

# METHODOLOGIES FOR CONSERVATION ASSESSMENTS OF THE GENETIC BIODIVERSITY OF AQUATIC MACRO-ORGANISMS

BERT, T. M., SEYOUM, S., TRINGALI, M. D. and McMILLEN-JACKSON, A.

Florida Marine Research Institute, 100 Eighth Avenue Southeast, St. Petersburg, Florida, 33701, USA

Correspondence to: Theresa M. Bert, Florida Marine Research Institute, 100 Eighth Avenue Southeast,  
St. Petersburg, Florida, 33701, USA, e-mail: [theresa.bert@fwc.state.fl.us](mailto:theresa.bert@fwc.state.fl.us)

Received October 20, 2001 – Accepted October 31, 2001 – Distributed August 31, 2002

## ABSTRACT

International organizations and biodiversity scientists recognize three levels of biodiversity: genetic, species, and ecosystem. However, most studies with the goal of assessing biodiversity collect data at only a single level – that of the species. Even when multiple levels of biodiversity are considered, usually only ecosystem diversity is also evaluated. Genetic diversity is virtually never considered. Yet, genetic diversity is essential for the maintenance of populations and species over ecological and evolutionary time periods. Moreover, because components of genetic diversity are independent of either species or ecosystem diversity, genetic diversity can provide a unique measure by which to assess the value of regions for conservation. Regions can be valuable for conservation of their genetic resources regardless of their levels of species or ecosystem uniqueness or diversity. In general, the same methods and statistical programs that are used to answer questions about population genetics and phylogenetics are applicable to conservation genetics. Thus, numerous genetic techniques, laboratory methods, and statistical programs are available for assessing regional levels of genetic diversity for conservation considerations. Here, we provide the rationale, techniques available, field and laboratory protocols, and statistical programs that can be used to estimate the magnitude and type of genetic diversity in regions. We also provide information on how to obtain commonly utilized statistical programs and the type of analyses that they include. The guide that we present here can be used to conduct investigations of the genetic diversity of regions under consideration for conservation of their natural resources.

*Key words:* aquatic, biodiversity, conservation, fish, genetics, phylogenetics, populations.

## RESUMO

### **Metodologias para avaliação da conservação da biodiversidade genética de macroorganismos aquáticos**

Organizações internacionais e pesquisadores da biodiversidade reconhecem três níveis de biodiversidade: da genética, de espécies e de ecossistemas. Entretanto, muitos estudos desenvolvidos com o objetivo de estimar a biodiversidade coletam dados somente em um único nível – o de espécies – e, mesmo quando diferentes níveis da biodiversidade são considerados, usualmente apenas a diversidade de ecossistema é avaliada, sendo que a diversidade genética raramente é avaliada. No entanto, o conhecimento da diversidade genética é essencial para a manutenção das populações e das espécies em períodos ecológicos e evolutivos. Além disso, como seus componentes são independentes de outras espécies ou da diversidade de ecossistemas, a diversidade genética pode fornecer uma medida pela qual pode-se estimar o valor das regiões para conservação. As regiões podem ser valiosas para a conservação de seus recursos genéticos independente de seus níveis de espécies, ecossistema ou diversidade. Em geral, o método e o programa estatístico utilizados para responder questões sobre genética de populações e sobre filogenética são aplicáveis para conservação genética. Assim, numerosas

técnicas genéticas, métodos laboratoriais e programas estatísticos estão disponíveis para estimar os níveis regionais da diversidade genética para conservação. Neste trabalho são apresentadas as razões, as técnicas disponíveis, os protocolos de campo e laboratório e os programas estatísticos que podem ser empregados para estimar a magnitude e o tipo de diversidade genética nas regiões. Também são dadas informações sobre como obter os programas estatísticos comumente utilizados e as formas de análises que eles incluem. O roteiro apresentado pode ser utilizado para conduzir investigações da diversidade genética de regiões em estudo visando à conservação de seus recursos naturais.

*Palavras-chave:* aquática, biodiversidade, conservação, peixe, genética, filogenética, população.

## INTRODUCTION

The conservation of genetic diversity is of paramount importance for long-term species survival because the maintenance of gene pools with sufficient variability is necessary for species to adapt to changing environments (Lande, 1995; Ayala, 1997; Meffe & Carroll, 1997). Recognizing this, international organizations (e.g., Convention on Biological Diversity [CBD] and DIVERSITAS) and scientists who study biodiversity (e.g., Gaston, 2000) identify three levels of biodiversity: genetic, species, and ecosystem. The Convention on Biological Diversity recommended the inclusion of genetic diversity, together with species diversity and ecosystem diversity, in assessments of the conservation value of a region (see [www.biodiv.org](http://www.biodiv.org)). Rapid-assessment evaluations of biological diversity are ongoing worldwide, at both international (e.g., the Rapid Assessment Program, Aquatic Rapid Assessment Program [see the websites of Conservation International, [www.conservation.org](http://www.conservation.org), or The Field Museum, [www.fieldmuseum.org](http://www.fieldmuseum.org)]), and national/regional levels (e.g., the Biodiversity Virtual Institute Program supported by the State of São Paulo Research Foundation (FAPESP) within the BIOTA/FAPESP [see [www.biotasp.org.br](http://www.biotasp.org.br)]), and they are gaining in importance. Unfortunately, genetic diversity is rarely included as an integral component in these types of studies.

Multiple, independent measures of a region's biotic systems provide greater insight into its history and processes and a better basis for predicting future biodiversity than do unilateral measures (McKinney *et al.*, 1998). Most importantly for preservation of regions for the conservation of their biodiversity, genetic diversity can be used as a measure of biodiversity that is independent from either species or ecosystem

biodiversity. Areas may be of critical importance for the preservation of genetic diversity but be unremarkable in their species number or ecosystem uniqueness. Thus, areas that otherwise may be excluded from conservation efforts using other measures of biodiversity can be conserved based on their genetic diversity.

Genetic analysis provides a unique assessment of the conservation value of a region because it facilitates the identification of areas that harbor unusually high levels of genetic diversity, that act as havens for ancient or novel genetic lineages, or that possess habitats conducive to speciation. Areas of unusually high genetic diversity in many lineages serve as potential sources of evolutionary importance because they are essentially genetic "banks". Areas can have any level of genetic diversity and yet be of importance because they are genetic repositories for ancestral or derived genotypes. Areas of high genetic diversity in particular lineages may be characterized by numerous faunal or floral hybrid zones, where coadapted gene complexes from different taxonomic lineages are recombining in novel ways. Locations of hybrid zones as well as regions possessing representatives of species-rich clades are thought to be sites of ongoing evolution and speciation (Harrison, 1990; Erwin, 1991). Those areas make exceptionally important contributions to the world's pool of evolutionary diversity (Vane-Wright *et al.*, 1994) and they are important for the generation of future biodiversity (Avice, 1996).

In addition, the pattern and type of genetic diversity in an area can be an indicator that the area constitutes an important ecological zone. For example, the co-occurrence of multiple hybrid zones in an area can be a strong indicator of the presence of an ecotone because hybrid zones both tend to be generated and to collect in ecotones

(Barton, 1985; Moore & Price, 1993). Because they are ecological transition zones, ecotones tend to harbor high numbers of species as well as unusual genetic representatives of those species. Such areas are sources of evolutionarily novel genetic combinations and are important to preserve for future biodiversity (Smith *et al.*, 1997).

Genetic analyses using standard population genetics and evolutionary genetics laboratory protocols and analytical procedures yield excellent information for evaluating an area for its genetic conservation potential. At both macrogeographic and microgeographic levels, these types of analyses provide information on not only the magnitude, type, and distribution of genetic variation sequestered within and between species but also on population subdivision and movement and mating patterns of individuals within species (see, e.g., Bowen *et al.*, 1992). Levels of gene flow, migration rates of individuals between populations, and effective population sizes (the theoretically based number of breeders in a population [Wright, 1969; Slatkin, 1985; Beerli & Felsenstein, 2001]) can be estimated. The evolutionary relationships among populations and species, populations harboring ancient or speciose lineages, and areas of hybridization can be identified. In addition, the genetic population structure of some species can serve as a model for other species with similar life histories, distributions, and dispersal characteristics. Thorough surveys of the population genetics of key species can provide information on the potential of numerous species for recolonization if those species are eliminated from a local area or a broad region.

Particularly for aquatic groups, the existing morphologically based taxonomy for a group may not reflect the underlying genetic diversity (Moritz, 1994), in part because many groups are poorly known. Thus, as well as providing unique information on the level of genetic importance of an area, biodiversity studies based on phylogenetic and population genetic methods can also be blended with morphological analyses to identify new species, lineages, and areas of hybridization. Moreover, because it is physically or logistically impossible to tag and track the many small species of fish and invertebrates that inhabit tropical and subtropical river systems, genetic analysis can provide the only window into the movement patterns and populational relationships of many

of these species. In addition, genetic analysis for conservation studies is invaluable because the samples are frequently collected and the analyses conducted when the aquatic system is relatively pristine and most species are at natural population levels. Genetic analyses can thereby provide critical baseline information on natural levels of genetic diversity and population structure of numerous species. Thus, these types of genetics studies are important from the perspectives of evolutionary biology, population ecology, and fisheries biology, as well as conservation biology.

Here, we describe various types of procedures that can be used to conduct a rapid assessment of the genetic diversity of an aquatic region. Studies of genetic variation, population genetic structuring, and phylogenetics of species can be conducted on any group of living organisms collected in the field. We refer principally to fish species because this class of organisms has been the focus of our studies. However, the methods presented here can be generally applied to most aquatic macro-organisms with little modification.

## FIELD METHODS

Rapid qualitative assessments to assess the importance and potential of a particular geographic region for the conservation of species' genetic diversity can be made by scientists knowledgeable in population genetics, evolutionary genetics, or conservation genetics. In addition to the obvious need to collect tissue samples for genetic analysis, the geneticist must have the opportunity to make detailed, first-hand observations of the habitat structure and hydrographic regime of the region being considered for conservation and also the species composition of the collections being made. Some knowledge of the population biology and reproductive strategy of as many of the species being collected as possible is also helpful. More detailed, quantitative assessments require a laboratory with sufficient equipment to conduct extensive genetic analyses and a full-time researcher for at least 6 months to perform the laboratory work.

Sampling for conservation assessments, particularly rapid assessments, is frequently multidisciplinary; involves people from several institutions, organizations, or museums; and has

several different objectives for both the sampling and the samples. Therefore, we describe the collection of samples for surveys of genetic diversity in the context of using those samples for multiple purposes and coordinating the collection of those samples with the collection of samples to be used for other purposes.

