Densities generally peak at 12 to 20 m depth, and often decline considerably below 25 m (Gerrodette 1979b; my own personal observations). In combination with the broad geographical range of *B. elegans*, this limited bathometric range results in an effectively one-dimensional coastal distribution. Local densities at depths of 10 to 15 m can exceed 500 per square meter (Fadlallah 1983; my own personal observations).

Sexes are separate. Females brood sexually-produced planulae that crawl over hard substrata for one to several days before attaching, usually very close to their mother (Gerrodette 1981; Fadlallah 1983). Males release sperm during a short period of time in the boreal fall (Fadlallah and Pearse 1982). The distance traveled by successful sperm of *Balanophyllia elegans*, and therefore neighborhood size, is unknown.

Collection of samples

I collected 75 individuals from each of 18 localities from along 54 km of the coast of Sonoma County, California, USA (Fig. 1). I chose this region because rocky habitats suitable for *Balanophyllia elegans* occur continuously (north of the Russian River), and the coastline is almost linear. All localities lie between Latitude 38° and 39° N and Longitude 123° and 124° W. Other than the northernmost (SIP) and southernmost (SBR) localities, adjacent localities were < 2 km apart.

At each locality, I gathered 15 individuals from each of 5 patches. Patches were separated by 4 to 8 m, and were located within an area of 50 to 200 m^2. Patches included all individuals found within a circle of 33 cm radius, the neighborhood distance calculated from direct measurements of larval dispersal (Fig. 2 of Gerrodette 1981). I collected only adult corals (those greater than $\simeq 6$ mm diam (Gerrodette 1979a) or ~ 150 mm^3 vol (Fadlallah 1983). All sampled patches were between 6 and 18 m below mean lower low water.

I sampled all localities between May and August of 1992. I removed encrusting algae and epifauna from coral samples before freezing them in liquid nitrogen in the field. I also removed large eggs and embryos found within females to avoid confounding larval and maternal genotypes. Samples were subsequently stored at -80 °C.

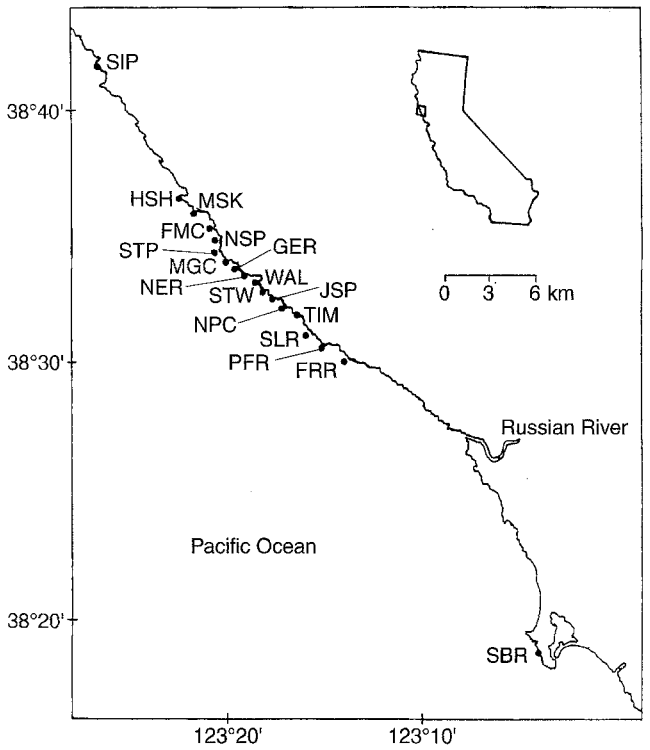

Fig. 1 Localities sampled for *Balanophyllia elegans* (*Inset* location of Sonoma County within state of California). Abbreviations and coordinates for sampled localities are presented in descending order of latitude [*SIP* (Sipple's Point, Sea Ranch, 38°41′31″ N; 123°31′12″ W); *HSH* (Horseshoe Point, 36′26″ N; 22′17″ W); *MSK* (Fish Triangle Point, 35′56″ N; 21′27″ W); *FMC* (south of Fisk Mill Cove, 35′18″ N; 20′44″ W) *NSP* (west of Stump Beach, 34′49″ N; 20′25″ W); *STP* (Salt Point, 34′10″ N; 20′24″ W); *MGC* (mouth of Gerstle Cove, 33′51″ N; 19′51″ W); *GER* (south end of Gerstle Cove, 33′47″ N; 19′32″ W); *NER* (*Nereocystis* Ridge, 33′34″ N; 18′58″ W); *WAL* (south side of Ocean Cove, 33′14″ N; 18′16″ W); *STW* (Stillwater Cove, 32′47″ N; 16′17″ W); *JSP* (south of Wilson Triangle, 32′03″ N; 17′44″ W); *NPC* (north of Phallus Cove, 32′12″ N; 17′04″ W); *TIM* (Timber Cove, 31′51″ N; 16′17″); *SLR* (*Gersemia* Rocks, 30′55″ N; 15′39″ W); *PFR* (Northwest Cape, Fort Ross, 30′37″ N; 14′58″ W); *FRR* (Fort Ross Reef, 30′00″ N; 13′57″ W); and *SBR* (South Bodega Reserve, 18′47″ N; 04′06″ W)]

Electrophoresis

I characterized eight locus genotypes for individuals collected from all patches at each of the 18 locations shown in Fig. 1. I ran all samples using 12%(w/v) starch gel electrophoresis. Enzyme stains and electrophoretic conditions were modified from Selander et al. (1971). I assayed four loci – hexokinase (*HK*, EC 2.7.1.1), phosphoglucomutase (*PGM*, EC 2.7.5.1), phosphoglucose isomerase (*PGI*, EC 5.3.1.9), and mannose-phosphate isomerase (*MPI*, EC 5.3.1.8) – using the pH 8.0 Tris-citrate buffer system of Selander et al. I added 4 ml of 2-mercaptoethanol to 300 ml of starch just before pouring to improve the resolution of *PGI* bands. Four other loci – triosephosphate isomerase (*TPI*, EC 5.3.1.1), peptidase (*PEP*, EC 3.4.11/13, leucyl alanine substrate), aspartate aminotransferase (*AAT*, EC 2.6.1.1), and leucine aminopeptidase (*LAP*, EC 3.4.11.1/2) – were assayed using the more dilute pH 8.0 Tris-citrate buffer system of Ward and Beardmore (1977).

