CONCEPTS & SYNTHESIS
EMPHASIZING NEW IDEAS TO STIMULATE RESEARCH IN ECOLOGY

Ecology, 87(5), 2006, pp. 1094–1103

THE USE OF GENETIC CLINES TO ESTIMATE DISPERSAL DISTANCES OF MARINE LARVAE

Erik E. Sotka[1,2] and Stephen R. Palumbi[1,3]

[1]*Stanford University, Hopkins Marine Station and the Department of Biological Sciences, Oceanview Boulevard, Pacific Grove, California 93950 USA*

Abstract. Many unresolved issues in the ecology and evolution of marine populations center on how far planktonic larvae disperse away from their parents. Genetic tools provide a promising way to define the spatial spread of larvae, yet their accurate interpretation depends on the extent to which genetic loci are under selection. Genetic clines, geographic zones in which genetically differentiated populations interbreed, provide opportunities to explicitly and simultaneously quantify the relative roles of selection and dispersal. Here, we review the theory and analysis of genetic clines and apply these techniques to published studies of multilocus clines in the sea.

The geographic width of a stable genetic cline is determined by a balance between the homogenizing effects of dispersal and the diversifying effects of selection. For marine researchers, the power of genetic clines is that, if selection and clinal width are quantified, then the average geographic distances that larvae move can be inferred. Measuring selection or dispersal through laboratory or field-based experimentation is possible, though logistically difficult, for pelagically dispersed organisms. Instead, dispersal may be more robustly quantified from the degree of linkage disequilibrium between two or more loci, because linkage disequilibrium integrates selection across multiple life stages and generations. It is also relatively insensitive to whether exogenous or endogenous selection operates. Even without quantifying linkage disequilibrium, the theory of genetic clines indicates that the average dispersal distance of larvae is a fraction (i.e., generally $<35\%$) of the clinal width. Because cline theory is based on several underlying assumptions, including near-equilibrium between selection and migration, the dispersal distances inferred from empirical data should be of the correct order but may not be precise. Even so, such estimates of larval dispersal are valuable, as they can be utilized to design appropriate scales for future investigations and provide some guidance to conservation efforts.

Key words: allele frequency; genetic clines; hybridization; larval dispersal; linkage or gametic disequilibrium; marine larvae; planktonic dispersal; selection.

Introduction

Genetic tools provide an independent description of the movement of marine larvae, and have proven useful in testing expectations of dispersal based on an organism's life history (Doherty et al. 1995, Bohonak 1999), oceanographic currents (Shulman and Bermingham 1995, Bernardi et al. 2003, Gilg and Hilbish 2003) and biogeography (see reviews in Cunningham and Collins 1998, Avise 2001, Grosberg and Cunningham 2001, Wares 2002). For instance, high-dispersal species often reveal little or no genetic differentiation (Palumbi 1994, Bohonak 1999, Kinlan and Gaines 2003), though there are published examples of surprisingly strong genetic structure among species that spend their entire lives in the plankton (de Vargas et al. 1999, Bucklin et al. 2000, Darling et al. 2000, Rynearson and Armbrust 2000, Bucklin et al. 2002) or have long planktonic phases (Barber et al. 2000, Taylor and Hellberg 2003). Strong genetic structure implies that larvae are not moved like passive particles in ocean currents (Armsworth et al. 2001), that larvae may not survive the transit of deep

Manuscript received 6 December 2004; revised 11 July 2005; accepted 26 July 2005; final version received 5 October 2005. Corresponding Editor: C. W. Fox.

[2] Present address: Grice Marine Laboratory and the Department of Biology, College of Charleston, 205 Fort Johnson Road, Charleston, South Carolina 29412 USA.

[3] E-mail: spalumbi@stanford.edu

water (Cowen et al. 2000), or that the mixing of water in distinct ocean gyres does not significantly mix populations of plankton (de Vargas et al. 1999, Bucklin et al. 2002, Sotka et al. 2004), among other possibilities.

The diversity of recent findings suggests that describing marine dispersal patterns will require a great deal of detailed attention to many individual species. Overall, we do not know dispersal distance for most marine species of commercial or conservation interest to within 1–3 orders of magnitude (Palumbi 2004). This lack of information has prompted increasing calls for development of new technical approaches to understanding marine dispersal (Swearer et al. 2002, Palumbi 2003).

In addition, the typical interpretation of genetic structure potentially suffers from a serious bias. With the exception of studies using allozyme loci (Koehn et al. 1976, 1980, Place and Powers 1979, McDonald 1987, Karl and Avise 1992, Johannesson et al. 1995, Schmidt and Rand 1999), genetic estimates of dispersal generally assume that the molecular markers are selectively neutral, and that any observed structure is largely due to reduced dispersal. F_{st}-based approaches typically assume equilibrium between dispersal and drift when in fact, divergence may be due to selection or the signature of historical events (Whitlock and McCauley 1999). For instance, when two populations differ in allele frequencies, researchers typically infer that dispersal is weak between them. If the loci are under selection or are linked to selected loci, then dispersal between the two populations may actually be strong and local genetic differences may be maintained by selection. Thus, inferring marine dispersal from molecular markers depends on the strength of the assumption about selective neutrality. In reality, genetic differentiation, dispersal, and selection are all intertwined; to infer any one of these processes requires knowledge of the other two.

Genetic clines, that is, geographic zones in which genetically differentiated populations interbreed, occur throughout the oceans (Gardner 1997, Avise 2001). Because the width and shape of genetic clines commonly represent an evolutionary balance between selection and dispersal (Slatkin 1973, Barton and Hewitt 1985, Barton and Gale 1993, Mallet 2001), clines provide researchers with the ability to powerfully analyze both evolutionary forces simultaneously.

The theory of genetic changes along clines in one and two dimensions has been carefully investigated empirically in several terrestrial and aquatic systems (for review, see Endler 1977, Barton and Hewitt 1985, Barton and Gale 1993, Harrison 1993, Arnold 1997, Butlin 1998). In this review, we briefly outline the theoretical frameworks that underlie modern analyses of genetic clines, and apply these methods to well-described genetic clines of marine organisms. Our goal is to ask what guidance current theory could potentially provide to estimates of larval dispersal in marine settings.