Prior to the field trip, considerable preparations should be made. Obtain as much knowledge as possible concerning the geography and aquatic community structure of the area to be sampled. Search in the published scientific literature for population genetics studies of species from the general region. Work with people knowledgeable about the flora and fauna of the region to identify locations that are potentially important for conserving genetic resources. Combine this information with other information provided by the scientists conducting other components of sampling to determine the specific sites to be sampled. Determine the distribution of the samples among the participating institutions and, if necessary, obtain the necessary permits to export samples from the host country and to import samples to the country at which the laboratory analyses will be conducted. Both the location of adequate facilities for the storage and maintenance of the samples and the priority rights of the host country should be considered when discussing the distribution of samples.

Typically, live fish and invertebrates collected for species-diversity counts are preserved directly in buffered formalin. Molecular genetics research is best performed on tissues that are maintained without preservatives or that are preserved without formalin. Thus, special considerations for preserving tissues without formalin (for example, preservation in alcohol or a Tris-salt buffer containing DMSO [dimethyl sulfoxide], freezing, or drying) should be made prior to the field trip. If possible, more than one method of preservation should be used and duplicate samples of each individual should be taken when sufficient tissue is available. This will increase the probability that at least one sample from each individual will successfully arrive at the institution where the laboratory analyses will be conducted. In addition, these samples can be archived in museums or tissue libraries for future use in not only genetic analyses, but also heavy metal or pesticide detection, deter-

minations of physiological properties, or other types of analyses that require specially preserved samples. These types of samples are invaluable because they can provide, for many disciplines of science, baseline data for regions where development and environmental alteration may be imminent but have not yet impacted the biota.

To ensure the broadest representation of species and intraspecific populations, collections should be made from as many different habitats as possible (e.g., in as many habitat types as possible in lakes, streams, rapids, flooded areas, river banks, main channels, estuaries, and in the ocean, nearshore and in open water). Careful notes on the ecology of each collecting site and a characterization of the habitat should always be made. If possible, photographs should be taken of each habitat, of individuals representative of each species, and of all unique or unusual morphotypes. These are also the goals of the scientists who collect specimens to estimate species diversity. Thus, the sampling strategies for collecting samples at these two levels of biodiversity should be compatible and the collection of organisms for genetic analysis can be done in conjunction with the collections obtained for assessing species diversity.

Tissue collections from species that can be readily or eventually taxonomically identified should be made at two levels. (1) At least one representative of as many morphologically distinguishable species as can be obtained should be sampled. These samples are useful for phylogenetically based analyses and, together with their formalin-preserved morphological counterparts, serve as valuable archived museum specimens. (2) Numerous individuals (optimally, 25-75) of species that are abundant in a field collection should be sampled. If possible, these population-level collections should contain individuals of different age groups within species. Usually this can be accomplished by collecting individuals over a broad size range. If this is possible, the estimated or measured size or the age class of each individual should be recorded. These samples serve as representatives of local populations in genetic analyses conducted to define intraspecific population genetic structure, gene flow, migration rates, and species- or population-level genetic variation. Different species may be common in different habitats. Thus, it is possible that by the end of the

field trip, only a single population-level sample has been obtained for some species. Nevertheless, these samples should be collected whenever possible and saved so that they can be compared to other population-level samples of the same species that may be collected during field trips to other locations.

We use a three-level approach for tissue collection and laboratory analysis in our studies of the genetic component of biodiversity. The equipment that we take in the field to obtain these collections is listed in Table 1.

(1) From each fish, we attempt to obtain a tissue sample of somatic muscle (and liver if the fish is large and gonad if the fish is reproductively mature) for freezing in liquid nitrogen or on dry ice or for preserving in alcohol, DMSO buffer, or

(for tissues destined for DNA analysis), a lysis buffer containing proteinase K. The frozen tissues are versatile; they can be used for allozyme electrophoresis and other procedures that require tissues that have not been exposed to preservatives. Frozen tissues also are easier to use than preserved tissues for more sophisticated DNA analyses. However, the logistics associated with collecting, transporting, and maintaining frozen tissues are substantial. Thus, this method may not be logistically possible.

Tissue biopsies such as blood samples, needle biopsies, fin clips, small tissue plugs, skin scrapings, or appendage components can usually be taken non-destructively from individuals of sufficient size. This type of tissue collection is particularly important for endangered, threatened, or rare and endemic species.

**TABLE 1**  
Standard field equipment for obtaining tissue samples of aquatic macroorganisms for conservation genetics analysis.

Task	Equipment
Dissecting the tissues	Waterproof paper; pencils or rapidographs and India ink for labeling samples 2-3 Pairs of surgical-quality scissors; scalpels; small scalpel blades (size 11), large scalpel blades (size 22) Forceps, toothed and blunt-tipped Tray for dissecting; plastic gloves (e.g., surgical gloves) Cloth towels; paper towels (Optional) Bunsen burner with fuel tank for sterilizing dissecting utensils
Freezing the tissues <sup>1</sup>	Large cryogenic (liquid nitrogen) flask containing liquid nitrogen or styrofoam coolers filled with dry ice String; women's stockings; labeling tape (for storing samples in the cryogenic container) Dry cryogenic shippers (shippers that can be cooled by filling with liquid nitrogen, which is then poured out for transport of samples via airplane) Clear plastic wrap (we prefer Handiwrap <sup>®</sup> ); heavy-grade aluminum foil
Preserving voucher specimens	Screw-cap plastic vials (5 ml-10 ml), or larger, as needed 70% Ethanol, isopropanol, or DMSO (the long-term utility of DMSO for tissue preservation has not yet been proven) Gallon-sized Nalgene plastic jars for storing vials containing the preserved tissues Wire for wiring labels to fish; wire cutters; Ziploc plastic bags for storage of voucher specimens; permanent markers Alcohol and formalin for preservation Airtight containers for storage and shipping

1. Optional, but freezing is necessary for protein electrophoresis.

(2) From each fish, we also attempt to obtain a tissue sample for preservation (in ethanol, isopropyl alcohol, or DMSO-buffer). These tissue samples can be used for almost any type of DNA analysis other than allozyme electrophoresis (e.g., DNA sequencing, restriction fragment length polymorphism [RFLP] analysis, random amplified polymorphic DNA [RAPD] analysis, denatured gradient gel electrophoresis [DGGE] analysis, microsatellite DNA analysis). They are easy to preserve, transport, and store; and they remain viable as usable samples for very long periods of time.

(3) After obtaining our tissues for genetic analyses, we preserve the remainder of each fish in buffered formalin if we have obtained tissue samples for both freezing and alcohol preservation, or we preserve the fish remains in an alcohol if we have obtained a tissue sample only for freezing, or if the fish is very small. We use these fish bodies as voucher specimens for later positive taxonomic identification because some fish cannot be unambiguously identified in the field. Therefore, body components critical for morphological species identification should be left intact.

Correctly and unambiguously labeling all samples from each individual is very important. Thus, before embarking on the expedition, a labeling system should be set up to expedite the process of tagging the frozen or solution-preserved tissues and the voucher samples from which the tissues were taken. A method of wrapping the samples to be frozen should be devised. If samples will be frozen, arrangements should be made in advance to procure liquid nitrogen or dry ice. If liquid nitrogen is to be used, the cryogenic flask and dry shippers should be shipped to the staging location ahead of time and, if possible, chilled with liquid nitrogen prior to the field trip; the other equipment should also be shipped in advance, if possible. A source of liquid nitrogen should be located prior to the departure date for the fieldwork and the cryogenic flask should be partially filled with liquid nitrogen just prior to departure for the field. If a method for freezing tissues is not available, all tissues can be preserved in alcohol, but this will eliminate allozyme electrophoresis as a candidate for genetic analysis.

Contamination of the sample in the field is a problem; it is particularly easy for small or loose pieces of tissue or body parts (e.g., skin flakes, hair, fish scales) or body fluids (e.g., blood,

mucous) to be transferred from one sample to another. Considerable care should be taken to clean the surface of the dissecting tray and the dissecting utensils after each dissection. A Bunsen burner can be used to sterilize the dissecting equipment after each dissection (optional). Storage limitations typically necessitate that small samples (approximately 1 sq. cm) are taken for freezing and for fluid preservation and that only one type of sample is collected from some individuals. If the specimens are very small and only one type of sample can be taken, we take only samples for freezing if possible. To obtain high quality frozen tissues, the individuals should be kept alive or cold (e.g., on ice) until dissection. (Tissue dissection is easier if the specimens have been cooled.) The tissues should be dissected as soon as possible and the tissues should be frozen within a few minutes after dissection. Long-term storage for frozen samples should be at  $-20^{\circ}\text{C}$  to  $-80^{\circ}\text{C}$ , until they are used in a laboratory analysis ( $-80^{\circ}\text{C}$  is essential for tissues designated for allozyme electrophoresis). The limitations on the storage capacity of the cryogenic containers in the field should also be considered when deciding how to prioritize samples for freezing versus other methods of preservation.

Each individual should be assigned a unique sample number; the tissue samples and the voucher specimen of each individual should be labeled with that number using waterproof paper or other method of permanent marking. We wrap our fish to be frozen first in a 7 cm x 7 cm square of plastic wrap (we prefer Handiwrap<sup>®</sup>) and then in an aluminum foil square of comparable size, and we place the sample label (labeled waterproof paper) between the two wrapping materials. We also wire the sample number to the fish's jaw whenever the fish is of sufficient size to keep as a voucher specimen.

Evaporation of the liquid nitrogen or dry ice is a serious problem in the field. Therefore, if possible, cool the samples on ice prior to placing them in the container used for freezing the tissues. Several samples can be held on ice and put collectively into the storage container. Because retrieving and sorting samples placed loosely in a large cryogenic container is problematic, we put the samples into women's stockings that have been connected to the handles of the flask via a labeled string.

For long-term storage, the alcohol- or DMSO-preserved samples in small vials filled with the preservative can be placed into large plastic

(e.g., Nalgene®) jars; the jars can then be filled with the same preservative. Frozen tissues should be stored at  $-80^{\circ}\text{C}$  in a freezer with an independent battery-operated alarm system. Formalin-preserved individuals should be changed into alcohol as soon as possible. All samples should be cross-referenced with the collection from which they came.