Laboratory crosses confirmed that electromorphs were allelic (Hellberg unpublished data). Alleles were numbered to indicate their percent mobility relative to the most common allele at that locus from Santa Cruz (Hellberg 1994). One allele (65) at the *Aat* locus which could not be consistently resolved previously was scored in this study. I included two internal standards at two positions within each gel.

Analysis

I used Wright's *F*-statistics (Wright 1978) to characterize overall genetic subdivision. These statistics use departures from levels of heterozygosity expected under complete panmixia to partition total inbreeding (F_{IT}) into components due to inbreeding within subpopulations (F_{IS}) and subdivision among populations (F_{ST}). For selectively neutral loci (assuming migration rates exceed mutation rates), estimates of F_{ST} based on different loci can be combined without assuming equal mutation rates (Slatkin 1985). However, historical events can contribute to the variance of F_{ST} among loci. Gene flow (Nm), the average number of genetically effective migrants exchanged between populations each generation, can be estimated from F_{ST} using the approximation $Nm = (1 - F_{ST})/4F_{ST}$. This provides an unbiased estimator of gene flow that is insensitive to moderate levels of selection (Slatkin and Barton 1989). Two alternative estimators of Wright's F_{ST}, G_{ST} of Nei (1973) and θ of Weir and Cockerham (1984), both accurately estimate Nm when genes move between all populations with equal likelihood, a moderate number of subpopulations (10) are sampled, and m (the rate of migration) is > 0.1 (Slatkin and Barton 1989; Cockerham and Weir 1993). Thus, results for hierarchical *F*-statistics for these two parameters should be nearly identical for the low migration rates expected in *Balanophyllia elegans*. Wright's parameters were calculated using the estimators of Weir and Cockerham (1984). I computed allele frequencies and hierarchical *F*-statistics from genotype frequencies using the program of Weir (1990). The program estimates standard errors by jackknifing over loci. I figured overall F_{PL} (subdivision of patches within localities) manually using the relationship $F_{PL} = (F_{IL} - F_{IP})/(1 - F_{IP})$, where F_{IL} denotes total heterozygote deficiency within localities and F_{IP} specifies heterozygote deficiencies within patches. I also calculated the jackknifed standard error for F_{PL} by hand using the formula of Weir. Bootstrapped confidence intervals were used to determine whether *F*-statistics departed significantly from panmictic expectations.

Slatkin and Maddison (1990) and Slatkin (1993) used $\hat{M}$, the estimate of Nm calculated separately for pairs of populations, to extract the geographical pattern of gene flow from spatial patterns of differentiation. $\hat{M}$ should not be viewed as the actual number of migrants moving between a pair of populations each generation, but rather as equivalent to the number of migrants required to account for observed genetic differences if migrants could move directly between populations. Simulations demonstrate (Slatkin and Maddison 1990; Slatkin 1993; Hellberg 1994), and analytical models confirm (Slatkin 1991), that in a one-dimensional array of stepping-stones the expected slope of a regression of $\log_{10}(\hat{M})$ vs $\log_{10}$ (distance of separation) (in number of steps) is -1.0. Under a two-dimensional stepping-stone, the expected slope of the regression is approximately -0.5. I computed $\hat{M}$ based on G_{ST} (Nei 1973) and θ (Weir and Cockerham 1984) for all pairwise combinations of populations using a program provided by M. Slatkin. Both of these calculations have advantages and disadvantages for estimating $\hat{M}$. θ is unbiased, but may become negative when Nm is large, resulting in undefined estimates of $\hat{M}$. G_{ST} is bound positive, but biased as a result. $\hat{M}$ based on G_{ST} should be twice that based on θ for pairwise comparisons (Slatkin 1993; Cockerham and Weir 1993). I reckoned the shortest nautical distance between localities from 24 000:1 scale maps.

The significance of the relationship between log ($\hat{M}$) and log (distance) cannot be evaluated using standard regression techniques because the regression would be based on non-independent pairwise comparisons. As an alternative, I used ordinary least-squares (OLS)

regression, with degrees of freedom adjusted to the number of populations (not the number of pairwise contrasts) to determine whether the slope of the regression was different from zero. The confidence intervals of these adjusted OLS regressions are identical (to two decimal places) to confidence intervals obtained using Mantel's test (Hellberg 1994).

I determined the slope of significant OLS regressions using reduced major-axis (RMA) regression. Asymmetric 95% confidence intervals about the slope of the RMA regressions were calculated following McArdle (1988), using degrees of freedom corresponding to the number of populations sampled. RMA regression better estimates the relationship between gene flow and the number of steps separating populations than OLS regression when geographical distance approximates the number of steps between populations (Hellberg 1994). OLS regression underestimates the slope of the regression in such cases because OLS regression assumes that the independent variable (the number of steps separating populations, in this case) is measured without error.

To determine whether one locus contributed disproportionately to the geographic pattern of genetic differentiation, I excluded one locus at a time (jackknifed over loci) and recalculated the slope of log ($\hat{M}$) vs log (distance) as well as the correlation between these two.