Using selection and linkage disequilibrium to estimate dispersal within a cline

On balance, clinal theory provides several valuable "yardsticks" for empiricists to compare with data from natural populations. The most fundamental of these predictions states that equilibrial width of a cline represents a balance between the diversifying effects of selection and the homogenizing effects of dispersal. When selection (s) represents the difference in fitness between genotypes at the center of the cline, and σ is the standard deviation of the distance from parents to offspring along a linear gradient (which is broadly proportional to the width of the dispersal cloud around parents and is linearly related to the distance that an offspring moves on average), then the width of the cline at equilibrium is proportional to $\sigma\ s^{-1/2}$ (Kimura and Weiss 1964, Slatkin 1973, Barton and Hewitt 1985, Mallet and Barton 1986, Barton and Gale 1993, Moore and Price 1993). Thus, narrow clines in highly dispersive organisms will be maintained only when there are high levels of selection, while narrow clines in poorly dispersed organisms can be maintained by weaker selection (Wright 1948).

Selection operates either because hybrids of the parental lines are generally less fit, or alternatively, parents or hybrids may be less fit in nonnative environments. Those different modes of selection have been called endogenous and exogenous selection, respectively, and they lead to similar consequences when clines occur between differentially adapted species or populations. Most commonly, clines are the result of some mix of both modes of selection (*Bombina* toads [Szymura and Barton 1986], *Heliconius* butterflies [Mallet et al. 1990], *Mercenaria* clams [Bert and Arnold 1995], *Mytilus* mussels [Bierne et al. 2002]). There is a scenario in which hybrid zones are maintained by selection favoring hybrids within a narrow zone of intermediate habitat (termed bounded hybrid superiority; [Moore 1977]). This scenario is thought to be extremely rare (Barton and Gale 1993), but when present, does not apply to the theoretical expectations.

The distinction between endogenous and exogenous selection is crucial for understanding the potential mobility of the hybrid zone. For example, if endogenous selection operates (selection against hybrids), then the transition zone tends to shift toward any region that had low population densities (Moore and Price 1993; see Barton [1986] for a discussion of the effects of population density on genetic clines). Alternatively, hybrid zones maintained by exogenous selection tend to remain stationary at a particular place on an ecological gradient, or shift in geographic position when the environment changes (e.g., Blum 2002, Dasmahapatra et al. 2002).

If the magnitude of selection were known, this could be used along with the width of a cline to estimate

dispersal. This is because, for a variety of types of selection, the balance between selection and dispersal is

$$w^2 = \frac{K\sigma^2}{s_e} \quad (1)$$

where w is the cline width, s_e is the "effective" selection coefficient, and K is a multiplier that depends on the type of selection. Strict application of this equation requires assumptions be met that might be rare in natural settings, including Gaussian dispersal, weak selection, and genetic equilibrium. The "effective" selection coefficient acting on the clinal locus includes direct selection on the locus that displays clinal variation in addition to the cumulative levels of indirect selection on linked loci (see Barton and Bengtsson [1986] for details). The multiplier K varies from ~3 in the case of exogenous selection across an ecotone to 4 in the case of heterozygote disadvantage at the center of a cline (Barton and Gale 1993, Moore and Price 1993). Frequency-dependent selection against rare genotypes can increase K to 8–12, because frequency-dependent selection is effectively weaker than heterozygote disadvantage (Mallet and Barton 1989). Because cline width is proportional to the square root of K, different types of selection give cline widths that are similar to within about a factor of 2. In general, however, there are substantial deviations from the theoretical expectation when any type of selection is strong ($s > 0.1$; Mallet and Barton 1989). In many cases, terrestrial and aquatic researchers have used this formulation to assess levels of selection acting on loci largely because empirical estimates of dispersal can be generated from direct censuses (Barton and Hewitt 1985, Harrison 1993, Arnold 1997). However, because these dispersal values are often severely underestimated (Barton and Hewitt 1985), selection estimates are probably far lower than the actual levels.

Estimation of the clinal width is relatively straightforward. Traditionally, cline width has been defined as the geographic distance between populations that contain 10% or 20% and 80% or 90% of the parental gene frequencies (e.g., Endler 1977, Bert and Arnold 1995), but in the theory of Eq. 1, cline width is the inverse of the maximum slope of the cline (Barton and Gale 1993). An estimate can be generated by eye but likelihood-based estimates are preferable (e.g., N. Barton and S. Baird' s software Analyse; *available online*).[4] An estimate of width may assume that allele frequencies vary between 0.0 and 1.0 along the cline, or alternatively, if the populations are not fixed on either side, cline width $w = \Delta p/\text{slope}$, where Δp is the change in gene frequencies among parental populations at the ends of the cline and slope is the slope at the center of the cline.

Determining a precise estimate of the selection coefficient (s_e) can be complicated (Endler 1986). Field-based experiments can sometimes detect rather strong selection coefficients ($s > 0.10$; Hoekstra et al. 2001), but weaker levels of selection are more difficult to measure. Koehn et al. (1980) described strong selection ($s = 0.2$–0.5) on *Lap* allotypes in the mussel *Mytilus edulis*, and a strong cline in *Lap* gene frequencies over 32 km of coastline. Using Eq. 1 would suggest average larval dispersal (σ) of 4.8–14.4 km if selection were measured at the center of the cline, but these figures depend strongly on the accuracy of measures of selection over space and time. Laboratory-based experiments can be more sensitive, but in many cases, their results may not be generalizable to more natural conditions. These complications in the experimental determination of selection reduce the applicability of Eq. 1 when used by itself, and have led many population geneticists to turn to other analyses.

Alternatively, the degree of linkage disequilibrium (LD), the nonrandom gametic association of alleles between two or more loci, provides researchers with another useful tool. Positive values of LD along a cline generally reflect an excess of parental gametic haplotypes and a reduction of hybrid gametic haplotypes within the cline. Linkage disequilibrium is generated when either endogenous or exogenous selection acts within the clines. The net effects are that hybrids are less readily generated or maintained, parental alleles do not readily recombine, and a greater than expected number of parental gametes or haplotypes are encountered within a hybrid zone (Fig. 1). Consequently, the higher the rate of migration across the clines of a given width, the larger the number of parental genotypes found within the clines, and the higher the degree of linkage disequilibrium (i.e., selection). Because linkage disequilibrium is generated by selection after migration in each generation, it is largely equivalent to "effective" selection when $s_e < 0.10$. When greater levels of selection maintain the clines, LD is not strictly equivalent to s_e, but rather approximates selection within an order of magnitude (see Barton and Gale [1993] for discussion). It is not strictly generated by epistatic interactions between loci.