## LABORATORY ANALYSES

### *Techniques*

Many molecular genetics techniques appropriate for studies of genetic diversity exist. These include techniques involving the differential migration of DNA through a gel based on length of DNA fragment (e.g., RFLP or amplified fragment length polymorphism [AFLP] analysis, RAPD analysis, microsatellite DNA analysis) or on molecular configuration, charge, or exact nucleotide sequence of DNA (e.g., DGGE, ASO [allele-specific oligonucleotide] probing); sequencing of DNA (performed on mitochondrial DNA [mtDNA] or various components of nuclear DNA); and analysis of DNA products (e.g., amino acid sequencing, protein [= allozyme] electrophoresis). Excellent descriptions of these and other techniques used for conservation genetics and some applications are provided in Smith & Wayne (1996) and Bert *et al.* (2002) and references therein. Regardless of the methods used, both nuclear and mitochondrial DNA should be surveyed if possible.

Although starch gel allozyme electrophoresis is the oldest and crudest of all DNA-based techniques (the technique was first applied to population genetics by Lewontin & Hubby [1966]), it is still among the cheapest, fastest, and easiest of all DNA-based population genetics techniques, and it requires the least money and very little specialized equipment to set up a laboratory and execute the protocol. A plethora of data is available from the scientific literature on the genetic variation and population genetic structure of numerous organisms as determined by the analysis of allozyme alleles. In addition, a large array of statistical analytical methods directed at determining population-level genetic variation, population genetic structure within species, effective population size, gene flow between populations, and phylogenetic relationships among species are compatible with allozyme data.

However, there are several notable limitations to allozyme electrophoresis. (1) Probably most importantly, the direct link between the expression of allozyme banding patterns on gels and the genes coding for those patterns has been established for only relatively few loci coding for allozymes and for only a limited number of species. Thus, for most loci for most organisms, the researcher must make the assumption that the output from electrophoresis (essentially colored bands on a gel) corresponds directly to genetic alleles at a single locus. Ideally, the bands can be interpreted as the direct products of single genes with codominant alleles that are expressed as genotypes that conform to Hardy-Weinberg genotype equilibrium expectations. Thus, homozygotes are seen as single bands and heterozygotes as multiple bands (e.g., two, three or five bands for monomeric, dimeric, or tetrameric loci, respectively). (2) Considerable divergence in allozyme (allele) frequencies at assayed loci must have occurred in order to detect genetic differences between populations by a statistical test. The genes that code for the proteins assayed by this technique typically do not mutate rapidly, not all genetic mutations result in a change in protein structure, and not all differences in protein structure result in differences in electrophoretic mobility. Therefore, some significant differences in allele frequencies that occur between populations may not be detected. Consequently, some subdivisions in population structure that exist may not be detected. (3) Fresh-frozen tissues are required for the analysis. The logistics of obtaining those tissues on some field trips can be formidable. (4) Typically the proteins with the most variation (alleles) are those found in the internal organs (e.g., liver, gonads). Thus, in most cases, the individuals must be sacrificed to obtain the tissue samples. (5) The bands on a gel require interpretation. The banding patterns for some loci can be complicated. They can be blurred or otherwise different from the banding patterns that are expected (e.g., "null alleles", alleles that are not seen as bands, can be present). In addition, a multitude of nongenetic causes (e.g., post-translational modification, unidentified protocol errors) can complicate or obscure the banding patterns or make them ambiguous.

Despite these problems, it is the consistently high level of conformation of most allozyme "loci" to Hardy-Weinberg genotype expectations and

straightforward interpretation of the banding patterns that has given researchers faith that the banding patterns that they are viewing on a gel are in fact the direct products of codominant alleles belonging to single genes. Allozymes can provide information on the levels and partitioning of genetic diversity that is valuable for resource conservation, particularly at the level of distinguishing species and detecting interspecific hybridization. Moreover, that information can be obtained within the time-frames imposed by rapid conservation assessments and the analysis can be performed even in laboratories with limited equipment, budgets, and personnel.

Alternative nuclear gene markers that are proving effective for discerning population structure are the nuclear DNA from introns (Palumbi & Baker, 1994; Friesen *et al.*, 1997; Chow & Takeyama, 2000; Huang & Bernardi, 2001), anonymous nuclear DNA (Karl, 1996; Bageley *et al.*, 1997), anonymous mtDNA markers (Seyoum *et al.*, in preparation), and microsatellite nuclear DNA (Bentzen *et al.*, 1996; McConnel *et al.*, 1995; O'Connell *et al.*, 1998). The genetic function of the anonymous markers is sometimes not known (i.e., they could be in coding or non-coding regions of DNA), but justifications for their status as non-coding are usually fairly convincing. Both introns and microsatellite DNA are thought to be in non-coding regions of nuclear DNA, and they both can be highly variable. Of all genetic markers, microsatellite DNA is generally considered to be the most sensitive (Goldstein & Pollack, 1997). Typically, the nucleotide sequence is obtained for introns and a length-based analysis is used for microsatellite DNA.

Alternatively, specific segments of nuclear DNA can be examined by the by the point-mutation-based RAPD analysis or by the conformation-based DGGE analysis. Both the RAPD and DGGE methods have limitations. The homology and repeatability of the banding patterns generated by RAPDs can be difficult to establish and many problems can arise in the interpretation and analysis of RAPD data. This technique is best suited for the analysis of pedigrees. For DGGE, if the technique has not been adequately refined for the specific species being analyzed, some single-nucleotide substitutions may not be detected. Thus genetic diversity would be underestimated.

The direct analysis of nuclear DNA is expensive and requires an extensive laboratory setup and, ideally, access to an automated DNA sequencer. In the case of microsatellite DNA analysis, an extensive, species-specific "library" of nuclear DNA fragments must be developed and assayed for microsatellites with sufficient levels of length variation before populations can be surveyed. However, all of the types of analysis that can be accomplished using allozyme loci can also be done using microsatellite DNA, and microsatellite DNA has a number of advantages over allozyme loci. Most importantly, it is the direct analysis of DNA rather than DNA products. In addition, presumably, microsatellites are neither used in nor related to the DNA coding process (but see, e.g., Streelman *et al.*, 1998) and therefore they are presumed to be free from the influences of selection. Similarly to allozyme loci, population-level microsatellite DNA genotype frequencies generally conform to Hardy-Weinberg expectations and thus, they can be used to assess genetic variation, population structure and subdivision, gene flow, and interspecific hybridization. They can serve as markers to detect selection against particular populational components (e.g., hybrid individuals), and several hypervariable loci can be used together to determine pedigrees.

The mitochondrial genome has been used extensively in conservation studies based on population genetics and systematics. MtDNA haplotype data can be used to determine genetic variation, population subdivision, gene flow, and phylogenetic relationships. From these data, information on the history of species or populations over ecological or evolutionary time periods can be inferred, as well as breeding structure and movement patterns.

Some of the properties of mtDNA that make it especially useful for genetic analyses are that it is non-recombinant (Hayashi *et al.*, 1985), has a faster rate of evolution than nuclear DNA (Brown *et al.*, 1982), and is easy to isolate and manipulate (Dowling *et al.*, 1990). Because it is effectively haploid and usually only maternally inherited (Wilson *et al.*, 1985; Meyer, 1993), the effective population size is reduced and, thereby, sensitivity to genetic drift is increased (Birky *et al.*, 1989). Since in most taxa the mitochondrial data obtained pertain to maternal lineages, both maternal and



biparental components of gene flow can be distinguished when data from an mtDNA analysis is used in combination with data from a nuclear DNA analysis, allowing for the identification of sex-specific migration behaviors (e.g., Bowen *et al.*, 1992; Karl *et al.*, 1992; Palumbi & Baker, 1994).

Another advantage to using mtDNA is that the genome is well characterized for fish and other aquatic and marine organisms, and primer sequences for the PCR amplification (a technique that uses the polymerase chain reaction [Innis *et al.*, 1990] to make many copies of the specific DNA segment to be analyzed) and analysis of many of the mitochondrial genes are readily available. A limitation of the analysis of mtDNA is that the entire mitochondrial genome is considered to be a single locus, whereas nuclear DNA can provide multiple loci.

For population-level analyses, a portion of the D-loop (control region), a rapidly mutating and therefore usually the most variable (Brown, 1985) portion of the mtDNA genome, can be evaluated by sequencing or some type of sensitive analysis capable of detecting single nucleotide differences between individuals. The D-loop has been extensively used to study equilibrium between gene flow and genetic drift (Avise, 1994), historical events such as population bottlenecks (Toline & Baker, 1995), and present day population structure (e.g., Stepien, 1995; Seyoum *et al.*, 2000). For analysis of higher level taxa, portions of more conserved mitochondrial genes (e.g., those coding for cytochrome oxidase, NADH dehydrogenase, or ribosomal RNA) should be chosen. Mitochondrial DNA can be analyzed by sequencing, RFLP, or DGGE.

### **Protocols**

The first analyses conducted should be on the samples for which population-level numbers have been collected. For allozyme and microsatellite analyses, as many loci as can be resolved should be assayed in as many individuals as available, to enhance the probability of detecting rare alleles and to provide adequate statistical power for analyses of genetic population structure and comparisons of measures of genetic variability for known species. The second priority should be to clarify the taxonomic status of cryptic or confusing morphological taxon groups and verify any suspected hybridization. Third, as the number

of collections from different regions accumulates, any combination of samples can be processed in the laboratory for phylogenetic analyses. However, evaluations based on phylogenetic inferences are incomplete until samples from a large number of species within the phylogenetic group(s) of interest and species that can serve as appropriate outgroups have been collected. The ability to assess the value of a region from both population genetics and evolutionary perspectives will increase as the number of collections from different areas within the region increases, thus allowing the cumulative data bases to be integrated or compared.

Using allozyme electrophoresis and having unrestricted access to a fully equipped laboratory, a single researcher can develop the protocols and assay approximately 1,000 individuals in six months; another 1-2 months would be needed for data entry and analysis. For DNA analysis, access to a PCR machine is necessary because nearly all techniques require amplification of one or more segments of DNA. To establish the protocols for most DNA analyses, a researcher typically needs approximately 2-6 months. The amount of time needed depends on the techniques chosen and the amount of information on directly applicable techniques and (for PCR) primers that has been previously published or is otherwise available. After the specific analytical procedures have been developed, an additional 4-6 months is usually needed to manually sequence (without an automated sequencer) an approximately 400-base-pair region of mtDNA for 200 individuals, and to analyze the data. For ease in reading sequences generated by hand, use a "block sequencing" method (on a given gel for a given species, run all adenosines together, all cytosines together etc.). Sequencing DNA with an automated sequencer is easier, faster, and usually less expensive. In our laboratory, a researcher can process DNA segments of approximately 450 nucleotides from raw tissue through editing the DNA sequence for approximately 400 individuals per month using an ABI Prism 310 Genetic Analyzer<sup>®</sup>. Thus, with the advantage provided by even a small, tabletop automated DNA sequencer, more than 2,000 individuals could be analyzed for a DNA segment of the length typically used for conservation genetics in a six-month time period.

