International Council for the
Exploration of the Sea

C.M.1994/M:5
Ref.:G+H+K

REPORT OF THE STUDY GROUP ON STOCK IDENTIFICATION PROTOCOLS FOR FINFISH AND SHELLFISH STOCKS

Copenhagen, 16-19 August 1994

This document is a report of a Working Group of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council. Therefore, it should not be quoted without consultation with the General Secretary.

*General Secretary
ICES
Palægade 2-4
DK-1261 Copenhagen K
DENMARK

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Main Tasks

At its 1993 Statutory Meeting, ICES resolved (C.Res. 1993/2:63) that A Study Group on Stock Identification Protocols for Finfish and Shellfish Stocks will meet under the chairmanship of Dr. K. Friedland (USA) in Lowestoft, UK from 15-18, August 1994. (The venue was subsequently changed to ICES Headquarters.) The Study Group would consider questions related to stock identification methodology and protocols for the application of stock identification techniques in fisheries management advice (Appendix 1).

The Study Group considered 9 papers submitted by participants (Appendix 2). References cited in the report are given in Appendix 3.

The Study Group considered a range of case studies relevant to its terms of reference and continued work on its Stock Identification Methodology. The application of allozymes and newly developed DNA methods, termed RAPD, were examined with examples from both invertebrate and vertebrate fisheries. In addition, the use of shape analysis of herring otoliths to separate stocks and the use of parasite as biologic tags for fish stocks were also discussed. Participants contributed six new sections to the Stock Identification Methodology. The new topics included genetic techniques, analysis of parasites, and statistical algorithms for stock composition analysis. It is significant to note that three of these new contributions were made via correspondence, thus demonstrating the importance and functionality of this mode of participation. As new material is prepared for the Methodology, an increasing emphasis will be placed on review and consensus building on the content of sections.

1.2 Participants

P. Abaunza	Spain
R. Fitzgerald	Ireland
K. Friedland (Chairman)	USA
O. Hagström	Sweden
D. Libby	USA
A. Sanjuan	Spain
P. Smith	New Zealand

A list of addresses is given in Appendix 4.

2 RECENT ADVANCES IN STOCK IDENTIFICATION METHODOLOGY AND CASE STUDIES

2.1 Allozyme Polymorphism in Cephalopods

Allozyme polymorphism in Galician cephalopods (Sepia)

The Study Group considered a comparative study of genetic variation in two species of cuttlefish from the Northwest Iberian Peninsula (Galician waters). The species have different distribution patterns with *Sepia officinalis* restricted to shallow water in rias and *S. orbignyana* with a more continuous distribution in deeper water on the shelf down to 200 m. For each species two samples of about 30 animals were collected from the north and south of the area and tested for 30 enzyme loci. A low level of genetic variation was found in both species. For the few polymorphic ($P<0.95$) loci (2 in *S. officinalis* and 4 in *S. orbignyana*) there were no significant differences between areas. It was concluded that allozyme markers are not useful genetic tools for discriminating populations of Galician cephalopods from the Iberian Peninsula. Other genetic methods, such as sequencing mtDNA, will be tested and samples collected from a wider part of the range of the two species.

Discussion by the Study Group suggested that there may be sufficient biological evidence to manage *S. officinalis* as separate stocks. The species has a restricted distribution in shallow water rias. Spawning occurs at the head of the rias and the demersal eggs hatch directly into the adult form with no planktonic stage. Fishing patterns suggest that adults do not make extensive movements. In addition there are different oceanographic conditions between the northern and southern rias.

Allozyme polymorphism in cephalopods

During the Study Group meeting, a review was made of allozyme studies on cephalopods. Allozymes have been studied in the families Loliginidae, Ommastrephidae, Gonatidae, Sepildae and Nautiloidae. Low levels of genetic variation have been found in most species when measured by the percentage loci polymorphic, mean heterozygosity and mean number of alleles per locus (Garthwaite *et al.*, 1989; Brierley *et al.*, 1993; Katugin 1993; Perez-Losada 1993; Yeatman and Benzie 1994, see Table 2.1.1), with the exception of *Nautilus* (Woodruff *et al.*, 1993) and *Berryteuthis* (Katugin 1993). The low level of genetic variability in most Cephalopods limits the power of allozyme techniques for stock identification studies. It is possible that more extensive allozyme studies utilising more enzymes and a wider range of tissues and buffer systems may detect genetic variation that would be applicable as population markers. Alternatively, the application of DNA techniques may provide genetic markers for stock identification of cephalopods.

Several examples of cryptic speciation, hitherto undetected species, have been revealed in allozyme studies of cephalopods. Cryptic species have been found in the genera *Photololigo* (Yeatman and Benzie 1993, 1994), *Loligo* (Augustyn and Grant 1988, Garthwaite *et al.*, 1989), *Illex* (Carvalho *et al.*, 1992) and *Nototodarus* (Smith *et al.*, 1981).

2.2 Random Amplified Polymorphic DNA

The Study Group reviewed the potential application of random amplified polymorphic DNA (RAPD) (ref. Section 116 in Appendix 5) in fisheries studies. RAPD markers generated with 10-base oligonucleotide primers were tested as a tool for stock and species identification. Amplification products were separated in agarose gels and stained with ethidium bromide. Results from RAPD analyses were compared with results from isoelectric focusing and allozyme electrophoresis.

Different DNA fragments were detected in two species of bluefin tuna (*Thunnus thynnus* and *T. maccoyii*) which could not be distinguished by conventional isoelectric focusing of muscle proteins. Out of 32 primers, 4 produced different DNA fragments between the two species. The average percent difference between pooled samples of the two species was 6.3%. Applying the diagnostic primers to identify northern bluefin tuna, *T. thynnus*, in New Zealand waters, it was shown that specimens recorded as northern bluefin tuna in the 1992 and 1994 fishery were *T. maccoyii*.

To evaluate RAPD's as a stock identification tool, 130 arbitrarily chosen primers were tested in samples of orange roughy *Hoplostethus atlanticus* from six geographically isolated sites off the east coast of New Zealand. No polymorphisms were detected with 118 primers. Fish from all six sites shared the same fragment patterns. A few primers (4/118) showed variation of weakly staining fragments. However, repeat tests on the same individuals did not produce repeatable DNA fragment patterns for these weakly staining fragments. The strongly staining fragments were consistent in repeat tests. Conventional allozyme techniques on orange roughy samples from the same sites revealed significant heterogeneity off the east coast of New Zealand.

The RAPD method is technically more demanding and more expensive than allozyme electrophoresis, but is easily accommodated in any laboratory experienced in RFLP studies. Pooling of DNA samples permits a rapid screening of a large number of primers and takes no more laboratory time than conventional allozyme screening. Using this approach, the RAPD technique clearly distinguishes closely related species of bluefin tunas that cannot be separated by conventional iso-electric focusing of proteins. Thus RAPD's are likely to be used as a tool for taxonomic problems and for identifying fish fillets and fish products. In addition, because the technique works with minute quantities of DNA, it will be applicable to identifying fish eggs and larvae.

From the experience in New Zealand, it appears the RAPD technique may have less resolving power than other DNA methods for stock separation with marine species. Screening of 130 primers found no genetic markers that would allow separation of orange roughy *Hoplostethus atlanticus* samples into stock units. In contrast conventional allozyme polymorphisms tested in larger samples from the same sites indicated several stock units. It is possible that screening a larger number of primers, or smaller size primers, may detect variation that must exist in the DNA as indicated by the allozyme markers.

2.3 Otolith Database

Fish otoliths are widely used in scientific studies including fisheries biology (taxonomy, population dynamics and stock assessment protocols), ecology (predator prey interactions), and even in archaeology. Because of these applications most marine and fisheries laboratories maintain large collections of otoliths including selected parts as reference records. However, establishing, maintaining and updating these collections is time consuming and labour-intensive and not particularly easy to excess. A computerized system holding this information could have useful application in fisheries science.

A prototype of a otolith database developed at the Aquaculture Development Center, Ireland, was presented to the Study Group. The system comprises computerized image processing and visual database utilizing commercially available software. The otoliths are shown as colour images, sometimes in multiple examples to provide a demonstration of variation between individual specimens. In addition, numerical data related to the otolith and the donor fish are very easily available. The system is operated via a user-friendly, interactive, menu-based format with a graphical user interface. This allows the user, when searching the database, to review both visual and numerical data by a variety of window-options. A related application has been developed to facilitate inter-laboratory calibration of fish age determination in Ireland.

The Study Group sees a large variety of applications of this type information display and strongly supports further development of the present database. It was also suggested that the developers of this system consider the parameters of image standardization relevant to ageing, image processing, and shape analysis. It is envisioned that this system and image database could form the basis of interdisciplinary studies and exchanges, thus, the widest possible input into image standardification will hopefully yield the widest use in future application.

2.4 Shape Analysis of Larval Herring Otoliths

The Study Group considered a study of shape analysis of herring larvae otoliths for stock discrimination in the Gulf of Maine. The herring spawning in the Gulf of Maine is dominated by autumn spawners. The spawning areas are well known, but the degree of mixing of her-

ring larvae from different spawning grounds is unknown. Information concerning the degree of mixing of larvae from these populations would provide valuable information on the relative importance of the different spawning populations in providing juvenile recruits to the stock complex.

Otolith morphology changes with the fish's growth and age. Larvae with major otolith diameter of 40 to 90 microns and 15-30 increment counts (daily rings) were selected to reduce the effect of ontogenic factors on the analysis. The analysis comprised both size measurements of the otoliths, like area, perimeter and length of major axis. Other variables collected were amplitude and phase angle from the Fourier transform. To describe all otoliths equally, the same landmark should be used as the origin for the perimeter trace. Larvae otoliths are round without obvious landmarks and two different methods were applied in the analysis, one based on the otolith's centroid and another on the nucleus. The point of the minimum axis was used to establish a common reference point and the shape was oriented and re-plotted. The analysis was performed on both left and right otoliths for each geographic area. Twenty-three variables for all otoliths were then compared between areas with the use of discriminant function analysis. About twenty otoliths from each area were used as the training set and the remaining otoliths were used in the test set.

The result showed no significant differences in the size variable or relative growth rate between areas. The discriminant function analysis gave a maximum classification success of 84% for test otoliths. The highest classification success was achieved using all variables derived from the right otolith of the larvae. All other combination of variables gave lower classification success. The results were insensitive to the node center location.

Although the study is based on a small number of otoliths, the results seems very promising and it was suggested that the methods should be tested on juvenile and adult herring if samples with known area of origin could be obtained. A stepwise approach to explore the discriminating power of the different variables could give more insight into which factors differentiate populations and/or areas. It was also stressed the importance to base the comparison on the same size or age of fish in this kind of analysis. The Study Group notes that other descriptors of shape, such as morphologic indices and complex Fourier shape analysis, were not attempted in this study and encourages their testing.

2.5 Further Consideration of Parasites as Biological Tags

The Study Group reviewed work on the incidence of the parasite *Anisakis simplex* in the horse mackerel, *Trachurus trachurus* (L), in the Bay of Biscay. The analysis showed differences in total abundance of *A. simplex* collected from specimens between three areas. It was pointed out that this difference could be spurious when samples are not analysed by age and/or length. Parasite numbers were shown to increase with age and length in horse mackerel. Statistical analysis involved testing the parasite occurrence distributions between areas using the Wilcoxon-Mann-Whitney non-parametric test. Significant differences in parasite abundances were found after stratifying samples by age and length. Age was determined to be the most important factor and length the lesser factor in the analysis. When using abundance of parasites as a means to differentiate groups of fish, care must be taken to filter out accumulating effects from large ranges in age or size, etc., in order to test only samples that are similar.

The Group discussed other considerations in the application of parasites in fish stock identification:

1) Further investigation into the life history and genetics of the parasite in question, if parts that are critical to the parasite/host relationship are not known. As in the case of *A. simplex*, what forms of zooplankton, and possibly other organisms, serve as intermediate host?

2) The use of multiple parasite species is believed to increase the resolution of the technique.

3) The benefits of finding "tag" parasites that only infest certain life stages of the host. This could be useful in discerning juvenile migratory behaviour, especially when coming into contact and mixing with other juvenile or adult groups.

4) Characterize parasite abundance throughout the hosts migratory range in order to correlate the variability of any changes in parasite abundance to the abundance of the host.

5) The addition of other biological (i.e. age, length, weight, condition, etc.) and morphological (shape, truss analysis, etc.) data on the host in relation to parasite infestation.

6) Under the present regime of techniques used in parasite-based stock identification studies, timeliness in respect to the assessment process remain an issue.

7) The Study Group encourages ongoing and new research to improve the efficiency of parasite assay methods.

The above considerations are suggested guidelines for the development of parasite biological "tag" information.

