Antarctic Science 12 (3): 288–296 (2000) © British Antarctic Survey Printed in the United Kingdom

Molecular tools for understanding population structure in Antarctic species

PATRICK M. GAFFNEY

College of Marine Studies, University of Delaware, Lewes, DE 19958, USA
pgaffney@udel.edu

Abstract: During the last decade, methods for detecting DNA polymorphisms have proliferated at a bewildering pace. Today the investigator must choose among various types of genetic markers as well as between a variety of methods for discovering and screening polymorphisms. Polymorphisms useful for the analysis of population structure are found in both mitochondrial and nuclear genomes. Marker development may proceed along two routes: 1) discovery of species-specific markers, and 2) application of universal methods. Species–specific markers are based on sequence data from the target species or close relatives, whereas universal markers are based on the use of polymerase chain reaction (PCR) primers targeted to regions highly conserved across diverse taxa. Markers commonly employed include mitochondrial DNA polymorphisms, microsatellites, anonymous nuclear loci and known genes (both coding and noncoding regions). Methods for detecting polymorphisms range from technically simple (RFLP analysis) to more sophisticated mutation-scanning methods. We review the application of these approaches to several key Antarctic species (the Patagonian toothfish *Dissostichus eleginoides*, the mackerel icefish *Champsocephalus gunnari*, and the squid *Martialia hyadesi* Rochebrune & Mabille, 1889) and present preliminary data on genetic polymorphisms in toothfish and icefish.

Received 18 October 1999, accepted 21 February 2000

Key words: *Champsocephalus*, *Dissostichus*, genetics, intron, *Martialia*, mtDNA, polymorphism, population structure

A brief history of genetic markers used to study population structure in marine species

The delineation of population structure in marine species has been an enduring passion of fisheries biologists for decades. For our purposes, population structure refers to the genetic architecture of a species, i.e. whether it exists as a single freely interbreeding population, or is subdivided into genetically distinct subpopulations. The implicit assumption behind the quest to understand genetic population structure is that the genetically distinct subpopulations are to some extent independent evolutionary units, with distinct biological properties relevant to their management. The role of such genetic information in the overall management of commercially or ecologically important species continues to be debated, but few would argue that it is wise to proceed in the absence of population genetic information.

Until recently, the range of genetic tools available for the study of population structure in any species was limited; for all practical purposes, the only genetic markers readily available for a variety of species were enzyme polymorphisms (allozymes). These were followed by the appearance of mitochondrial DNA (mtDNA) polymorphisms, which provided a new and often different perspective on evolutionary population biology by combining phylogenetic and geographic information (phylogeography, *sensu* Avise *et al.* 1987). A third category of genetic markers, nuclear DNA polymorphisms, has received a great deal of attention in recent years, and promises to enrich our understanding of population structure in a variety of species well beyond the expectations of previous decades. Each of these classes of genetic markers is discussed briefly below in terms of its utility for the study of Antarctic marine resource species.

Allozymes

Allozymes enjoyed a long reign as the primary type of genetic marker used to assess population structure. Allozyme surveys often were able to clarify major boundaries to gene flow in aquatic organisms, as well as occasionally revealing the existence of previously undetected species. Despite being superseded by molecular methods, this technique still possesses a number of attractive features. Standard starch gel electrophoresis is inexpensive in terms of both reagents and equipment, and technically straightforward. Many of the loci examined are physiologically well-characterized, allowing the possibility that the genetic variation detected may be directly associated with fitness variation in the organisms examined. Although this aspect of allozyme polymorphisms has inspired evolutionary geneticists, the possibility that allelic frequencies might be determined by factors other than random genetic drift alone has raised questions about their utility for inferring population structure (Karl & Avise 1992).

Other drawbacks to the use of allozyme markers include the need to use fresh or well-frozen tissue samples, the inability to detect a significant proportion of underlying genetic variation, and the absence of sufficient polymorphism in some species. However, it was not until other types of markers became available that their relative merits could be evaluated.

mtDNA

With the advent of DNA techniques, vast repositories of genetic information became available to the researcher. Among the first of these to be exploited by population geneticists was the mitochondrial genome. Mitochondrial DNA (mtDNA) tends to evolve rapidly, typically exhibits substantial polymorphism, and is particularly sensitive to the effects of random genetic drift owing to its uniparental inheritance. Initially, mtDNA variation was examined by restriction enzyme analysis of entire mitochondrial genomes, which was only possible when moderate amounts of (preferably mitochondria-rich) tissue in good condition were available. The subsequent development of the polymerase chain reaction (PCR) coupled with the use of universal primers have greatly expanded the use of mtDNA analyses, allowing studies to be conducted with minute amounts of poorly preserved tissue.

Although mitochondrial polymorphisms have been widely used to examine population structure in a great number of species, they are not without drawbacks. Because the mitochondrial genome does not undergo recombination, it acts as a single locus, whose genealogy may not be synonymous with the population genealogy. In addition, the use of PCR and other methods for the subsequent detection of DNA sequence variation in the amplified targets is more expensive and require more technical expertise than conventional protein electrophoresis. Levels of mtDNA polymorphism may vary considerably among species. Many marine teleosts exhibit little mtDNA sequence diversity (Ovenden *et al.* 1989), although some species possess hypervariable mtDNA (e.g. menhaden, Avise *et al.* 1989). Cases of extremely low diversity are typically ascribed to historical bottlenecks in population abundance, from which species with enormous population sizes (such as menhaden) may be exempt.

A final caveat regarding the use of mtDNA variation to infer population structure stems from the uniparental inheritance of the mitochondrial genome. When patterns of gene flow differ between males and females, a biased picture of population structure will result from reliance on either mitochondrial or nuclear markers alone (Palumbi & Baker 1994). Although departures from the model of uniparental inheritance are known to occur, they are either uncommon (e.g. heteroplasmy, the occurrence of more than one mtDNA sequence type within a single individual) or restricted to certain taxa (e.g. the doubly uniparental inheritance of mitochondrial genomes in mussels).