I calculated expected heterozygosities at each locality using allele-frequency data for each locus and then averaging over the eight loci surveyed using BIOSYS (Swofford and Selander 1989). Expected, rather than observed, heterozygosities (Nei 1978) were used so that comparisons between locations would not be obscured by inbreeding or subdivision within localities. As an alternative measure of genetic variation, I also calculated the number of alleles per locus for each locality. Leberg (1992) showed experimentally that the number of alleles per locus may be a more sensitive indicator of past genetic bottlenecks than multi-locus measures of heterozygosity.

Simulations

I used a simulation model modified from one originally developed by Slatkin (1993) to determine the expected regression slopes and coefficients of determination (r^2) between log (gene flow) and log (distance) when propagules disperse according to the stepping-stone model and gene flow is inferred from eight mendelian nuclear genetic markers. The simulations generated and analyzed allelic frequencies for diploid organisms in populations of 10 000 individuals, with gene flow passing only between adjacent populations. Each replicate (= locus) of the simulation consisted of two parts. First, the history of each sampled gene was traced back to a single common ancestor using a coalescence process. Once the ancestry of the sampled genes had been established, the second part of the simulation assigned allelic identities to them.

A single iteration consisted of eight such replicates, corresponding to the eight allozyme loci I used in my empirical study. I adjusted the average heterozygosity of these simulated loci (= 0.28) to closely match those surveyed in *Balanophyllia elegans* by setting the product of population size and mutation rate (Nu) = 0.005 (Hellberg 1994). In all of my simulations, I sampled 25 diploid individuals from each of 18 populations. Attempts to run larger simulations failed, although limited results suggest that when $n \geq 25$, changing the number of individuals sampled from each population matters less than changing the number of populations sampled. Sampled populations were centered in a one-dimensional array of 101 populations to moderate edge effects. Each sampled population was separated from the nearest sampled population by two steps (i.e. I sampled populations 32, 34, 36, ... 66). Each generation, a fraction, $1/m$, of the population consisted of immigrants drawn equally from the two adjacent populations. The value of m was set so that Nm would be 1 or 30, corresponding to the levels of gene flow between neighboring populations observed at rangewide and 1 to 50 km spatial scales, respectively.

Resulting gene-frequency data were analyzed by regressing $\log_{10}(\hat{M})$ versus $\log_{10}$(distance of separation) as described above. As with the empirical data, $\hat{M}$ was calculated using either θ (Weir and Cockerham 1984) or G_{ST} (Nei 1973). I computed the distribution of regression slopes from 100 iterations. The narrower the frequency distributions, the greater the power of particular conditions (estimator used, level of gene flow between neighbors) to reveal patterns established by gene flow between adjacent populations.

Results

The eight loci surveyed in *Balanophyllia elegans* were highly polymorphic at most of the localities sampled (Table 1). Average heterozygosity within localities was 0.35 (SE = 0.00702). Neither expected heterozygosities ($F = 0.665$, $P > 0.83$) nor the average number of alleles per locus ($F = 0.649$, $P > 0.85$) differed significantly among localities, although statistical tests for such differences have low power (Archie 1985). F-statistics suggest significant heterozygote deficiencies at all levels of hierarchical spatial structure (as determined by bootstrapped 95% confidence intervals which excluded zero) except within patches (Table 2). Nonsignificant F_{IP} suggests no departure from panmictic expectations within patches. The mean F_{LT} (for subdivision between localities) corresponded to an Nm between locations of ~ 4.5 immigrants per generation.

For populations of *Balanophyllia elegans* from different localities, the slope of the log-log RMA regression between $\hat{M}$ (calculated using G_{ST}) and distance was -0.89 and the coefficient of determination (r^2) was 0.54 (Fig. 2). The asymmetric 95% confidence interval around this RMA regression coefficient ranged from -0.63 to -1.27, thus including -1.0, the theoretically expected slope for a linear array of stepping-stones at equilibrium. The OLS regression between log($\hat{M}$ based on G_{ST}) and log (distance of separation) was significant (slope = -0.66, $p < 0.001$).

When $\hat{M}$ was calculated using θ, the slope of the log-log RMA regression was -1.28 (Fig. 2), and again included -1.0 in its confidence interval ($-0.83 \geq$ slope ≥ -1.91). Distance explained less of the variation in $\hat{M}$ than when G_{ST} was used to calculate $\hat{M}$ ($r^2 = 0.41$). Fig. 2 shows three outlier points ($\hat{M} > 350$) for comparisons based on θ. In addition, another three values were negative, so that their logarithms were undefined and excluded from the analysis. The slope of the OLS regression based on θ was significant (slope = -0.82, $P < 0.001$). The intercept of the OLS regression suggested that the equivalent of 37 migrants moved between populations separated by 1 km. Assuming a perfect linear stepping-stone and figuring the average distance between localities (6.7 km), this value matches the average number of migrants estimated by hierarchical F-statistics above.