3 STOCK IDENTIFICATION METHODOLOGY

The Study Group considered six new contributions to the Stock Identification Methodology (see Appendix 5). The sections are in various levels of completion, and will hopefully attract the attention and input of other experts in the field. In addition, the Table of Contents of the Methodology has been modified to reflect the addition of new sections and some minor reorganization. The Study Group desires the widest possible participation in this effort and encourages experts to contact the Chairman if they can contribute to the content or review of an existing section, or have suggestions of new sections that may be appropriate for inclusion in this compendium.

4 RECOMMENDATIONS

4.1 Recommendations

The Study Group makes the following recommendations.

1. The Study Group encourages the continued development of the database of otolith images and suggests that the widest possible sampling of potential users be surveyed so that image and database formats may have broad application. The Study Group chairman will communicate the names of prospective reviewers to the database designers.

2. It is anticipated that this Study Group will require a series of meetings to complete its goal of assembling a comprehensive Stock Identification Methodology. The Study Group encourages participants to continue to work with the Study Group by correspondence and to track the progress of the Group for meetings of special interest.

3. The Study Group recognizes that a comprehensive Stock Identification Methodology will have application outside the ICES area and will draw upon the findings of experts not directly involved in ICES research or assessments. The Study Group suggests cooperation on the Stock Identification Methodology be explored with other organizations, such as PICES in the Pacific area, to facilitate the exchange of information and to draw upon the widest base of expertise.

4.2 Suggested Terms of Reference for Next Meeting

The Study Group on Stock Identification Methodology for Finfish and Shellfish Stocks (Chairman: Dr. K. Friedland, USA) date and location to be decided, should:

a) continue to describe methodology for the application of stock identification and classification data, as they may vary by species, fisheries, and life history characteristics. During its next meeting, the Study Group should concentrate its efforts on the areas of genetics, parasites as biological tags, classification statistics, and life history traits.

d) report to the Demersal Fish, Pelagic Fish, Baltic Fish, Anadromous and Catadromous Fish, and Shellfish committees at the 1996 Statutory Meeting.

4.3 Future Meeting

The Study Group recommends a third meeting during the fall of 1995, and that the possibility of meeting at the Fisheries Laboratory in Aberdeen, U.K. be explored.

Table 2.1.1 Genetic variation in Cephalopod species measured as the percentage of loci polymorphic at 1% level (P), the mean heterozygosity (H), and the mean number of alleles per locus (Na).

Taxon	No. loci	No. individ.	P	H	Na	Ref
Loliginidae						
Loligo opalescence	30	45	17	0.037	1.06	1
L. pealei	19	40-994	5	0.006	1.01	2
L. plei	9	8	0	0.000	1.00	2
L. vulgaris reynaudii	30	44	23	0.030	1.05	1
L. vulgaris vulgaris	30	15	7	0.011	1.01	1
L. gahi	21	277-1031	29	0.059	1.09	3
Lolliguncula brevis	9	8	0	0.000	1.00	2
Ommastrephidae						
Illex illecebrosus	11	10-156	9	0.005	1.01	4
Ommastrephes bartramii	35	41	46	0.066	1.08	5
Nototodarus sloani	9	36-100	22	0.060	1.11	5
Gonatidae						
Berryteuthis magister	23	440	43	0.131	1.33	5
Sepiidae						
Sepia orbignyana	30	60	23	0.043	1.25	6
S. officinalis	30	60	7	0.029	1.07	6

1)Augustin and Grant 1988; 2) Garthwaite *et al.*, 1989; 3) Carvalho and Loney 1989; 4) Romero and Amaratunga 1981; 5) Katugin 1993; 6) Sanjuan *et al.*, data reported this meeting.

APPENDIX 1

Study Group on Identification Protocols for Finfish and Shellfish

Terms of Reference

2:27 A Study Group on Stock Identification Protocols for Finfish and Shellfish Stocks (Chairman: Dr. K. Friedland, USA) will meet in Lowestoft, England, UK from 15-18 August 1994 to:

a) review, describe, and evaluate established methodologies to discriminate or define finfish and shellfish stocks and, in particular, evaluate different methods used to separate herring stocks and apply recent mathematical advances in pattern recognition and classification on available meristic and morphometric data;

b) continue to describe protocols for the application of stock identification and classification data, as they may vary by species, fisheries, and life history characteristics;

c) evaluate the usefulness of the data derived from different genetic methods for stock composition analysis;

d) report to the Demersal Fish, Pelagic Fish, Baltic Fish, Anadromous and Catadromous Fish, and Shellfish committees at the 1994 Statutory Meeting.

APPENDIX 2

Study Group on Identification Protocols for Finfish and Shellfish

DOCUMENTS SUBMITTED TO THE STUDY GROUP

W.Doc.1 *Sanjuan, A., Perez-Losada, M. and Guerra, A.* Preliminary results on allozyme polymorphism in *Sepia officinalis* and *S. Orbignyana* (Mollusca: Cephalopoda) populations from Galician Waters (NW of the Iberian Peninsula).

W.Doc.2 *Abaunza, P. and Villamor, B.* Further considerations into the usefulness of parasites as biological tags for marine fish stock discrimination, with special reference on use of *Anisakis simplex* (L3)

W.Doc.3 *Smith P.J., McVeagh, S.M., Benson, P.G. and Winn, R.F.* Random amplified polymorphic DNA: a potential fisheries tool for species and stock identification.

W.Doc.4 *Friedland, K.D.* Analyses of Calcified Structures-Texture and Spacing Patterns.

W.Doc.5 *Campana, S.* Chemical Composition of Body Parts and Tissues-Otolith Elemental Fingerprinting.

W.Doc.6 *Koljonen, M.-L.* Genetic Analysis-Allozymes.

W.Doc.7 *Prager M.H.* Comments on Statistical Algorithms for Stock Composition Analysis.

W.Doc.8 *Fitzgerald, R.* Prototype Otolith Database.

W.Doc.9 *Libby, D.* Discriminating Larval Herring Otoliths Using Shape Analysis.

APPENDIX 3

Study Group on Identification Protocols for Finfish and Shellfish

REFERENCES

Augustyn, C.J. and Grant, W.S. 1988. Malacologia Vol. 29, pp. 215-233

Brierley, A.S., Thorpe, J.P., Clarke, M.R., Martins, H.R. 1993. In Recent Advances in Cephalopod Fisheries Biology, Okutani, T., O'Dor, R.K., and Kubodera, T. (eds.), Tokai Univ. Press, Tokyo, pp. 61-69

Carvalho, G.R., and Loney, K.H. 1989. J. Exp. Mar. Biol. Ecol., Vol. 126, pp. 231-241

Carvalho, G.R., and Pitcher, T.J. 1989. J. Exp. Mar. Biol. Ecol., Vol. 126, pp. 243-258

Carvalho, G.R., Thompson, A., and Stoner, A.L. 1992. J. Exp. Mar. Biol. Ecol., Vol. 158, pp. 105-121

Garthwaite, R.L., Berg, Jr. C.J., and Harrigan, J. 1989. Biol. Bull. Mar. Biol. Vol. 177, pp. 287-294

Katugin, O.N. 1993. In Recent Advances in Cephalopod Fisheries Biology, Okutani, T., O'Dor, R.K., and Kubodera, T. (eds.), Tokai Univ. Press, Tokyo, pp. 201-213

Kubodera, T. (eds.), Tokai Univ. Press, Tokyo, pp. 641-652

Pérez-Losada, M., 1993. Polimorfismos alozímicos y relaciones filogeneticas de nueve especies de cefalópodos de las familias Sepiidae, Ommastrephidae y Loliginidae, Fac. Ciencias, Univ. Vigo, 130 pp.

Smith, P.J., Roberts, P.E., Hurst, R.J. 1981. N. Z. J1 mar. Freswat. Res., Vol. 15, pp. 247-253

Romero, M.C.L., and Amaratunga, T. 1981. NAFO SCR Doc. 81/IX/103, 13 pp.

Woodruff, D.S., Mulvey, M., Saunders, W.B., Carpenter, M.P. 1993. Proc. Acad. Nat. Sci. Philad., Vol. 135, pp. 147-153

Yeatman, J. & Benzie, J.A.H. 1993. In Recent Advances in Cephalopod Fisheries Biology, Okutani,T., O'Dor, R.K. & Kubodera, T. (eds.), Tokai Univ. Press, Tokyo, pp. 641-652

Yeatman, J., and Benzie, J.A.H. 1994. Mar. Biol., Vol. 118, pp. 7987

APPENDIX 4

LIST OF PARTICIPANTS

Study Group on Identification Protocols for Finfish and Shellfish

15 - 18 August 1994

Name	Address	Telephone	Fax	E-mail
Pablo Abaunza	Inst. Español de Oceanografía Centro de La Coruña PO Box 130 15080 La Coruña Spain	+34 81 205 362 or +34 81 205 366	+34 81 229077	
Richard Fitzgerald	Aquaculture Development Centre University College Cork Ireland	+353 21 276871 Ext. 4936/4955	+353 21 277 922	nadc@ucc.ie
Dr K. Friedland **(Chairman)**	Northeast Fisheries Science Center NMFS/NOAA Woods Hole, MA 02543 USA	+1 508 548 5123	+1 508 548 1158	kfriedla@whsunl.wh.whoi.edu
Olle Hagström	Institute of Marine Research Box 4 S-45321 Lysekil Sweden	+46 523 14180	+46 523 13977	
David Libby	Maine Department Marine Resources W. Boothbay Harbour, ME 04575 USA	+1 207 633 9500	+1 207 633 9579	
Andres Sanjuan	Laboratorio de Xenetica Facultade de Ciencias Apdo 874 Universidade de Vigo 36200 Vigo Spain	+34 86 81 2273 or +34 86 81 2317	+34 86 81 23 82	sanjuan@cesga.seinv.es
Peter Smith	MAF Fisheries P.O. Box 297 Wellington New Zealand	+ 04 3861 029	+04 3860 572	sfpjs@frc.maf.govt.nz

APPENDIX 5

STOCK IDENTIFICATION METHODOLOGY

TABLE OF CONTENTS

2 Introduction

There have been many excellent reviews of marking techniques and stock identification (Kumpf *et al.*, 1985; Parker *et al.*, 1990; Anon., 1993), but what is still unavailable to fisheries scientists is a synthetic overview of these subjects with a bent towards application. Many of the reports on stock identification are result-oriented case studies that are too narrowly focused or, conversely, overview perspectives lacking the detail needed to guide researchers. The Study Group's Terms of Reference ask that a protocol for stock identification be developed. To attempt to address that request, the Study Group will, over the necessary course of its meetings, assemble a series of reports organized around the central theme of defining protocols for doing stock identification research. The authors of these protocol reviews will attempt to explain the application of the methodology as currently accepted by the scientific community while providing worked examples as appropriate and a listing of the important literature references.

When the Study Group has completed its work, recipients of the reports will have a compendium of reviews that will hopefully serve as a valuable source document on stock identification for a number of years. If the Study Group has worked productively, it may be possible to have the series published in some other media. In any event, what should be encouraged is widespread participation and aggressive review and revision of these reports.

References

Anon. 1993. Biogeographical identification of English Channel fish and shellfish stocks. Report by IMFREMER Centres and MAFF Directorate to the EC Commission DG XIV, March 1993, 191 pp.

Kumpf, H.E., R.N. Vaught, C.B. Grimes, A.G. Johnson, and E.L. Nakamura. 1985. Proceedings of the Stock Identification Workshop. NOAA Tech. Mem. NMFS-SEFC-199, Panama City, Florida.

Parker, N.C., A.E. Giorgi, R.C. Heidinger, D.B. Jester, E.D. Prince, and G.A. Winans. 1990. Fish-Marking Techniques. American Fisheries Society Symposium 7, Bethesda, Maryland.

3 Organization of the Report

This report will be organized in an open format. The format is intended to invite and facilitate additions and revisions over time. Part of this open format will require that each section has its own numbering sequence so that new sections can be added in the future without renumbering following sections. Because this is a "source of information" document, it is important to attract the broadest range of expertise available for input on specific sections. This may be best accomplished by correspondence in some cases. At the same time, those of us most involved in the Study Group can greatly facilitate the effectiveness of this project by distributing information about the protocols manual to potential contributors.

It is important to remember that sections of the protocol are part of the Study Group report and subject to review and revision at the discretion of the Study Group. Section authors do not have special controls over these sections and must be prepared to defer to the consensus view of the Study Group.

Each section should be a stand-alone source of information including all the text, tables, figures and references for that section. Each section should begin with a chronology of revisions, and if appropriate, include an explanation of the revision changes.

The Study Group expects that there will be revisions made to the Table of Contents and does not suggest that the topic subheadings are an exhaustive list. The topics included are intended to express the range of topics the Study Group would hope to cover during its tenure. The Study Group is open to suggestions for additional topics and modification of existing topics.