Nuclear DNA

The most recent additions to the population geneticist's "molecular toolbox" derive from the primary genetic material, the nuclear genome. Useful DNA polymorphisms are found in both coding and noncoding regions; three major categories will be discussed here.

Coding regions. For most species, little DNA sequence information is available in public databases, and it is often restricted to partial gene sequences derived from cDNA libraries. For example, a search of GenBank using the keyword *Dissostichus* retrieved 18 nuclear DNA entries (17 for *D. mawsoni* Norman, one for *D. eleginoides* Smitt), while a search for *Champsocephalus* retrieved two for *C. gunnari* Lönnberg. Often cDNA sequences reported include untranslated flanking regions, which may yield more polymorphisms than the translated sections (Levitt 1991).

When coding region sequence data are available for target species or their close relatives, the processes of PCR primer design, target region amplification and screening for polymorphisms can be made efficient. The nature of the sequence variation observed (i.e. synonymous vs non-synonymous substitutions, patterns of amino acid replacement) is not only relevant to primer design, but may be important in interpreting observed patterns of variation.

This approach is clearly limited by the amount of information available in genetic databases. In addition, coding regions will tend to exhibit less polymorphism than noncoding regions, with occasional single nucleotide polymorphisms being the rule.

Introns. One important type of noncoding region is the intron, a common feature of eukaryotic genes. Introns typically display higher levels of polymorphism than coding regions; because the intron sequence is removed during mRNA processing, there appear to be fewer selective constraints on sequence variation. Nucleotide substitutions, insertions and deletions, and sometimes stretches of repetitive DNA are commonly found in introns.

Because intron positions are typically conserved within particular taxonomic groups, it is possible to design PCR primers anchored in the exons straddling a target intron, to achieve "exon-primed, intron-crossing" (EPIC) PCR (Palumbi & Baker 1994). A number of vertebrate EPIC primers are available (Lessa 1992, Slade *et al.* 1993); however, there is no guarantee that they will perform reliably for a species of interest. In addition, EPIC primer sets often amplify multiple bands, representing introns in duplicated loci and/or pseudogenes (Palumbi 1996). Locus-specific intron amplification can be achieved by refined primer design, but to date only a few such primer sets are available (Quattro & Jones 1999).

Microsatellites. The most recent major addition to the

molecular toolbox is the class of simple sequence repeat loci known as microsatellites. Allelic variation occurs both as variation in size (i.e. the number of tandem repeats of the microsatellite motif) and nucleotide composition. Microsatellite loci are often extremely polymorphic, and the number of such loci is in principle almost unlimited (Wright & Bentzen 1994).

The primary drawback to microsatellites as molecular markers for population studies is the effort required to develop a set of loci for a particular species. Although some microsatellite-amplifying primer sets can be used across a wide range of related taxa (Rico *et al.* 1996), for the most part marker development is species-specific, and not a technically trivial task. An additional concern relevant to Antarctic species is the possibility that, in species with large effective population sizes, the spectrum of alleles may be determined more by mutational dynamics than by patterns of gene flow. Under some models, isolated subpopulations may possess identical allelic frequency distributions; in such cases the use of microsatellite loci alone to infer population structure would be misleading (Nauta & Weissing 1996). In addition, microsatellite loci are often far more polymorphic than traditional markers; the high heterozygosity and large number of alleles can complicate the statistical analysis of population structure (Hedrick 1999).

Screening PCR products for polymorphisms

Once a target region has been amplified by PCR, an effective method of detecting sequence variation must be applied. With the exception of microsatellite loci, most amplified regions will display sequence variation in the form of nucleotide substitutions rather than insertions/deletions. How can this variation be detected? The number of methods available for mutation detection is sizeable and still growing; we discuss here several that have proven useful for us.

One primary determinant of the mutation screening method to be employed is the size of the amplified fragment. For small fragments (200–500 bp), several high-resolution techniques are commonly used. Single-strand conformation polymorphism (SSCP) relies on the tendency for nucleotide substitutions to alter the electrophoretic mobility of single-stranded DNA. While this method is readily applied to any PCR product with the 200–500 bp size range, resolution is less than perfect: typically, 70–80% efficiency is reported for fragments up to 400 bp in size, with higher (up to 90%) efficiency for fragments < 200 bp in size (Hayashi & Yandell 1993, Prosser 1993).

Denaturing gel electrophoresis (DGE) of double-stranded DNA (dsDNA) molecules likewise relies on differences in electrophoretic mobility among sequence variants. dsDNA molecules tend to denature in discrete domains when exposed to chemical denaturants or increasing temperature; detectable changes in the melting temperature of a given domain can result from a single base change. Once a dsDNA molecule begins to denature, its electrophoretic mobility decreases greatly, allowing visualization of sequence differences as differential migration of PCR products on a denaturing acrylamide gel. The detection of single base substitutions can be enhanced greatly by the generation of heteroduplexes, i.e. dsDNA molecules that contain one or more mismatched bases. These will arise during the course of PCR when a template containing more than one allele is amplified (e.g. in a heterozygote for a nuclear gene), or can be generated deliberately by mixing two or more PCR products and subjecting them to a round of heat denaturation and reannealing. Heteroduplex mismatches greatly decrease the stability of a dsDNA molecule, causing it to denature earlier than homoduplex dsDNA. The use of an outgroup as a heteroduplex generator is a powerful means of enhancing mutation detection (Campbell *et al.* 1995).