For clines at equilibrium, the balance between selection and dispersal can be represented at the center of a cline by

$$w^2 \approx \frac{\sigma^2}{D_{AB}r} \quad (2)$$

where r is the rate of recombination and D is the maximum level of linkage disequilibrium between loci A and B (Barton 1982, 1986). The basic formulation of D is the two-locus deviation from random expectation, or

$$D_{AB} = p_{AB} - p_A p_B \quad (3)$$

where p_A is the frequency of allele A at locus 1, p_B is the frequency of allele B at locus 2, and p_{AB} is the frequency of AB. If one assumes the loci are unlinked, then the rate of recombination is $r = 0.5$. Because changes in allele frequencies (p) affect the maximum potential genetic

[4] ⟨http://helios.bto.ed.ac.uk/evolgen/Mac/Analyse/⟩

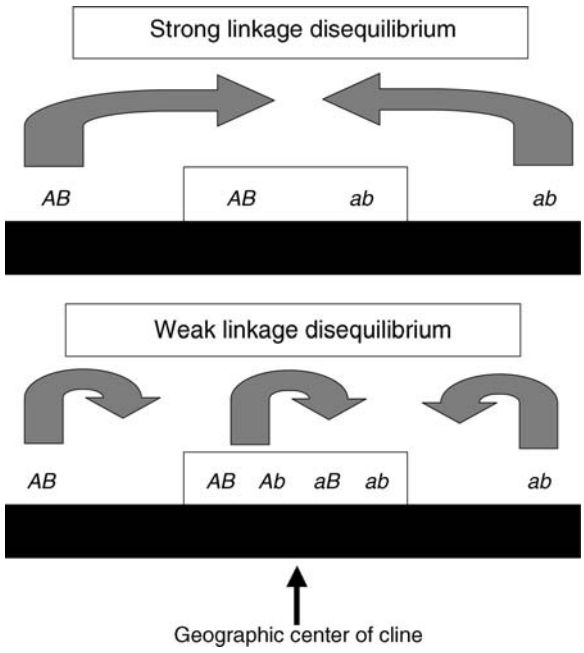

FIG. 1. A graphical depiction of the gametic (=linkage) disequilibrium among two loci (*Aa* and *Bb*) at the center of a cline. The arrows indicate strength and direction of larval dispersal, and letters indicate gametic haplotypes. For a given clinal width (shown as an outlined box), a signal of strong gametic disequilibrium (i.e., high frequency of parental gametes *AB* and *ab*) in the center of the cline is generated when parental genotypes readily disperse from the edges and hybridization is relatively infrequent because of selection. A signal of weak gametic disequilibrium (i.e., hybrid gametes *Ab* and *aB*) occurs when dispersal is weak and hybridization is common.

disequilibrium (Hedrick 1987), Eq. 2 can be replaced with one that uses the correlation coefficient between loci, R_{AB}:

$$R_{AB} = \frac{D_{AB}}{\sqrt{p_A p_B (1 - p_A)(1 - p_B)}} \quad (4)$$

so that at the center of a perfect cline when $p_A = p_B = 0.5$, then (see Barton 1982; see also Mallet 2001)

$$w^2 \approx \frac{4\sigma^2}{R_{AB} r} \quad (5)$$

The theory is simplest for populations at the center of the cline from which these formulas were derived. In practice, researchers should generate data on linkage disequilibria for many populations across the cline. In part, this will ensure that the cline is conforming to theoretical expectations of the geographic distribution of LD (i.e., it is unimodal, symmetric about the center, and tails to zero on the edges because there is no opportunity for LD within parental populations). Deviations from these expectations provide valuable insight (e.g., Barton 1986). Further, the geographic distribution of LD values can be used by the software Analyse to simultaneously assess the balance between LD, dispersal, and cline width at several points within the cline (Barton 2000).

In cases where the cline width can be measured, there are several reasons why estimating LD is an attractive method for estimating larval dispersal. First, the LD approach contains a great deal of statistical power in the detection of weak levels of selection ($s < 0.10$), levels that are logistically difficult to detect through selection experiments. Because the LD approach also infers dispersal distances when strong selection operates, we may infer dispersal across a broad range of disequilibrium values. Second, the LD approach essentially integrates the recent history of selection and gene flow that occurs at any and all life-history stages of an organism. In contrast, selection experiments are commonly designed to detect selection on a limited subset of possible life stages (Hoekstra et al. 2001). Third, the LD approach detects any type of selection that may operate to maintain the cline (Barton and Gale 1993, Kruuk et al. 1999).

Finally, the genetic information needed to detect levels of LD is not formidably large. A survey of the recent animal literature on linkage disequilibrium suggests that, on average, four polymorphic loci (range 2–9 loci) genotyped for ~50 animals per location have been used to detect an LD of ~0.16 (range of significant results 0.02–0.35: see Szymura and Barton 1986, Mallet et al. 1990, Bert and Arnold 1995, Duggins et al. 1995, Hare and Avise 1996, Planes and Doherty 1997, Lenormand et al. 1998, Rawson et al. 1999, Bierne et al. 2002, 2003, Dasmahapatra et al. 2002, Bronson et al. 2003, Morgan-Richards and Wallis 2003, Nielsen et al. 2003, Rawson et al. 2003, Vines et al. 2003). In general, it is more difficult to quantify gametic disequilibrium among systems that have low levels of population differentiation. Because studies tend to reveal significant linkage disequilibrium when $D_{AB} \approx 0.02$ or greater, this appears to be a reasonable lower limit of detectability. For example, Bierne et al. (2002) used length polymorphisms from diploid adults at five loci and $n \approx 50$–90 individuals to show a significant $D_{AB} = 0.03$ within a hybrid *Mytilus* mussel zone.

However, difficulties in measuring multilocus LD remain serious. LD is a gametic disequilibrium and, as such, can be difficult to estimate in natural populations without large variances (Hedrick 1987). The main problem is that one cannot discern whether the gametes that formed a double heterozygote were coupled (i.e., *AA*/*BB*) or repulsed (i.e., *AB*/*AB*), though this is not a concern for cytonuclear disequilibrium (e.g., LD between a mitochondrial and nuclear locus). In addition, there is variance associated with determining "true" allele frequencies because of sampling error, and LD is difficult to estimate when there is a heterozygote deficit (Barton 2000).