4 Section Authorship

Each section should have a listing of the contributor(s) for each version or revision. It is suggested that these contributor lists form the basis of authorship if the report is developed into a cooperative report or book. Authorship should first depend on substantive contribution to the development and writing of the section. The order of authorship should depend on commonly recognized criteria; for example, those that do the most work are usually higher in the authorship order. Obviously, for sections authored by an individual or only a few people, authors will probably find it easier to form a consensus on authorship. A number of sections may undergo revision over time as new information relevant to the section becomes available. Some sections may, therefore, have a number of different authors with differing views on their relative contribution. If a group of authors fails to find a consensus on authorship order, they may consider alphabetical ordering or a random draw.

5 Preparation of Text

The ICES Secretariat currently uses WordPerfect 5.1 for DOS as its word processing system. It would be desirable for contributions to the methods protocol to be submitted in hard copy and as WordPerfect files. This does not prevent contributors from using other software systems as long as their text can be translated into an electronic format that can be imported by WordPerfect (i.e., documents could originate in MS Word and be translated to WordPerfect, or be translated into simple ASCII).

6 Preparation of Tables

Tables intended for inclusion in methods protocol sections can be prepared as text tables directly included in the text. If tables are complex and if they may require revision at some later date, a version in spreadsheet form (either .WK1 or . XLS formats) can accompany the text form.

7 Preparation of Figures

Figures intended for inclusion in the methods protocol sections can be submitted as original hardcopy. As formatting of these documents may evolve into a Windows-based format some time in the future, authors are also encouraged to submit versions of their hardcopy figures as metafile (.WMF) or bitmap (.BMP) formats as appropriate.

Figures can be an effective way of expressing ideas and concepts. The generous use of figures is enthusiastically encouraged.

8 Participation by Correspondence

The Study Group encourages the widest possible participation of experts in stock identification. Recognizing that not all potential contributors will be able to, or desire to, participate in Study Group meetings, contribution by correspondence is encouraged. Contribution by correspondence can be contributions to section content or review comments on sections already developed. Corresponding participants will be afforded the same authorship considerations as those that have participated directly in Study Group meetings.

The current Study Group chairman is Dr. Kevin Friedland and he can be contacted at the address below:

Dr. Kevin Friedland
National Marine Fisheries Service
166 Water Street
Woods Hole, Massachusetts 02543
USA

kfriedla@whsun1.wh.whoi.edu

Teleph: 508-548-5123
FAX: 508-548-1178

9 Guidance to Authors on Section Content

Sections are intended to be focused explanations of the salient features of selected topics related to stock identification methodology. The target audience is expected to have a basic understanding of fishery and resources management science, but it should be assumed that the audience is new to the specific topic addressed in the section. The sections should be written in sufficient detail for a scientific investigator to use the information as a point of departure on the use of the method. However, the sections are not intended to be in such detail that they would be used as an exact guide to the implementation of the technique or method. The sections should also serve as a source of reference for those interested in evaluating the application of stock identification data. For example, members of ACFM or other assessment oversight committees would hopefully find the reviews presented here useful in technical evaluations of a range of different stock assessments.

103 Analyses of Calcified Structures-Texture and Spacing Patterns

First Draft: Copenhagen, August, 1994

K. D. Friedland
National Marine Fisheries Service
166 Water Street
Woods Hole, Massachusetts 02543

Requires completion; discussion is preliminary, herring otolith references needed, expand methodology section to include development of variables.

Introduction

Texture and spacing patterns found in hard body parts is a well established technique of stock separation utilizing traditional laboratory techniques (Clutter and Whitesel, 1956; Henry, 1961; Mosher, 1963; Anas and Murai, 1969; Lear and Sandeman, 1980; Antere and Ikonen, 1983; Reddin et al. 1988; Lund and Hansen, 1991) and newer imaging methods (Cook, 1982; Barlow and Gregg 1991; Schwartzberg and Fryer 1989; Ross and Pickard 1990; Friedland *et al.* 1994). The approach is in part dependant upon the correlation between the growth of the animal and the calcified structure to be analyzed. For example, fish scale growth and the rate of circuli deposition are related (Doyle *et al.* 1987; Barber and Walker, 1988; Fisher and Pearcy, 1990). The variation of these features throughout the species' range, as a product of both genetic and environmental influences, become the discrimination characters that allow their use in stock identification.

Spacing patterns can be measured with conventional measurement techniques or with enhanced technologies such as image processing. Simple features, such as the distance from one life history transition zone to another, can be recognized visually and the distance recorded. For example, the distance from an otolith focus to a check marking the transition from juvenile to adult life stages could serve as a stock separation character. Spacing patterns may also be a more complex set of measurements requiring the use of image processing to achieve measurement accuracy and to deal with the volume of data generated. For example, the spacings or distances from a landmark to a large number of structures can be extracted. When this approach is applied to fish scales, the location of upwards of fifty circuli may be measured, thus generating forty-nine spacing variables per individual.

The texture of calcified body parts can be measured with physical probes or by the analysis of optical density profiles. The idea of developing a two or three dimensional map of a scale or otolith using a micro-stylus has been considered and appears technically feasible (*note: insert references*), however, it does not appear the approach has been attempted in stock identification research. On the other hand, using optical density profiles to represent texture has been applied. The idea was first proposed by Major et al. (1972) and simply involves the use of transmitted light from a specimen as a representation of surface features. For example, the circuli on a fish scale appears dark because they are surface features which are optically denser than inter-circuli spaces. The periodicity of circuli, or features of the optical profile, can be analyzed with frequency domain statistics.

Methodology

Spacing and Texture from Optical Density Profiles

A specimen is prepared and examined with an image processor equipped with an optics subsystem appropriate for the specimen. For small specimens, this may entail the use of a compound microscope, whereas for larger specimens, the use of a dissecting microscope may be sufficient. A path or line is identified for data extraction (Figure 103-1). The data path is selected to strategically represent a growth phase or life history feature comparable to other specimens. From this line, a luminescence profile of transmitted light is extracted which will be processed further. Manual methods or automated algorithms can then be used to identify features along the data path; in the case of fish scales, where circuli or annuli occur. The raw measurements can be saved for use as spacing indices or landmark distances. Alternatively, the luminesce profile can be treated as a frequency domain time series. A Fourier transform of the luminescence pattern can be calculated and various time series statistics from the transform could serve as classification model variables. For example, textural features of a scale could be expressed as the magnitudes of the Fourier transform (the magnitude is calculated as the square root of the sum of the squared coefficient for each harmonic). Other frequency domain statistics may also be used.

Discussion

One advantage of image processing spacing and texture data is that it is quantitative and objective in that measurements are automated and without the problems associated with manual scale readings (Lund, et al. 1989). Douglas, Minckley and Tyus (1989) suggested that qualitative characters are excellent features upon which to base group separations, and presented extensive data that demonstrated that even untrained observers show a high degree of feature interpretation. However, the automated approach does offer improvements since it addresses sources of procedural inaccuracy, such as those associated with reader fatigue, and removes any doubt of reader subjectivity from potentially sensitive management data. Manual scale reading and image processing techniques use essentially the same features of the scale or otolith to form information databases to classify the samples to origin. With image processing, the classification algorithm can be defined explicitly for review and there is a complete quantitative audit trail for each decision.

Spacing and texture data are undoubtedly influenced by the environment, therefor, it is essential to test the robustness of the variables based on these data for annual or long term sources of variability. For example, annual variation in climate and food resources are known to affect circuli deposition in Atlantic salmon (Reddin, et al. 1988). This has necessitated the reparameterized of classification models with that species (Reddin, Verspoor and Downton 1990). It is probably wise with any scale-based discrimination procedure to maintain reference collections so that classification models can be updated. Annual or longer term sources of variability will be irrelevant if the proper reference samples are collected and applied.

Imaging techniques offer to the investigator the ability to produce vast quantities of data. A number of investigators have begun to explore ways of reducing the number of variables generated by these sorts of analyses. One such approach has been to employ variable averaging, or expressing data as means of pairs or quadruplets of adjacent variables. In studies using circuli spacing data, Barlow and Gregg (1991) reported that model efficiency was similar or only slightly higher for averaged data. In a similar study, Friedland *et al.* (1994) reported similar or only slightly lower model classification efficiency. The appeal of averaged data lies in the anticipated robustness of models to the potential problem that information content may be dispersed over a number of adjacent variables. However, it not clear there is any advantage to treating the data in this manner and there may be a loss of information content when averaging is performed. Friedland et al. (1994) suggested there may be important information in individual imaging derived variables (spacing or magnitude values), and that pre-processing of variables, like averaging, should be applied on a case by case basis where it improves classification efficiency.

Friedland *et al.* (1994) identified a potential source of bias while using image processing techniques to extract circuli spacing data. When automated procedures are used to mark a circulus it is relatively infrequent that the circulus would not be marked or left unmarked after manual correction. However, circuli are often marked more than once and the double mark could escape manual correction. This has been observed to occur with circuli that are very wide or are of complex morphology. This error will tend to add to the total number of circuli for a specimen and decrease the circuli spacing for the adjacent circuli pairs. The direction of the bias would be to reduce circuli spacing for the specimen.

Precision with image processing techniques is very high, which is in contrast to approaches dependent on a scale reader. It is well known that fatigue, pattern of prior observations, and long term familiarity can affect the precision of scale readers (Lund, et al. 1989). Feature extraction with the image processor is identical regardless of stage of the analysis and it is also insensitive to problems created by changes in project personnel.

References

Anas, R.E. and Murai, S. (1969) Use of scale characteristics and a discriminant function for classifying sockeye salmon *Oncorhychus nerka* by continent of origin. International Pacific Salmon Fisheries Commission Bulletin 26.

Antere I. and Ikonen E. (1983) A method of distinguishing wild salmon from those originating from fish farms on the basis of scale structure. Council Meeting, International Council for the Exploration of the Sea. 1983/M:26, 6pp.

Barber, W. E. & Walker, R. J. (1988). Circuli spacing and annulus formation: is more than meets the eye? The case for sockeye salmon, *Oncorhynchus nerka*. Journal of Fish Biology **32**, 237-245.

Barlow C.G. and Gregg B.A. (1991) Use of circuli spacing on scales to discriminate hatchery and wild barramundi, *Lates calcarifer* (Bloch). Aquaculture and Fisheries Management **22**, 491-498.

Clutter, R. and Whitesel, L. (1956) Collection and interpretation of sockeye salmon scales. International Pacific Salmon Fisheries Commission Bulletin 9.

Cook. R.C. (1982) Stock identification if sockeye salmon (*Oncorhychus nerka*) with scale pattern recognition. Canadian Journal of Fisheries and Aquatic Sciences 39, 611-617.

Douglas M.E., Minckley W.L. and Tyus H.M. (1989) Qualitative characters, identification of Colorado River chubs (Cyprinidae: Genus *Gila*) and the "art of seeing well". Copeia 89(3), 653-662.

Doyle, R. W., Talbot, A. J. & Nicholas, R.R. (1987). Statistical interrelation of length, growth, and scale circulus spacing: appraisal of growth rate estimator for fish. Canadian Journal of Fisheries and Aquatic Sciences 44, 1520-1528.

Fisher, J.P. & Pearcy, W.G. (1990). Spacing of scale circuli versus growth rate in young coho salmon. Fishery Bulletin 88, 637-643.

Friedland, K.D., C. Esteves, L. P. Hansen, and R. A. Lund. (1994) Discrimination of Norwegian Farm, Ranch, and Wild Origin Atlantic Salmon by Image Processing. Fisheries Management and Ecology 1, 000-000. (in press)

Henry, K.A. (1961) Racial identification of Farser River sockeye salmon by means of scales and it application to salmon management. International Pacific Salmon Fisheries Commission Bulletin 12.

Lear, W.H. and Sandeman, E.J. (1980) Use of scale characteristics and discriminant functions for identifying continental origin of Atlantic salmon. Rapp. P.-v. Reun. Cons. int. Explor. Mer. 176:68-75.

Lund R.A. and Hansen L.P. (1991) Identification of wild and reared Atlantic salmon, *Salmo salar* L., using scale characters. Aquaculture and Fisheries Management 22, 499-508.

Lund R.A., Hansen L.P. and Järvi T. (1989) Identification of reared and wild salmon by external morphology, size of fins and scale characteristics. NINA Forskningsrapport 1, 1-54. (In Norwegian with English summary.)

Major, R.L., Mosher, K.H., and Manson, J.E. (1972) Identification of stocks of Pacific salmon by means of scale features. p.209-231. in: R.C. Simon and P.A. Larkin (eds.) The stock concept in Pacific salmon, H.R. MacMillan Lectures in Fisheries, Univ. of British Columbia, Vancouver, BC.

Mosher, K.H. (1963) Racial analysis of red salmon by means of scales. International Pacific Salmon Fisheries Commission Bulletin 11.

Reddin D.G., Stansbury D.E. and Short P.B. (1988) Continent of origin of Atlantic salmon (*Salmo salar* L.) at West Greenland. Journal du Conseilles 44, 180-188.

