ANALYSIS OF RAPIDLY EVOLVING MOLECULES AND DNA SEQUENCE VARIANTS: ALTERNATIVE APPROACHES FOR DETECTING GENETIC STRUCTURE IN MARINE POPULATIONS

JOSEPH E. NEIGEL
Department of Biology
University of Southwestern Louisiana
Lafayette, Louisiana 70504

ABSTRACT

Rapidly evolving genetic markers, such as mitochondrial DNA (mtDNA) sequence variants, provide novel opportunities for the study of natural populations, but not without presenting special challenges in analysis and interpretation. For many species of marine organisms, it has been difficult to detect geographic structure in the distribution of genetic markers. In allozyme studies, geographic structure was sought in the frequencies of alleles among populations or locales. However, theoretical considerations suggest that nearly complete demographic isolation over long periods of time may be required to produce detectable differences in allele frequencies. Thus, past connections or demographically insignificant levels of migration may have obscured ecologically significant divisions within species.

Genetic markers that evolve rapidly are now being used, with the expectation that they will provide greater sensitivity for the detection of genetic structure. In some cases, mtDNA surveys appear to have fulfilled this expectation by revealing genetic differentiation on a finer geographic scale than allozyme surveys. However, in addition to evolving at a faster rate, mtDNA can be analyzed in terms of genealogical relationships among sequences. In this paper, theoretical models are used to evaluate methods that use rapidly evolving markers and genealogical information for the analysis of genetic population structure.

RESUMEN

Los marcadores genéticos que evolucionan rápidamente, tales como secuencias variables de ADNmt, proveen oportunidades novedosas para el estudio de poblaciones naturales, aunque su análisis e interpretación representan un reto. Para muchas especies de organismos marinos ha sido difícil detectar una estructura geográfica de la distribución de marcadores genéticos. En estudios de alocimos, la estructura geográfica se ha buscado en la frecuencia de alelomorfos entre poblaciones o lugares. Sin embargo, consideraciones teóricas sugieren que para generar diferencias en frecuencias de alemorfos que sean detectables, es necesario un aislamiento demográfico casi completo durante largos periodos de tiempo. De esta manera, conexiones en el pasado o niveles demográficos insignificantes de migración podrían haber opacado divisiones ecológicas importantes entre especies.

Hoy en día se usan marcadores genéticos que evolucionan rápidamente con la expectativa de que provean una mayor sensibilidad para la detección de patrones genéticos. En algunos casos, estudios de ADNmt parecen llenar esta expectativa, revelando diferencias genéticas en una escala geográfica mas fina que las reveladas por los alocimos. Además de evolucionar rápidamente, el ADNmt se puede analizar en términos de relaciones genealógicas entre secuencias. En este artículo, se usan modelos para evaluar métodos que usan marcadores genéticos que evolucionan rápidamente e información genealógica para el análisis de la genética de la estructura de la población.

INTRODUCTION

Species are generally viewed as composed of loosely connected populations. Populations are thought of as demographic units, and most individuals enter the population as the immediate descendants of others from the same population. But some migration between populations is expected, and it is natural to consider the magnitude of migration in terms of the proportion of individuals that are migrants. This proportion is defined as the migration rate. Thus we might consider a population in which less than one percent of the individuals arrive as migrants (a migration rate of less than 0.01) as a well-defined demographic unit; another population, which receives half of its individuals as migrants (a migration rate of 0.5) might simply be considered as part of a larger demographic unit. Populations are also thought of as genetic units, in the sense that most of the gene pool originates from within the population.

Recently developed techniques, such as DNA fingerprinting, have received much attention because they can be used to determine first-order relationships (e.g., paternity) with a high degree of confidence (Jeffreyes et al. 1985). A simple thought experiment suggests that it should be possible to extend this approach to identify demographic units and assess levels of migration. First, using genetic markers, individual paternity could be established with a high degree of confidence. Once paternity relationships were known, it would be a simple matter to distinguish individuals that were migrants from the progeny

of the residents. In practice, however, we usually cannot trace individual paternity in natural populations. Instead, we sample individuals from multiple populations, and attempt to relate the distribution of genetic variation among these samples to patterns of migration among populations. This approach demands the use of models that predict how demographic processes, such as migration, will affect the distribution of genetic variation.