The regression slopes and correlations reported here for *Balanophyllia elegans* fall within the range of values for simulated linear stepping-stones in which gene flow

Table 1 *Balanophyllia elegans.* Allele frequencies for 18 localities from Sonoma County, California, USA. Alleles numbered to indicate mobility relative to most common allele from Santa Cruz locality of a previous study (Hellberg 1994). All frequencies based on samples of total of 75 individuals from each locality (15 individuals from each of 5 patches). Localities listed from north to south. Locality abbreviations as in Fig. 1

Locus allele	Locality																	
	SIP	HSH	MSK	FMC	NSP	STP	MGC	GER	NER	WAL	STW	JSP	NPC	TIM	SLR	PFR	FRR	SBR
HK																		
113								0.027	0.020						0.007	0.013		
107	0.060	0.027			0.020	0.007	0.060	0.007	0.020				0.007	0.007	0.027	0.027		0.007
100	0.940	0.973	1.000	1.000	0.980	0.973	0.927	0.960	0.960	0.847	0.787	0.787	0.766	0.906	0.886	0.847	0.953	0.833
93						0.020	0.013	0.007		0.153	0.213	0.213	0.227	0.087	0.080	0.113	0.047	0.160
PGM																		
100	0.487	0.740	0.720	0.733	0.720	0.773	0.827	0.720	0.800	0.673	0.633	0.640	0.653	0.773	0.787	0.860	0.567	0.533
91	0.513	0.260	0.280	0.267	0.280	0.227	0.173	0.280	0.200	0.327	0.367	0.360	0.347	0.227	0.213	0.140	0.433	0.467
PGI																		
100	0.413	0.280	0.273	0.207	0.300	0.173	0.220	0.240	0.353	0.327	0.353	0.307	0.260	0.367	0.307	0.467	0.360	0.387
85	0.587	0.720	0.727	0.793	0.700	0.827	0.780	0.760	0.647	0.673	0.647	0.693	0.740	0.633	0.693	0.533	0.640	0.593
68																		0.020
TPI																		
100	0.973	0.927	0.867	0.813	0.867	0.827	0.887	0.820	0.893	0.960	0.973	0.906	0.927	0.860	0.867	0.927	0.967	0.780
93	0.027	0.073	0.133	0.187	0.133	0.173	0.113	0.180	0.107	0.040	0.027	0.087	0.073	0.140	0.133	0.073	0.020	0.220
86												0.007					0.013	
PEP																		
107	0.167	0.073	0.053	0.073	0.020	0.087	0.073	0.067	0.027	0.147	0.193	0.160	0.120	0.053	0.100	0.113	0.033	0.253
100	0.633	0.833	0.833	0.760	0.893	0.773	0.740	0.806	0.920	0.706	0.667	0.660	0.613	0.833	0.773	0.847	0.960	0.394
92	0.200	0.094	0.113	0.167	0.087	0.140	0.187	0.127	0.053	0.147	0.140	0.180	0.267	0.113	0.127	0.040	0.007	0.353
AAT																		
100	0.060	0.113	0.140	0.027	0.033	0.020	0.060	0.027	0.013	0.007	0.013	0.040	0.007	0.087	0.067	0.053	0.100	0.013
92		0.167	0.067	0.107	0.054	0.100	0.127	0.067	0.067							0.007		
82	0.940	0.720	0.793	0.866	0.913	0.833	0.780	0.886	0.907	0.953	0.873	0.787	0.880	0.907	0.873	0.940	0.840	0.987
65						0.047	0.033	0.020	0.013	0.040	0.114	0.173	0.113	0.007	0.060		0.060	
LAP																		
115											0.007	0.013	0.013					
111	0.247	0.400	0.427	0.306	0.427	0.313	0.427	0.447	0.480	0.473	0.600	0.560	0.560	0.473	0.400	0.433	0.533	0.427
106	0.720	0.413	0.333	0.427	0.313	0.447	0.460	0.353	0.220	0.313	0.220	0.240	0.254	0.360	0.466	0.447	0.347	0.020
100	0.033	0.180	0.194	0.260	0.246	0.187	0.093	0.193	0.254	0.180	0.120	0.140	0.133	0.147	0.100	0.093	0.120	0.100
98			0.013			0.020	0.013							0.007	0.007			
96		0.007	0.033	0.007	0.007	0.033	0.007	0.007	0.033	0.027	0.053	0.047	0.040	0.013	0.020	0.027		0.447
90					0.007				0.013	0.007					0.007			0.007
MPI																		
111	0.133	0.453	0.407	0.440	0.300	0.347	0.353	0.460	0.553	0.413	0.340	0.367	0.427	0.480	0.567	0.547	0.233	0.013
106	0.020								0.020									0.013
100	0.840	0.540	0.586	0.553	0.700	0.653	0.640	0.540	0.427	0.587	0.660	0.633	0.573	0.520	0.433	0.453	0.760	0.967
96	0.007	0.007	0.007	0.007			0.007										0.007	0.007

Table 2 *Balanophyllia elegans.* Weir and Cockerham (1984) estimates of Wright's *F*-statistics calculated separately for each locus. Means and standard errors were obtained by jackknifing over loci (F_{IT} individuals within total; F_{LT} among localities within total; F_{IL} individuals within localities; F_{PL} among patches within localities; F_{IP} inbreeding within patches) F_{PL} was calculated as $F_{PL} = (F_{IL}(-F_{IP})/(1-F_{IP})$

Locus	F_{IT}	F_{LT}	F_{IL}	F_{PL}	F_{IP}
HK	0.0886	0.0729	0.0170	0.0315	−0.0150
PGM	0.1062	0.0425	0.0666	0.0294	0.0383
PGI	0.0279	0.0207	0.0073	0.0252	−0.0184
TPI	0.0641	0.0294	0.0357	0.0161	0.0199
PEP	0.1056	0.0650	0.0435	0.0343	0.0095
AAT	0.0812	0.0431	0.0398	0.0310	0.0091
LAP	0.0800	0.0601	0.0212	0.0182	0.0031
MPI	0.1730	0.0764	0.1046	0.0197	0.0866
Mean	0.0943[a]	0.0529[a]	0.0436[a]	0.0257[a]	0.0195
SE	0.0183	0.0075	0.0140	0.0023	0.0145