Although denaturing gradient electrophoresis typically offers higher resolution than SSCP for PCR products of the same size range, its efficacy is sequence-dependent. For optimal resolution, the PCR product must possess a suitable "melting profile", and the electrophoresis protocol is usually designed to detect variation in a single domain. For known DNA sequences, PCR primers can be designed to yield products with desirable melting properties, by using software that calculates the melting profile of any given DNA sequence. When the amplification target is optimized in this way, DGE is capable of detecting 100% of single-base substitutions. When the sequence of a PCR product is unknown, as is usually the case when universal primers are used, denaturing gel electrophoresis protocols must be determined empirically, and may in some cases defy optimization (e.g. when the lowest melting temperature domain is flanked by more stable domains).

For medium-sized PCR products (0.5–1.0 kb), SSCP and DGE are less effective. We have found three other approaches to be useful: restriction fragment length polymorphism analysis (RFLP), RFLP coupled with denaturing gel electrophoresis (RFLP/DGE), and low-stringency, single-primer PCR (LSSP-PCR).

RFLP analysis is a familiar and simple method, in which a set of PCR products is screened by digestion with a panel of restriction enzymes, after which sequence variation manifested as differences in the presence/absence of restriction sites is visualized on an agarose or acrylamide gel. For products of intermediate size, enzymes with four-base recognition sites are most useful, along with degenerate five- and six-base enzymes.

RFLP/DGE is simply a combination of two methods, in which set of PCR products digested with a particular restriction enzyme is run on a denaturing gel. When the product sequence is known, it is possible to select a restriction enzyme that yields fragments with desirable sizes and melting profiles. This approach typically detects 50–80% of sequence variants (Sheffield *et al.* 1990).

LSSP-PCR is a method that has not been widely used but offers a quick means of determining the amount of sequence

variation an amplified region possesses. In this procedure, the PCR product is purified and then re-amplified using only one of the original PCR primers, under conditions of low stringency. This second-round amplification usually yields reproducible fingerprint-like patterns (Fig. 1) that may differ as the result of even a single nucleotide difference (Pena *et al.* 1994). The technique is probably best suited to haploid templates such as mtDNA, and may be most useful as a quick method for determining the amount of sequence variation a fragment possesses.

For large (> 1 kb) PCR products, RFLP analysis remains the simplest method of assessing sequence variation. Other methods for screening large fragments have recently become available from commercial suppliers, e.g. Cleavase® Fragment Length Polymorphism (Third Wave Technologies, Inc.) and Base Excision Sequence Scanning (Epicentre Technologies Corporation). These methods offer higher sensitivity than RFLP analysis and may be particularly advantageous for products with modest levels of variability that the RFLP approach might miss entirely.

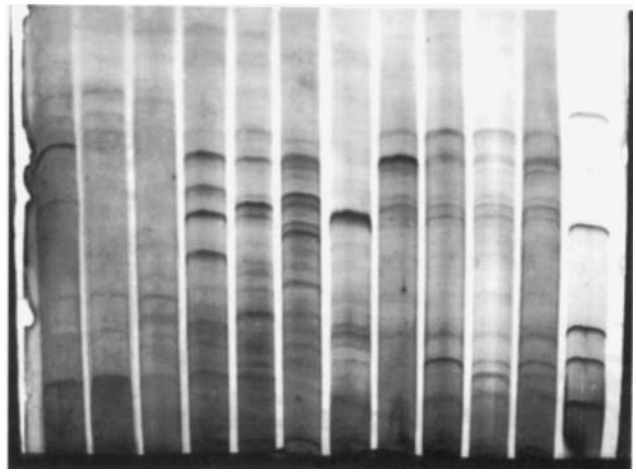

Fig. 1. LSSP-PCR of a ≈1.2 kb PCR product amplified from the mitochondrial control region of the Atlantic silverside *Menidia menidia* (Linnaeus, 1766). Lane 1: pBR322/*Bst*N I ladder; lanes 2–3: replicate amplifications from the same individual; lanes 4–12: different individuals. All individuals show distinct profiles, reflecting nucleotide substitutions confirmed by direct sequencing (unpublished data).

A suggested approach for developing genetic markers

Two governing principles are important in the development of genetic markers for a given species. First, the use of various types of markers increases the probability of finding informative polymorphisms, and ensures a minimally biased view of population structure. Second, technically simple approaches should be explored before undertaking more demanding techniques. We report here some preliminary data illustrating the application of this approach to three Antarctic species: Patagonian toothfish, mackerel icefish and seven star flying squid.

Patagonian toothfish

DNA was extracted from ethanol-preserved skeletal muscle samples using a DNA FastPrep apparatus (Bio101, Inc.). A panel of universal nuclear and mtDNA primers was tested against a range of annealing temperatures (46–60°C) using a gradient thermal cycler. For these loci yielding consistent and clear single amplification products, a set of 18 individuals (three individuals from each of six collection sites) was screened by digestion with a panel of restriction enzymes followed by visualization of digests on agarose gels stained with ethidium bromide.

Mitochondrial DNA

Universal primers CO1a-H and CO1f-L (Palumbi *et al.* 1991) amplified a ≈0.7 kb fragment of the cytochrome oxidase subunit I (COI) gene. RFLP screening with 16 enzymes (*Alu* I, *Alw*26 I, *Bfa* I, *Bsr* I, *Dde* I, *Dpn* II, *Hae* III, *Hha* I, *Hinf* I, *Hpa* II, *Mse* I, *Msp* I, *Nla* III, *Rsa* I, *Sau*96 I, *Tsp*509 I) revealed no polymorphisms. Although this product appears to exhibit little variability, it could be further screened by RFLP/ DGE of one or more of the digests that yielded fragments of suitable sizes.