In this paper, I will briefly review some theoretical models that relate migration to the distribution of genetic variation, and then consider some of the special problems that can be anticipated when we attempt to put these models to use in marine systems. I will then explore alternative approaches with rapidly evolving genetic markers, and offer some predictions on what we can expect from them.

BASIC THEORY

The basic theory that relates migration to the distribution of genetic variation was developed by Wright (1951, 1965). Detailed reviews of more recent extensions of this theory, and their relevance to fisheries management are included in Ryman and Utter (1987). Fertilization is viewed as a process in which pairs of gametes are sampled to form zygotes. If each gamete is represented as a random variable with a value determined by its ancestry, then a correlation coefficient can be defined for the pairs of gametes that combine to form zygotes. Only gametes that trace to a common ancestor, and are thus identical by descent, have the same value. When gametes combine at random, the correlation coefficient is zero. Positive correlation coefficients result when gametes of common ancestry combine more frequently than expected. Wright (1951) defined a set of correlation coefficients ("F-statistics") in terms of the correlations between gametes. Of particular interest here is F_{ST}, defined as "the correlation between random gametes within a population, relative to gametes of the total population" (Wright 1965). Thus if gametes drawn from the same population are more likely to have a common ancestor than gametes drawn from different populations, F_{ST} is positive. It is important to note that "the correlation between gametes" makes no reference to genetic variation, it is based on the ancestry of gametes.

Before the development of molecular genetic markers, the usual way to calculate F_{ST} was from pedigree data (e.g., Wright 1965). But in many situations, including studies of natural populations, pedigree information is not available, and genetic markers are used as an indication of ancestry. If gametes can be distinguished by the alleles they carry, then gametes with a common ancestry will be more likely to carry the same alleles. This is the basis for estimating F_{ST} from genetic data. If allele frequencies vary among populations, this implies that gametes within an individual population are correlated, and F_{ST} has a positive value.

A variety of theoretical models indicate a very robust relationship between F_{ST} and two other parameters: effective population size and migration rate (the proportion of individuals that enter a population as migrants). Here I'll follow the usual convention, and consider the migration rate to be based on individuals that not only physically move between populations, but also are successful at reproduction as well. In these models, the equilibrium value of F_{ST} represents a balance between the process of genetic drift and migration.

Genetic drift occurs when some gametes are by chance overrepresented among those that form zygotes, and the pairing of these correlated gametes increases F_{ST}. The magnitude of this sampling error is inversely related to the "effective population size." The effective population size is generally smaller than the actual number of individuals in the population, because differences in the reproductive contributions of individuals increase the sampling variance for gametes.

The effect of migration is opposite to that of genetic drift. By adding gametes that originated from outside the population to the sample, migration lowers F_{ST}. If both genetic drift and migration are taking place among a large group of populations, F_{ST} will approach an equilibrium value, $\hat{F}_{ST}$:

$$\hat{F}_{ST} = \frac{1}{4N_e m+1} \qquad (1)$$

where N_e is the effective size of each population, and m is the proportion of individuals in each population that are migrants.

There has been some confusion over how F_{ST} should be estimated from actual data, and how it should be interpreted (see Weir and Cockerham 1984). For example, real populations are unlikely to be of uniform size or to exchange migrants equally; thus an "average" F_{ST} must be defined to accommodate this heterogeneity. A more workable quantity, G_{ST}, was introduced by Nei (1973), with provisions for multiple alleles and multiple loci. However, as exemplified by Cockerham and Weir's analysis (1993), it is useful to retain the somewhat abstract definition of F_{ST}, so that it remains independent of the method used for its estimation. In this way, F_{ST} can provide a standard basis for comparison of different methods and approaches. If F_{ST} is defined as a purely demographic parameter—i.e., Wright's (1951) "correlation between gametes"—then, in theory, an accurate estimate of F_{ST} from allozyme data should agree with an estimate based on DNA sequence data. If these estimates did not agree, then it would indicate that at least one of the estimates was wrong. F_{ST} is thus widely used because it should be possible to compare independent