Analysis of microsatellites could take substantially longer because of the extensive time

required to develop the DNA fragment library mentioned above. Developing a battery of 5-15 microsatellite loci to use for a single taxon or closely related taxa can take up to 6 months. Fortunately, for many aquatic macroorganisms, microsatellite loci have been developed and the time needed to develop a taxon-specific protocol is rapidly diminishing because, frequently, published protocols can be adapted for the taxa to be analyzed. If an automated sequencer is used, the sequencing can proceed very rapidly. In our laboratory, a single researcher can analyze approximately 100 individuals for 10 loci per month, without multiplexing loci (i.e. running several loci simultaneously in the DNA sequencer). By multiplexing loci, the number of individuals analyzed per month can be doubled, tripled, or multiplied by an even greater amount. Thus, after the laboratory procedures are established, the entire laboratory analysis of a few hundred individuals for numerous microsatellite DNA loci could take only a few months.

For a single rapid-assessment, conservation-oriented expedition, the sample size for genetic assessment probably will not exceed 400-600 individuals. (Various field-associated limitations usually determine the maximum number of individuals that can be sampled.) If all individuals are used for allozyme analysis, and half, or less, are used also for more sophisticated DNA techniques, one researcher devoted to this analysis could complete the assessment within six to twelve months. Obviously, any type of analysis other than allozyme electrophoresis can be accomplished far more rapidly through the use of a DNA sequencing machine. More individuals could be surveyed, and both mtDNA and nuclear genes could be assayed.

Laboratory procedures for conducting the laboratory analyses described here can be found in any publication describing research in which these techniques were used and in the references within those publications. Standard procedures for setting up a laboratory or conducting allozyme electrophoresis are described in Shaw & Prasad (1970), Selander *et al.* (1971), Harris & Hopkinson (1976), Richardson *et al.* (1986), Murphy *et al.* (1990), Hillis *et al.* (1996). Standard procedures for setting up a laboratory or conducting DNA RFLP analysis or DNA sequencing are described in Sambrook *et al.* (1989), Erlich (1992), and Hillis *et al.* (1996). PCR

techniques, procedures and applications are described in Innis *et al.* (1990) and Erlich (1992). Techniques for conducting DGGE and RAPD analysis can be found in Myers *et al.* (1986) and Williams *et al.* (1990) respectively, and references therein, and by searching published papers in which those techniques are described. However, the most efficient way to learn any of these techniques is to serve as an apprentice in an established laboratory.

## STATISTICAL METHODS

Although traditionally, phylogenetic methods have been used to determine relationships among species and higher taxonomic categories, these methods have been commonly applied to populations within species for all types of studies in which the relationships and interactions among populations and identification of lineages undergoing rapid speciation and ancestral lineages is needed. A common method for identifying cryptic species is through phylogenetic analysis. Thus, species richness can be increased above that detected by morphological identification or other means. Alternatively, phenotypically distinct individuals or groups that are genetically similar or identical can be identified. Morphologically differentiated individuals or groups may belong to single, highly polymorphic species with habitat-specific phenotypic variation. In this case, the preservation of those habitats may be important for the maintenance of the gene pool of the polymorphic species. In these ways, phylogenetic analysis can directly link genetic diversity to both species diversity and ecosystem diversity.

Population genetics analyses can also be integrated with biodiversity at the species and ecosystem levels. These types of analyses provide information on the interactions among populations and allow identification of populations that harbor unique genetic variation. Estimates of levels of interbreeding between populations and movement of individuals among populations can be obtained, as well as information on breeding structure and differences between sexes in breeding strategies. If sufficient genetic variation is present and the DNA techniques can be established, even parentage can be determined in some cases. Thus, species that have highly subdivided populations can be identified and critical populations preserved.

Habitats that are used as breeding grounds or corridors for the exchange of individuals among populations can be protected.

At least preliminary population-level assessments can be made for many species for which there are multiple, temporally or spatially separated samples of five individuals or more. Of course, the larger the sample sizes, the greater the level of statistical resolution and therefore the greater the ability to detect population subdivision when it is present. The statistics performed to assess population structure and genetic variability can also elucidate cryptic species or hybridization. Those computer programs with components that evaluate the relationships of population samples or species samples to each other can be used not only to understand the relative connection or isolation among populations but also for phylogenetic analysis.

To integrate the genetics work with other components when determining the conservation significance of a region, the known geographic ranges of the species collected should be considered in all phases of the genetics studies, and marginal areas of species ranges should be sampled if possible. Species experiencing declines in numbers may still be distributed throughout their former ranges, including peripheral areas (Channell & Lomolino, 2000). Compared to the usually much larger core population more centrally located within the species range, these areas can harbor populations with unique allele frequencies at loci undergoing selection for the species' marginal habitat conditions. These peripheral populations are adapted for the extreme environmental conditions within the species' tolerance ranges and may be important from a conservation perspective (Lesica & Allendorf, 1995). The loss of locally adapted populations within species and of their genetic material reduces the resilience of species to environmental change.

For both phylogenetic and population genetic applications, any of the genetic markers described above can be used. In the past, we have used allozymes and mtDNA (see, e.g., Bert, 1986; Tringali & Bert, 1996; Seyoum *et al.*, 2000; McMillen-Jackson *et al.*, in review). Microsatellite loci are typically too variable for use in interspecific phylogenetic analysis other than that of closely related species, but they are highly utilitarian for intraspecific analyses such as determining genetic

relationships among populations, conducting pedigree analysis, and tracking specific genetically unique individuals or populations.

For data obtained from nearly all types of genetic analysis, large arrays of statistical methods are available for determining population-level genetic variation, intraspecific population genetic structure, effective population size, gene flow between populations, relative age of lineages, evolutionary relationships among species, and a number of other populational and species characteristics. Many of the commonly used statistics for the types of data that are derived from allozyme or microsatellite analysis, restriction fragment analysis, or DNA sequencing are described in Table 2. In addition, Appendix 1 provides, for a variety of these data types, data management programs and statistical programs that each include some or most of these statistics, the general types of applications of these programs, citations for the sources of these programs (when known), and their procurement sources (usually websites). Nearly all of these programs are free. The website listings and email addresses to obtain these programs are current as of the date of publication of this article. Websites that are good sources of phylogenetic and population genetic analytical programs are also listed in Appendix 1.

For DNA nucleotide sequence data, computer software packages have been developed to provide information analogous to that used for allozymes. Many of the statistics used for allozymes are also used for microsatellite DNA data. Similarly, a number of computer software programs have been developed to analyze RFLP data. A general description of the types of analyses that we typically perform and the programs that we use follow. Both Arlequin and MEGA can be used to generate nucleotide base composition, counts of transitions and transversions, number of base substitutions, haplotype frequencies, and genetic distances between haplotypes and between populations using various methods. For phenetic and phylogenetic analyses, MEGA has more clustering methods than does Arlequin, which has only the minimum spanning tree method. The programs ESEE or CLUSTAL can be used to align DNA sequences; ESEE is limited to manual alignment of sequences in pairs whereas CLUSTAL is an auto-alignment program that can align multi-

ple sequences simultaneously. Methods for calculating haplotype diversity and nucleotide sequence diversity (measures of within-population genetic variability) are available in Arlequin, DnaSP (this program considers only two populations at a time), MEGA, and REAP. Phenetic cluster analyses can be performed using MEGA, NJTREE, PAUP, or PHYLIP. Cladistic grouping methods (which are phylogenetically based) are available in PHYLIP and PAUP. Standard errors for the branch lengths of UPGMA phenograms based on haplotypes can be calculated using NJTREE; thus their statistical significance can be assessed. MEGA, PAUP, and PHYLIP estimate statistical significance of clustering methods using bootstrap and other types of analyses such as jackknifing, permutation tests, or interior branch tests.

In general, the geographic distribution of, and significant differences between, haplotypes in a tree are often used as evidence of population structure. Distances between populations can be used in cluster analyses, but to our knowledge, only the NJTREE program assesses the statistical significance of a population-based tree, and that is only for the UPGMA tree. Geographic partitioning of haplotype distributions can be identified using  $\chi^2$ , log-likelihood tests ( $G$ -tests), or the exact test of population differentiation (in Arlequin). To minimize the effect of large numbers of empty cells typical of population-level sequence data, Monte Carlo randomization tests should be used (e.g., the MONTE program in REAP). The  $V$ -test may be used to examine geographic partitioning in the percentage occurrence of specific haplotypes. The partitioning of molecular variance ( $\phi$ ) within and between user-defined hierarchical levels can also be examined using the AMOVA procedure in the AMOVA and Arlequin programs. The  $\phi$  statistics (analogous to  $F$ -statistics) incorporate pairwise genetic distances between haplotypes. The significance of  $\phi$  can be determined by using the randomization program in AMOVA or Arlequin. The effective number of migrants per generation (overall gene flow)  $w$  can be calculated from the  $\phi$  statistics by using the classical  $F_{ST}$  equation. If mtDNA is used,  $N_{efm}$  (female gene flow) can be estimated. To assess the relationship between genetic distance and geographic distance, use the Mantel test in Arlequin.

The analytical procedures for RFLP data are similar to those for sequence data; however,

different computer programs are sometimes used. Within-population haplotype diversity and nucleotide diversity and between-haplotype and between-population divergences can be estimated using REAP or RESTSITE. Phenetic cluster analysis of the pairwise matrices of haplotype divergences or interpopulational sequence divergences can be performed using PHYLIP, MEGA, or NJTREE. Testing for geographic variation in haplotype frequencies and the hierarchical analysis of molecular variance can be accomplished using the methods and programs described for analysis of sequence data.  $F$ -statistics can also be calculated for RFLP data (e.g., see Gold *et al.*, 1993).

All analyses that are performed as multiple tests of a single hypothesis should be corrected for that problem by appropriately adjusting significance levels. We use the sequential Bonferroni correction.