Analytical methods to quantify linkage disequilibrium are under constant development, but at least three methods are available currently for clinal analysis. First,

maximum likelihood estimates of LD as described by Barton (2000) can be generated using the program Analyse. This program can utilize genotype data from a single population at the center of the cline or, preferably, across several populations within the hybrid zone. Second, a series of pairwise linkage disequilibrium comparisons can be calculated using the basic formulation (Eq. 3). Disequilibrium among nuclear and mitochondrial loci (Asmussen et al. 1987) also can be readily estimated. Third, the variance of the hybrid index, a measure of the number of introgressed alleles that occur within a single individual, can be used to estimate linkage disequilibrium (Szymura and Barton 1986, Barton and Gale 1993). For four nuclear loci, the maximum hybrid index is 8, where 0 and 8 are "parental" lines, and 1–7 represent genotypes containing some hybrid loci. This approach can be corrected for shared polymorphisms among populations at the endpoints of the cline. Zapata et al. (2001) also developed a sign-based estimator for D_{AB} and $D_{AB\max}$ effective across multiple loci and alleles.

Once estimates of LD are available, they are given confidence intervals that are then translated to confidence intervals of dispersal distances. Such variance can be generated using Monte Carlo or Markov chain resampling procedures as implemented by a number of population genetic programs (Asmussen et al. 1987, Asmussen and Basten 1994, Basten and Asmussen 1997), including Analyse and C. J. Basten's software CND (*available online*).[5] One limitation to some of these software packages is that it remains unclear how to combine confidence intervals for a series of pair-by-pair estimates of LD (N. Barton, *personal communication*).

The clinal theory outlined here is explained in much greater detail by its architects (for detailed reviews, see especially Barton and Hewitt [1985], Barton and Gale [1993], Harrison [1993], Arnold [1997], and Mallet [2001]), and includes a host of factors that complicate its applicability to particular data sets. For example, the equations largely assume an evolutionarily static balance between selection and dispersal. If the cline is new because a new population is invading a region after human introduction, the relationships between selection, width, and dispersal will be very different. For most clines, however, stabilization of clines occurs very rapidly. It has been estimated, for example, that if $s = 0.1$, clines stabilize within $\sim$10 generations (Barton and Gale 1993).

A second complication is that clinal theory is directly relevant for loci under direct selection, such as some allozymes (e.g., Koehn et al. 1980, Powers 1988) or morphological traits (Mallet et al. 1990). However, many hybrid zones are detected using molecular markers (e.g., microsatellites) that are less likely to be under direct selection. If these neutral genes are not physically linked with loci under direct selection, then recombination quickly breaks up linkage disequilibria between loci, and the introgression of the neutral alleles across the cline will occur unimpeded (Barton 1979, Takahata and Slatkin 1984). Thus, genetic differences among populations may be extremely strong immediately after secondary contact (i.e., after historically separated populations reconnect) and begin to weaken after hundreds or thousands of generations of gene flow (see Sotka et al. [2004] for one simulation of clinal collapse). As a consequence, such neutral clines often look like a "staircase" of several steps of allele frequency across the hybrid zone. Neutral markers are rarely fixed on both sides of the hybrid zone (Barton 1979, Barton and Bengtsson 1986, Barton and Gale 1993), and the introgression will eventually homogenize allele frequencies. The net effect of the flattening of the cline is a rather weak, but artificial, increase in estimated clinal width. Any given empirical estimate of selection in a snapshot of time will infer a rate of dispersal that is somewhat lower than the actual dispersal. In fact, a neutral cline will flatten at a rate proportional to a predictable product of dispersal and time. The width of a cline of neutral genes t generations after two differentiated populations come together is expected to be $\sim 2.51\sigma(t)^{-1/2}$ (Endler 1977, Barton and Gale 1993), assuming equal population sizes. On the other hand, clines at neutral genes can be stable if these genes are physically linked to genes under selection (Barton 1979). Linkage of neutral markers to many genes under selection results in an effective selection (s_e) that helps sculpt neutral gene clines in a manner analogous to the action of non-neutral clines (Barton 1986).

Third, these relationships are largely based on an underlying diffusion approximation, and may be violated by rare long-distance dispersal. In theory, examination of higher order recombinants can allow estimation of long-distance dispersal (Barton 2000), though this has not been attempted for any marine system to date.

Fourth, the theory considers hybrid zones between two parental populations and bi-allelic loci. When more than two alleles are involved, researchers often pool alleles according to their population sources. Such multiallelic ($\geq$3 alleles) measures of linkage disequilibrium tend to underestimate LD.

Fifth, the rate of recombination will be <0.5 if loci are physically linked, but the rates can be difficult to ascertain. In general, however, the chances of physical linkage among small numbers of loci (e.g., $<$10) are probably low.

Sixth, shifts in population densities, physical barriers to dispersal, and asymmetric gene flow can slightly alter the relationships between cline width, dispersal, and selection values (Barton 1979, 1986, Barton and Bengtstron 1986).

The consequence of these complications is that estimates of dispersal based on selection or linkage disequilibrium will be robust and of the right order, but