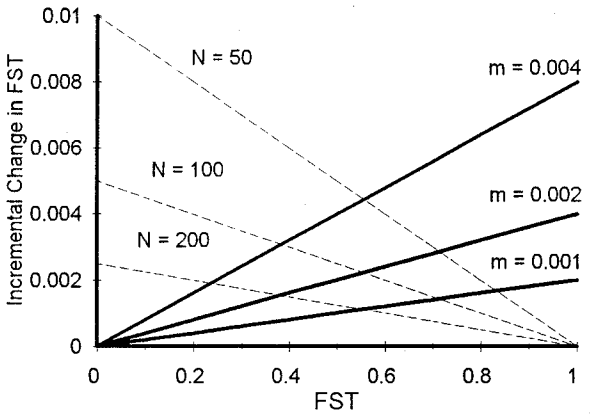

Figure 1. An equilibrium for F_{ST} occurs at the point where the increase due to genetic drift (*dashed lines*) is balanced by the decrease due to migration (*solid lines*). The lines cross over the equilibrium values of F_{ST} for particular combinations of effective population size and migration rate. Several of these combinations yield an equilibrium F_{ST} of 0.56. Because genetic drift is dependent on population size, less migration is needed to achieve this value in larger populations.

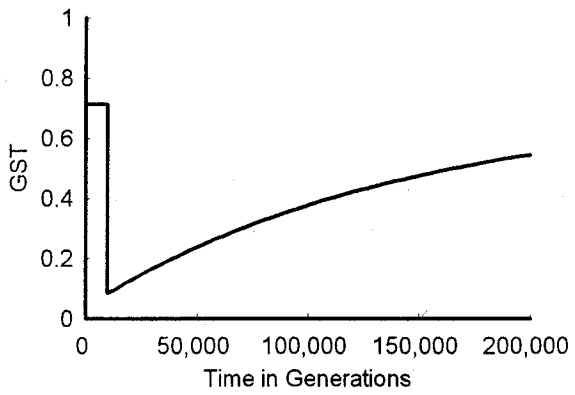

Figure 2. Among large populations, a brief episode of migration can have a long effect on G_{ST}. In this numerical example, the migration rate among populations of 100,000 was raised from zero to 0.1 at generation 10,000, then returned to zero after ten generations. In this example, the equilibrium value of G_{ST} reflects a balance between genetic drift and a mutation rate of 10^{-6}.

measurements of F_{ST} and relate them all to underlying demographic processes. In this paper, I will use F_{ST} to refer to the theoretical demographic parameter, and G_{ST} to refer to statistics that are based on genetic data. Although this convention is not universally applied, it reflects the original definitions of these quantities and avoids the introduction of additional terminology.

F_{ST} AND ESTIMATES OF MIGRATION

Equation 1 implies that the migration rate among a group of populations can be estimated from F_{ST}. However, there are several limitations to this approach. First, the equilibrium value of F_{ST} represents a balance between the opposing effects of genetic drift and migration. But whereas the effect of migration is dependent only on the migration rate, the effect of drift is inversely proportional to population size. Thus in larger populations, the effect of genetic drift is reduced, and a lower migration rate can achieve the same equilibrium value of F_{ST} as would be reached in smaller populations with higher migration rates (figure 1). For example, populations of 100 individuals that exchanged 50% of their individuals as migrants would reach the same F_{ST} as populations of one million with only 0.005% migration. As a consequence of this dependency on population size, very little migration is needed between large populations to keep F_{ST} close to zero.

In the above example, the equilibrium value of F_{ST} would be 0.005. Such small values of F_{ST} are difficult to estimate accurately from population genetic data because they require the detection of small differences in allele frequencies. To further complicate matters, the effective population size, N_e, cannot simply be equated with the number of individuals in the population, but must be adjusted for various factors that would affect the process of genetic drift. These include the sex ratio, the variance in number of offspring per individual, and fluctuations in population size over time. In practice, effective population size can seldom be estimated with much confidence.

A second problem with the interpretation of F_{ST} is the possibility that the populations have not reached an equilibrium between genetic drift and migration. Crow and Aoki (1984) showed that the number of generations, t, for F_{ST} to be near an equilibrium value is:

$$t \approx \frac{1}{2m + \frac{1}{2N_e}} \qquad (2)$$

In most instances, one of the two terms in the denominator will tend to be the major determinant of t. If the migration rate, m, is small relative to $1/2N_e$, populations will slowly diverge under the process of genetic drift, and F_{ST} will approach its maximum value of one. It takes genetic drift approximately $2N_e$ generations to approach this equilibrium. If, on the other hand, the migration rate is relatively high, migrants and residents will quickly become mixed, and F_{ST} will approach zero. Migration requires only about $1/2m$ generations to approach this equilibrium. This behavior implies that although F_{ST} can be rapidly lowered by a high migration rate, the elevation of F_{ST} by a reduction in migration will occur very slowly (see figure 2). Thus observations based on F_{ST} cannot distinguish between the immediate effects of ongoing migration and the residual effects of past migration.