## DISCUSSION

Estuaries, swamps and floodplains, seagrass/algae flats, and wetlands are by far the most valuable environments in terms of natural capital and provision of ecosystem services (Costanza *et al.*, 1997). In the freshwater components of these aquatic environments and riverine systems, the loss of biodiversity is far worse than in forests, grasslands, or coastal ecosystems (Johnson *et al.*, 2001). Genetic diversity necessarily accompanies species diversity. Thus, when species diversity is lost in these habitats, genetic diversity is lost also. However, as high as they are, the losses of species understate the magnitude of the loss of genetic diversity (Vitousek *et al.*, 1997). Moreover, losses of genetic diversity can be expected to continually increase because the normal increasingly wide range of conditions that ecosystems experience over time (Chapin *et al.*, 1997) is presently exacerbated by anthropogenically related land usage. Humans have the highest level of population growth in regions of exceptionally high biodiversity (biodiversity hotspots, as measured principally by species counts; Myers *et al.*, 2000), and a disproportionately high percentage of the world's human population lives in these biodiversity hotspots (Cincotta *et al.*, 2000). Particularly in developing countries, in regions where the roads are the rivers, streams, or coastal waters, the impact

of this increasing population is most pronounced in the aquatic environment. Unless drastic measures are enacted, the rate of environmental change will continue to accelerate as the human population increases. Unless environmental usage patterns change, we can expect unprecedented and continually accelerating losses in the world's "seedbank" of aquatic genetic diversity.

The maintenance of genetic diversity is essential to facilitate species' adaptations to this environmental change in order to increase the probability of long-term sustainability of ecosystem structure. It has long been well established, principally through studies of the effects of inbreeding and small population size on fitness components on commercially valuable aquatic species, that decreases in population genetic diversity can render the population less genetically fit in many ways, including those of importance for survival of the population over ecological or evolutionary time (e.g., reproductive capability and survival rate). Thus, preserving the genetic diversity of species can enhance species' ability to adapt to new environmental conditions and thus influence their survival over these time frames (Levin, 1995).

Yet only a few biodiversity assessments have included even rudimentary studies of genetic diversity (e.g., Bert, 1999). Nearly all biodiversity studies have been limited to species counts (Knowlton, 2001) because morphologically differentiated forms (which usually constitute different species) are easily recognized and the evolutionary significance of species is understood by a broad spectrum of humans (Gaston, 2000). On a global basis, the numerous biodiversity and taxonomic studies that have been conducted in a wide variety of ecosystems form a strong database for comparing new species counts with those previously published over the past century. This substantial database provides a framework of reference for studies of species diversity and enables global comparisons to be made.

For genetic diversity studies based on molecular population genetic and phylogenetic methods, the database itself is sparser and more diverse because molecular population genetics and evolutionary genetics are new fields compared to traditional, morphologically based taxonomy and because multiple genetic techniques are available to perform these studies. Nevertheless, there is a

framework for understanding genetic variation and genetic population structure in aquatic systems. Numerous allozyme studies have been conducted in all types of aquatic environments; the database of genetic diversity using mtDNA RFLP data is large and still growing; microsatellite DNA studies are becoming the standard for investigating population genetics using nuclear DNA; and the database for DNA sequence data is expanding exponentially. From a global conservation perspective, some of these databases are not standardized (e.g., allozymes), but nevertheless, they are of value for comparisons, particularly when spatial or temporal evaluations of intraspecific genetic diversity among populations is needed.

Because regions of overall importance genetically are not necessarily those regions of highest species diversity and because genetically based assessments yield estimates of biodiversity with components that are independent of the estimates determined using species counts or species/area relationships, areas for which genetic diversity assessments have been made may be designated as worthy of conservation efforts regardless of their absolute species diversity. Genetic assessments can be used to evaluate the potential contribution of these areas to the maintenance of regional, continental, or worldwide gene pools. Areas in which many of the resident species harbor high levels of genetic variation are repositories of regional genetic diversity essential for adaptation to global change. Areas with relatively high percentages of ancestral lineages or relic species are strongholds of ancestral genes for evolving lineages. Areas with high concentrations of recently evolved species or hybridizing species may represent regions where evolution is occurring particularly rapidly or uniquely and where the potential for increasing overall levels of genetic diversity is high.

Because genetic variation is the raw material of natural selection (Fisher, 1930), biodiversity ultimately *is* genetic diversity (Avice, 1996). Populations are natural entities that should be preserved for their own sakes and because they are essential components of ecosystems that provide goods, basic life-support systems, and enjoyment for humans; their long-term production and usefulness ultimately depend on preserving genetic variation both among and within species (Burger & Lynch, 1995; Currens & Busack, 1995; Lande & Shannon, 1996).

TABLE 2

Utility of some population genetic and evolutionary genetic statistics for studies of the conservation of genetic diversity. For further explanation of these measures and additional measures, see a basic population genetics or systematics textbook (e.g., Hartl & Clark, 1997; Hillis *et al.*, 1996), biological statistics book (e.g., Sokal & Rohlf, 1995), comprehensive conservation genetics book (e.g., Smith & Wayne, 1996), or the explanations and references in the computer programs listed in Appendix 1. Numbers in the computer programs columns refer to the programs described in Appendix 1. Randomly amplified polymorphic DNA (RAPD) data has not been included in the "Applicable data type" category because that data must be converted before using in any of the statistical programs listed here (see Appendix 1, number 16 for further explanation). RFLP = restriction fragment length polymorphisms.

Category	Use	Statistics	Applicable data type	Computer programs
<b>Tests of genetic diversity</b>				
Basic measures of genetic variation	Comparisons among populations in levels of genetic variation (peripheral and founder populations can have decreases in these measures compared to other populations and populations with exceptionally high levels of genetic variation are valuable sources of genetic variation for the species)	$P$ (percentage of polymorphic loci), $H$ (average percentage of individuals heterozygous per locus), $n_a$ (average number of alleles per locus)  $h$ (haplotype diversity), $\pi$ (nucleotide diversity)	Allozyme, microsatellite DNA  DNA sequence, RFLP	2, 8  2, 6, 12, 17, 18
Allele/genotype/haplotype frequencies	Comparisons among populations (localities, regions) (significant differences indicate population structuring); identification of peripheral or founder populations (these may have significantly different allele or haplotype frequencies or may be missing rare alleles or haplotypes compared to mainstream populations); investigations of clinal variation (i.e., geographically structured changes in allele frequencies, suggesting some type of selection, restricted gene flow, or hybridization; Endler, 1977)	Percentages	Allozyme, microsatellite DNA, DNA sequence, RFLP	2, 8, 17
Conformation to Hardy-Weinberg genotype expectations	Indicators of straightforward genetic basis (single gene, codominant alleles, devoid of selection); nonconformance is an indicator of selection (e.g., at a single locus) or hybridization (e.g., at multiple loci) that is accompanied by selection (usually against hybrids, as indicated by a deficit of heterozygotes) or nonrandom breeding.	$\chi^2$ , $G$ -test (Sokal & Rohlf, 1995), $V$ -test (DeSalle <i>et al.</i> , 1987)	Allozyme, DNA sequence (nuclear DNA only)	2, 8
<b>Tests of relationships among individuals, collections, or taxa</b>				
Measurement of hierarchical levels of genetic variation	Estimation of genetic similarity among individuals within populations, among individuals within species, among populations within species, or among species; location of source of genetic differences among groups	$F$ -statistics and analogues (see Gold <i>et al.</i> , 1993)  AMOVA	Allozyme, DNA sequence, microsatellite DNA, RFLP  Allozyme, microsatellite DNA, DNA sequence, RFLP	1, 2, 8, 19  1, 2

TABLE 2 (Continued)

Measurements of genetic distance	Provide information on evolutionary relationships among lineages, populations, or taxonomic groups; essential calculation for analysis of phylogenetic relationships.	Inter-individual, inter-population, and inter-haplotype distances (many measures [e.g., Nei's $D$ , $P$ , $F_{ST}$ , $R_{ST}$ ]; see Hillis <i>et al.</i> , 1996)	Allozyme, microsatellite DNA, DNA sequence, RFLP	2, 3, 5, 6, 8, 12, 14, 15, 17, 18
Phylogenetic cluster analysis	Generation of "trees" or networks that illustrate an estimation of evolutionary relationships among genotypes or haplotypes, populations, or species; identification of ancient or recently derived lineages, or lineages undergoing rapid speciation	UPGMA, neighbor-joining, parsimony, maximum likelihood, minimum spanning (all are algorithms); minimum evolution	Allozyme, microsatellite DNA, DNA sequence, RFLP	2, 5, 12, 13, 14, 15, 18
Gene flow analyses	Estimation of effective number of migrants among populations per generation ( $N_m$ ), can be subsetted by sex ( $N_{em}$ ); determination of degree of isolation of populations	Private alleles method  $F_{ST}$ method (Slatkin, 1985)  Likelihood analysis (Beerli & Felsenstein, 2001)	Allozyme, microsatellite DNA, DNA sequence, RFLP Allozyme, microsatellite DNA, DNA sequence, RFLP DNA sequence	Slatkin, 1985  2, 8, 19  10
Independence of location and allele frequencies	Detailed evaluation of population genetic structure, performed locus by locus; supportive evidence for genetic isolation among populations or for selection	Exact test, $G$ -test (Sokal & Rohlf, 1995), $V$ -test (used when numerous cells in allele frequency matrix equal zero; DeSalle <i>et al.</i> , 1987)	Allozyme, microsatellite DNA, DNA sequence, RFLP	2, 8 ( $R \times C$ exact test only)
<b>Supportive Statistics</b>				
Selective neutrality	Determination of presence or absence of selection operating on genetic markers used for other analyses.	$D$ (Tajima, 1989; mtDNA only), $F_S$ (Fu, 1997), and others	DNA sequence	2, 6, 12
Statistical reliability	Determination of statistical confidence in a result, e.g., a grouping, tree, or network derived from a cluster analysis.	Bootstrap, confidence interval, Monte Carlo (minimizes the effect of a large number of empty cells in a data matrix), randomization, jackknife	Allozyme, microsatellite DNA, DNA sequence, RFLP	2, 12, 14, 15, 17
Linkage disequilibrium (Weir, 1996)	Determination that two distinct gene pools are mistakenly considered as a single gene pool; identification of the existence of hybridization.	Single locus, multiple loci	Allozyme, microsatellite DNA, DNA sequence, RFLP	2, 6, 8
Physical factor-population genetic relationships	Detection of geographically arrayed selection (e.g., allele frequencies vs. longitude or latitude, from small streams to large rivers, or over salinity or temperature gradients)	Mantel test, correlation analysis, isolation by distance (Slatkin, 1993)	Allozyme, microsatellite DNA, DNA sequence, RFLP	2 (Mantel test only), BIOM* (or, for correlations, any good standard statistics program)
<b>Post-hoc tests</b>				
Grouping tests	Location of significant differences in allele frequencies between populations	Simultaneous Test Procedure (STP)	Allozyme, microsatellite DNA, DNA sequence, RFLP	BIOM*
Repeated tests statistical correction	Correction for multiple tests of a single hypothesis (it is necessary tests to clearly formulate the hypothesis to determine the exact level of correction needed)	Sequential Bonferroni correction	Allozyme, microsatellite DNA, DNA sequence, RFLP	Rice, 1989

\*BIOMstat: Statistical Software for Biologists, version 3.3; available from Exeter Software ([www.exetersoftware.com/cat/biomstat.html](http://www.exetersoftware.com/cat/biomstat.html)) for \$300 US (\$230 US for educational and governmental institutions).