A small portion (*c.* 0.5 kb) of the control region was amplified using the primers L-Pro and H16498 (Meyer *et al.* 1990). Of 19 restriction enzymes used (*Alu* I, *Bam*H I, *Bfa* I, *Bst*U I, *Dde* I, *Dpn* II, *Hae* III, *Hha* I, *Hinf* I, *Hpa* II, *Mse* I, *Msp* I, *Nla* III, *Rsa* I, *Sau*96 I, *Scr*F I, *Taq* I, *Tas* I, *Tsp*509 I), six (*Hinf* I, *Hpa* II, *Msp* I, *Scr*F I, *Tas* I, *Tsp*509 I) revealed a total of nine restriction site polymorphisms.

A larger (≈1.2 kb) portion of the control region downstream of the L-Pro/H16498 fragment was amplified using primers GC-L16498 (Gaffney, unpublished data) and 12SAR-H (Palumbi *et al.* 1991). Of 18 restriction enzymes used (*Alu* I, *Alw*26 I, *Bam*H I, *Bfa* I, *Bst*U I, *Dde* I, *Dpn* II, *Hae* III, *Hha* I, *Hinf* I, *Hpa* II, *Mse* I, *Msp* I, *Nla* III, *Rsa* I, *Scr*F I, *Taq* I, *Xba* I), nine (*Alw*26 I, *Hae* III, *Hpa* II, *Mse* I, *Msp* I, *Nla* III, *Rsa* I, *Scr*F I) revealed polymorphisms not yet fully characterized.

The primers CB2-H and CB1-L (Palumbi *et al.* 1991) amplified a ≈0.3 kb cytochrome *b* product. Although too small for effective RFLP screening, it is a suitable candidate for DGE screening, as demonstrated by Orbacz & Gaffney (2000).

In summary, several regions of the toothfish mitochondrial genome can be readily amplified using universal primers, and several of these possess useful levels of polymorphism. Examination of geographic variation in the polymorphisms discovered, as well as in additional amplified regions not yet screened, will be a powerful tool for evaluating global population structure in this species.

Nuclear DNA

Five EPIC primer pairs were found to amplify toothfish successfully, although one set (CK6-5' and CK7-3', Palumbi *et al.* 1991) yielded three bands, presumably representing introns from multiple creatine kinase loci. Of the remaining four, one - LDH A intron 6, amplified using primers LDHA6F1 and LDHA6R (Quattro & Jones 1999) - yielded a small product suitable for DGE screening. The primers CMX4F and CMX5R of Chow (1998) amplified a ≈450 bp product suitable for RFLP or DGE/RFLP screening. Primers S7RPX2F and S7RPX3R yielded a ≈750 bp product from intron 2 of the S7 ribosomal protein (Chow & Hazama 1998), which will be screened by RFLP. Primers RPEX1F and RPEX2R (Chow & Hazama 1998) yielded a ≈850 bp product from intron 1 of the S7 ribosomal protein. This product was screened with ten restriction enzymes (*Alu* I, *Ava* II, *Bam*H I, *Bfa* I, *Bst*U I, *Bsu*R I, *Csp*6 I, *Hha* I, *Taq* I, *Tas* I); three (*Bfa* I, *Taq* I, *Tas* I) revealed polymorphisms yet to be fully characterized.

These preliminary results suggest that introns amplified using universal EPIC primers are likely to contain informative polymorphisms. When coupled with the analysis of mtDNA variation, these markers should provide a robust picture of population structure in *D. eleginoides*.

Candidate loci for marker development in toothfish

Although the mitochondrial and nuclear loci examined to date offer a number of potentially informative polymorphisms, additional loci are always welcome. EPIC primers for toothfish could be designed from the genomic sequence of the *D. mawsoni* trypsinogen gene, which contains several introns (Chen *et al.* 1997). Venkatesh *et al.* (1999) amplified several introns in *D. mawsoni* using universal primers: introns 6a and 10a in the dystrophin gene (458 bp and 1333 bp), growth hormone intron 4a (1070 bp), mixed lineage leukaemia intron 25a (299 bp), and MHC class II B-chain intron 2a (124 bp). Judging from the near-identity (99.3% of 993 nucleotides) of nucleotide sequences coding for lactate dehydrogenase A (LDH-A) in *D. eleginoides* and its congener *D. mawsoni*, it is

```
          1                                                  ∇    50
Dele      MST-KEKLIS HVMKEEPVGS RNKVTVVGVG MVGMASAISI LLKDLCDELA
Dmaw      ...-...... .......... .......... .......... ..........
Cgun      ...-...... .......... .S........ .......... ..........
Fhet      ...-Q..... .........C .......... .........V ..........
Mmus      .A.L.DQ..V NLL...QAP- Q..I...... A....C.... .M...A....

          51                                   ∇                100
Dele      MVDVMEDKLK GEVMDLQHGS LFLKT-KIVG DKDYSVTANS KVVVVTAGAR
Dmaw      .......... .......... .....-.... .......... ..........
Cgun      .......... .......... .....-.... .......... ..........
Fhet      L......... ..A......A .....H...A .......... ..........
Mmus      L......... ..M....... .....P...S S...C..... .L.II.....

          101                                             ∇      150
Dele      QQEGESRLNL VQRNVNIFKF IIPNIVKYSP NCILMVVSNP VDILTYVAWK
Dmaw      .......... .......... .......... .......... ..........
Cgun      .......... .......... .......... .......... ..........
Fhet      .......... .......... .......... ....L..... ..........
Mmus      .......... .......... .......... H.K.LI.... ..........

          151                                                  200
                                                               ∇
Dele      LSGFPRNRVI GSGTNLDSAR FRHLIGEKLH LHPSSCHAWI VGEHGDSSVP
Dmaw      .......... .......... .......... .......... ..........
Cgun      ......H... .......... .......... .......... ..........
Fhet      ......H... .......... ....M...F. .......G.. .........A
Mmus      I....K.... ...C...... ..Y.M..R.G V.AL...G.V L.........

          201                                        ∇         250
Dele      VWSGVNVAGV SLQGLNPQMG TEGDGENWKA IHKEVVDGAY EVIKLKGYTS
Dmaw      .......... .......... .......... .......... ..........
Cgun      .......... .......... .......... .......... ..........
Fhet      ......I... ...T...N.. AD..S....E L..K...... ..........
Mmus      .......... ..KS...EL. .DA.K.Q..E V..Q...S.. ..........

          251                                   ∇                300
Dele      WAIGMSVADL VESIIKNMHK VHPVSTLVQG MHGVKDEVFL SVPSVLGNSG
Dmaw      .......... .......... .......... .......... ..........
Cgun      .......... .......... .......... .......... ..........
Fhet      .......... ....V..L.. .......... .......... .I........
Mmus      ....L..... A...M..LRR ...I..MIK. LY.INED... ...CI..QN.

          301                                  333
Dele      LTDVIHMTLK AEEEKQLQKS AETLWGVQKE LTL
Dmaw      .......... .......... .......... .V.
Cgun      .......... .......... .......... ...
Fhet      .......... P......V.. .......... ...
Mmus      IS..VKV..T P...AR.K.. .D....I... .QF
```