F_{ST} AND MARINE POPULATIONS

Allozyme surveys of marine species with nektonic or extended planktonic life stages have typically shown very

little genetic divergence between geographic locales (for example, see the reviews by Gyllensten 1985, on non-anadromous marine fish, and Burton 1983, on marine invertebrates). Furthermore, reports of genetic divergence with low F_{ST} (or G_{ST}) values should be treated with caution.

There are two components to the variance in allele frequencies among population samples. One is the actual variance in allele frequencies among the populations; the other is the variance due to sampling a limited number of individuals from each population. If corrections are not applied for the latter component of variance, the estimate of F_{ST} will be upwardly biased, and it will appear that there is some population structure when there is actually none.

For example, imagine a group of populations, among which there is a very high level of migration, so that allele frequencies are exactly the same in each population. Samples of 15 individuals are taken from each population, and the frequencies of two alleles at one locus are determined for the samples. The actual allele frequencies in the populations are 0.5, but because of sampling error, the allele frequencies in the samples will vary around 0.5. The expected variance in allele frequencies among the samples would be about 0.008, corresponding to an uncorrected G_{ST} value of 0.03. With equation 1, the estimate for $N_e m$ would be 7.25. Thus regardless of how much migration was actually occurring, it would appear that migration was limited to about 7 individuals per generation. For large populations, this would erroneously suggest a very low migration rate (7.25 divided by the population size).

Methods for obtaining unbiased estimates of F_{ST} and related quantities have been developed by Nei and Chesser (1983) and Weir and Cockerham (1984). The latter method also corrects for the sampling error associated with a small number of populations. Unless such methods have been used, small values of G_{ST} cannot be considered good evidence of population structure.

An obvious explanation for the absence of discernible population structure in a marine species is that the potential for dispersal, provided by either passive transport in currents or active swimming, allows sufficient migration to prevent divergence among populations. Marine invertebrates without extended pelagic stages often do exhibit significant population structure (Hedgecock 1994). But lack of population structure can also be explained as an effect of large effective population sizes. Among large populations, the level of migration required to prevent genetic divergence is lowered. Furthermore, the effects of a past migration event would last longer in a large population, again because genetic drift would operate more slowly. Thus when we examine large populations that have the potential for high rates of gene flow, even if this occurs only rarely, it is not surprising to find little evidence of genetic divergence. Unfortunately, we cannot assume that such populations are strongly linked in a demographic sense. For example, consider a group of isolated populations with effective sizes of 100,000. An episode of migration, with a migration rate of ten percent for only ten generations, would eliminate nearly all of the divergence among them. However, if these populations then became completely isolated again, with no migration, it would take hundreds of thousands of generations for genetic divergence to be restored (figure 2). Thus the absence of genetic evidence for stock structure does not imply that such structure does not exist.

ALTERNATIVE APPROACHES

With the availability of new and more powerful molecular tools, more sensitive methods for detecting population genetic structure should be found. We cannot assume, however, that detection of more variation will, by itself, reveal more population structure. For although more sensitive molecular techniques can reveal additional genetic differences between individuals from different populations, the background level of genetic differences between individuals from the same population will also be increased. Our ability to detect population structure depends on finding, on average, more differences between individuals from different populations than between individuals from the same population.

The basic theory described above implies that if the correlation between gametes within a population is weak, it should be equally weak when measured with either moderately polymorphic or highly polymorphic genetic markers. Thus if we cannot differentiate populations with moderately polymorphic markers, highly polymorphic markers may do no better. This is not to say that detecting additional genetic variation isn't of any help at all. For those cases in which earlier attempts have failed to find any genetic polymorphisms, or for those in which the polymorphisms that were found were inadequate for statistical reasons, more suitable markers can probably be found by using molecular techniques.