Therefore, it is essential to consider investigations of genetic diversity as integral components of studies for the conservation of biodiversity, particularly in the highly threatened aquatic environment.

## APPENDIX 1

Some population genetic and phylogenetic analytical packages useful for assessing the genetic component of biodiversity for conservation assessments. Both analytical and data management programs are listed. The numbers assigned to the programs are cross-referenced in Table 2, which describes the statistics that can be found in those programs. The statistics that are generally used for the analysis of biological data can be found in standard biological statistics programs (e.g., BIOM [StatSoft, Inc., 1999], SAS<sup>®</sup> [available from SAS Institute, Inc., Cary, North Carolina; <http://www.sas.com>], Statistica<sup>®</sup> [StatSoft, Inc., 1999]; all of these programs must be purchased). Programs that we have used are noted by asterisks. Further information on most programs can be found on their websites. Additional websites with extensive lists of various kind of relevant software are <ftp://ftp.ebi.ac.uk/pub/software/dos/> and <http://genamics.com/software/index.htm>. A web service that links to biodiversity, evolution, systematics, and several free software packages is <http://darwin.eeb.uconn.edu/molecular-evolution.html#software>. A comprehensive list of phylogenetic software is in Hillis *et al.* (1996) and in <http://evolution.genetics.washington.edu/phylip/software.html>.

### 1. \*AMOVA (Excoffier *et al.*, 1992): <http://lgb.unige.ch/software/win/amova/>

The Analysis of Molecular Variance program performs an analysis of population genetic structure at the molecular level. It can be used to partition molecular variance at different, user-defined, hierarchical levels. AMOVA is similar to other hierarchical procedures for compartmentalizing variance among groups in gene frequencies (e.g., *F*-statistics) except that it incorporates inter-haplotype nucleotide substitution data. Tests for statistical differences among frequencies are conducted within the program by non-parametric permutation procedures.

*Note added in press:* AMOVA is no longer supported and there will be no future releases of

this program. It is currently replaced by the more powerful and versatile software program Arlequin.

### 2. \*Arlequin (Schneider *et al.*, 2000): <http://lgb.unige.ch/arlequin/>

Arlequin is a multifaceted population genetics software package that may be used for the analysis of population genetic structure at the molecular level. Data can be in the form of allelic or genotypic frequencies (e.g., allozymes or microsatellites), or haplotypes (e.g., RFLPs or DNA sequences). Arlequin is one of the few programs that analyses population demographic features and computes parameters and confidence intervals under an expansion model. The program includes routines for the following analyses: AMOVA; haplotype, allele, and population frequency testing; linkage disequilibrium; Hardy-Weinberg genotype frequency equilibrium; selective neutrality; population assignment for genotypes, and the Mantel test for the correlation of matrices. An important advantage of this program is its capacity to accommodate large data sets.

### 3. BioEdit (Hall, 1999): <http://www.mbio.ncsu.edu/RNaseP/info/programs/BIOEDIT/bioedit.html>

The Biological Editor is a general-purpose program for manual pair-wise alignment of sequences or multiple automated alignment of up to 50 sequences with up to 20,000 nucleotides. It provides a large number of sequence manipulation and editing features and several fully automated links to local and WWW based sites and softwares. It also has several accessory applications including maximum likelihood, Fitch-Margoliash, and least-squares distance methods with or without evolutionary clock assumptions. It provides an integrated working environment to view, align, and analyze with simple point-and-click operations.

### 4. \*CLUSTAL Thompson *et al.* (1994; DOS version): <ftp://ftp-igbmc.u-strasbg.fr/pub/clustalW/>; Thompson *et al.* (1997) and Jeanmougin *et al.* (1998) (Windows version): <ftp://ftp-igbmc.u-strasbg.fr/pub/clustalX/>

This program is a general-purpose, multiple-alignment program for DNA or protein sequence. It has a window interface that allows the researcher to see the sequences before and after alignment. The output file from this program can be converted



into several different formats using an internal FORCON program.

**5. DAMBE (Xia, 2000): <http://web.hku.hk/~xxia/software/software.htm>**

The Data Analysis in Molecular Biology and Evolution program is an integrated software package for retrieving, converting, manipulating, aligning, and analyzing molecular sequence data. For nucleotides and amino acids, the program also performs phylogenetic reconstruction using distance, maximum parsimony, and maximum likelihood methods. One important feature of the program is that it can reconstruct ancestral states using either maximum parsimony or maximum likelihood methods to which discrete- or gamma-distributed substitution models can be fit.

**6. \*DnaSP (Rozas & Rozas, 1999): <http://www.bio.ub.es/~julio/DnaSP.html>**

The DNA Sequence Polymorphism program is an easy-to-use software package for the analysis of DNA variation from nucleotide sequence data. With this program, one can conduct analyses of nucleotide variation within and between populations for noncoding, synonymous, or nonsynonymous sites, linkage disequilibrium, recombination, gene flow, gene conversion, and selective neutrality (e.g., Fu and Li's; Hudson, Kreitman and Aguadé's; McDonald and Kreitman's; and Tajima's tests). DnaSP can also be used to conduct population dynamic analyses based on mismatch distributions under constant or population expansion models.

**7. \*ESEE (Cabot & Beckenbach, 1989): <ftp://ftp.ebi.ac.uk/pub/software/dos>**

The Eyeball Sequence Editor (ESEE) is a manual, multi-sequence, editing program for nucleotide and amino acid sequences. It is a convenient alignment program that can also be used to estimate percentage of matches between two nucleotide or amino acid sequences. Several other software programs are also available through this site.

**8. GENEPOP (Raymond & Rousset, 1995)**

This Population Genetics program includes all of the standard tests for performing basic, population genetics analysis on nuclear genes and nuclear gene products (e.g., allozymes). It also has many options available for determining genetic

differentiation among populations, linkage disequilibrium, and the partitioning of genetic variation among populations.

**9. \*FORCON (written by J. Raes and Y. Van de Peer, Department of Biochemistry, University of Antwerp (UIA), Universiteitsplein 1, B-2610 Antwerpen, Belgium): [http://www2.no.embnet.org/embnet.news/vol6\\_1/ForCon/](http://www2.no.embnet.org/embnet.news/vol6_1/ForCon/)**

The Format Conversion program can be used to automatically convert nucleotide or amino acid sequence "input files" from one format to another for use in many popular software packages. It allows one to avoid the tedious reformatting of data sets via manual editing.

**10. \*LAMARC (Beerli & Felsenstein, 2001): <http://evolution.genetics.washington.edu/lamarc.html>**

The Likelihood Analysis with Metropolis Algorithm using Random Coalescence software is a package composed of four programs that can be used to estimate population genetic parameters such as effective population size, population growth rate, and migration rate. It is executable in the Windows NT/95 and PowerMac systems.

**11. \*MacClade (Maddison & Maddison, 1992): <http://phylogeny.arizona.edu/macclade/macclade.html>**

The Macintosh Cladistic program facilitates the examination of character evolution and manipulation of phylogenetic trees constructed using other software (e.g., PAUP). It is available for purchase only for Macintosh computers.

**12. \*MEGA (Kumar *et al.*, 2001): <http://www.megasoftware.net/>**

The Molecular Evolutionary Genetics Analysis program is very useful for the analysis of sequence data. It can be used to estimate evolutionary distances under various substitution models, reconstruct phylogenetic trees (including minimum evolution and maximum parsimony trees), compute statistical quantities such as bootstrap values and interior branch lengths for trees, and conduct tests for selection. It can also be used to compute nucleotide divergences between populations for input into other tree-making programs. This program can accommodate large sample sizes.

**13. NJTREE (Jin & Ferguson, 1990)**

This Neighbor-joining Tree and UPGMA tree program will generate these two types of phenetic trees from populational genetic distance matrices. Importantly, standard errors are estimated for the UPGMA tree branch nodes. Within the program, TDRAW produces the very basic dendrograms (trees).

**14. \*PAUP (Swofford, 1998): <http://paup.csit.fsu.edu/>**

The Phylogenetic Analysis Using Parsimony program has been released in Macintosh, PowerMac, Windows, and Unix/nVMS versions. It is being continually edited and updated (the latest version is 4.0 beta). This comprehensive maximum parsimony program is compatible with MacClade. It also includes non-cladistic methods for determining relationships among species or populations, such as minimum evolution, the method of least squares, Lake's method of invariants, and maximum likelihood. It also performs phenetic analyses such as generating neighbor-joining and UPGMA trees on distance matrices and calculates many other indices and statistics.

**15. \*PHYLIP (Felsenstein, 1993): <http://evolution.genetics.washington.edu/phylip.html>**

The Phylogeny Inference Package is widely used program for inferring phylogenies. It can analyze data in the form of molecular sequences, gene frequencies, restriction sites, distance matrices, and discrete characters. Its methodologies include those for phenetics, clustering, parsimony, distance matrices, and maximum likelihood. Several methods for determining the confidence in these trees or networks are also included.

**16. RAPDistance (Armstrong *et al.*, 1994): <http://life.anu.edu.au/molecular/softward/RAPDistance/>**

The Randomly Amplified Polymorphic DNA Distance program is principally an editing program designed to record and facilitate the analysis of RAPD data (and also RFLP data, if some care is taken). The primary data is encoded for the presence or absence of shared bands. Several options exist for editing the resulting file. The data may then be used in other programs to calculate pairwise distances between the samples, build phenograms, or perform multivariate analysis. The

analytical programs with which the edited data will be compatible are listed in the program.

**17. \*REAP (McElroy *et al.*, 1992): <http://bioweb.wku.edu/faculty/mcelroy>**

The Restriction Enzyme Analysis Package is designed to facilitate manipulation and phylogenetic analysis of restriction fragment or restriction site data. REAP can be used to create discrete character data sets for phenetic or cladistic analysis and to estimate distances among OTUs from fragment, restriction site, or sequence data. It is especially good for estimating nucleotide divergence within and among populations and evaluating the level of heterogeneity in population frequency distributions by Monte Carlo simulation. This option of REAP has now been expanded to include up to 50 populations having up to 200 haplotypes. The expanded version is available by request from D. McElroy.