[5] ⟨http://statgen.ncsu.edu/sisg/software_BRC.php⟩

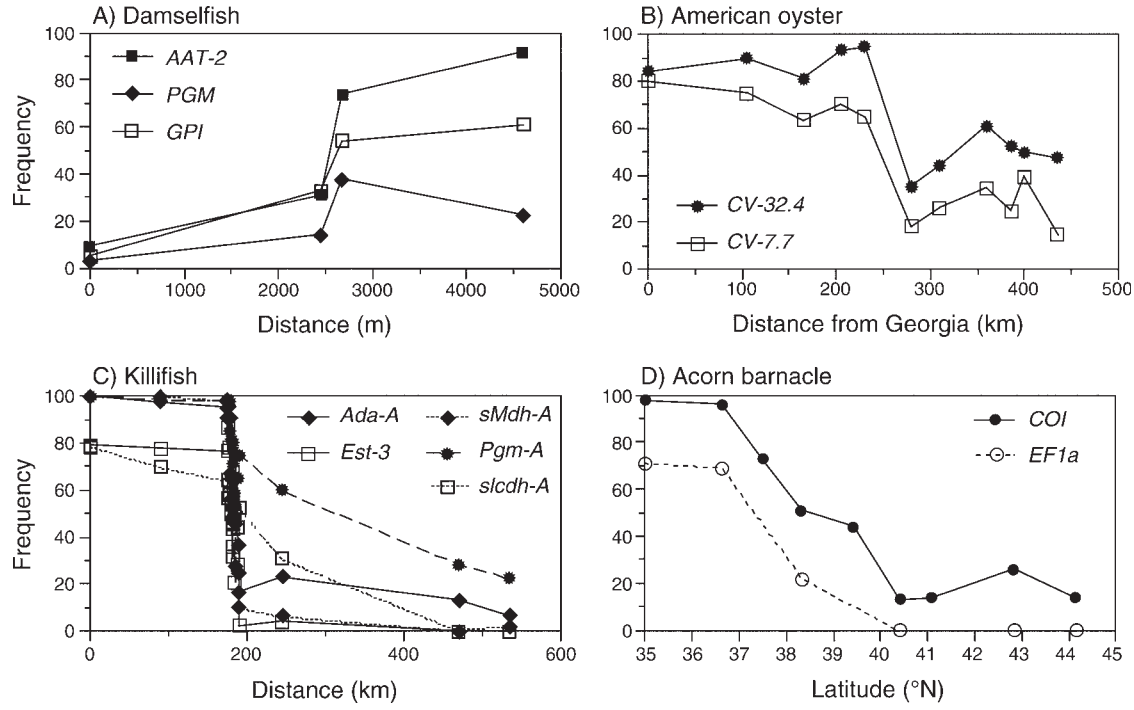

FIG. 2. Multilocus genetic clines (A) in the damselfish *Acanthochromis polyacanthus* on the Great Barrier Reef (Planes and Doherty 1997); (B) in the American oyster *Crassostrea virginica* between the Atlantic Ocean and Gulf of Mexico (Hare and Avise 1996); (C) between *Fundulus majalis* and *F. similes* (Duggins et al. 1995); and (D) in the acorn barnacle *Balanus glandula* along the west coast of North America (Sotka et al. 2004).

may not be precise. Further, the dispersal estimates reflect the distances traveled within the clinal region only. Geographic variation in gene flow and dispersal is probably the rule rather than the exception among pelagically dispersed marine animals (Sotka et al. 2004). Estimates of dispersal from one location should be extended throughout the species' geographic range only with extreme caution. In response to these complications, researchers should attempt to place bounds of confidence on the estimates of LD and dispersal using detailed genetic analysis, and subsequently link these estimates with laboratory or field-based experiments on the nature and strength of selection (s_e). Despite these caveats, these data might add substantially to our understanding of marine dispersal patterns because so few other sources of data are available.

Clines in marine ecosystems

The marine literature is filled with genetic descriptions of hybrid zones between species (Gardner 1997) and differentiated populations (Avise 2001, Grosberg and Cunningham 2001). However, we know of only one marine study that utilized cline theory to explicitly estimate dispersal distances. Planes and Doherty (1997) describe a steep cline between two color morphs in a single species of tropical damselfish, *Acanthochromis polyacanthus,* on the Great Barrier Reef (GBR) (Fig. 2A). Across huge swaths of habitat in the southern GBR, damselfishes are uniformly black and fixed at three allozyme loci (*AAT-2*, *GPI-1*, and *PGM*). Across the northern GBR, damselfishes are bicolored and nearly fixed for alternative alleles at the same loci. There is a relatively thin zone of contact between these two extremes, characterized by fishes of intermediate morphs and allozyme frequencies. Genotypes of fishes at the center of this cline display strong linkage disequilibrium averaged across pairwise comparison of the three loci (average $R_{AB} = 0.168$; $D \approx 0.04$), indicating that gametes contain an excess of parental gametes (and conversely a deficit of recombinant gametes) relative to conditions under linkage equilibrium. Given that the cline width is ~1.3 km and assuming the loci are unlinked ($r = 0.5$), the authors use the formula (Eq. 4) to infer that average dispersal is 0.189 km per generation, or 15% of the cline width. This dispersal distance appears appropriate given the uniquely philopatric life-history of this fish. Females lay demersal eggs, and the larvae develop directly to the juvenile stage before hatching. As a result, offspring do not move far away before settlement.

The clinal framework potentially can provide insight into the spread of marine larvae within other species. For example, the American oyster *Crassostrea virginica* displays striking differences between Atlantic and Gulf of Mexico populations that shift within a ~340-km stretch of the eastern Florida coastline at Cape

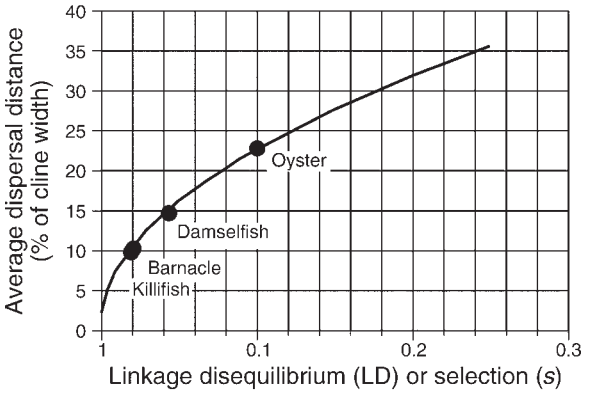

FIG. 3. The correlation between dispersal distance within a stable cline and linkage disequilibrium (LD) or selection (s). The curve is based on the two formulas Eqs. 1 and 2. Dispersal is given as the proportion of the clinal width (e.g., 20% of a 100-km clinal width is 20 km). Under reasonable levels of selection, the dispersal distance is a fraction of the clinal width.

Canaveral (Fig. 2B). The geographic differences probably arose in allopatry when the two ocean basins were disengaged during glacial maxima, but genetic differences have maintained themselves on a fine spatial scale in this oyster as well as many other estuarine and freshwater taxa (Avise 2001). In *C. virginica,* Hare and Avise (1996) quantified linkage disequilibrium among two nuclear and one mitochondrial RFLPs and found a statistically significant amount of linkage disequilibrium, several kilometers south of the center of the strongest portion of the genetic cline. At other geographic locations, LD values were nonsignificant and spatially variable. The authors conclude from the relatively weak and highly variable LD that if selection against hybrids occurs in this cline, it is probably minimal. This suggests that low dispersal has helped maintain the cline since the reconnection of historically allopatric populations (i.e., secondary contact; Hare and Avise 1996). We can use the LD framework to estimate exactly how low this dispersal may be. Assuming the cline is 340 km (as delineated by the predecessor of the program Analyse; Hare and Avise 1996), then weak levels of linkage disequilibrium at $D_{AB} \approx 0.05$ and 0.10 (or $R_{AB} \approx 0.2$ and 0.4) at the center of the cline would yield average dispersal distances of 54 km and 76 km per generation, or 16–22% of the clinal width.