The best hope for the development of more sensitive methods for analyzing population structure may lie not with the detection of more genetic variation, but rather with the detection of different kinds of variation. The first indication of this came about with studies of mitochondrial DNA variation in animal populations. In some cases, surveys of mtDNA appeared to provide a much more detailed picture of geographic variation than had allozymes (Avise et al. 1987; Moritz et al. 1987). This could be due in part to the maternal inheritance of mtDNA. Because females pass on a single mtDNA genotype to progeny, the effective population size for mtDNA is probably smaller than that for nuclear genes,

including those encoding most allozymes, which are present in two copies per individual and subject to biparental inheritance. A smaller effective population size would accelerate genetic drift, and thus increase divergence among populations. But even without this presumed effect of maternal inheritance, the very nature of mtDNA data would provide a different view of genetic population structure than that seen with allozyme data. This is because the polymorphisms detected in mtDNA molecules can be used to infer genealogical relationships (Avise et al. 1987).

The genealogical dimension of mtDNA data makes it possible to apply methods of analysis that could not be used with allozyme data. For example, the principles of cladistics can be applied to mtDNA variation within a species or population. In the cladistic approach to systematics, "characters" shared among groups of organisms are used to place them into hierarchical groups, or "clades." This hierarchy of groups is represented as a phylogenetic tree. One of the major tools of cladistics, parsimony analysis, identifies the phylogenetic tree that requires the least number of character transformations.

Slatkin and Maddison (1989) developed a method to estimate the product of effective population and migration rate (N_em) from a phylogenetic tree of the mtDNA within a species. This method treats the geographic location of each individual as a character, and uses a parsimony analysis to determine the minimum number of migration events that could reconcile this location "character" with the mtDNA phylogeny. Computer simulations provide a simple function to convert the minimum number of migration events to an estimate of N_em.

Neigel et al. (1991) observed that in many animal species, mtDNA variants (referred to as haplotypes) do not appear to have reached equilibrium distributions across the species' range. Groups of related mtDNA haplotypes, which represent maternal lineages, are often clustered geographically. Whereas geographic clustering of individual haplotypes could be explained by drift, the clustering of multiple related haplotypes implies that the association is historical. A model was proposed in which each new mtDNA lineage begins with a single individual at a specific location, and then disperses from that point over multiple generations. Dispersal is not limited to movements between distinct populations, but is limited by absolute distance. Individual haplotypes represent the youngest lineages; older lineages are composed of groups of related haplotypes. The model predicts that if dispersal is constant over time, the variance of a lineage's geographic distribution should be proportional to its age.

Furthermore, if the ages of mtDNA lineages can be estimated in generations, a standardized single-generation dispersal distance can be estimated from mtDNA data.

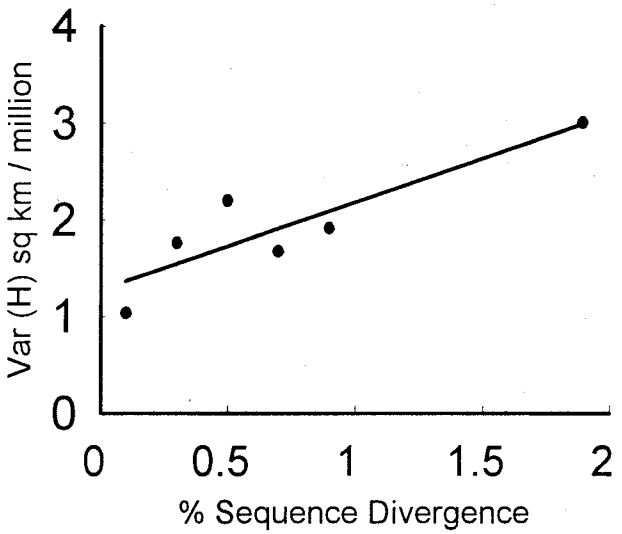

Figure 3. The relationship between the ages of mtDNA lineages and the variances of their geographic distributions in the American oyster, *Crassostrea virginica*. The positive slope of this relationship indicates that younger lineages have not achieved equilibrium distributions, and corresponds to a single-generation dispersal distance of 3.3 km.

This concept was tested further by Neigel and Avise (1993), who analyzed nine mtDNA data sets, including data from the American oyster, *Crassostrea virginica*; the American eel, *Anguilla rostrata*; and the hard-head catfish, *Arius felis*. Of these three marine species, only the American oyster showed evidence of nonequilibrium distributions. As shown in figure 3, the geographic distributions of older lineages have higher variances. The slope of this relationship corresponds to a single-generation dispersal distance of 3.3 km.