Fig. 2. LDH-A amino acid alignment from several species. Dele = *Dissostichus eleginoides* (GenBank accession AF170027); Dmaw = *D. mawsoni* (AF079827); Cgun = *Champsocephalus gunnari* Lönnberg, 1905 (AF079824); Fhet = *Fundulus heteroclitus* (Linnaeus, 1766) (L42535); Mmus = *Mus musculus* Linnaeus, 1758 (Y00309). Gaps are indicated by '.'; dots indicate identity with the reference sequence (Dele). Inverted triangles indicate the location of introns in the mouse (Mmus).

https://doi.org/10.1017/S0954102000000353 Published online by Cambridge University Press

likely that improved EPIC primers for toothfish could be designed using sequence data from the latter species in the event the degenerate primers of Venkatesh *et al.* (1999) do not amplify *D. eleginoides* templates successfully. When the genomic structure of a gene is not known for the species of interest, EPIC primer design relies on putative intron locations inferred from other species, e.g. LDH-A (Fig. 2).

Several microsatellite loci have also been developed for toothfish (R.D. Ward, personal communication 1998; P.J. Smith, personal communication 1998). A comparison of patterns of geographic variation in mtDNA, introns and microsatellites promises to be very informative, and should provide the level of resolution necessary for management decisions regarding stock structure in this heavily exploited species.

Mackerel icefish

The stock structure of *C. gunnari* is of considerable interest in view of its extensive exploitation and the biological evidence for separate stocks (Duhamel *et al.* 1995). Previous allozyme surveys revealed negligible polymorphism (Duhamel *et al.* 1995), whereas RFLP analysis of whole mtDNA molecules revealed moderate nucleotide diversity but little evidence of population subdivision (Williams *et al.* 1994). Additional studies employing larger sample sizes and additional genetic markers are clearly needed, particularly if population subdivision has occurred relatively recently (since the last ice age).

DNA was extracted from ethanol-preserved skeletal muscle samples obtained during the GeneFlow Cruise (Rodhouse 1997) as described above. PCR amplification and RFLP screening were performed as previously described.

Mitochondrial DNA

Several universal primer sets yielded consistent amplifications. These included a ≈0.6 kb fragment of the 16S ribosomal DNA gene amplified by primers 16SAR and 16SBR (Palumbi *et al.* 1991), which was not screened in view of its typically low intraspecific variability; a ≈0.3 kb portion of the cytochrome *b* gene amplified with primers CB2-H and CB1-L (Palumbi *et al.* 1991), to be screened by DGE; and a ≈0.7 kb fragment of the COI gene (to be screened by RFLP and RFLP/DGE).

A ≈1.5 kb fragment of the control region was amplified using primers L-Pro (Meyer *et al.* 1994) and 12SAR–H (Palumbi *et al.* 1991). Digestion with 22 enzymes (*Acc* I, *Aci* I, *Afl* III, *Alu* I, *Alw*26 I, *Apa* I, *Apa*L I, *Ase* I, *Ava* II, *Bam*H I, *Bfa* I, *Bsa*H I, *Bsr* I, *Dde* I, *Dpn* II, *Dra* I, *Hae* III, *Hin*f I, *Msp* I, *Rsa* I, *Tsp*509 I, *Stu* I) revealed polymorphisms for *Afl* III, *Apa* I, *Alu* I, *Alw*26 I, *Bam*H I, *Bsa*H I, *Dra* I and *Hae* III. Some of these may be attributed to size polymorphism, a common feature of the teleost control region, but our preliminary screening with agarose mini-gels was inadequate to determine fragment sizes accurately. A smaller (≈550 bp) portion of this region amplified with primers L-Pro and H16498 will be screened by RFLP and RFLP/DGE.

A ≈1.2 kb fragment of the NADH dehydrogenase subunit 2 (ND2) gene was amplified using primers c-Trp and t-Met (Park *et al.* 1993). To date, no polymorphisms have been found with *Alu* I, *Alw*26 I, *Apa* I, *Ase* I, *Ava* II, *Bsr* I, *Dpn* II, *Dra* I or *Eco*R I.