Another example comes from a very narrow hybrid zone between two fish species, *Fundulus majalis* and *F. similes*, in northeastern Florida (Duggins et al. 1995). Five allozymes clearly distinguish the two species and the region of overlap (Fig. 2C). The authors calculate a series of pairwise gametic disequilibrium (Δp_{AB}), but conclude that these values are not significantly different from zero. Interestingly, however, closer examination of their distributions (in Duggins et al. 1995:Fig. 3) shows that at the center of the cline, there are higher values of pairwise gametic disequilibrium than on the edges, as expected from cline theory. The cline averages 27 km in width across loci (excluding the *sIcdh-A* locus, which is less diagnostic than the others). If disequilibrium is about $D_{AB} \approx 0.02$ ($R_{AB} = 0.08$), then the average dispersal of *Fundulus* spp. in the center of the cline is ~3 km or 10% of the cline width.

Finally, a marine cline was recently discovered among populations of the acorn barnacle *Balanus glandula* (Sotka et al. 2004), one of the most intensively studied and abundant members of upper-intertidal communities of the northeast Pacific Ocean (Fig. 2D). Among sequences of the mitochondrial locus cytochrome oxidase I (*CO I*) and the nuclear locus elongation factor 1-alpha (*EF1a*), populations from southern California (latitudes 35°–37° N) and northern populations from Cape Mendocino, California to Vancouver Island, Canada (latitudes 40.5°–44° N) are dominated by distinct haplotype groups. The shift between northern and southern genetic types occurs along a 450-km section of the central California coast. The evolutionary forces that maintain the barnacle clines are uncertain, but indirect support for a role for environmentally based selection comes from the concordance between the barnacle hybrid zone and several transitions in biotic and abiotic conditions across central California (Sotka et al. 2004). The inverse of the maximal slopes of the clines between northern and southern populations is ~700 km. Our current estimate of cytonuclear linkage disequilibrium in the center of the cline is $D_{AB} \approx 0.02$ (or $R_{AB} \approx 0.08$; $n = 48$ individuals). This estimate was generated from software program CND (E. E. Sotka and S. R. Palumbi, *unpublished data*). If further sampling confirms this, then the average dispersal distances of larvae near the center of the cline is estimated to be on the order of ~70 km per generation, or 10% of the clinal width.

Average dispersal distance is a fraction of the clinal width

This review emphasizes that larval dispersal and local selection are intertwined and require simultaneous examination. Linkage disequilibrium is one powerful way to infer the action of selection along a cline, and information on linkage and cline geography can produce useful insights into dispersal. However, in the absence of detailed information on linkage or selection, it is possible to produce a first approximation of larval dispersal distance using clinal theory. Even under high levels of selection (e.g., $s \approx 0.25$) and at equilibrium, the average geographic distance dispersed by marine larvae, as measured by the variance in distance between parent and offspring or neighborhood size (sensu Wright 1948), is less than about one-third of the cline width ($\sigma < 0.35w$; Fig. 3). Empirical studies support these levels of dispersal. Estimates of LD from four marine studies infer an average dispersal distance of ~20% or less of the clinal width.

Although this is a crude approximation, it suggests that, in general, populations on the endpoints of marine

clines do not typically disperse larvae across the entire cline width in one generation. Instead, typical propagules may require 3–5 generations to traverse the cline. Only if selection were very large (e.g., hybrids were infertile) relative to rates of recombination (see Barton and Gale 1993) would a marine cline be maintained by a dispersal distance that was as long as the cline was wide. In cases where selection was measurable but ecologically moderate ($s \approx 0.1$), then $\sigma \approx 0.11w$. In other words, for selection that ranges from moderate to strong, clines are generally several times wider than average dispersal distance. For clines subject to weak selection, the subsequent clines can be an order of magnitude wider than dispersal, or more.

There are likely to be conspicuous instances in which clines are far from equilibrium because of recent invasions or range shifts. Apart from these cases, clinal stabilization is generally so rapid that the spatial extent of many natural genetic clines will provide an invaluable initial range of larval dispersal distances that could be used to design appropriate scales of genetic or experimental analyses.

Acknowledgments

Funding was supported by the Andrew W. Mellon Foundation and the National Science Foundation. We thank J. Avise, N. Barton, T. Bert, J. Coyne, C. Duggins, C. Fox, H. Galindo, M. Hare, D. Hartl, R. Nielsen, M. Slatkin, P. Rawson, S. Vollmer, and J. Wakeley for reading previous versions of the manuscript. We also indebted to N. Bierne and J. Mallet for their detailed and thoughtful reviews.

Literature Cited

Armsworth, P. R., M. K. James, and L. Bode. 2001. When to press on or turn back: dispersal strategies for reef fish larvae. American Naturalist **157**:434–450.

Arnold, M. 1997. Natural hybridization and evolution. Oxford University Press, New York, New York, USA.

Asmussen, M. A., J. Arnold, and J. C. Avise. 1987. Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. Genetics **115**: 755–768.

Asmussen, M. A., and C. J. Basten. 1994. Sampling theory for cytonuclear disequilibrium. Genetics **138**:1351–1363.

Avise, J. C. 2001. Phylogeography: the history and formation of species. Harvard University Press, Cambridge, Massachusetts, USA.

Barber, P. H., S. R. Palumbi, M. V. Erdmann, and M. K. Moosa. 2000. Biogeography—a marine Wallace's line? Nature **406**:692–693.

Barton, N. H. 1979. Gene flow past a cline. Heredity **43**:333–339.

Barton, N. H. 1982. The structure of the hybrid zone in *Uroderma bilobatum* (Chiroptera: Phyllostomatidae). Evolution **36**:863–866.

Barton, N. H. 1986. The effects of linkage and density-dependent regulation on gene flow. Heredity **57**:415–426.

Barton, N. H. 2000. Estimating multilocus linkage disequilibria. Heredity **84**:373–389.

Barton, N. H., and B. O. Bengtsson. 1986. The barrier to genetic exchange between hybridising populations. Heredity **56**:357–376.

Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pages 13–45 *in* R. Harrison, editor. Hybrid zones and the evolutionary process. Oxford University Press, New York, New York, USA.

Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. Annual Review of Ecology and Systematics **16**:113–148.

Basten, C. J., and M. A. Asmussen. 1997. The exact test for cytonuclear disequilibria. Genetics **146**:1165–1171.

Bernardi, G., L. Findley, and A. Rocha-Olivares. 2003. Vicariance and dispersal across Baja California in disjunct marine fish populations. Evolution **57**:1599–1609.

Bert, T. M., and W. S. Arnold. 1995. An empirical-test of predictions of 2 competing models for the maintenance and fate of hybrid zones—both models are supported in a hard-clam hybrid zone. Evolution **49**:276–289.

Bierne, N., P. Borsa, C. Daguin, D. Jollivet, F. Viard, F. Bonhomme, and P. David. 2003. Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *M. galloprovincialis*. Molecular Ecology **12**:447–461.

Bierne, N., P. David, A. Langlade, and F. Bonhomme. 2002. Can habitat specialization maintain a mosaic hybrid zone in marine bivalves? Marine Ecology Progress Series **245**:157–170.

Blum, M. J. 2002. Rapid movement of a *Heliconius* hybrid zone: evidence for phase III of Wright's shifting balance theory? Evolution **56**:1992–1998.

Bohonak, A. J. 1999. Dispersal, gene flow, and population structure. Quarterly Review of Biology **74**:21–45.

Bronson, C. L., T. C. Grubb, and M. J. Braun. 2003. A test of the endogenous and exogenous selection hypotheses for the maintenance of a narrow avian hybrid zone. Evolution **57**: 630–637.

Bucklin, A., O. S. Astthorsson, A. Gislason, L. D. Allen, S. B. Smolenack, and P. H. Wiebe. 2000. Population genetic variation of *Calanus finmarchicus* in Icelandic waters: preliminary evidence of genetic differences between Atlantic and Arctic populations. ICES [International Council for the Exploration of the Sea] Journal of Marine Science **57**:1592–1604.

Bucklin, A., P. H. Wiebe, S. B. Smolenack, N. J. Copley, and M. E. Clarke. 2002. Integrated biochemical, molecular genetic, and bioacoustical analysis of mesoscale variability of the euphausiid *Nematoscelis difficilis* in the California Current. Deep-Sea Research Part I. Oceanographic Research Papers **49**:437–462.

Butlin, R. K. 1998. What do hybrid zones in general, and the *Chorthippus parallelus* zone in particular, tell us about speciation? Pages 367–378 *in* D. J. Howard and S. H. Berlocher, editors. Endless forms: species and speciation. Oxford University Press, New York, New York, USA.

Cowen, R. K., K. M. M. Lwiza, S. Sponaugle, C. B. Paris, and D. B. Olson. 2000. Connectivity of marine populations: open or closed? Science **287**:857–859.

Cunningham, C. W., and T. Collins. 1998. Beyond area relationships: Extinction and recolonization in marine molecular biogeography. Pages 297–322 *in* R. DeSalle and B. Schierwater, editors. Molecular approaches to ecology and evolution. Birkhauser Verlag, Basel, Switzerland.

Darling, K. F., C. M. Wade, I. A. Stewart, D. Kroon, R. Dingle, and A. J. L. Brown. 2000. Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers. Nature **405**:43.

Dasmahapatra, K. K., M. J. Blum, A. Aiello, S. Hackwell, N. Davies, E. P. Bermingham, and T. Mallett. 2002. Inferences from a rapidly moving hybrid zone. Evolution **56**:741–753.

de Vargas, C., R. Norris, L. Zaninetti, S. W. Gibb, and J. Pawlowski. 1999. Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. Proceedings of the National Academy of Sciences (USA) **96**:2864–2868.

Doherty, P. J., S. Planes, and P. Mather. 1995. Gene flow and larval duration in 7 species of fish from the Great-Barrier-Reef. Ecology **76**:2373–2391.

Duggins, C., Jr., A. Karlin, T. A. Mousseau, and G. Relyea. 1995. Analysis of a hybrid zone in *Fundulus majalis* in a northeastern Florida ecotone. Heredity **74**:117–128.

Endler, J. 1977. Geographic variation, speciation, and clines. Princeton University Press, Princeton, New Jersey, USA.

Endler, J. 1986. Natural selection in the wild. Princeton University Press, Princeton, New Jersey, USA.

Gardner, J. P. A. 1997. Hybridization in the sea. Advances in Marine Biology **31**:1–78.

Gilg, M. R., and T. J. Hilbish. 2003. Patterns of larval dispersal and their effect on the maintenance of a blue mussel hybrid zone in southwestern England. Evolution **57**: 1061–1077.

Grosberg, R. K., and C. W. Cunningham. 2001. Genetic structure in the sea: from populations to communities. Pages 61–84 *in* M. D. Bertness, S. Gaines, and M. E. Hay, editors. Marine community ecology. Sinauer Associates, Sunderland, Massachusetts, USA.

Hare, M. P., and J. C. Avise. 1996. Molecular genetic analysis of a stepped multilocus cline in the American oyster (*Crassostrea virginica*). Evolution **50**:2305–2315.

Harrison, R., editor. 1993. Hybrid zones and the evolutionary process. Oxford University Press, New York, New York, USA.

Hedrick, P. 1987. Gametic disequilibrium measures: proceed with caution. Genetics **117**:331–341.

Hoekstra, H. E., J. M. Hoekstra, D. Berrigan, S. N. Vignieri, S. N. Hoang, C. E. Hill, P. Beerli, and J. G. Kingsolver. 2001. Strength and tempo of directional selection in the wild. Proceedings of the National Academy of Sciences (USA) **98**: 9157–9160.

Johannesson, K., E. Rolan-Alvarez, and A. Ekendahl. 1995. Incipient reproductive isolation between two sympatric morphs of the intertidal snail *Littorina saxatilis*. Evolution **49**:1180–1190.

Karl, S. A., and J. C. Avise. 1992. Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. Science **256**:100–102.

Kimura, M, and G. H. Weiss. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics **49**:561–576.

Kinlan, B. P., and S. D. Gaines. 2003. Propagule dispersal in marine and terrestrial environments: a community perspective. Ecology **84**:2007–2020.