Reddin D.G., Verspoor E., Downton P.R. (1990) An integrated phenotypic and genotypic approach to stock discrimination of Atlantic salmon. Journal du Conseilles 47, 83-88.

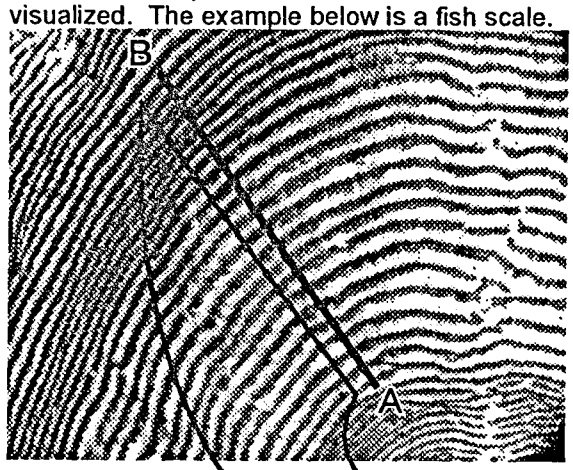
Ross W.R. and Pickard A. (1990) Use of scale pattern and shape as discriminators of wild and hatchery striped bass stocks in California. In: N.C. Parker, A.E. Giorgi, R.C. Heidinger, D.B. Jester, E.D. Prince and G.A. Winans (Eds) Fish-Marking Techniques. American Fisheries Society Symposium 7. Bethesda, Maryland. American Fisheries Society, pp 71-77.

Schwartzberg M. and Fryer J. (1989) Experiments in identifying hatchery and naturally spawning stocks of Columbia Basin spring chinook salmon using scale pattern analysis. Technical Report 89-3, Columbia River Inter-Tribal Fish Commission. 25pp.

Figure 103-1. Analyses of Calcified Structures-Texture and Spacing Patterns

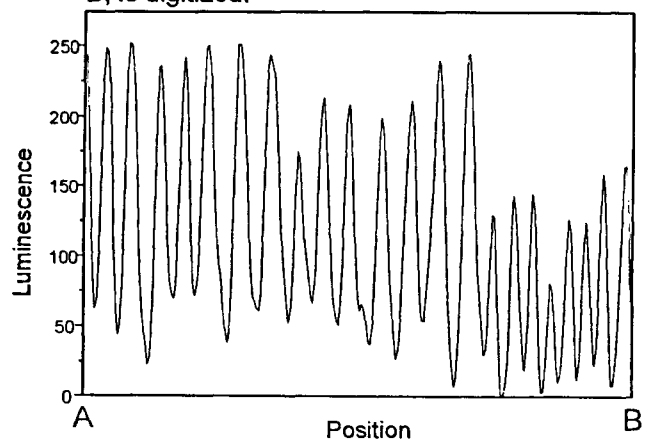
Image of Structure

Any structure can be analysed if annular, circuli, or other periodic marks can be visualized. The example below is a fish scale.

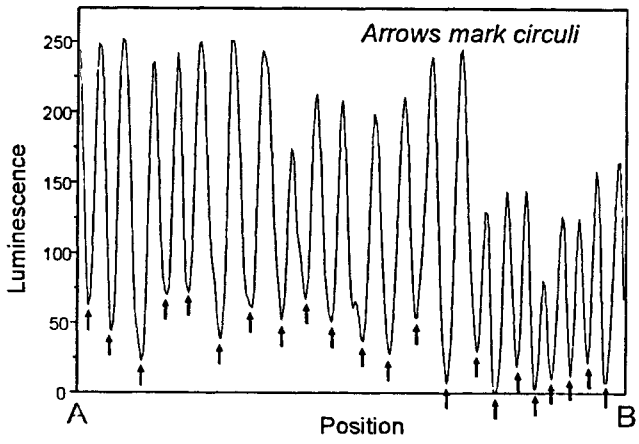


Luminescence Profile

The luminescence along a specific path, here represented by the line between points A and B, is digitized.

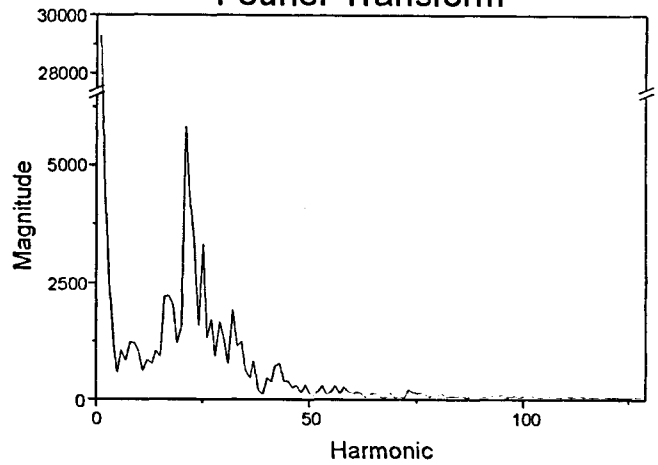


Circuli Spacing and Distance



Features from the luminescence profile are extracted. For example, valleys of luminescence or dark areas, which correspond to where circuli occur, are marked. The distance between circuli, or the distance from a reference point to a circulus, can be extracted.

Fourier Transform



The luminescence profile can be treated as a continuous time series. A Fourier transform of the series is performed and the frequency domain features of the time series extracted. For example, the magnitude or phase angle of each harmonic can be analyzed.

109 PARASITES AS BIOLOGICAL TAGS

First Draft: Copenhagen, August 1994

P. Abaunza
Inst. Espanol de Oceanografía
Centro de la Coruna
La Coruna, Spain

The methodology and discussion section of this chapter should be expanded. Additional references are also required

Parasites have been widely used as biological indicators of the population biology, migrations and phylogenetics of fish. One of these fields of application is that concerned with stock identity. The use of parasites has some advantages compared with other methods for stock discrimination such as artificial tagging (Williams *et al.*, 1992):

- They are most appropriate for studies of delicate or deepwater species or crustaceans where artificial tags cannot be applied.
- Samples may be obtained from routine sampling programmes and so are less expensive.
- Eliminate doubts concerning the possible abnormal behaviour of artificially tagged hosts.

The close relationship that exists between the parasite and its host makes it possible to use the former to learn more about the biology of the latter, but limitations exist as pointed out by Sinderman (1983).

- Parasite life cycles are often complex or unknown.
- Parasite ecology which involves one or more hosts and the environmental conditions is even more complex.
- Parasite identification is often uncertain or subject to taxonomic disagreement.

In addition:

- In some cases it is necessary to know the age of the host and there are several fish species in which the age reading technique is not validated.
- An ideal natural tag may be described as possessing the following characteristics taken from Kabata (1963), Sinderman (1983), MacKenzie (1987) and Williams *et al.* (1992):
- The parasite should have significantly different levels of infestation in the subject host in different parts of the study area. The following criteria: prevalence, mean intensity and abundance, (see Margolis *et al.*, 1982), must be taken into account at least when infestation levels are analyzed.
- The parasite should persist in the host for at least the duration of the study period and preferably longer.

- The life cycle of the parasite should preferably involve only one host.

- The infestation level of a tag parasite should remain relatively stable from season to season and from year to year. However, seasonal variations can determine seasonal migrations of the subject host.

- The parasite should be detected and easily identified. The examination of the host for a tag parasite should involve the minimum of dissection.

- Pathogenic parasites which affect host behaviour should be avoided.

A natural tag with all of the above attributes is rarely achieved, and compromises must be made. Departure from the ideal may be offset by the use of several different parasites simultaneously (Sinderman, 1983).

Different parasite species have been used as tags in freshwater, anadromous, and marine fish and in marine invertebrates as described in the review of Williams *et al.* (1992). Protozoans, larval helminths and adults (Platyhelminthes, Nematoda, Acanthocephala) and crustaceans, have been used as tags in fish species, such as herring, *Clupea harengus* (Chenoweth *et al.*, 1986; MacGladdery, 1987); mackerel, *Scomber scombrus* (MacKenzie, 1990); Pacific mackerel, *Scomber australicus* (Rohde, 1987); Sockeye salmon, *Oncorhynchus nerka* (Bailey *et al.*, 1989); Greenland halibut, *Reinhardtius hippoglossoides* (Arthur and Albert, 1993). The most commonly used parasites are larval stages of anisakid nematodes, (Sinderman, 1983; Williams *et al.* 1992), probably because they are amongst the most common helminths of marine teleosts (MacKenzie, 1987, Sinderman, 1990, Quinteiro, 1990).

There are some considerations to take in account in the data processing and analysis when parasites are used.

- Summary statistics of infestation levels (prevalence, mean intensity, abundance) must be expressed by the mean plus and minus the standard deviation followed by the range (Margolis *et al.*, 1982).

- The application of appropriate statistical methods is recommended to discern the parasites that are useful as tags, to test the differences in infestation levels and to identify the stock composition of the host population. Multivariate models seems to be the more powerful ones: Bailey *et al.*, 1988), used a multivariate maximum likelihood model to resolve the stock composition of Sockeye salmon; non parametric discriminant analyses was employed by Arthur and Albert (1993) in Greenland halibut. Various statistical methods have been used for testing the differences between samples. For example, Brattey and I-Hsun Ni (1992) tested the differences in prevalence and abundance of parasites in harp seals, *Phoca groenlandica*, using G tests and the nonparametric Kruskal-Wallis and Mann-Whitney tests. Methods are summarised in Elliot (1979) and Sokal and Rohlf (1981).

The distribution of observed frequencies or abundances of parasites, provide information about population dynamics of the parasite and helps to understand the infestation levels observed in the host, especially when host age is taken into account (Anderson, 1978; Anderson and May 1978, Pacala and Dobson, 1988; Brattey and I-Hsun Ni, 1992).

Special care must be take when analyses are made with long lived parasites such as anisakids worms (Arthur and Albert, 1993). Rohde (1993) defines the age of the host as one of the main factors to be considered in analysing infestation levels of long lived parasites. In fact when infestation is proved to be age dependent, all analyses carried out with the intention of finding geographical differences in the host population must be performed on groups or similar age (Abaunza *et al.*, *in press*; Abaunza and Villamor, 1994). The length of the host has more uncertainties than age (the environmental conditions have an influence on fish growth) and of lesser importance when explaining infestation of long lived parasites.

Genetics of parasites:

Genetic analysis of the parasite population may provide a tool for stock identification Beverley Burton (1978) first used allozymes of parasites. Allozyme studies have shown that many of the ascaridoid nematode genera infecting fish and marine mammals in the North Atlantic, consist of several morphologically similar but generally distinct species (Orecchia *et al.*, 1986; Nascetti *et al.*, 1986, 1990; Paggi *et al.*, 1991). Some recent papers incorporate this information to the analysis of infestation levels (Brattey and Bishop, 1992; Brattey and I-Hsun Ni, 1992), but there are no references about the success on the genetic analysis of their parasites for fish stock identification. However, genetic studies have been used to confirm parasite species composition of the host and to ensure uniformity in the data analyses.

REFERENCES:

ABAUNZA P., and B. VILLAMOR.- 1994. Further considerations into the usefulness of parasites as biological tags for marine fish stock discrimination, with special reference on use of *Anisakis simplex* (L3). *Working Document* presented to the 1994 ICES Study Group on Stock Identification Protocols for Finfish and Shellfish Stocks.

ABAUNZA, P., B. VILLAMOR and J.R. PEREZ.- (Submitted). Infestation by larvae of *Anisakis simplex* (Nematoda: Ascaridata) in horse mackerel, *Trachurus trachurus* (L.), and Atlantic mackerel, *Scomber scombrus* L., in ICES Divisions VIIIb, VIIIc and IXa (N-NW of Spain). *Scientia Marina* (Submitted).

ANDERSON, R.M.- 1978. The regulation of host population growth by parasitic species. *Parasitology*, 76, 119-158.

ANDERSON, R.M. and MAY, R.M.- 1978. Regulation and stability of host-parasite population interactions, I. Regulatory processes. *Journal of Animal Ecology*, 47, 219-247.

ARTHUR J.R., and E. ALBERT.- 1993. Use of parasites for separating stocks of Greenland halibut (Reinhardtius hipoglossoides) in the Canadian northwest Atlantic. *Can. J. Fish. Aquat. Sci.*, 50, 2175-2181.

BAILEY, R.E., MARGOLIS, L., and GROOT, C.- 1988. Estimating stock composition of migrating juvenile Fraser River (British Columbia) sockeye salmon, *Oncorhynchus nerka*, using parasites as natural tags. *Can. J. Fish. Aquat. Sci.*, 45, 586-591.

BAILEY, R.E., MARGOLIS, L., and WORKMAN, G.D. - 1989. Survival of certain naturally acquired freshwater parasites of juvenile sockeye salmon, *Oncorhynchus nerka* (Walbaum), in hosts held in fresh and sea water, and implications for their use as population tags. *Can. J. Zool.*, 67, 1757-1766.