In retrospect, it is not surprising that the genealogical nature of mtDNA variation has suggested new approaches for the analysis of genetic population structure. The basic theory represented by the use of F_{ST} as a population structure parameter has now been extended to include maternally transmitted mtDNA (Takahata and Palumbi 1985) as well as genealogical relationships among DNA sequences (Slatkin 1991). However, these alternative methods of analysis still require that the frequencies of genetic variants have diverged between populations, and thus the methods are of no help if divergence is lacking.

What is more surprising is that some populations that exhibited very little divergence in allozyme allele frequencies have been found to differ sharply in mtDNA. One of the most dramatic examples is provided by studies of the American oyster, *Crassostrea virginica*. The contrast between the results of an allozyme study conducted by Buroker (1983) and a mtDNA study conducted by Reeb and Avise (1990) are shown in figure 4. Although allozyme frequencies at most loci are similar among these populations, there is a sharp division in mtDNA variation that separates Atlantic and

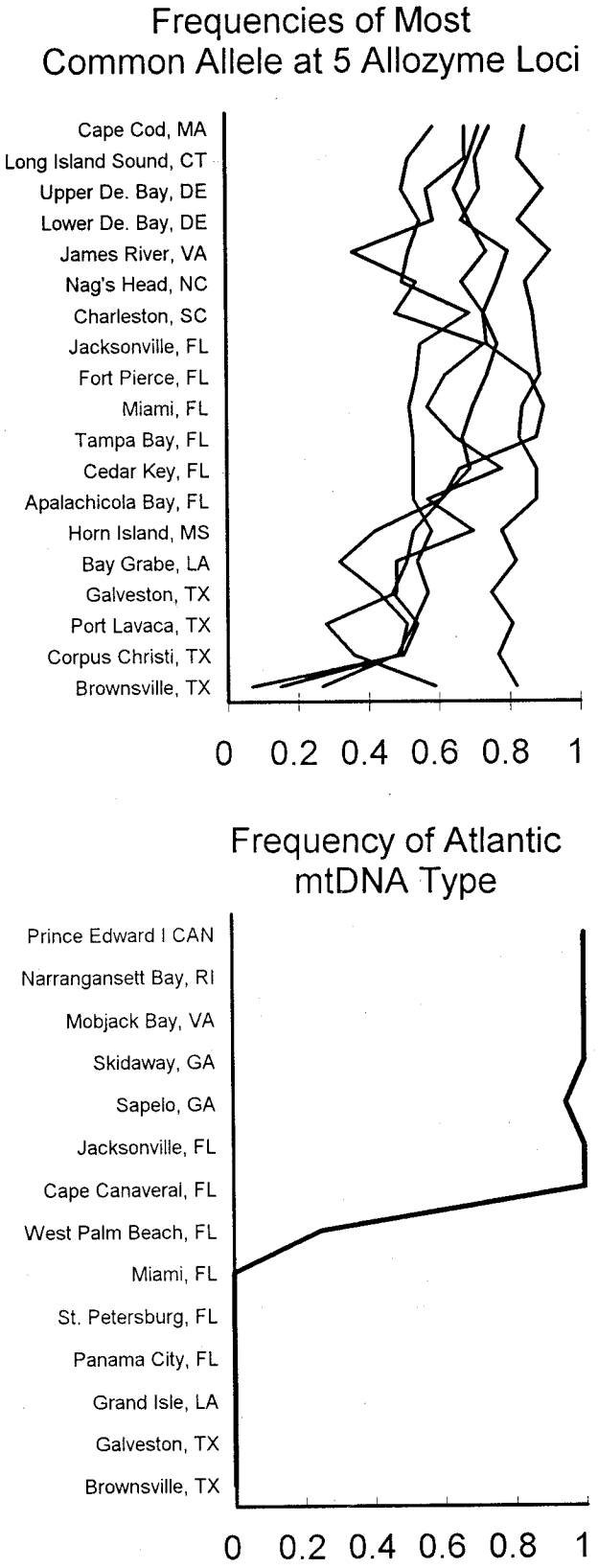

Figure 4. Geographic variation in frequencies of allozyme alleles and mtDNA haplotypes in the American oyster, *Crassostrea virginica*. *Upper panel*, frequencies of most common allele at 5 allozyme loci (Buroker 1983). *Lower panel*, frequency of Atlantic mtDNA type (Reeb and Avise 1990).