Nuclear DNA

Four EPIC primer pairs were found to amplify icefish successfully. LDH A intron 6 amplification yielded a small product suitable for DGE screening. The primers CMX4F and CMX5R amplified a ≈ 450 bp product. A panel of 16 individuals was screened with *Alu* I, *Dde* I, *Dpn* II, *Hae* III, *Hha* I, *Hpa* II, *Msp* I, *Nla* III, *Rsa* I, *Sau*96 I, *Taq* I, and *Tsp*509 I. No polymorphisms were apparent, but several of the digests with multiple cut sites were poorly resolved and should be examined further on a higher-resolution medium such as acrylamide. Primers S7RPX2F and S7RPX3R yielded a ≈ 750 bp product, which will be screened by RFLP. Primers RPEX1F and RPEX2R yielded a ≈ 0.8 kb product, for which a panel of 16 individuals was screened with 12 restriction enzymes (*Alu* I, *Bfa* I, *Dde* I, *Hha* I, *Dpn* II, *Hae* III, *Hpa* II, *Mse* I, *Msp* I, *Nla* III, *Rsa* I, *Sau*96 I). No polymorphisms were

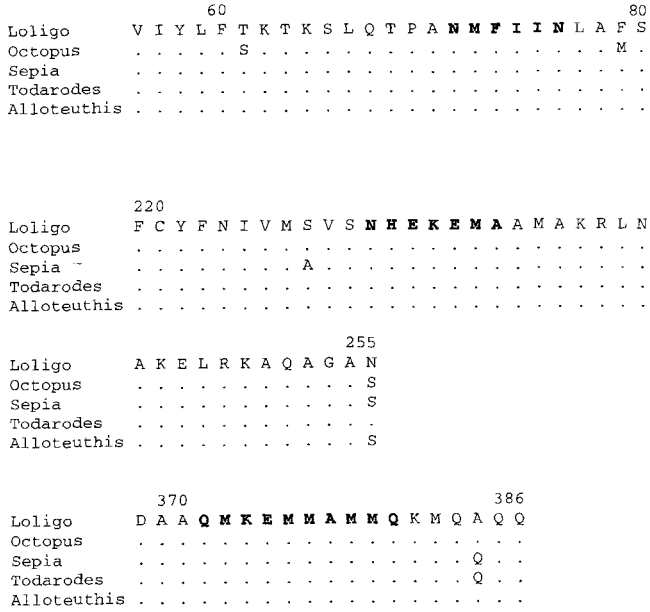

Fig 3. Aligned segments of cephalopod rhodopsin amino acid sequence showing similarity among species. Dots indicate identity with the reference sequence (Lforbesi). Portions of sequence in bold face represent areas with minimally degenerate codons useful for degenerate primer design. Key to species: Lforbesi = *Loligo forbesi* Steenstrup, 1856 (SWISS-PROT Accession P24603; Lsubulata = *Loligo subulata* Lamarck, 1798 (Q17094); Octopus = *Octopus dofleini* (P09241); Sepia = *Sepia officinalis* Linnaeus, 1758 (O16005); Todarodes = *Todarodes pacificus* (P31356).

https://doi.org/10.1017/S0954102000000353 Published online by Cambridge University Press

detected.

Candidate loci for marker development in icefish

Several published sequences are available for *C. gunnari*. Sequences from portions of two mitochondrial genes (cytochrome *b*, 752 bp and control region, 365 bp; Chen *et al.* 1998) could be used to design sensitive DGE protocols for detecting single-base substitutions in these regions. Promising nuclear genes include the myoglobin gene (Small *et al.* 1998), which contains three characterized introns, and LDH-A (Fig. 2). For the former, the known genomic sequence of the introns and flanking exonic sequences would allow design of EPIC primers for high-resolution DGE analysis.

Seven star flying squid

To date, genetic analysis of population structure in *M. hyadesi* has been restricted to one allozyme survey, which found evidence indicating population subdivision, despite the low level of genetic variability typical of other squid species (Brierley *et al.* 1993). Additional genetic markers are necessary to allow a more precise and reliable delineation of stock structure in this species.

Mitochondrial DNA

Although considerable work has been done on mtDNA sequence variation in cephalopods, largely for purposes of phylogenetic analysis, no DNA sequence data from *M. hyadesi* are available. Universal primers have been used for cephalopods to amplify ≈0.4 kb of 16S rDNA (Bonnaud *et al.* 1994), ≈0.7 kb of cytochrome oxidase subunit III and ≈0.4 kb of cytochrome oxidase subunit II (Bonnaud *et al.* 1997), and ≈0.7 kb of cytochrome oxidase subunit II (Carlini & Graves 1999). These regions are good candidates for RFLP and DGE/RFLP, or DGE analysis if sequence data are available to allow design of primers for amplification of smaller target regions. Norman *et al.* (1994) demonstrated the utility of DGE to detect variation in a ≈0.6 kb fragment of the COIII region of *Loligo forbesi*, although not all sequence variants detected by direct sequencing were distinct in their DGE analysis, presumably because of the large size of the product.

Nuclear DNA

A large number of nuclear genes in cephalopods have been the subject of molecular genetic analysis. Comparative amino acid alignments can be used to identify highly conserved

```
            1                                                  50
Lymnaea     MACCIPDELK EQKRINQEIE RQLKRDKRDA RRELKLLLLG TGESGKSTFI
Patinopecten ....LSE.A. ......C... KE.RKA.... .......... ..........
Loligo      ....LSE.A. .......... K..R...... .......... ..........

            51                                                 100
Lymnaea     KQMRIIHGAG YSDEDKRSHI KIVYQNIFMA MHAMIRAMDT LNIQYINPAN
Patinopecten ........T. ..E....GF. .......... ..S....... IK.SFEVAD.
Loligo      ........S. ..E..RKGFE ........S. IQTL.A..E. .SLE.KD.S.

            101                                                150
Lymnaea     RENGNMIRQI DYETVTTFDK PCVDAIISLW NDDGIQECYD RRREYQLTDS
Patinopecten E..AI....V .......L.S QS.E..L... A.A....... ..........
Loligo      N.HAEFLNS. .ADSADI.ED GH.T..KGC. T.P.M..... ..........

            151                                                200
Lymnaea     AKYYLDSVER ISQQDYLPTL QDILRVRVPT TGIIEYPFDL DSIIFRMVDV
Patinopecten ......A.D. .AEPN..... .......... .......... ..........
Loligo      ......D... .HEPG.I... .......... .......... ..........

            201                                                250
Lymnaea     GGQRSERRKW IHCFENVTSI MFLVALSEYD QVLVESDNE- NRMEESKALF
Patinopecten .......... .......... .......... .........- ..........
Loligo      .......... .......... .......... .........E ..........

            251                                                300
Lymnaea     RTIITYPWFQ NSSVILFLNK KDLLEEKIMH SHLVDYFPEF DGPKKEASTA
Patinopecten .......... .......... .......... .......... ..Q..D.QG.
Loligo      .......... .......... .........T ...A....DY ....CDYEA.

            301                                                350
Lymnaea     REFILKMFVE LNPDPDKIIY SHFTCATDTE NIRFVFAAVK DTILQLNLKE
Patinopecten .....R...D .......... .......... .......... ..........
Loligo      ...MMDSYMD ..E.KE.ML. Y.Y....... .......... ..........

            351
Lymnaea     YNLV
Patinopecten ....
Loligo      ....
```