[a] Bootstrapped 95% confidence interval does not include zero

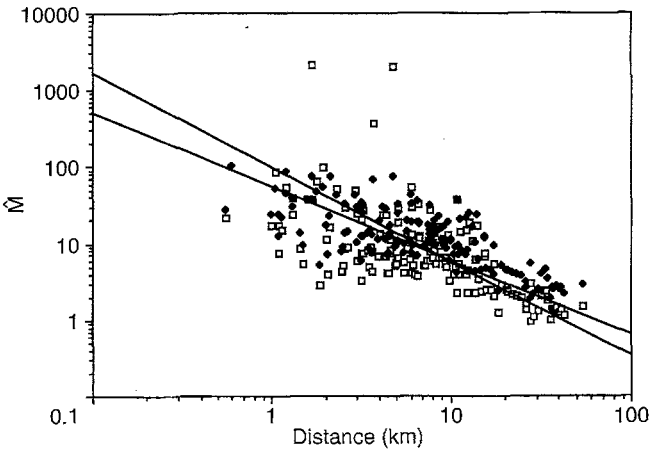

Fig. 2 *Balanophyllia elegans.* Reduced major axis regressions of $\log_{10}$ gene flow ($\hat{M}$, individuals/generation) vs $\log_{10}$ distance of separation (km) for all pairwise combinations of 18 localities examined in present study. $\hat{M}$ based on average of 8 electrophoretically polymorphic enzyme loci as estimated by G_{ST} (◆, $y = 1.80 - 0.89x$; $r^2 = 0.54$; Nei 1973) and θ (□, $y = 1.94 - 1.28x$, $r^2 = 0.41$, Weir and Cockerham 1984)

is inferred using eight polymorphic nuclear loci (Fig. 3b, d, f). Most notably, Fig. 3f shows that when 30 migrants move between neighboring populations each generation, the coefficient of determination (r^2) between log $\hat{M}$ and log separation calculated using Weir and Cockerham's θ tends to be lower than that for Nei's G_{ST}. In contrast, for lower levels of gene flow (Fig. 3a, c and e), the expected regression slopes and correlations matched very closely for the two calculations. Analytically, the variance in estimates of gene flow based on θ should exceed those based on G_{ST}, especially for pairwise comparisons when Nm is large (Cockerham and Weir 1993).

To assess whether any single locus determined the overall negative correlation between log($\hat{M}$) and log (distance), I calculated slopes of the log-log RMA regression (using G_{ST}) after excluding loci one at a time from the analysis. All slopes were significantly negative, and in every case the confidence intervals included −1.0 and were less than −0.5 (Table 3). The correlations between $\hat{M}$ and distance likewise remained high and similar to that for the full eight-locus data set (Table 3).

Discussion

Determining whether and how the distance of separation between populations acts as a barrier to gene flow is critical both to understanding how species respond evolutionarily to environmental variation, and to evaluating alternative models of speciation. To do this requires (1) inferring patterns of genetic exchange between populations from raw data on geographical patterns of genetic variation, and (2) comparing inferred patterns of gene flow to patterns expected under different dispersal regimes.

Given the limited larval dispersal capabilities of *Balanophyllia elegans* and its broad, linear coastal distribution, I expected the relationship between $\hat{M}$ and geographic separation in this coral to match the expectations of the stepping-stone model. In the present study, I inferred levels of gene flow in *B. elegans* at a spatial scale spanning distances of separation between 600 m and 54 km. As at both smaller (among patches separated by 4 to 40 m) and rangewide (40 to 3000 km) spatial scales (Hellberg 1994), I found significant genetic subdivision among localities at this intermediate spatial scale ($F_{LT} = 0.053$, Table 2). In contrast to these two other spatial scales, however, pairwise estimates of inferred levels of gene flow correlated strongly with distance (Fig. 2). In addition, and unlike the case at the two other spatial scales, the RMA slopes of gene flow (estimated using either G_{ST} or θ) vs distance included −1.

These regression slopes and correlations closely match those generated by simulated linear stepping-stone models employing the same number of variable nuclear markers (Fig. 3). When $Nm = 30$ (the approximate level of gene flow between adjacent sampled populations of *Balanophyllia elegans*), inferring gene flow using Nei's G_{ST} results in both a larger coefficient of determination (Fig. 3f) and a greater, less-biased RMA slope (Fig. 3d) than when these same values are estimated using Weir and Cockerham's θ. These expected differences are evident in the data presented here for *B. elegans*. Both r^2 and RMA slope based on G_{ST} (0.54 and −0.89, respectively) are larger than the corresponding values based on θ (0.41 and −1.28). When $Nm = 1$, the differences between G_{ST} and θ were not