BEVERLEY-BURTON, M.- 1978. Population genetics of *Anisakis simplex* (Nematoda: Ascaridoidea) in Atlantic salmon (*Salmo salar*) and their use as biological indicators of host stocks. *Env. Biol. Fishes.*, 3, 369-378.

BRATTEY, J., and I-H. Ni.- 1992. Ascaridoid nematodes from the stomach of harp seals, *Phoca groenlandica*, from Newfoundland and Labrador. *Can. J. Fish. Aquat. Sci.*, 49, 956-966.

BRATTEY, J. and C.A. BISHOP.- 1992. Larval *Anisakis simplex* (Nematoda: Ascaridoidea) infection in the musculature of Atlantic cod, *Gadus morhua*, from Newfoundland and Labrador. *Can. J. Fish. Aquat. Sci.*, 49, 2635-2647.

CHENOWETH, J.F., McGLADDERY, S.E., SINDERMANN, C.J., SAWYER, T.K. and J.W. BIER.- 1986. An investigation into the usefulness of parasites as tags for herring (*Cuplea harengus*) stocks in the western north Atlantic, with emphasis on use of the larval nematode

Anisakis simplex. *J. Northw. Atl. Fish. Sci.*, Vol 7, 25-33.

ELLIOT, J.M.- 1979. Some methods for the anlysis of samples of benthic invertebrates. 2nd ed. Freshwater Biol. Assoc. Sci. Publ. No. 25, 160 pp.

KABATA, Z.- 1963. Parasites as biological tags. *ICNAF Spec. Publ.* No. 4, 31-37.

McGLADDERY, S.E.- 1987. Potential of *Eimeria sardinae* (Apicomplexa: Eimeridae) oocysts for distinguising between different spawning groups and between first and repeat-spawning Atlantic herring (*Cuplea harengus harengus*). *Can. J. Fish. Aquat. Sci.*, 44, 1379-1385.

McKENZIE, K.- 1987. Parasites as indicators of host populations "Parasitology - Quo Vadit?" . Proc. 6th Int. Congr. Parasitology, Brisbane, Australia. *Int. J. Parasitol.*, 17, 345-352.

McKENZIE, K.- 1990. Cestode parasites as biological tags for mackerel (*Scomber scombrus* L.) in the Northeast Atlantic. *J. Cons. perm. int. Explor. Mer.*, 46 155-166.

MARGOLIS, L., ESCH, G.W., HOLMES, J.C., KURIE, A.M., and G.A. SCHAD.- 1982. The use of ecological terms in parasitology (report of an ad hoc committee of the American Society of Prasitologists). *J. Parasitol.*, 68(1), 131-133.

NASCETTI, G., BERLAND, L. BULLINI, S. MATTIUCCI, P. ORECCHIA and L. PAGGI.- 1986. Due specie gemelle in *Contracaecum osculatum* (Ascaridida: Anisakidae): isolamento riprodutivo e caratteri diagnostici a livello elettoforetico. *Ann. Ist. Super. Sanita* 22, 349-352.

NASCETTI, G., L. BULLINI, R. CIANCHI, L.PAGGI, P. ORECCHIA, S. MATTIUCCI, S. D' AMELIO, and B. BERLAND.- 1990. Genetic relationships among aniskid species belonging to the genera *Contracaecum and Phocascaris*. *Bull. Fr. Soc. Parasitol.* 8 (Suppl. 1): 721 pp.

ORECCHIA, P.,B. BERLAND, L. BILLINI, S. MATTIUCCI, AND L. PAGGI. 1986. Divergenza genetica di specie dei generi Contracaecum e Phocascaris (*Ascaridia, Aniskidae*) con diverso ciclo biologico. Ann. Ist. Super. Sanita 22:345-348.

PACALA, S.W., and A.P. DOBSON.- 1988. The relation between the number of parasites/host and host age: population dynamic causes and maximum likelihood estimation. *Parasitology*. 96, 197-210.

PAGGI, L., G. NASCETTI, R. CIANCHI, P. ORECCHIA, S. MATTIUCCI, S. D' AMELIO, B. BERLAND, J.BRATTEY, J.W. SMITH, and L. BULLINI.- 1991. Genetic evidence for three sealworm species within Pseudoterranova decipiens (Nematoda: Ascaridida, Ascaridoidea) in the North Atlantic, and Norwegian and Barents seas. Int. *J. Parasitol*, 21, 195-212.

QUINTEIRO, P. 1990. Helmintos parasitos de peces de las costas gallegas. Tesis Doctoral, Universidad de Santiago de Compostela, 243 pp.

ROHDE, K., 1987 Differente populations of Scomber australasicus in New Zealand and south eastern Australia, demostrated by a simple method using monogenean sclerites. *J. Fish Biol.*, 30, 651-658

ROHDE, K. 1993. *Ecology of Marine Parasites.* 2nd ed., CAB International, Wallingford, 298 pp.

SINDERMANN, C.J. 1983. Parasites as Natural Tags for Marine Fish: a Review. *NAFO Sci. Coun. Studies*, 6, 63-71

SINDERMANN, C.J. 1990. *Principal Diseases of Marine Fish and Shellfish*, Vol 1. 2nd ed., Academic Press, Inc. San Diego, 521 pp.

SOKAL, R.R. and F.J. ROHLF 1981. *Biometry*, 2nd ed. W.H.Freeman and Company, New York, 859 pp.

WILLIAMS, H.H., K. MaCKENZIE and A.M. McCARTHY, 1992. Parasites as biological indicators of the population biology, migrations, diet, and phyllogenetics of fish. *Reviews in Fish Biology and Fisheries*, 2, 144-

110 Chemical Composition of Body Parts and Tissues-Otolith Elemental Fingerprinting

First Draft: Copenhagen, August, 1994

S. Campana
Bedford Institute of Oceanography
P.O. Box 1006
Darmouth, N.S.
Canada B2Y 4A2

Requires description and/or amplification concerning factors (physiological and environmental) modifying elemental uptake into otolith, ontogenetic and growth-related effects, whole-otolith versus beam-based techniques, advantages and disadvantages of each of the analytical options, statistical approaches for analyzing the multi-elemental fingerprint, temporal and spatial stability of a fingerprint, use of elemental isotopes, sources of contamination (and its avoidance), and the differentiation of sample versus stock differences. First draft is simply introduction to topic. The methodology and discussion sections of this chapter should be expanded. Additional references are also required.

Introduction

Recent studies have pointed to the potential of the otoliths (ear stones), found in all fishes, as natural population markers (Edmonds et al. 1989; Campana and Gagné 1994). While otoliths are well known for the formation of the annual (Casselman 1987) and daily growth rings (Campana and Neilson 1985) used in their age determination, it is their elemental composition which has attracted attention as a potential means to track and identify fish populations. The potential is based on two observations: a) otoliths grow throughout the life of the fish, and unlike bone, are metabolically inert; once deposited, otolith material is unlikely to be resorbed or altered (Campana and Neilson 1985); and b) the calcium carbonate and trace elements that make up 90% of the otolith appear to be mainly derived from the water (Simkiss 1974). Accordingly, the elemental composition of the otolith reflects that of the water in which the fish lives, although not necessarily in a simplistic fashion (Kalish 1989; Fowler et al. 1995). Since the elemental composition of seawater varies from place to place (Johnson et al. 1992), fish populations occupying different water masses should contain otoliths with different exposure histories to the ocean environment, despite any periodic inter-mixing.

113 Genetic Analysis-Allozymes

First Draft: Copenhagen, August, 1994

M.-L. Koljonen
Finnish Game and Fisheries Research Institute
P.O. Box 202, FIN 00151 Helsinki, Finland

Requires completion; discussion is preliminary. The methodology and discussion sections of this chapter should be expanded. Additional references are also required.

Introduction

Allozymes are genetically different forms of functionally similar enzymes, produced by alternative gene forms in the same gene or *gene locus*. The different forms of one gene are called *alleles*. Other forms of a same enzyme, such as products of different loci, are called iso-enzymes or *isozymes*. By studying the occurrence of allozymes, it is possible to get gene level information from the genetic structure of populations or species. The use of several gene loci simultaneously has greatly increased knowledge about the genetic differentiation within species (Utter 1991 and references therein, and several others ?).

Allele frequency differences offer one potential method for the mixed stock analysis. However, the amount of genetic divergence between natural populations is related to the degree of reproductive isolation between breeding populations. As a result, the discriminatory power of genetic differences varies between species depending on their life history strategies. Anadromous salmonids is thus a promising marine fish group for genetic mixed stock analysis. In artificial reproduction, allele frequency differences are nevertheless possible to increase to some extent. Allele frequency data has until now been most successfully used in the mixed stock analysis of Pacific salmonids (Oncorhynchus spp) (Milner et al. 1983, Gall et al. 1989, Shaklee et al 1990, Marshall et al. 1991).

Material and methods

Electrophoresis

The main method for the study of allozyme variation is enzyme electrophoresis (Shaw and Prasad 1970, Harris and Hopkinson 1976, Siciliano and Shaw 1976, Aebersold et al. 1987). Electrophoresis is based on the separation of differently charged molecules in an electrical field. Mutation in a gene locus coding a protein causes a change in the DNA nucleotide order, which in part changes the amino acid composition of the enzyme. If these alterations change the electrical charge of the molecule, a new allozyme can be observed in the electrophoresis. In practice about one third of DNA nucleotide changes are noticesable as charge changes at the enzyme level.

The material used for electrophoresis is comprised of either fresh or frozen tissues of the species concerned, because in present methods the enzymes must be "alive" and their degradation begins immediately after the death of the organism. The enzymes are usually extracted from tissue samples, using the muscle, liver or heart. The enzyme solution from each individual is applicated in a gel of starch, cellulose acetate or other medium (Figure 113-1). An electrical field is applied over the gel and the enzyme molecules migrate through the gel at a rate related to their electrical charges. The shape and weight of the molecules can affect, to some extent, the migration rate.

After running the gel (4-8 h), it is sliced horizontally into thin slabs (usually 3-5). Each slab can be individually stained for a different enzyme. After staining the enzyme-specific banding pattern appears and can be read and interpreted for each individual. For interpretation of the banding patterns, the molecular structure of the enzyme (possible interbands) and the locus structure should be known (Utter et al. 1974, Utter et al. 1987). A standardized system for interpretation and locus nomenclature has been developed (Shaklee et al. 1990). The end product of the electrophoresis is multilocus genotype data for individual organisms.

Allele frequencies

On the basis of the occurrence of individual allozyme forms, the relative proportion of each allele in a population, that is, the allele frequencies in a population, are calculated (Figure 113-1). Locus-specific allele frequencies are usually the basic data for genetic analysis. Allele frequencies are generally used to measure the genetic differentiation between populations, genetic distances and the amount of genetic divergence within populations, i.e., mean heterozygosites and levels of polymorphisms (Roughgarden 1979, others ?). In addition, allele frequency data is used to identify and describe populations in their original environments and in some cases to analyse mixtures of populations.

In the most simple case, two populations have diagnostic differences, which means that they have different alleles fixed in the same locus and every individual differs in this respect from all the individuals of an other population. Diagnostic differences are, however, very rare when studied populations are from the same species. Diagnostic differences usually occur only after long and complete reproductive isolation and in most cases diagnostic differences indicate species level differentiation.

A more useful application than looking for diagnostic differences is to use the relative differences of allele frequencies for the distinction of populations. Although individual fish cannot be classified into different populations, relatively reliable estimates of the proportions of contributing stocks can be obtained in many mixed stock analyses (Pella and Milner 1987, Wood et al. 1987, Brodziak et al. 1992).

When using genetic data for mixed stock analysis, the number of variables depends on the number of variable loci available. The accuracy of the proportion estimates depends on the number of loci used. One clear advantage is, that multilocus genotypes and their distributions can be used, which greaely increses the number of variables available. In cases where two alleles occur in one gene locus, three different genotypes are possible (AA, AB, BB, Figure 113-1). The number of possible combinations in multilocus genotypes increases according to formula 3^L, when several loci are analysed (for five loci with two alleles, different multilocus genotypes $3^5 = 243$, AAAAAAAAAA, AAAAAAAAAB, ...).

The expected distributions of multilocus genotypes can effectively be used in the maximum likelihood-based method (see Pella and Milner 1987, and chapter 504 Maximum Likelihood Stock Discrimination Models). The maximum likelihood-based genetic method developed for mixed stock analyses is originally called GSI (genetic stock identification) (Milner et al. 1985, 1987), but that name has been condidered partly misleading and thus the more general term, mixed stock analyses (MSA), (Utter and Ryman 1993) is recommended.