**18. \*RESTSITE (Nei & Miller, 1990): <http://www.genome.wi.mit.edu/~jmiller/restsite/progs/>**

The Restriction Site program is useful for conducting phenetic analyses based specifically on RFLP (restriction site and restriction fragment) data. Accounting for the sizes of recognition sequences of restriction enzymes, this program can be used to estimate the average number of nucleotide substitutions within and between populations and to construct neighbor-joining and UPGMA trees from genetic distance matrices.

**19. RSTCALC (Goodman, 1997): <http://helios.bto.ed.ac.uk/evolgen>**

The  $R_{ST}$ -Calculation program is a PC-based software package useful for the analysis of microsatellite data. It generates estimates of  $R_{ST}$ , which is an analogue of  $F_{ST}$  that is unbiased with respect to differences in sample sizes and differences in variance between loci, to analyze population structure and genetic differentiation and to estimate gene flow. Statistical significance can be assessed using bootstrap and permutation tests.

*Acknowledgments* — The authors gratefully acknowledge B. Chernoff for providing the impetus to write this manuscript; N. Menezes for guiding us to the *Revista Brasileira de Biologia* and for reviewing this manuscript; B. Chernoff, A. Machado, H. Lopez-Rojas, and A. Bonilla for assistance with the development of the field collecting methods; B. Ballard for

a valuable review and the suggestion to include the sources for the statistical programs; and C. Crawford for additional references. The preparation of this manuscript was supported by funding from the State of Florida and the U.S. Department of the Interior, Fish and Wildlife Service, Federal Aid for Sport Fish Restoration Grant Number F-69 to T. M. Bert.

## REFERENCES

- ARMSTRONG, J. S., GIBBS, A. J., PEAKALL, R. & WEILLER, G., 1994, *The RAPDistance Package*. Research School of Biological Sciences, Institute of Advanced Studies, Canaberra, Australia.
- AVISE, J. C., 1994, *Molecular markers, natural history and evolution*. Chapman & Hall, London.
- AVISE, J. C., 1996, Introduction: the scope of conservation genetics, pp. 1-9. *In: J. C. Avise & J. L. Hamrick (eds.), Conservation genetics: case histories from nature*. Chapman & Hall, New York.
- AYALA, F. J., 1997, Molecular genetics and evolution, pp. 1-20. *In: F. J. Ayala (ed.), Molecular evolution*. Sinauer Associates, Inc. Sunderland, Massachusetts.
- BAGELEY, M. J., MEDRANO, J. F. & GALL, G. A. E., 1997, Polymorphic molecular markers from anonymous nuclear DNA for genetic analysis of populations. *Molec. Ecol.*, 6: 309-320.
- BARTON, N., 1985, Analysis of hybrid zones. *Ann. Rev. Ecol. Syst.*, 16: 113-148.
- BEERLI, P. & FELSENSTEIN, J., 2001, Maximum likelihood estimation of a migration matrix and effective population sizes in subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci.*, 98: 4563-4568.
- BENTZEN, P., TAGGART, C., RUZZANTE, D. & COOK, D., 1996, Microsatellite polymorphism and the population structure of Atlantic cod (*Gadus morhua*) in the Northwest Atlantic. *Can. J. Fish. Aquat. Sci.*, 53: 2706-2721.
- BERT, T. M., 1986, Speciation in western Atlantic stone crabs (genus *Menippe*): the role of geological processes and climatic events in the formation and distribution of species. *Mar. Biol.*, 93: 150-170.
- BERT, T. M., 1999, Landscape-based genetic assessment of the fishes of the Rios Tahuamanu and Manuripi Basins, Bolivia. *In: B. Chernoff & P. W. Willink (eds.), A Biological Assessment of the Aquatic Ecosystems of the Upper Rio Othon Basin, Pando, Bolivia. Bulletin of Biological Assessment (Conservation International)*, 15: 68-72.
- BERT, T. M., TRINGALI, M. D. & SEYOUM, S., 2002, Development and application of genetic tags for ecological aquaculture, pp. 47-76. *In: B. Costa-Pierce (ed.), Ecological aquaculture*. Blackwell Scientific, New York.
- BIRKY, C., FUERST, W. P. & MARUYAMA, T., 1989, Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, 121: 613-627.
- BOWEN, B. W., MEYLAN, A. B., ROSS, J., LIMPUS, C. J., BALAZS, G. & AVISE, J. C., 1992, Global population structure and natural history of the green turtle (*Chelonia mydas*) in term of matriarchal phylogeny. *Evolution*, 46: 865-881.
- BROWN, W. M., 1985, The mitochondrial genome of animals, pp. 95-130. *In: R. J. MacIntyre (ed.), Molecular evolutionary genetics*. Plenum Press, New York.
- BROWN, W. M., PRAGER, E. M., WANG, A. & WILSON, A. C., 1982, Mitochondrial DNA of primates: tempo and mode of evolution. *J. Molec. Evol.*, 18: 225-239.
- BURGER, R. & LYNCH, M., 1995, Evolution and extinction in a changing environment: a quantitative genetic analysis. *Evolution*, 49: 151-163.
- CABOT, E. L. & BECKENBACH, A. T., 1989, Simultaneous editing of multiple nucleic acid protein sequences with ESEE. *CABIOS*, 5: 233-234.
- CHANNELL, R. & LOMOLINO, M. V., 2000, Dynamic biogeography and conservation of endangered species. *Nature*, 403: 84-86.
- CHAPIN III, F. S., WALKER, B. H., HOBBS, R. J., HOOPER, D. U., LAWTON, J. H., SALA, O. E. & TILMAN, D., 1997, Biotic control over the functioning of ecosystems. *Science*, 277: 500-504.
- CHOW, S. & TAKEYAMA, H., 2000, Nuclear and mitochondrial DNA analyses reveal four genetically separated breeding units of the swordfish. *J. Fish Biol.*, 56: 1087-1098.
- CINCOTTA, R. P., WISNEWSK, I. J. & ENGELMAN, R., 2000, Human population in the biodiversity hotspots. *Nature*, 404: 990-992.
- COSTANZA, R., D'ARGE, R., DE GROOT, R., FARBER, S., GRASSO, M., HANNON, B., LIMBERG, K., NAEEM, S., O'NEILL, R. V., PARUELO, J., RASKIN, R. G., SUTTON, P. & VAN DEN BELT, M., 1997, The value of the world's ecosystem services and natural capital. *Nature*, 387: 253-260.
- CURRENS, K. P. & BUSACK, C. A., 1995, A framework for assessing genetic vulnerability. *Fisheries*, 20: 24-31.
- DeSALLE, R., TEMPLETON, A., MORI, I., PLETSCHER, S. & JOHNSTON, J. S., 1987, Temporal and heterogeneity of mtDNA polymorphisms in natural populations of *Drosophila mercatorum*. *Genetics*, 116: 215-223.
- DOWLING, T. E., MORITZ, C. & PALMER, J. D., 1990, Nucleic acids II: restriction site analysis, pp. 250-317. *In: D. M. Hillis & C. Moritz (eds.), Molecular systematics*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- ENDLER, J. A., 1977, *Geographic variation, speciation, and clines*. Princeton Univ. Press, Princeton, New Jersey, USA.
- ERLICH, H. A., 1992, *PCR technology: principals and applications for DNA amplification*. W. H. Freeman & Company, New York.
- ERWIN, T. L., 1991, An evolutionary basis for conservation strategies. *Science*, 253: 750-752.

- EXCOFFIER, L., SMOUSE, P. E. & QUATTRO, J. M., 1992, Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA data. *Genetics*, *131*: 479-491.
- FELSENSTEIN, J., 1993, *PHYLIP (Phylogenetic Inference Package), version 3.5c*. Distributed by the author, Department of Genetics, University of Washington, Seattle.
- FISHER, R. A., 1930, *The genetical theory of natural selection*. Clarendon Press, Oxford, England.
- FRIESEN, V. L., CONGDON, B. C., WALSH, H. E. & BIRT, T. P., 1997, Intron variation in marbled murrelets detected using analyses of single stranded conformational polymorphisms. *Molec. Ecol.*, *6*: 1047-1058.
- FU, Y. X., 1997, Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*: 915-925.
- GASTON, K. J., 2000, Global patterns in biodiversity. *Nature*, *405*: 220-227.
- GOLD, J. R., RICHARDSON, L. R., FURMAN, C. & KING, T. L., 1993, Mitochondrial DNA and population structure in red drum (*Sciaenops ocellatus*) from the Gulf of Mexico and Atlantic Ocean. *Mar. Biol.*, *116*: 175-185.
- GOLDSTEIN, D. B. & POLLACK, D. D., 1997, Launching microsatellites: a review of mutation processes and method of phylogenetic inference. *J. Hered.*, *88*: 335-342.
- GOOLDMAN, S. J., 1997, RstCalc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molec. Ecol.*, *6*: 881-885.
- HALL, T. A., 1999, Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucl. Acids. Symp. Ser.*, *41*: 95-98.
- HARRIS, H. & HOPKINSON, D. A., 1976, *Handbook of enzyme electrophoresis in human genetics*. Oxford University Press, North Holland, United Kingdom.
- HARRISON, R. G., 1990, Hybrid zones: windows on evolutionary processes, pp. 69-128. In: J. Antonovics & D. Futuyma (eds.), *Oxford surveys in evolutionary biology*, 7<sup>th</sup> vol. Oxford University Press, Oxford, England.
- HARTL, D. L. & CLARK, A. G., 1997, *Principles of Population Genetics, Second Edition*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- HAYASHI, J. I., TAGASHIRA, Y. & YOSHIDA, M. C., 1985, Absence of extensive recombination between inter- and intraspecies mitochondrial DNA in mammalian cells. *Exp. Cell. Res.*, *160*: 387-395.
- HILLIS, D. M., MORITZ, C. & MABLE, B. K., 1996, *Molecular systematics*, 2<sup>nd</sup> ed. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- HUANG, D. & BERNARD, I. G., 2001, Disjunct Sea of Cortez-Pacific Ocean *Gillichthys mirabilis* populations and the evolutionary origin of their Sea of Cortez endemic relative. *Gillichthys seta*. *Mar. Biol.*, *138*: 421-428.
- INNIS, M. A., GEFLAND, D. H., SNINSKY, J. J. & WHITE, T. J. (eds.), 1990, *PCR protocols: a guide to methods and applications*. Academic Press, Inc., San Diego, California.
- JEANMOUGIN, F., THOMPSON, J. D., GOUY, M., HIGGINS, D. G. & GIBSON, T. J., 1998, Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.*, *23*: 403-405.
- JIN, L. & FERGUSON, J. W. H., 1990, *Neighbor-joining tree and UPGMA tree software*. Center for Demographic and Population Genetics. University of Texas Health Center, Houston, Texas, USA.
- JOHNSON, N., REVENGA, C. & ECHEVERRIA, J., 2001, Managing water for people and nature. *Science*, *292*: 1071-1072.
- KARL, S. A., BOWEN, B. W. & AVISE, J. C., 1992, Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. *Genetics*, *131*: 163-173.
- KARL, S., 1996, An application of anonymous nuclear loci to conservation biology, pp. 38-53. In: T. B. Smith & R. K. Wayne (eds.), *Molecular genetic approaches in conservation*. Oxford University Press, New York.
- KNOWLTON, N., 2001, Coral reef biodiversity – habitat size matters. *Science*, *292*: 1493-1495.
- KUMAR, S., TAMURA, K., JAKOBSEN, I. B. & NEI, M., 2001, *MEGA 2: Molecular Evolutionary Genetics Analysis*. Arizona State University, Tempe, Arizona, USA.
- LANDE, R., 1995, Breeding plans for small populations based on the dynamics of quantitative genetic variance, pp. 318-340. In: D. Ballou, M. Gilpin & T. J. Foose (eds.), *Population management for survival and recovery*. Columbia University Press, New York.
- LANDE, R. & SHANNON, S., 1996, The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, *50*: 434-437.
- LESICA, P. & ALLENDORF, F. W., 1995, When are peripheral populations valuable for conservation? *Conservat. Biol.*, *9*: 753-760.
- LEVIN, D. A., 1995, Metapopulations: an arena for local speciation. *J. Evol. Biol.*, *8*: 635-644.
- LEWONTIN, R. C. & HUBBY, J. L., 1966, A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. *Genetics*, *54*: 595-609.
- MADDISON, W. P. & MADDISON, D. R., 1992, *MacClade, version 3.0*. Sinauer Associates, Inc., Sunderland Massachusetts, USA.
- McCONNELL, S., HAMILTON, L., MORRIS, D., COOK, D., PAQUET, D., BENTZEN, P. & WRIGHT, J. M., 1995, Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. *Aquaculture*, *13*: 19-30.