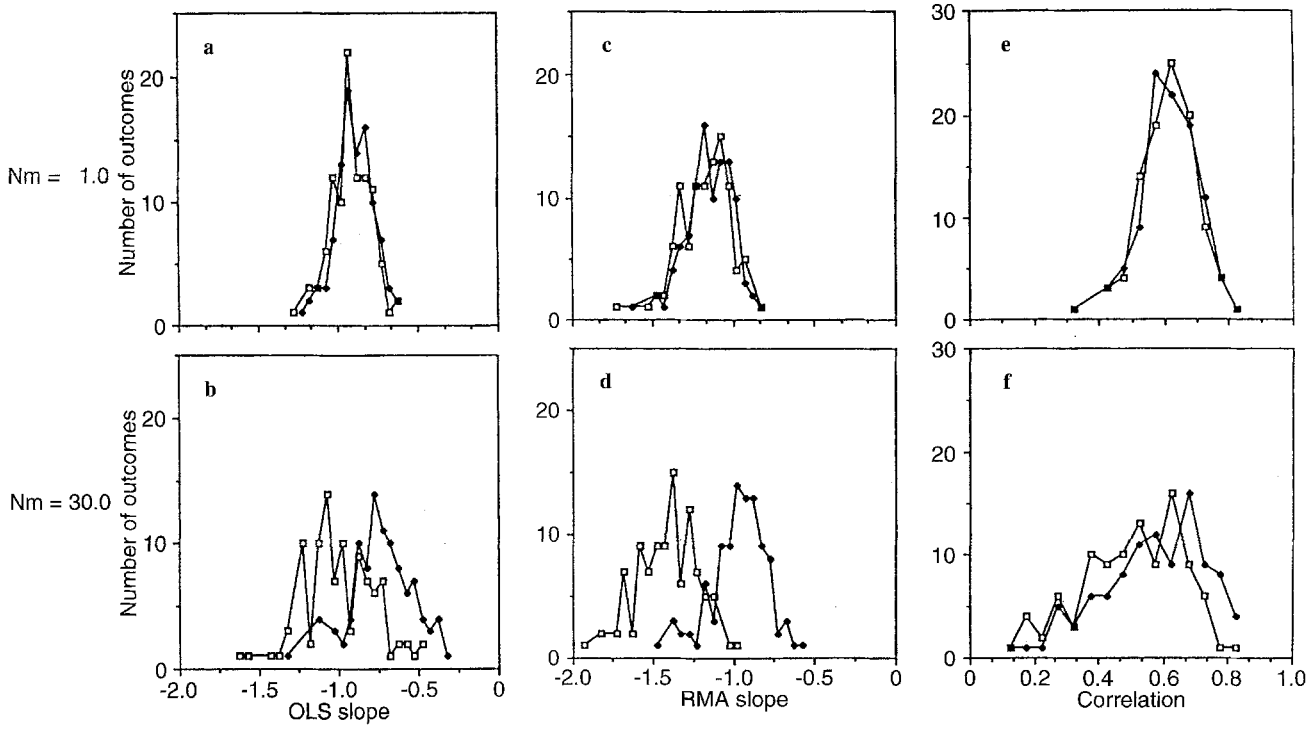

Fig. 3 *Balanophyllia elegans*. Distributions of slopes of ordinary least-squares regressions (**a**, **b**,), reduced major axis regressions (**c**, **d**), and correlations (r^2, **e** and **f**) between $\log_{10} \hat{M}$ (individuals/generation) and $\log_{10}$ (number of steps separating populations) in a one-dimensional stepping-stone as a function of gene flow (*Nm*) between adjacent populations. Values of *Nm* (1.0 and 30.0) correspond to those observed between adjacent populations of *Balanophyllia elegans* at spatial scales spanning > 3000 km (Hellberg 1994) and 54 km (present study), respectively. $\hat{M}$ calculated using either G_{ST} (◆) or θ (□) using eight simulated loci. Theoretically expected regression slope is −1.0. Samples of 25 diploid individuals were taken from each of 18 central populations, spaced two steps from nearest sampled population, within a 101×1 array of populations (●). Distributions for 100 simulated data sets, $N = 10\,000$ for all populations; *OLS* ordinary least-squares regression; *RMA* reduced major-axis division

Table 3 *Balanophyllia elegans*. Slopes of RMA regression of $\log_{10}\hat{M}$ (individuals/generation) vs $\log_{10}$ nautical distance (km) based on data sets from which one locus has been excluded. $\hat{M}$ estimated using Nei's gene-diversity statistic G_{ST} (Nei 1973) (*CI* confidence interval)

Locus excluded	Slope	95% CI of slope	r^2
HK	−0.92	$-0.79 \geq b \geq -1.12$	0.54
PGM	−0.87	$-0.75 \geq b \geq -1.00$	0.56
PGI	−0.94	$-0.81 \geq b \geq -1.07$	0.53
TPI	−0.91	$-0.78 \geq b \geq -1.04$	0.55
PEP	−0.89	$-0.77 \geq b \geq -1.01$	0.56
AAT	−0.98	$-0.84 \geq b \geq -1.12$	0.50
LAP	−0.87	$-0.73 \geq b \geq -1.02$	0.48
MPI	−0.86	$-0.72 \geq b \geq -1.01$	0.49

apparent (Fig. 3a, c, e). These simulation results suggest that when levels of gene flow are expected to be high, the correlation between gene flow and distance will be stronger and the RMA slope less biased (closer to its analytical expectation) when gene flow is estimated using G_{ST} rather than θ. Regardless of the estimator used, the power to detect relationships between gene flow and distance varies considerably with the number and heterozygosity of the genetic markers employed (Fig. 3 in Hellberg 1994), making simulations a necessary component of any such investigation.

Natural selection acting directly on allozyme markers may generate allele-frequency differences at spatial scales similar to those considered in this study (Koehn et al. 1980; Burton 1986; Johannesson and Johannesson 1989). However, in *Balanophyllia elegans*, no single allozyme marker disproportionately affects the regression coefficient or correlation between gene flow and distance at either the intermediate spatial scales analyzed in this paper (Table 3), or at larger spatial scales (Hellberg 1994). This correspondence between eight presumably unlinked genetic markers seems unlikely had natural selection, acting independently on each of the allozyme loci, generated the pattern. Barring natural selection, gene flow in *B. elegans* at a spatial scale of hundreds of meters to tens of kilometers apparently proceeds primarily between neighboring populations.

At the larger, rangewide spatial scale (Hellberg 1994), the slope of the regression between $\hat{M}$ and distance is less than −1 (RMA slope = −0.60, $\hat{M}$ based on θ), and distance explains < 9% of the variation in gene flow, far less than expected for a linear stepping-stone at equilibrium (cf. Fig. 3e). Three different mechanisms may account for this weak relationship: (1) continuing long-distance genetic exchange, (2) disequilibria between gene flow and drift, and (3) barriers to gene flow in addition to distance of the separation itself.