Discussion

The usefulness of allele frequency data for a mixed stock analysis depends on the amount of divergence between contributing stocks. Genetic differentiation between stocks and populations is only possible in reproductive isolation, thus, anadromous salmonids are probably the most promising marine fish group for the genetic mixed stock analysis. In addition, stocks from clearly different lakes or separate hatchery strains are probably more likely to differ genetically sufficiently from each other than, for example, species with pelagic spawning behaviour.

Until now, genetic mixed stock analysis has been used mainly for Pacific salmonids (Oncorhynchus spp.) (Milner et al. 1983, Gall et al. 1989, Shaklee et al 1990, Marshall et al. 1991, Koljonen 1994?). However, the possible range of useage might be wider.

If sufficient genetic differences exist between the populations studied, genetic analysis in general (not just allozymes) has several advantages compared to other natural or artificial marks (see Milner et al. 1985, Pella and Milner 1987).

- Allozymes are inherited characteristics and thus the environment does not change their manifestation.
- In general allozymes are also regarded neutral in relation to natural selection, however some exceptions are known (e.g. the Malic enzyme in Atlantic salmon, Verspoor et al. 1991).

- Allozymes are discrete characters and thus interpretation of their existence is relatively unambiguous and on the basis of numerical data obtained, differences can be comparably quantified.

- Genetic differences are relatively stable from year to year and from generation to generation (see Waples 1990).

- Alleles occur naturally, so that no marking process is needed.

- Allele frequencies are stable throughout the life cycle and occur in all populations so all life stages and also unmarked wild populations can be studied.

- The costs of the analysis are reasonable.

- Allele frequency differences do not change the viability and catchability of fishes, at least not to the extent that Carlin-tags do for instance.

- A specific advantage compared to other identification methods is that alleles, as well as other genetic marks, are inherited in a Mendelian manner and thus it is possible to analyze the gene flow to the next generation.

- In a controlled breeding system it is also possible to change the allele frequencies of certain populations for experimental purposes (Seeb et al. 1986, 1990).

Figure 113-1. Standard steps for obtaining allele frequency data from electrophoresis (Modified from Utter et al.1987). A: Crude protein is extracted from tissue. B: Extract from each fish is introduced individually to gel by filter paper inserts. C: Different allozymes move different distances in an electrical field. D: Allozymes are made visible with specific stains, the genotypes (AA,AB,BB) are interpreted and the allele frequencies are calculated.

References

Aebersold, P.B., G.A. Winans, D.J. Teel, G.B. Milner and F.M. Utter. 1987. Manual for starch gel electrophoresis: A method for the detection of genetic variation. NOAA Technical Report NMFS 61, Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service. 19 pp.

Allendorf, F.W., N. Mitchell, N. Ryman and G. Ståhl. 1977. Isozymelo/ci in brown trout (Salmo trutta L.): detection and interpretation from population data. Hereditas 86: 179-190.

Brodziak, J., B. Bentley, D. Bartley, G.A.E. Gall, R. Gomulkiewicz and M. Mangel. 1992. Test of genetic stock identification using coded wire tagged fish. Can. J. Fish. Aquat. Sci. 49: 1507-1517.

Gall, G.A.E., B. Bentley, C. Panattoni, E. Childs, C. Qi, S. Fox, M. Mangel, J. Brodziak and R. Gomulkiewicz. 1989. Chinook mixed fishery project 1986-1989. University of California, Davis. 192 p.

Harris, H. and D.A. Hopkinson, 1976. Handbook of enzyme electrophoresis in human genetics. American Elsevier, New York.

Koljonen, M.-L. 1994. Distinquishing between local and migrating Atlantic salmon (Salmo salar L.) stocks by genetic stock composition analysis. Manuscript.

Marshall, A.R., M. Miller, C. Busack and S.R. Phelps. 1991. Genetic stock identification analysis of three 1990 Washington Ocean and Strait of Juan de Fuca chinook salmon fisheries. GSI Summary report

91-1. Washington Department of Fisheries, Salmon Research and Development Olympia, Washington, 48 pp.

Milner, G.B., D.J. Teel and F.M. Utter. 1983. Genetic stock identification study. National Oceanic and Atmospheric Administration, Norhwest and Alaska Fisheries Center, Seattle, WA 33 p. (and Appendix).

Milner, G.B., D.J. Teel, F.M. Utter, and G.A. Winans. 1985. A genetic method of stock identification in mixed populations of Pacific salmon, Oncorhynchus spp. Marine Fisheries Review 47:1-8.

Pella, J.J. and G.B. Milner. 1987. Use of genetic marks in stock composition analysis, p 247-276. In N. Ryman, N. and F. Utter (ed.) Population Genetics and Fishery Management. Washington Sea Grant Program, Seattle. University of Washington Press.

Roughgarden, J. 1979. Theory of population genetics and evolutionary ecology. An introduction. - Macmillan, New York.

Seeb, J.E., L.W. Seeb and F.M. Utter. 1986. Use of genetic marks to assess stock dynamics and management programs for chum salmon. Trans. Am. Fish. Soc. 115:448-454.

Seeb, L.W., J.E. Seeb, R.L. Allen and W.K. Hershberger. 1990. Evaluation of adalts returns of genetically marked chum salmon, with suggested future applications. American Fisheries Society Symposium 7: 418-425.

Shaw, C.R. and R. Prasad. 1970. Starch gel electrophoresis of enzymes - a compilation of recipes. Biochem. Genet. 4: 297-320.

Shaklee, J.B., C. Busack, A. Marchall, M. Miller and S.R. Phelps. 1990. The electrophoretic analysis of mixed-stock fisheries of Pasific salmon, p. 235-265. In Z.I.Ogita, and C.L. Markert (ed.) Isozymes: Structure, Function, and Use in Biology and Medicine. Progress in Clinical and Biological Research. Vol. 344. Wiley-Liss, Inc. New York
973 p.

Shaklee, J., F.W. Allendorf, D.C. Morizot, G.S. Whitt. 1990. Gene-nomenclature for protein-coding loci in fish. Trans. Am. Fish. Soc. 119: 2-15.

Siciliano, M.R. and C.R. Shaw. 1976. Separation and visualization of enzyme gels. In: Smith, I. (ed.), Cromatographic and electrophoretic techniques, Vol. 2, p.217-238. Zone electrophoresis. Interscience Publishers, John Wiley and Sons, Inc., New York.

Utter, F.M., H.O. Hodgins and F.W. Allendorf. 1974. Biochemical genetic studies on fishes: potentials and limitations. In D. Malins (ed.) Biochemical and biophysical perspective in marine biology, Vol.1, p. 213-237. Academic Press, San Francisco.

Utter, F.M. 1991. Biochemical gentics and fishery management: an historical perspective. J.Fish.Biol. 39:1.20.

Utter, F., P. Aebersold and G. Winans. 1987. Interpreting genetic variation detected by electrophoresis, p 21-45. In N. Ryman and F. Utter (ed.) Population Genetics and Fishery Management. Washington Sea Grant Program, Seattle. University of Washington Press.

Utter, F. and N. Ryman. 1993. Genetic markers and mixed stock fisheries. Fisheries, Vol 18:11-21.

Verspoor, E., N.H.C. Fraser and A.F. Youngson. 1991. Protein poly-morphism in Atlantic salmon within Scottish river: evidence for selection and estimates of gene flow between tributaries. Aquaculture 98. 217-230.

Waples, R.S. 1990. Temporal changes of allele frequency in Pasific salmon: implications for mixed-stock fishery analysis. Can. J. Fish. Aquat.Sci. 47: 968-976.

Whitmore, D.H. (ed.). 1990. Electrophoretic and isoelectric focusing techniques in fisheries management. CRC Press, Inc. 350 pp. (not yet in the text)

Wood, C.C., S. McKinnel, T.J. Mulligan and D.A. Fournier. 1987. Stock identification with the maximum-likelihood mixture model: sensitivity analysis and application to complex problems. Can. J. Fish. Aquat. Sci. 44: 866-881.

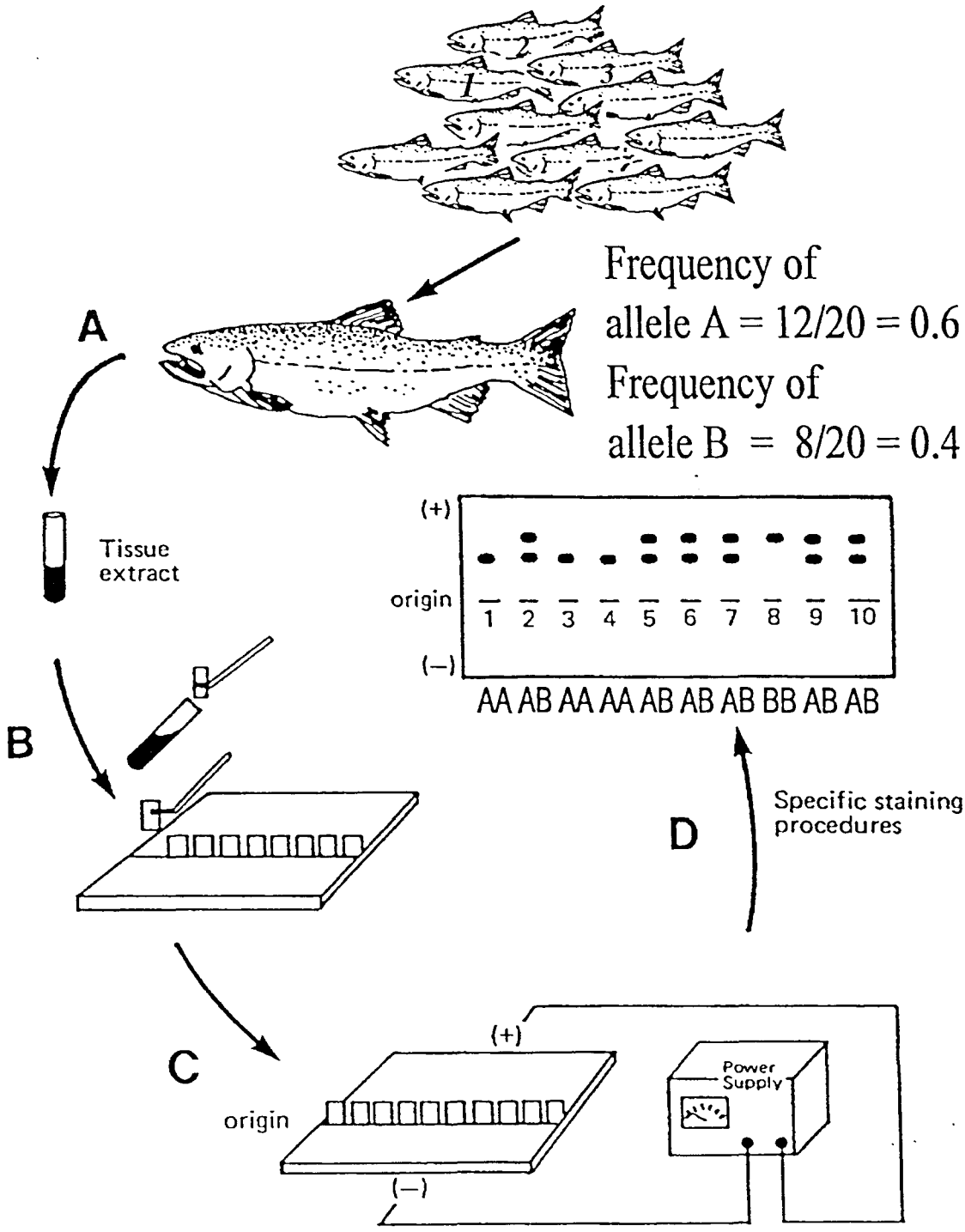


Figure 113-1. Standard steps for obtaining allele frequency data from electrophoresis (Modified from Utter et al.1987). A: Crude protein is extracted from tissue. B: Extract from each fish is introduced individually to gel by filter paper inserts. C: Different allozymes move different distances in an electric field. D: Allozymes are made visible with specific stains, the genotypes (AA,AB,BB) are interpreted and the allele frequencies are calculated.

116 Genetic analyses- Random amplified polymorphic DNA

First Draft: Copenhagen, August 1994.

Peter Smith
MAF Fisheries
PO Box 297
Wellington
New Zealand

The methodology and discussion sections of this chapter should be expanded. Additional references are also required.

Introduction

The development of the polymerase chain reaction (PCR) which amplifies DNA enables analyses to be carried out on very small samples and offers a new range of tools for the detection of genetic variation, without the need for cloning and sequencing. PCR amplification of mtDNA and subsequent digestion with restriction enzymes has been suggested as a potential fisheries tool but initial demonstration has been with the separation of well described fish species (Chow et al 1993, Chow and Inoue 1993). PCR amplification and subsequent sequencing of the mitochondrial cytochrome b genes has been used to distinguish closely related tuna species including species pairs which could not be separated by conventional isoelectric focusing of proteins (Bartlett & Davidson 1991).