- McELROY, D., MORAN, P., BERMINGHAM, E. & KORNFIELD, I., 1992, REAP: the restriction enzyme analysis package, version 4.0. *J. Hered.*, 83: 157-158.
- McKINNEY, F. K., LIDGARD, S., SEPKOSKI, J. J., Jr. & TAYLOR, P. D., 1998, Decoupled temporal patterns of evolution and ecology in two post-Paleozoic clades. *Science*, 281: 807-809.
- McMILLEN-JACKSON, A. L., BERT, T. M., SEYOUM, S. ORSOY, T., CRUZ-LOPEZ, H. & CRABTREE, R. E., in review, Molecular genetic variation in tarpon (*Megalops atlanticus* Valenciennes) in the northern Atlantic Ocean.
- MEFFE, G. K. & CARROLL, C. R., 1997, Genetics: conservation of diversity within species, pp. 161-202. In: G. K. Meffe & C. R. Carroll (eds.), *Principles of conservation biology*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- MOORE, W. S. & PRICE, J. T., 1993, The nature of selection in the Northern Flicker hybrid zone and its implications for speciation theory, pp. 196-225. In: R. G. Harrison (ed.), *Hybrid zones and the evolutionary process*. Oxford University Press, New York.
- MORITZ, C., 1994, Defining 'Evolutionarily Significant Units' for conservation. *Trends Ecol. Evol.*, 9: 373-375.
- MURPHY, R. W., SITES, J. W., Jr., BUTH, D. G. & HAUFLE, C. H., 1990, Proteins I: Isozyme electrophoresis, pp. 45-126. In: D. M. Hillis & C. Moritz (eds.), *Molecular systematics*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- MEYER, A., 1993, Evolution of mitochondrial DNA in fishes, pp. 1-33. In: P. W. Hochachka & T. P. Mommsen (eds.), *Biochemistry and molecular biology of fishes*, 2<sup>nd</sup> vol. Elsevier Science Publishers, Amsterdam.
- MYERS, N., MITTERMEIER, R., MITTERMEIER, C. G., DA FONSECA GUSTAVO, A. B. & KENT, J., 2000, Biodiversity hotspots for conservation priorities. *Nature*, 403: 853-858.
- MYERS, R. M., MANIATIS, T. & LERMAN, L. S., 1986, Detection and localization of single base changes by denaturing gradient gel electrophoresis. *Methods Enzymol.*, 155: 501-527.
- NEI, M. & MILLER, J. C., 1990, A simple method for estimating average number of nucleotide substitutions within and between populations from restriction data. *Genetics*, 125: 873-879.
- O'CONNELL, M., DILLON, M. C., WRIGHT, J. M., BENTZEN, P., MERKOURIS, S. & SEEB, J., 1998, Genetic structuring among Alaska herring (*Clupea pallasii*) populations identified using microsatellite variability. *J. Fish. Biol.*, 5: 150-63.
- PALUMBI, S. R. & BAKER, C. S., 1994, Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molec. Biol. Evol.*, 11: 426-435.
- RAYMOND, M. & ROUSSET, F., 1995, An exact test for population differentiation. *Evolution*, 49: 1280-1283.
- RICE, W. R., 1989, Analyzing tables of statistical tests. *Evolution*, 43: 223-225.
- RICHARDSON, B. J., BAVERSTOCK, P. R. & ADAMS, M., 1986, *Allozyme electrophoresis: a handbook for animal systematics and population studies*. Academic Press, Sydney, Australia.
- ROZAS, J. & ROZAS, R., 1999, DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, 15: 174-175.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T., 1989, *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- SCHNEIDER, S., ROESSLI, S. & EXCOFFIER, L., 2000, *Arlequin: a software for population genetic data analysis, version 2.000*. Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva, Switzerland.
- SELANDER, R. K., SMITH, M. H., YANG, S. Y., JOHNSON, W. E. & GENTRY, J. G., 1971, Biochemical polymorphisms and systematics in the genus *Peromyscus*. I. Variation in the old-field mouse (*Peromyscus polionotus*). *Univ. Texas Publ.*, 7103: 40-90.
- SEYOUM, S., TRINGALI, M. D., BERT, T. M., MCELROY, D. & STOKES, R., 2000, An analysis of genetic population structure in red drum (*Sciaenops ocellatus*) based on mtDNA control region sequences. *Fish. Bull.*, 98: 127-138.
- SHAW, C. R. & PRASAD, R., 1970, Starch gel electrophoresis of enzymes—a compilation of recipes. *Biochem. Genet.*, 4: 297-320.
- SLATKIN, M., 1985, Gene flow in natural populations. *Ann. Rev. Ecol. Syst.*, 19: 393-430.
- SLATKIN, M., 1993, Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, 47: 264-279.
- SMITH, T. B. & WAYNE, R. K., 1996, *Molecular genetic approaches in conservation*. Oxford University Press, New York.
- SMITH, T. B., WAYNE, R. K., GIRMAN, D. J. & BRUFORD, M. W., 1997, A role for ecotones in generating rainforest biodiversity. *Science*, 276: 1855-1857.
- SOKAL, R. R. & ROHLF, F. J., 1995, *Biometry: the principles and practice of statistics in biological Research, Third Edition*. W.H. Freeman and Company, San Francisco, California, USA.
- STATSOFT, INC., 1999, *STATISTICA for Windows* (computer program manual; \$797 US, \$995 US overseas), <http://www.statsoft.com>. Tulsa, Oklahoma, USA.
- STEPIEN, C., 1995, Population genetic divergence and geographic patterns from DNA sequences: examples from marine and freshwater fishes. *Am. Fish. Soc. Symp.*, 17: 263-287.
- STREELMAN, J. T., ZARDOYA, R., MEYER, A. & KARL, S. A., 1998, Multilocus phylogeny of cichlid fishes (Pisces: Perciformes): evolutionary comparison of microsatellite and single-copy nuclear loci. *Molec. Biol. Evol.*, 15: 798-808.
- SWOFFORD, D. L., 1998, *PAUP, and other methods. Phylogenetic Analysis Using Parsimony, version 4*. Sinauer Associates, Sunderland, Massachusetts, USA.

- TAJIMA, F., 1989, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.
- THOMPSON, J. D., GIBSON, T. J., PLEWNIAK, F., JEANMOUGIN, F. & HIBBINS, D. G., 1997, The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nuc. Acids Res.*, 24: 4876-4882.
- THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J., 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nuc. Acids Res.*, 22: 4673-4680.
- TOLINE, C. A. & BAKER, A. J., 1995, Mitochondrial DNA variation and population genetic structure of the northern redbelly dace. *Molec. Ecol.*, 4: 734-753.
- TRINGALI, M. D. & BERT, T. M., 1996, The genetic stock structure of common snook (*Centropomus undecimalis*). *Can. J. Fish. Aquat. Sci.*, 53: 974-984.
- VANE-WRIGHT, R. I., LEHMAN, C. L. & NOWAK, M. A., 1994, Habitat destruction and the extinction debt. *Nature*, 371: 65-66.
- VITOUSEK, P. M., MOONEY, H. A., LUBCHENCO, J. & MELILLO, J. M., 1997, Human domination of earth's ecosystems. *Science*, 277: 494-499.
- WEIR, B. S., 1996, *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- WILLIAMS, J. G. K., KUBELIK, A. R., LIVAK, K. J., RAFALSKI, J. A. & TINGEY, S. V., 1990, DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, 18: 6531-6535.
- WILSON, A. C., CANN, R. L., CARR, S. M., GEORGE, M., GYLLENSTEN, U. B., HELM-BYCHOWSK, I. K. M., HIGUCHI, R. G., PALUMBI, S. R., PRAGER, E. M., SAGE, R. D. & STONEKING, M., 1985, Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linnean Soc.*, 26: 375-400.
- WRIGHT, S., 1969, *Evolution and genetics of populations. iv. the theory of gene frequencies*. Univ. Chicago Press, Chicago, Illinois, USA.
- XIA, X., 2000, *DAMBE: Data Analysis in Molecular Biology and Evolution*. Dept. Ecol. & Biodiv., University of Hong Kong, P. R. China.