Koehn, R. K., R. Milkman, and J. B. Mitton. 1976. Population-genetics of marine pelecypods. 4. Selection, migration and genetic differentiation in blue mussel *Mytilus edulis*. Evolution **30**:2–32.

Koehn, R. K., R. I. E. Newell, and F. Immerman. 1980. Maintenance of an aminopeptidase allele frequency cline by natural selection. Proceedings of the National Academy of Sciences (USA) **77**:5385–5389.

Kruuk, L. E. B., S. J. E. Baird, K. S. Gale, and N. H. Barton. 1999. A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. Genetics **153**:1959–1971.

Lenormand, T., T. Guillemaud, D. Bourguet, and M. Raymond. 1998. Evaluating gene flow using selected markers: a case study. Genetics **149**:1383–1392.

Mallet, J. 2001. Gene flow. Pages 337–360 *in* I. P. Woiwood, D. R. Reynolds, and C. D. Thomas, editors. Insect movement: mechanisms and consequences. CABI Publishing, New York, New York, USA.

Mallet, J., and N. H. Barton. 1989. Inference from clines stabilized by frequency-dependent selection. Evolution **122**: 967–976.

Mallet, J., N. Barton, G. Lamas, J. Santisteban, M. Muedas, and H. Eeley. 1990. Estimates of selection and gene flow from measures of cline width and linkage disequilibrium in *Heliconius* hybrid zones. Genetics **124**:921–936.

McDonald, J. H. 1987. Repeated geographic-variation at 3 enzyme loci in the amphipod *Platorchestia platensis*. Evolution **41**:438–441.

Moore, W. S. 1977. An evaluation of narrow hybrid zones in vertebrates. Quarterly Review of Biology **52**:263–278.

Moore, W. S., and J. T. Price. 1993. Nature of selection in the northern flicker hybrid zone and its implications for speciation theory. Pages 196–225 *in* R. Harrison, editor. Hybrid zones and the evolutionary process. Oxford University Press, New York, New York, USA.

Morgan-Richards, M., and G. P. Wallis. 2003. A comparison of five hybrid zones of the weta *Hemideina thoracica* (Orthoptera : Anostostomatidae): Degree of cytogenetic differentiation fails to predict zone width. Evolution **57**:849–861.

Nielsen, E. E., M. M. Hansen, D. E. Ruzzante, D. Meldrup, and P. Gronkjaer. 2003. Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. Molecular Ecology **12**:1497–1508.

Palumbi, S. R. 1994. Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics **25**:547–572.

Palumbi, S. R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. Ecological Applications **13**:S117–S137.

Palumbi, S. R. 2004. Marine reserves and ocean neighborhoods: the spatial scale of marine populations and their management. Annual Reviews in Environmental Resources **29**:31–68.

Place, A. R., and D. A. Powers. 1979. Genetic variation and relative catalytic efficiencies: lactate dehydrogenase-b allozymes of *Fundulus heteroclitus*. Proceedings of the National Academy of Sciences (USA) **76**:2354–2358.

Planes, S., and P. J. Doherty. 1997. Genetic and color interactions at a contact zone of *Acanthochromis polyacanthus*: A marine fish lacking pelagic larvae. Evolution **51**:1232–1243.

Powers, D. A. 1988. A multidisciplinary approach to the study of genetic variation within species. Pages 102–129 *in* M. F. Feder, A. F. Bennett, W. W. Burggren, and R. B. Huey, editors. New directions in ecological physiology. Cambridge University Press, New York, New York, USA.

Rawson, P. D., V. Agrawal, and T. J. Hilbish. 1999. Hybridization between the blue mussels *Mytilus galloprovinialis* and *M. trossulus* along the Pacific coast of North America: evidence for limited introgression. Marine Biology **134**:201–211.

Rawson, P. D., C. Slaughter, and P. O. Yund. 2003. Patterns of gamete incompatibility between the blue mussels *Mytilus edulis* and *M. trossulus*. Marine Biology **143**:317–325.

Rynearson, T. A., and E. V. Armbrust. 2000. DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom *Ditylum brightwellii*. Limnology and Oceanography **45**:1329–1340.

Schmidt, P. S., and D. M. Rand. 1999. Intertidal microhabitat and selection at Mpi: Interlocus contrasts in the northern acorn barnacle, *Semibalanus balanoides*. Evolution **53**:135–146.

Shulman, M. J., and E. Bermingham. 1995. Early-life histories, ocean currents, and the population genetics of Caribbean reef fishes. Evolution **49**:897–910.

Slatkin, M. 1973. Gene flow and selection in a cline. Genetics **75**:733–756.

Sotka, E. E., J. P. Wares, J. B. Barth, R. K. Grosberg, and S. R. Palumbi. 2004. Strong genetic clines and geographic variation in gene flow in the rocky intertidal barnacle *Balanus glandula*. Molecular Ecology **13**:2143–2156.

Swearer, S. E., J. S. Shima, M. E. Hellberg, S. R. Thorrold, G. P. Jones, D. R. Robertson, S. G. Morgan, K. A. Selkoe, G. M. Ruiz, and R. R. Warner. 2002. Evidence of self-

recruitment in demersal marine populations. Bulletin of Marine Science **70**:S251–S271.
Szymura, J. M., and N. H. Barton. 1986. Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *Bombina variegata*, near Cracow in southern Poland. Evolution **40**:1141–1159.
Takahata, N., and M. Slatkin. 1984. Mitochondrial gene flow. Proceedings of the National Academy of Sciences (USA) **81**: 1764–1767.
Taylor, M. S., and M. E. Hellberg. 2003. Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. Science **299**:107–109.
Vines, T. H., S. C. Kohler, A. Thiel, I. Ghira, T. R. Sands, C. J. MacCallum, N. H. Barton, and B. Nurnberger. 2003. The maintenance of reproductive isolation in a mosaic hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*. Evolution **57**:1876–1888.
Wares, J. P. 2002. Community genetics in the Northwestern Atlantic intertidal. Molecular Ecology **11**:1131–1144.
Whitlock, M. C., and D. E. McCauley. 1999. Indirect measures of gene flow and migration: Fst $\neq 1/(4Nm + 1)$. Heredity **82**: 117–125.
Wright, S. 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution **2**: 279–291.
Zapata, C., S. Rodríguez, G. Visedo, and F. Sacristán. 2001. Spectrum of nonrandom associations between microsatellite loci on human chromosome. Genetics **158**:1235–1251.