Random amplified polymorphic DNA is a term applied to a method described independently by Williams et al (1990) and Welsh and McClelland (1990) for the identification of plant cultivars. The RAPD technique allows the detection of DNA polymorphisms by randomly amplifying regions of DNA by PCR with single arbitrary primers. Any section of DNA flanked by a pair of primer sites, and less than 5 000 base pairs apart, will be amplified by the technique. The amplification products are separated by agarose gel electrophoresis and detected by direct staining with ethidium bromide. The RAPD technique requires no specific probes and does not use radio-isotopes. The comparative ease with which a large number of DNA primers can be screened for polymorphism make the RAPD technique an attractive population tool.

Primers detect polymorphisms due to point mutations, which allow or disallow primer binding to the sample DNA, and due to insertions/deletions between primer binding sites. Different primers bind to different DNA sites and thereby detect different RAPD polymorphisms.

The majority of RAPD applications have been to distinguish between breeds of farm livestock (e.g., Gwakisa *et al* 1994), plant cultivars (e.g., Hu and Quiros 1991), and strains of mice (Welsh *et al* 1991). The technique has potential to provide genetic markers for fisheries studies but there are few published reports as of August 1994. Fisheries applications have been limited to a demonstration of the technique for distinguishing 12 species of freshwater fishes from 7 families (Dinesh *et al* 1993), and distinguishing *Tilapia* species (Skibinski *et al* 1994). Anon (1994) found different DNA fragments in two species of bluefin tuna, *Thunnus thynnus* and *T.maccoyii*, which could not be separated by conventional isoelectric focusing. The two tuna species showed an average difference of 6.3% measured with 32 primers.

A conference abstract on winter flounder reported no significant differences between onshore and offshore populations of winter flounder using both RFLP's in mtDNA and RAPD's (Spinka *et al* 1993). In the orange roughy *Hoplostethus atlanticus* no genetic polymorphism and no unique fragments were found in 6 geographically isolated populations tested with 130 primers, whereas conventional allozyme polymorphisms revealed significant genetic heterogeneity in samples from the same areas (Anon).

DNA extraction, amplification and separation

DNA can be extracted from fresh, frozen and ethanol preserved fish tissues including muscle and liver following standard protocols. DNA samples are amplified separately with oligonucleotide primers. Usually a random selection

of 10-base primers, with a G+C content of 60–70%, are used for amplifications. More than 1000 different 10-base primers are available ex stock from Operon Technologies, California.

Amplification reactions are carried out in 50μl volumes in a thermocycler. Serial dilutions of DNA samples are tested initially to determine optimum DNA volumes for amplification. Each reaction mixture contains approximately 50ng genomic DNA in 10mM Tris HCl pH8.3, 30ng single 10-base primer, 50mM KCl, 2mM MgCl2, 100 μM each of *d*ATP, *d*CTP, *d*GTP, and *d*TTP, and 1 unit Taq DNA polymerase in Perkin Elmer PCR buffer. Amplification is performed in a thermocycler programmed for 45 cycles of 1 min at 94°C, 1 min at 36°C, and 2 min at 72°C after Williams et al (1990). Amplification conditions may need to be adjusted to suit different thermocyclers.

The amplification products are separated in 1.4% agarose gels and detected with ethidium bromide under a UV light (312nm). A DNA size ladder is included in each gel. In addition it is usual to include a control sample of reaction excluding the genomic DNA sample to test for artifacts in the amplification.

DNA fragment scoring and analyses

The number of different amplification products, or DNA fragments, for each primer depends upon the primer sequence and the genome size of the test organism. Assuming that primer sites are randomly distributed throughout the genome, then for a typical vertebrate there will be between 2 and 10 fragments produced by each primer. The size of the fragments ranges from around 0.5 to 3.0 kb. Fragments are numbered with the primer code number and the fragment size, determined from the size ladder in the gel. Only distinct well stained fragments are scored. Sometimes weakly staining fragments are detected. These fragments may appear in some individuals but not others prepared at the same period. Frequently the weakly staining fragments appear inconsistently between repeat extractions of the same DNA samples. It is possible that these irregular fragments result from poor discrimination by a primer between slightly different nucleotide sequences. Sometimes a smear of DNA stain is observed in lanes in the gel, which may indicate an excess concentration of sample template DNA in the amplification reaction.

An index of similarity (or differences) can be calculated as the number of shared fragments (or unique fragments) between pairs of individuals divided by the number of fragments scored for both individuals (eg Gilbert *et al* 1990, Jeffreys and Morton 1987, Wetton *et al* 1987). The difference value (D) calculated after Gilbert *et al* (1990) is:

$$D = Nab / Na + Nb$$

where Nab is the number of fragments that differ between individuals a and b, and Na and Nb are the number of fragments in individuals a and b respectively. The average percentage difference is the average of all D values multiplied by 100 (Gilbert *et al* 1990).

For the majority of primers fragment patterns produced in intraspecific individuals are similar. Typically only a few primers reveal polymorphism. Thus a large number of primers have to be screened in order to find informative genetic markers. The pooling of DNA samples from different individuals of the same species permits a rapid screening of a large number of primers and takes no more laboratory time than conventional allozyme screening.

Discussion

The RAPD technique has been used to identify and distinguish closely related species of teleosts, some of which cannot be separated by conventional iso-electric focusing of proteins. Thus RAPD's are likely to be used as a tool for taxonomic problems and for identifying fish fillets and fish products. In addition, because the technique works with minute quantities of DNA, it should be applicable to identifying fish eggs and larvae.

The RAPD technique is technically more demanding than allozyme electrophoresis, but easily accommodated in a laboratory experienced in RFLP studies. Perhaps the only disadvantage is that RAPD markers are 5–10 times more costly to test, in terms of chemicals, than allozyme and IEF markers.

The application of the RAPD technique is less promising for stock separation. Screening of 130 primers found no genetic markers that would allow separation of orange roughy samples into stock units, whereas conventional allozyme polymorphisms tested in larger samples from the same sites revealed a significant heterogeneity and an indication of several stock units (Anon 1994). It is possible that screening a larger number of primers, or smaller size primers, may detect variation that must exist in the DNA as indicated by the allozyme markers.

References

Anon. 1994. The report of the Study Group (C.M.1994/M:5).

Bartlett, S.E. and Davidson, W.S. 1991. Identification of *Thunnus* tuna species by the polymerase chain reaction and direct sequence analysis of their mitochondrial cytochrome *b* genes. Can. J. Fish. Aqua. Sci. 48: 309-317.

Chow, S., Clarke, M.E. and Walsh, P.J. 1993. PCR-RFLP analysis on thirteen western Atlantic snappers (subfamily Lutjanidae): a simple method for species and stock identification. Fish. Bull. 91: 619-627.

Chow, S. and Inoue, S. 1993. Intra- and interspecific restriction fragment length polymorphism in mitochondrial genes of *Thunnus* tuna species. Bull. Nat. Res. Inst. Far Seas Fish., 30: 207-225.

Dinesh, K.R., Lim, T.M., Chua, K.L., Chan, W.K. and Phang, V.P.E. 1993. RAPD analysis; an efficient method of DNA fingerprinting in fishes. Zoo. Sci. 10: 849-854.

Gilbert, D.A., Lehman, N., O'Brien, S.J. and Wayne, R.K. 1990. Genetic fingerprinting reflects population differentiation in the California Channel Island fox. Nature 344: 764-766.

Gwakisa, P.S., Kemp, S.J. and Teale, A.J. 1994. Characterization of Zebu cattle breeds in Tanzania using random amplified polymorphic DNA markers. Animal Genetics 54: 89-94.

Hu, J. and Quiros, C.F. 1991. Identification of broccoli and cauliflower cultivars with RAPD markers. Plant Cell reports 10: 505-511.

Jeffreys, A.J. and Morton, D.B. 1987. DNA fingerprints of dogs and cats. Animal genetics 18: 1-15.

Skibinski et al 1994. RAPD's. Heredity in press

Spinka, T.I., Procaccini, D., Rupprecht, J. and Goddard, K. 1993. Analysis of winter flounder population subdivision by differences in random amplified polymorphic DNA. J. Pennsylvania Academy of Science 66: 204.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18: 7213-7218

Welsh, J., Peterson, C. and McClelland, M. 1991. Polymorphisms generated by arbitrarily primed PCR in the mouse: application to strain identification and genetic mapping. Nucleic Acids Research 19: 303-306.

Wetton, J.H., Carter, R.E., Parkin, D.T. and Walters, D. 1987. Demographic study of a wild house sparrow population by DNA fingerprinting. Nature 327: 147-149.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A. and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful genetic markers. Nucleic Acids Research 18: 6531-6535.

501 Statistical Algorithms for Stock Composition Analysis

First Draft: Copenhagen, August, 1994

Michael H. Prager
Southeast Fisheries Science Center
National Marine Fisheries Service
75 Virginia Beach Drive, Miami, FL 33149

Desirable revision of the present material might include the following two points, among others: (1) inclusion of more equations to parallel and make more precise the English descriptions included here, and (2) a more complete and up-to-date literature review. The citation of strongly relevant papers with precedence in the primary fisheries literature and also recent references (such as books) that summarize knowledge in statistics would be useful. Also, a stronger introduction.

1. Introduction

While much of the effort in stock identification research is applied to the development of features or characters to separate stocks, it is important to remember that the actual results of stock identification reserach are the classification models produced from these data. The choice of model and many aspects of model development are critical to the overall success of a stock identification project. This section provides an overview of the general types of classification models used in stock identification research.

2. The problem and its terminology

What is often called "stock identification" in fisheries is more precisely called "stock composition analysis." As used here, that means estimation of the stock composition of a mixed-stock sample with a known number J of component stocks (sometimes called "classes"). The proportion of the mixed sample (often representing the harvest in a fisheries context) that comes from stock j can be represented P_j. Note that $\sum_{j=1}^{J} P_j = 1$, and within this constraint, any particular P_j may equal zero. The process by which the P_j are estimated constitutes the stock composition analysis. "Estimated stock composition" and "estimated mixture proportions" are used here as synonyms.

The data used for such analyses are observations on characteristics of individuals; typical characteristics may include morphometrics, meristics, or genetic characters. This discussion assumes the availability of a training sample of individuals whose stock membership is known. The training sample is used to fit a model by means of the investigator's choice of *algorithm*, a word used here to denote a statistical method or group of related methods. The model is then used to estimate the stock composition of a mixed-stock sample. Such estimation can, but need not always, involve estimating the probability of stock membership of the individuals in the mixed-class sample.

3. Algorithms

Several algorithms have been used or proposed for this problem. This paper omits reference to algorithms based on only one characteristic, and considers instead algorithms with a statistical basis and that consider more than one characteristic. Also omitted is discussion of the many variants of each algorithm that have been introduced; e.g., polynomial discriminant analysis as used by Cook and Lord (1978), or age-invariant discriminant analysis as used by Fabrizio (1987).

3.1 Discriminant analysis (DA)

Linear discriminant analysis (LDA) seems to have been the first formal statistical method used for stock composition analysis (Hill 1959). Among the assumptions of LDA is that the characteristics follow a multivariate normal distribution with common variance-covariance structure among classes. This method has

been used many times and in many variations. One of its major advantages is the wide availability of reliable and flexible commercial software to perform discriminant analyses.

Quadratic discriminant analysis (QDA) has been proposed as preferable for problems in which the variance-covariance structure differs by class (Misra 1985), as QDA does not assume equality of variance among classes (Kendall et al. 1983). However, estimates from QDA generally are of higher variance than those from LDA, because of the additional parameters that must be estimated. Prager, Jones, and Fabrizio (in prep.) found that, on several test data sets of fisheries data of known composition, performance of QDA was substantially worse than that of LDA, even when variance-covariance matrices were unequal by class.

In using discriminant analysis to estimate stock composition, one can proceed in two slightly different ways. The first, which will be termed *discrete classification,* is to classify each individual in the (mixed-stock) sample into the class in which its membership is estimated as most likely. The estimate of stock composition is then formed from the relative numbers of individuals classified into each class. The second way to proceed will be termed *non-discrete classification.* This procedure is to sum across individuals the probability of membership in each class. The estimate of stock composition is then obtained from the relative sums by class.

As an simple example, consider a two-stock problem in which 3 fish are in the mixed-stock sample. Let the estimated probabilities of membership in Class I for the 3 fish be 0.55, 0.45, and 0.8. The discrete estimate of mixing proportions would be 2/3 from Class I and 1/3 from Class II. The non-discrete estimate would be 0.6 from Class I and 0.4 from Class II. The former estimate is obtained because 2 of the 3 fish are thought more likely to belong to Class I; the latter is the mean of the three probabilities given.

While discrete classification and non-discrete classification produce similar estimates, it seems logical to prefer non-discrete classification. There is no necessity to round estimated membership probabilities to whole numbers, as done in discrete classification, when the objective is to estimate mixing proportions.