Gulf of Mexico populations. This difference in resolution is clearly not due to the greater variability of mtDNA. Even if only the two major lineages of mtDNA were distinguished, the pattern would still be very clear. It also seems unlikely that a difference in effective population size could produce such a dramatic difference between allozyme and mtDNA distributions. The difference in F_{ST} for these data sets corresponds to over a hundredfold difference in Nm. In the case of the American oyster, it appears that either allozymes, mtDNA, or both are not behaving as predicted by the basic theory that links migration to genetic divergence. Because this theory assumes that the genetic markers in question are not subject to natural selection, one possibility is that the frequencies of allozyme alleles are held constant by natural selection, while mtDNA variants are subject to genetic drift.

Support for this interpretation was recently provided by a study of nuclear DNA sequence polymorphisms in the American oyster. Karl and Avise (1992) examined variation in "anonymous" nuclear single copy DNA sequences, which for the most part do not appear to encode proteins, and would therefore not be subject to the same forms of selection that would act upon genes that encode allozymes. Although these nuclear DNA sequences were only moderately polymorphic, they clearly separated Gulf of Mexico and Atlantic populations and therefore corroborated the mtDNA pattern.

The complete replacement of one form of mtDNA with another, as well as the extent of mtDNA sequence divergence between Atlantic and Gulf of Mexico populations of American oyster, suggests a long period of complete demographic isolation, such as would be expected of distinct species. This is a crucial point, because although mtDNA proved to be a better tool than allozymes, it was not by virtue of detecting weak population structure. Apart from the major Gulf/Atlantic division, the mtDNA survey of Reeb and Avise (1990) revealed very little population structure. For example, it did not distinguish oysters from individual embayments. This is consistent with other surveys of mtDNA in marine species, many of which have failed to reveal any significant population structure (Ovenden 1990).

CONCLUSIONS

Two basic approaches can be followed in future applications of molecular markers to detect and quantify the weak genetic population structure expected in marine systems. First, more refined estimates of F_{ST} could be made by increasing both the number of polymorphic loci and the number of individuals sampled. As discussed above, it is not enough to find differences in the frequencies of genetic markers among population samples;

these differences must be shown to be statistically significant. For this approach, nuclear DNA markers may be superior to mtDNA. Although it is not a trivial matter to identify nuclear DNA polymorphisms (see Karl and Avise 1993), the sampling of multiple independently segregating loci is only possible with nuclear markers. Furthermore, there is no clear benefit to using an extremely polymorphic molecule. Excessive polymorphism can actually obscure genetic population structure if the markers fail to reveal similarities as well as differences. Development of new technologies for screening DNA polymorphisms, such as denaturing gradient gel electrophoresis (Myers et al. 1986), should allow larger numbers of individuals to be sampled, perhaps in the range of several thousand. Such large sample sizes would extend the statistical power of conventional approaches to the detection of genetic population structure.

A second, more daring approach is to experiment with novel genetic markers, with the hope that some may prove especially good at detecting genetic population structure. If mtDNA has occasionally revealed previously unseen population structure, other markers may do so as well. Recently, Turner and co-workers (1991) reported a highly repetitive DNA sequence in pupfishes (*Cyprinodon variegatus*) that exhibits extreme uniformity within populations, but marked divergence between populations along the Atlantic coast of North America. In contrast, allozyme surveys have not differentiated these populations (Darling 1976; Duggins et al. 1983). One possibility is that, as suggested for the American oyster, allozyme alleles in pupfish populations are maintained at nearly constant frequencies by selection. The divergence of repetitive sequences may then simply represent normal genetic drift. Another, more tantalizing, possibility also exists. Other repetitive sequences, such as those that code for globin genes, are known to exhibit a phenomenon called "concerted evolution" (Zimmer et al. 1980), in which copies of a sequence within a lineage diverge more slowly than homologous sequences in different lineages. If concerted evolution occurs at the level of populations, it is conceivable that an unusually high proportion of variation in repetitive sequences may be partitioned among populations.

It has often been difficult to detect genetic differences between populations of marine species that have planktonic larvae. If genetic differences are in fact present, molecular methods that can increase sample sizes and identify novel kinds of markers may help to surmount this hurdle. Once genetic differences between populations are detected, the new statistical tools that have been developed from theoretical population genetics can be used to analyze these differences, and should provide new insights into the biology and history of marine populations.