Ongoing long-distance gene flow in *Balanophyllia elegans* could be mediated by adults attached to small boulders or abalone and rafted by kelp. In order to agree with the patterns I found at intermediate and large spatial scales, the magnitude of such gene flow would have to be sufficiently great to disrupt patterns over large scales, yet not so great as to hinder equilibration at the intermediate scale. Given the probable sporadic nature of rafting, this seems unlikely. In addition, the weak relationship at the broadest spatial scale reflects low levels of gene flow between proximate localities, rather than high levels between distant localities, suggesting that continuing long-distance gene flow is unlikely (Hellberg 1994).

If the weak correlation at larger spatial scales is due to insufficient time to reach equilibrium following historical changes in gene flow, then the relationship between $\hat{M}$ and distance should be greater at smaller spatial scales. This is because the equilibration time for F_{ST} (on which $\hat{M}$ is based) varies inversely with the migration rate among populations (Crow and Aoki 1984) and, under a stepping-stone model, migration increases with proximity. Using Crow and Aoki's approximation for the time required for F_{ST} to go halfway to equilibrium ($t_{1/2} = \ln 2/(2m + 1/N)$, I estimated that *Balanophyllia elegans* would require $> 40\,000$ yr to do this over its entire range (Hellberg 1994). This exceeds the interval between successive climatic fluctuations which may substantially shift the latitudinal range of *B. elegans*. At the scale of the present study, using the same demographic parameters as before but with an average of 4.5 immigrants per generation (calculated from F_{LT}, Table 2), it would take 4300 yr to go halfway to equilibrium, less than the interval between major late Pleistocene climatic fluctuations. Furthermore, the region sampled here lies at the center of the geographic range of *B. elegans*, perhaps buffering it from latitudinal range shifts.

Genetic subdivision within marine species can be strongly altered by coastal barriers (Johnson and Black 1991). Although such physical barriers to dispersal could decrease the association between gene flow and drift at a rangewide spatial scale, they should also increase the magnitude of the regression between $\hat{M}$ and distance to absolute values > 1. The sampled region in Sonoma County was chosen especially for its absence of any obvious barriers to dispersal (such as headlands or extensive stretches of uninhabitable soft substrate) other than geographic distance. Disentangling the effects of simple geographic isolation-by-distance and differentiation brought about by barriers to dispersal would require sampling replicate regions which impose and lack specific barriers to dispersal.

Taken together, the patterns of gene flow I found here and previously (Hellberg 1994) in *Balanophyllia elegans* suggest that different processes act to produce geographical genetic variation at different spatial scales. Within localities, continuing gene flow between patches separated by 5 to 40 m stifles any relationship between gene flow and distance. At the intermediate spatial scale presented here, moderately low levels of continuing gene flow dictate that the relationship between gene flow and drift is at equilibrium. The process of stepping-stone gene flow at this spatial scale should facilitate the evolution of local adaptive differences (Janson 1983; Janson and Ward 1984; Behrens Yamada 1989), as populations exchanging genes should face similar selective regimes (Endler 1973; Bell 1992). Finally, at larger spatial scales, gene flow and drift will not have equilibrated and genetic similarity will be determined by historical patterns of association between populations and by random drift. Historical changes in patterns of genetic exchange could arise from regional extinctions produced by changes in climate. Such climatic changes may place serious limitations on the likelihood of parapatric speciation in marine taxa with limited larval dispersal. However, these same climatic changes may transiently isolate peripheral populations in relatively warm- or cold-water refugia at the northern and southern extremes of the range, facilitating peripatric speciation (Valentine and Jablonski 1983; Vermeij 1989; Reid 1990; Lindberg 1991).

Acknowledgements I thank all who assisted me in collecting, especially S. Wing, H. Fasteneau, and A. Glusencamp. I thank R. Grosberg, R. Burton, J.E. Duffy, Y. Fadlallah, D. Levitan, P. Marko, A.R. Palmer, H.B. Shaffer, M. Strathmann, G. Vermeij and an anonymous reviewer for helpful discussions and criticisms of drafts of the manuscript. Special thanks to N. Taylor for assistance with modification of computer programs. M. Slatkin generously provided programs for both the pairwise calculations and simulations (developed with the support of NIH Grant GM40282). This research was funded by grants to the author from Sigma Xi, the Lerner-Gray Fund of the American Museum of Natural History, the Graduate Research Committee of the University of California at Davis, the Bodega Marine Laboratory, the UCD Center for Population Biology, the San Diego Supercomputer Center, and by NSF grants to R.K. Grosberg.

References

Archie JW (1985) Statistical analysis of heterozygosity data: independent sample comparisons. Evolution 39: 623–637

Ayre DJ (1991) Population subdivision in Australian temperate marine invertebrates: larval connections versus historical factors. Aust J Ecol 15: 403–411

Behrens Yamada S (1989) Are direct developers more locally adapted than planktonic developers? Mar Biol 103: 403–411

Bell G (1992) Five properties of the environment. In: Grant PR, Horn HS (eds) Molds, molecules, and evolution. Princeton, pp 33–56

Burton RS (1986) Evolutionary consequences of restricted gene flow among natural populations of the copepod *Tigriopus californicus*. Bull mar Sci 39: 526–535

Cockerham CC, Weir BS (1993) Estimation of gene flow from *F*-statistics. Evolution 47: 855–863

Crow JF, Aoki K (1984) Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proc natn Acad Sci USA 81: 6073–6077

Durham JW, Barnard JL (1952) Stony corals of the Eastern Pacific collected by the Velero III and Velero IV. Allan Hancock Pacif Exped 16: 1–110