Correction matrices are frequently used to improve mixture estimates from discriminant analyses (Cook and Lord 1978; Pella and Robertson 1978), and might be used to correct estimates from other classification-based methods as well. Millar (1987) has shown that the use of classification with correction is a special case of maximum likelihood FMP methods (described below).

3.2 Logistic regression

Logistic regression (Aldrich and Nelson 1984; Hosmer and Lemeshow 1989) was suggested for this application by Prager and Fabrizio (1990), who found the method promising. Its chief theoretical advantage is that it assumes neither multivariate normality nor equality of variances, and is appropriate for a wide variety of distributions (Kendall et al. 1983).

3.3 Neural networks

The term "neural networks" is not a precise one, but refers to a group of empirical methods that sift through and combine many models to arrive at a model of optimum (in some sense) complexity. Neural networks that use a separate data set for internal cross-verification are sometimes called "genetic algorithms." Such methods have proved valuable in such fields as nondestructive testing, flight control, terrain classification, and signal processing.

I believe that the first published application of these methods to stock composition analysis was made by Prager (1984; 1988), who used the so-called Group Method of Data Handling (Ivakhnenko and Ivakhnenko 1974), a genetic algorithm in today's terminology, to estimate stock composition of striped bass and American shad on data sets of known composition. The method performed at least as well as linear discriminant analysis by the criteria used in the study.

3.4 Finite mixture distribution (FMD) methods

The preceding three algorithms are often considered to be classification-based methods, as they estimate the probabilities of membership of each individual in each class of the mixed-class sample. The desired estimates of composition are derived from the classification probabilities of the individuals.

The final group of methods discussed here includes maximum-likelihood methods based on the theory of finite mixture distributions (Wolfe 1970; Everitt and Hand 1981). In this context, "finite" refers to the number of classes in the mixture. Such methods were introduced to fisheries literature by Fournier et al.

(1984) and Millar (1987). These methods are applicable to a wide variety of distributions, but are simplest to apply if one can assume that the characteristics follow a multivariate normal distribution with a common covariance matrix among classes, the same assumption that is used in linear discriminant analysis.

4. The importance of prior knowledge

Although discriminant analysis is not usually considered a Bayesian method, its estimates of class membership are conditional on prior estimates of stock composition. Perhaps because it is easy to specify these priors when using standard software, their importance is often overlooked; however, the reliance on priors is a major shortcoming of classification-based methods. This section describes the priors required in DA and how similar information is used in the other classification-based algorithms, which may not seem to require priors.

4.1 What are priors?

Priors, as used by discriminant analyses, are estimates of the probabilities that an individual is a member of each component class when the individual has been chosen at random and nothing further is known about it (i.e., its measured characteristics are not considered). If we assume that the classes are present in the mixed-class sample in proportion to their presence in the mixture under consideration (except for sampling error), the paradox involved in using discriminant analysis for this problem becomes clear. The priors, which are required to make an estimate, are precisely what we are trying to estimate: the relative class composition of fish in the mixture.

This paradox does not occur in some other fields that use classification methods because the structure of their questions is fundamentally different. For example, a typical medical application might be estimating the probability that a patient has a certain disease, conditional on facts about general health and family history. In that case, the proportion of individuals afflicted with the disease in the general population (the prior) is well known, and the focus of the study is on the individual. In similar applications, reliable priors are readily available, and classification methods are very useful. Notably, they do not display the fundamental drawback that they present in stock-composition studies.

4.2 Priors and discriminant analysis

In discriminant analyses, priors are specified in a formal way, and composition estimates cannot be made without them. In the absence of other information, priors are often made equal, so that if there are three stocks, for example, the prior probability for each stock is set to 1/3. If the priors are incorrect, the estimates of mixture composition are biased towards the priors. The purpose of correction matrices is to reduce such biases.

Frequently, studies using discriminant analysis include estimates of error rates. It is important to note that the error rates of an uncorrected discriminant estimator depend on the actual (and unknown) mixed-stock composition, and are generally smallest when the priors are correct.

4.3 Priors and logistic regression

In logistic regression, explicit priors are not specified by the analyst. However, a logistic regression estimator is derived from the distributions observed in the training sample. It therefore provides the least biased estimates when the mixed-class sample is of the same composition as the training sample. (This was unfortunately overlooked by Prager and Fabrizio 1990.) When this effect was removed, Prager et al. (in prep.) found the performance of logistic regression to be comparable to that of linear discriminant analysis, even on nonnormal data.

The use of correction matrices to reduce bias in mixing estimates from logistic regression seems feasible. However, I do not believe that this specific subject has been studied.

4.4 Priors and neural networks

Estimation by these methods is also conditional on the composition of the training sample. Bias of these estimators should behave similarly to those from logistic regression. The ultimate performance of such

a method also depends on the particular data and specific method involved. My personal opinion is that these methods are unlikely to provide a significant, repeatable, improvement over established FMD methods in fisheries applications. As with logistic regression, I believe that the use of correction matrices with neural networks has not yet been studied, at least in fisheries.

4.5 Priors and the FMD (mixture) methods

The FMD methods do not require priors, implicit or explicit. It stands to reason that, in most cases, they should produce estimates with reduced bias, compared to uncorrected classification-based algorithms. This may in some cases be accompanied by increased variance.

5. Evaluation of Methods

The preceding discussion implies that evaluation of methods and applications is not a simple task, whether undertaken on simulated or real data. The factor that has at times been underemphasized in such evaluations is how an estimator performs as the composition of the mixed-class sample varies from that of the training sample. One would expect the error of uncorrected discriminant analyses to become considerably worse as the true composition varies from the priors. A similar deterioration in performance of uncorrected LR or NN estimates will occur as the true composition varies from that of the training sample. The error rate of FMD methods may deteriorate as the composition of the mixed-stock sample becomes quite different from that of the training sample, but variation in the composition of the mixed-stock sample does not constitute violation of a major assumption, as it does with the classification-based methods.

6. Discussion

From the preceding, it can be seen that, for the stock-composition problem in fisheries, uncorrected methods are the least desirable, and FMD methods appear more appropriate than methods based on classification. Millar (1990) concluded that *corrected* classification estimators are as useful as FMD methods when the number of stocks is small (two or three), but recommended use of direct maximum likelihood estimation (FMD methods) for more complex problems.

A review of all studies discussing these methods is beyond the scope of the present paper. Relevant fisheries references beyond those already cited include Cook (1982, 1983), Pella and Milner (1986), Wood et al. (1987), and Gray (1994).

One reason for the relatively slow adoption of FMD methods may be the wide availability of standard commercial software for fitting discriminant functions, logistic regressions, and even neural networks. Software for mixture problems appears less readily available. However, R. Millar (pers. comm., 1992) has developed Fortran software that he is willing to share. That software implements the FMD method under the assumption of multivariate normality with constant variance.

It may not be obvious that commercial software for discriminant analysis can also be used to obtain maximum-likelihood FMD estimates of stock composition. Thus any analyst with access to standard statistical packages can explore the properties of FMD estimates.

To proceed in this way, it is necessary to assume that the measured characteristics are assumed multivariate normal with a common covariance matrix across stocks. Under this assumption, the equations for LDA and FMD are identical. What we call the "priors" in LDA are in FMD the stock proportions to be estimated. By using an EM algorithm (Dempster et al. 1977), maximum likelihood estimates of these proportions can be obtained under the FMD model. As stated by Millar (1987): "... in constructing a classification rule, one is actually doing all of the work required to construct the likelihood function, so from there it would be a matter of simply running a maximization program to obtain the maximum likelihood estimates." The procedure is as follows:

1. Fit a linear discriminant function to the training sample.
2. Obtain a starting guess for the priors. In the absence of other information, one can use equal priors.
3. Using the discriminant function from step 1 and the current priors, make a non-discrete estimate of the mixture proportions of the mixed-stock sample.
4. Revise the priors to equal the current estimated mixture proportions.
5. Repeat steps 3 and 4 until the composition estimates converge.

Without doubt, this procedure is more tedious than using software written specifically for the estimation of mixing proportions. However, if such software is not readily available, or if an investigator wishes to take the first steps into using FMD methods, this iterative procedure may prove useful.

7. Acknowledgments

Correspondence and conversations with J. Pella helped to clarify my understanding of the relationship between discriminant analysis and the finite mixture problem. Insight was also gained during an unpublished simulation study conducted with M. Fabrizio and C. D. Jones. I thank R. Millar for sharing his HIGHSEAS software. Partial support for this work was provided by the Virginia Sea Grant Consortium; the Rosenstiel School of Marine and Atmospheric Science, University of Miami; and the Miami Laboratory, Southeast Fisheries Science Center, NMFS, NOAA.

8. Sources Cited

Aldrich, J. H., and F. D. Nelson. 1984. Linear probability, logit, and probit models. Sage Univ. Paper series on Quantitative Applications in the Social Sciences, 07-045. Sage Publications, Beverly Hills, CA.

Cook, R. C. 1982. Stock identification of sockeye salmon (*Oncorhynchus nerka*) with scale pattern recognition. Can. J. Fish. Aquat. Sci. 39: 611-617.

Cook, R. C. 1983. Simulation and application of stock composition estimators. Can. J. Fish. Aquat. Sci. 40: 2113-2118.

Cook, R. C., and G. E. Lord. 1978. Identification of stocks of Bristol Bay sockeye salmon, *Oncorynchus nerka*, by evaluating scale patterns with a polynomial discriminant method. Fish. Bull. U.S. 76: 415-423.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood from incomplete data via the *EM* algorithm. J. Royal Statist. Soc., Series B, 39: 1-38.

Everitt, B. S., and D. J. Hand. 1981. Finite mixture distributions. Chapman and Hall, London. 143 p.

Fabrizio, M. C. 1987. Growth-invariant discrimination and classification of striped bass stocks by morphometric and electrophoretic methods. Trans. Am. Fish. Soc. 116: 728-736.

Fournier, D. A., T. D. Beacham, B. E. Riddell, and C. A. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Can. J. Fish. Aquat. Sci. 41: 400-408.

Gray, G. 1994. Bias in misspecified mixtures. Biometrics 50: 457-470.

Hill, D. R. 1959. Some uses of statistical analysis in classifying races of American shad (*Alosa sapidissima*). Fish. Bull. U.S. 59: 269-286.

Hosmer, D. H., and S. Lemeshow. 1989. Applied logistic regression. John Wiley & Sons, New York, NY. 307 p.

Ivakhnenko, A. G., and N. A. Ivakhnenko. 1974. Long-term prediction of random processes by GMDH algorithms using the unbiasedness criterion and balance-of-variables criterion. Sov. Autom. Control 7(4): 40-45.

Kendall, M., A. Stuart, and J. K. Ord. 1983. The advanced theory of statistics, vol. 3, 4th ed. Charles Griffin, London, UK. 780 p.

Millar, R. B. 1987. Maximum likelihood estimation of mixed stock fishery composition. Can. J. Fish. Aquat. Sci. 44: 583-590.

Millar, R. B. 1990. Comparison of methods for estimating mixed stock fishery composition. Can. J. Fish. Aquat. Sci. 47: 2235-2241.

Misra, R. K. 1985. Quadratic discriminant analysis with covariance for stock delineation and population differentiation: a study of beaked redfishes (*Sebastes mentella* and *S. fasciatus*). Can. J. Fish. Aquat. Sci. 42: 1672-1676.

Pella, J. J., and G. B. Milner. 1986 Use of genetic marks in stock composition analysis. *In* N. Ryman and F. Utter, eds. Population genetics & fishery management. University of Washington Press.

Pella, J. J., and T. L. Robertson. 1978. Assessment of composition of stock mixtures. Fish. Bull. U.S. 77: 387-398.

Prager, M. H. 1984. The group method of data handing: applications in oceanography and fishery science. Doctoral dissertation, University of Rhode Island.

Prager, M. H. 1988. Group method of data handling: a new method for stock identification. Trans. Am. Fish. Soc. 117: 290-296.

Prager, M. H., and M. C. Fabrizio. 1990. Comparison of logistic regression and discriminant analyses for stock identification of anadromous fish, with application to striped bass (*Morone saxatilis*) and American shad (*Alosa sapidissima*). Can. J. Fish. Aquat. Sci. 47: 1570-1577.

Prager, M. H., C. D. Jones, and M. C. Fabrizio. In preparation. Comparison of six variants of discriminant analysis, five of logistic regression, and a model based on finite mixture distributions for stock composition analysis: an empirical study on five data sets. Draft available from the M.H.P.

Wolfe, J. H. 1970. Pattern Clustering by multivariate mixture analysis. Multivar. Behav. Res. 5: 389-398.

Wood, C. C., S. McKinnell, T. J. Mulligan, and D. A. Fournier. 1987. Stock identification with the maximum-likelihood mixture model: sensitivity analysis and application to complex problems. Can. J. Fish. Aquat. Sci. 44: 866-881.