ACKNOWLEDGMENTS

Discussions with John Avise, Marty Ball, and Bob Zink have contributed to my ideas about gene flow and mitochondrial DNA. I am also indebted to Monty Slatkin for valuable advice, criticism, and encouragement. Comments by Dennis Hedgecock and an anonymous reviewer about an earlier draft of this paper were most helpful. Support for my work has been provided by grants from the NSF EPSCoR program, and the Louisiana Education Quality Support Fund.

LITERATURE CITED

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reed, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Ann. Rev. Ecol. Syst. 18:489–522.

Buroker, N. E. 1983. Population genetics of the American oyster *Crassostrea virginica* along the Atlantic coast and the Gulf of Mexico. Mar. Biol. 75:99–112.

Burton, R. S. 1983. Protein polymorphisms and genetic differentiation of marine invertebrate populations. Mar. Biol. Lett. 4:193–206.

Cockerham, C. C., and B. S. Weir. 1993. Estimation of gene flow from F-statistics. Evolution 47:855–863.

Crow, J. F., and K. Aoki. 1984. Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. PNAS USA 81:6073–6077.

Darling, J. D. 1976. Electrophoretic variation in *Cyprinodon variegatus* and systematics of some fishes of the subfamily Cyprinodontinae. Ph.D. dissertation, Yale Univ.

Duggins, C. F., Jr., A. A. Karlin, and K. G. Relyea. 1983. Electrophoretic comparison of *Cyprinodon variegatus* Lacepede and *Cyprinodon hubbsi* Carr, with comments on the genus *Cyprinodon* (Atheriniformes: Cyprinodontidae). Northeast Gulf Sci. 6:99–107.

Gyllensten, U. 1985. The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous and freshwater fishes. J. Fish. Biol. 26:691–699.

Hedgecock, D. 1994. Temporal and spatial genetic structure of marine animal populations in the California Current. Calif. Coop. Oceanic Fish. Invest. Rep. 35 (this volume).

Jeffreys, A. J., V. Wilson, and S. L. Thein. 1985. Hypervariable 'minisatellite' regions in human DNA. Nature 314:67–73.

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

———. 1993. PCR-based assays of Mendelian polymorphisms from anonymous single-copy nuclear DNA: techniques and applications for population genetics. Mol. Biol. Evol. 10:342–361.

Moritz, C., T. E. Dowling, and W. M. Brown. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Ann. Rev. Ecol. Syst. 18:269–292.

Myers, R. M., T. Maniatis, and L. S. Lerman. 1986. Detection and localization of single base changes by denaturing gradient gel electrophoresis. Meth. Enzymol. 155:501–527.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. PNAS USA70:3321–3323.

Nei, M., and R. K. Chesser. 1983. Estimation of fixation indices and gene diversities. Ann. Hum. Gen. 47:253–259.

Neigel, J. E., and J. C. Avise. 1993. Application of a random walk model to geographic distributions of animal mitochondrial DNA variation. Genetics 135:1209–1220.

Neigel, J. E., R. M. Ball, and J. C. Avise. 1991. Estimation of single generation migration distances from geographic variation in animal mitochondrial DNA. Evolution 45:423–432.

Ovenden, J. R. 1990. Mitochondrial DNA and marine stock assessment: a review. Aust. J. Mar. Freshwater Res. 41:835–853.

Reeb, C. A., and J. C. Avise. 1990. A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, *Crassostrea virginica*. Genetics 124:397–406.

Ryman, N., and F. Utter. 1987. Population genetics and fishery management. Seattle: Univ. Wash. Press, 420 pp.

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

Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123:603–613.

Takahata, N., and S. R. Palumbi. 1985. Extranuclear differentiation and gene flow in the finite island model. Genetics 109:441–457.

Turner, B. J., J. F. Elder, Jr., and T. F. Laughlin. 1991. Repetitive DNA sequences and the divergence of fish populations: some hopeful beginnings. J. Fish Biol. 39:131–142.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evol. 38:1358–1370.

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

———. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. Evol. 19:395–420.

Zimmer, E. A., S. L. Martin, S. M. Beverley, Y. W. Kan, and A. C. Wilson. 1980. Rapid duplication and loss of genes coding for the α chains of hemoglobin. PNAS USA 77:2158–2162.