Fig 4. Amino acid sequences of G(Q)a (GTP-binding protein 1, subunit a) from the gastropod *Lymnaea stagnalis* (Linnaeus, 1758) (SWISS-PROT accession P38411), the bivalve *Patinopecten yessoensis* (Jay, 1857) (O15975), and the cephalopod *Loligo forbesi* Steenstrup, 1856 (P38412). Portions of sequence in bold face represent areas with minimally degenerate codons useful for degenerate primer design.

https://doi.org/10.1017/S0954102000000353 Published online by Cambridge University Press

regions suitable for degenerate primer design (Fig. 3). Primers can be designed to amplify large products for RFLP analysis, or smaller targets for DGE analysis. In the latter case, if the PCR product proves to be larger than expected due to the presence of one or more introns, it may still be examined by RFLP.

There are also cases when a gene of interest may show only limited sequence conservation, yet still be worth exploring. For example, in the squid *Nototodarus sloanei* (Gray, 1849), glutathione S-transferase appears to be encoded by a single locus with an exon-intron structure similar to the vertebrate class π GST gene (Tomarev *et al.* 1993). The *N. sloanei* exon sequences immediately flanking target introns could be used to design degenerate primers for amplifying the homologous introns in *Martialia*, a member of the same family (Ommastrephidae), although successful amplification is not guaranteed.

Even genes that have not been extensively studied in cephalopods may be useful targets. For example, the Gqα gene (GTP-binding protein 1, subunit α) exhibits considerable amino acid sequence conservation among diverse molluscan taxa (Fig. 4). In cases such as this, it should be possible to design universal primers with wide taxonomic utility.

Conclusions

Rapid advances in molecular biology have had profound ramifications for other branches of biology, including the evolutionary biology of Antarctic organisms. For many years the only readily available tool for the genetic analysis of population structure was allozyme electrophoresis. Now it is possible to examine several classes of genes, both mitochondrial and nuclear, coding and noncoding regions, to arrive at a more robust understanding of how contemporary natural populations are structured. Expanding genetic databases coupled with versatile techniques such as PCR make it possible to undertake sophisticated genetic analyses in species for which no previous genetic information is available. The primary challenge now to population geneticists is to integrate information derived from a suite of molecular markers with an understanding of the evolutionary history of the species, as well as the physical and biotic forces affecting its demography.

Acknowledgements

I thank Paul Rodhouse for the introduction to problems of population structure in Antarctic species, for the opportunity to participate in the GeneFlow Cruise, and for samples of squid and icefish. Bob Ward and Peter Smith kindly provided toothfish samples. Stacy DeRuiter and JinPing Wang provided excellent technical assistance in PCR and RFLP analyses. Thanks also to the referees, T. Patarnello, J. Fenton and A. Rogers for their help. This work was supported by NSF grant OPP-9714439.

References

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. & Saunders, N.C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Avise, J.C., Bowen, B.W. & Lamb, T. 1989. DNA fingerprints from hypervariable mitochondrial genotypes. *Molecular Biology and Evolution*, **6**, 258–269.

Bonnaud, L., Boucher-Rodoni, R. & Monnerot, M. 1994. Phylogeny of decapod cephalopods based on partial 16S rDNA nucleotide sequences. *Comptes Rendus Academie de Science Paris Série III*, **317**, 581–588.

Bonnaud, L., Boucher-Rodoni, R. & Monnerot, M. 1997. Phylogeny of cephalopods inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **7**, 44–54.

Brierley, A.S., Rodhouse, P.G., Thorpe, J.P. & Clarke, M.R. 1993. Genetic evidence for population heterogeneity and cryptic speciation in the ommastrephid squid *Martialia hyadesi* from the Patagonian Shelf and Antarctic Polar Frontal Zone. *Marine Biology*, **116**, 593–602.

Campbell, N.J.H., Harriss, F.C., Elphinstone, M.S. & Baverstock, P.R. 1995. Outgroup heteroduplex analysis using temperature gradient gel electrophoresis: high resolution, large scale, screening of DNA variation in the mitochondrial control region. *Molecular Ecology*, **4**, 407–418.

Carlini, D.B. & Graves, J.E. 1999. Phylogenetic analysis of cytochrome *c* oxidase I sequences to determine higher-level relationships within the coleoid cephalopods. *Bulletin of Marine Science*, **64**, 57–76.

Chen, L., DeVries, A.L. & Cheng, C.-H.C. 1997. Evolution of antifreeze glycoprotein gene from a trypsinogen gene in Antarctic notothenioid fish. *Proceedings of the National Academy of Sciences of the USA*, **94**, 3811–3816.

Chen, W.-J., Bonillo, C. & Lecointre, G. 1998. Phylogeny of the Channichthyidae (Notothenioidei, Teleostei) based on two mitochondrial genes. *In* di Prisco, G., Pisano, E. & Clarke, A., *eds. Fishes of Antarctica: a biological overview*. Milan: Springer Verlag, 287–298.