Endler JA (1973) Gene flow and population differentiation. Science, NY 179: 243–250

Fadlallah YH (1983) Population dynamics and life history of a solitary coral, *Balanophyllia elegans*, from Central California. Oecologia 58: 200–207

Fadlallah YH, Pearse JS (1982) Sexual reproduction in solitary corals: overlapping oogenic and brooding cycles, and benthic planulas in *Balanophyllia elegans*. Mar Biol 71: 223–231

Gerrodette T (1979a) Ecological studies of two temperate solitary corals. Ph.D. thesis. Scripps Institution of Oceanography, University of Calif, San Diego

Gerrodette T (1979b) Equatorial submergence in a solitary coral, *Balanophyllia elegans*, and the critical life stage excluding the species from shallow water in the south. Mar Ecol Prog Ser 1: 227–235

Gerrodette T (1981) Dispersal of the solitary coral *Balanophyllia elegans* by demersal planular larvae. Ecology 62: 611–619

Hansen TA (1978) Larval dispersal and species longevity in Lower Tertiary gastropods. Science, NY 199: 885–887

Hellberg ME (1994) Relationships between inferred levels of gene flow and geographic distance in a philopatric coral, *Balanophyllia elegans*. Evolution 48: 1829–1854

Jablonski D (1986) Larval ecology and macroevolution in marine invertebrates. Bull mar Sci 39: 565–587

Janson K (1983) Selection and migration in two distinct phenotypes of *Littorina saxatilis* in Sweden. Oecologia 59: 58–61

Janson K, Ward RD (1984) Microgeographical variation in allozyme and shell characters in *Littorina saxatilis* Olivi (Prosobranchia: Littorinidae). Biol J Linn Soc 22: 289–307

Johannesson K, Johannesson B (1989) Differences in allele frequencies of *Aat* between high- and mid-rocky shore populations of *Littorina saxatilis* (Olivi) suggest selection in this enzyme locus. Genet Res 54: 7–11

Johnson MS, Black R (1991) Genetic subdivision of the intertidal snail *Bembicium vittatum* (Gastropoda: Littorinidae) varies with habitat in the Houtman Abrolhus Islands, Western Australia. Heredity, Lond 67: 205–213

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics, Austin, Tex 49: 561–576

Knowlton N, Jackson JBC (1993) Inbreeding and outbreeding in marine invertebrates. In: NW Thornhill (ed) The natural history of inbreeding and outbreeding: theoretical and empirical perspectives. University of Chicago, Chicago, Illinois, pp 200–249

Koehn RK, Newell RIE, Immermann F (1980) Maintenance of an aminopeptidase allele frequency cline by natural selection. Proc natn Acad Sci USA 77: 5385–5389

Larson A, Wake DB, Yanev KP (1984) Measuring gene flow among populations having high levels of genetic fragmentation. Genetics, Austin, Tex 106: 293–308

Leberg PL (1992) Effects of population bottlenecks on genetic diversity as measured by allozyme electrophoresis. Evolution 46: 477–494

Lindberg DR (1991) Marine biotic interchange between northern and southern hemisphere. Paleobiology 17: 308–324

McArdle BH (1988) The structural relationship: regression in biology. Can J Zool 66: 2329–2339

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc natn Acad Sci USA 70: 3321–3323

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, Austin, Tex 89: 583–590

Reid DG (1990) Trans-arctic migration and speciation induced by climatic change: the biogeography of *Littorina* (Mollusca: Gastropoda). Bull mar Sci 47: 35–49

Scheltema RS (1977) Dispersal of marine invertebrate organisms: paleobiogeographic and biostrategraphic implications. In: Kaufman EG, and Hazel JE (eds) Concepts and methods of biostratigraphy. Dowden, Hutchinson & Ross, Stroudsburg, Pennsylvania, pp 73–108

Selander RK, Smith MH, Yang SY, Johnson WE, Gentry JB (1971) Biochemical polymorphism and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*). Stud Genet, Austin, Tex 6: 49–90

Shuto T (1974) Larval ecology of prosobranch gastropods and its bearing on biogeography and paleontology. Lethaia 7: 239–256

Slatkin M (1985) Gene flow in natural populations. A Rev Ecol Syst 16: 393–430

Slatkin M (1991) Inbreeding coefficients and coalescence times. Genet Res 58: 167–175

Slatkin M (1993) Isolation by distance in equilibrium and nonequilibrium populations. Evolution 47: 264–279

Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43: 1349–1368

Slatkin M, Maddison WP (1990) Detecting isolation by distance using phylogenies of genes. Genetics, Austin, Tex 126: 249–260

Swofford DL, Selander RK (1989) BIOSYS-1: a computer program for the analysis of allelic variation in population genetics and biochemical systematics. University of Illinois, Urbana, Illinois

Valentine JW, Jablonski D (1983) Speciation in the shallow sea: general patterns and biogeographic controls. In: Sims RW, Price JH, Whalley PES (eds) Evolution, time and space: the emergence of the biosphere. Academic Press, New York, pp 201–226

Vermeij GJ (1989) Geographical restriction as a guide to the causes of extinction: the case of the cold northern oceans during the Neogene. Paleobiology 15: 335–356

Ward RD, Beardmore JA (1977) Protein variation in the plaice (*Pleuronectes platessa*). Genet Res 30: 45–62

Weir BS (1990) Intraspecific differentiation. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer, Sunderland, Massachusetts, pp 373–410

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38: 1358–1370

Wright S (1943) Isolation by distance. Genetics, Princeton 28: 114–138

Wright S (1951) The genetical structure of populations. Ann Eugen 15: 323–354

Wright S (1978) Evolution and the genetics of populations. Vol 4. Variability within and among natural populations. University of Chicago, Press, Chicago