Chow, S. 1998. Universal PCR primer for calmodulin gene intron in fish. *Fisheries Science*, **64**, 999–1000.

Chow, S. & Hazama, K. 1998. Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology*, **7**, 1255–1256.

Duhamel, G., Ozouf-Costaz, C., Cattaneo-Berrebi, G. & Berrebi, P. 1995. Interpopulation relationships in two species of Antarctic fish *Notothenia rossii* and *Champsocephalus gunnari* from the Kerguelen Islands: an allozyme study. *Antarctic Science*, **7**, 351–356.

Hayashi, K. & Yandell, D.W. 1993. How sensitive is PCR-SSCP? *Human Mutation*, **2**, 338–346.

Hedrick, P.W. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

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

Lessa, E.P. 1992. Rapid surveying of DNA sequence variation in natural populations. *Molecular Biology and Evolution*, **9**, 323–330.

Levitt, R.C. 1991. Polymorphisms in the transcribed 3' untranslated region of eukaryotic genes. *Genomics*, **11**, 484–489.

Meyer, A., Kocher, T.D., Basasibwaki, P. & Wilson, A.C. 1990. Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, **347**, 550–553.

Nauta, M.J. & Weissing, F.J. 1996. Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics*, **143**, 1021–1032.

NORMAN, J., MURPHY, J., PIERCE, G. & BOYLE, P. 1994. Preliminary molecular genetic analysis of stock structure in the squid *Loligo forbesi*. *ICES Council Meeting 1994/K*:23.

ORBACZ, E.A. & GAFFNEY, P.M. 2000. Genetic structure of tautog (*Tautoga onitis*) populations assayed by RFLP and DGGE analysis of mitochondrial and nuclear genes. *Fishery Bulletin*, **98**, 336–334.

OVENDEN, J.R., SMOLENSKI, A.J. & WHITE, R.W.G. 1989. Mitochondrial DNA restriction site variation in Tasmanian populations of orange roughy (*Hoplostethus atlanticus*), a deep-water marine teleost. *Australian Journal of Marine and Freshwater Research*, **40**, 1–9.

PALUMBI, S., MARTIN, S., ROMANO, S., MCMILLAN, W.O., STICE, L. & GRABOWSKI, G. 1991. *The simple fool's guide to PCR*, 2nd ed. Department of Zoology, University of Hawaii, 47 pp.

PALUMBI, S.R. 1996. Nucleic Acids II: the polymerase chain reaction. *In* HILLIS, D.M., MORITZ, C. & MABLE, B.K., *eds*. *Molecular systematics*. Sunderland: Sinauer Associates, 205–247.

PALUMBI, S.R. & BAKER, C.S. 1994. Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution*, **11**, 426–435.

PARK, L.K., BRAINARD, M.A., DIGHTMAN, D.A. & WINANS, G.A. 1993. Low levels of intraspecific variation in the mitochondrial DNA of chum salmon (*Oncorhynchus keta*). *Molecular Marine Biology and Biotechnology*, **2**, 362–370.

PENA, S.D.J., BARRETO, G., VAGO, A.R., DE MARCO, L., REINACH, F.C., NETO, E.D. & SIMPSON, A.J.G. 1994. Sequence-specific "gene signatures" can be obtained by PCR with single specific primers at low stringency. *Proceedings of the National Academy of Sciences of the USA*, **91**, 1946–1949.

PROSSER, J. 1993. Detecting single-base mutations. *Trends in Biotechnology*, **11**, 238–246.

QUATTRO, J.M. & JONES, W.J. 1999. Amplification primers that target locus-specific introns in actinopterygian fishes. *Copeia*, **1999**, 191–196.

RICO, C., RICO, I. & HEWITT, G. 1996. 470 million years of conservation of microsatellite loci among fish species. *Proceedings of the Royal Society of London*, **B263**, 549–557.

RODHOUSE, P.G. 1997. *RRS* James Clark Ross *cruise 26, 16 Nov–15 December 1997. Gene Flow cruise*. Cambridge: British Antarctic Survey, 120 pp.

SHEFFIELD, C.V., COX, D.R. & MYERS, R.M. 1990. Identifying DNA polymorphisms by denaturing gradient gel electrophoresis. *In* INNIS, M.A., GELFAND, D.H., SNINSKY, J.J. & WHITE, T.J. *PCR protocols: a guide to methods and applications*. San Diego: Academic Press, 206–218.

SLADE, R.W., MORITZ, C., HEIDEMAN, A. & HALE, P.T. 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. *Molecular Ecology*, **2**, 359–373.

SMALL, D.J., VAYDA, M.E. & SIDELL, B.D. 1998. A novel vertebrate myoglobin gene containing three A+T-rich introns is conserved among Antarctic teleost species which differ in myoglobin expression. *Journal of Molecular Evolution*, **47**, 156–166.

TOMAREV, S.I., ZINOVIEVA, R.D., GUO, K. & PIATIGORSKY, J. 1993. Squid glutathione S-transferase. Relationships with other glutathione S-transferases and S-crystallins of cephalopods. *Journal of Biological Chemistry*, **268**, 4534–4542.

VENKATESH, B., NING, Y. & BRENNER, S. 1999. Late changes in spliceosomal introns define clades in vertebrate evolution. *Proceedings of the National Academy of Sciences of the USA*, **96**, 10 267–10 271.

WILLIAMS, R., SMOLENSKI, A.J. & WHITE, R.W.G. 1994. Mitocondrial DNA variation of *Champsocephalus gunnari* Lönnberg (Pisces: Channichthyidae) stocks on the Kerguelen Plateau, southern Indian Ocean. *Antarctic Science*, **6**, 347–352.

WRIGHT, J.M. & BENTZEN, P. Microsatellites: genetic markers for the future. *Reviews in Fish Biology and Fisheries*, **4**, 384–388.