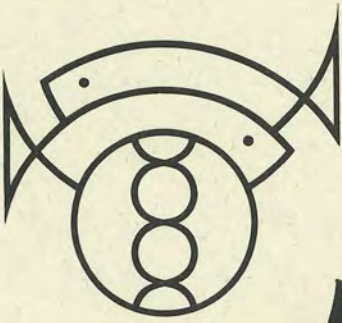


25~27 JUNE

1990



POPULATION GENETICS  
& ITS APPLICATIONS  
TO FISHERIES  
MANAGEMENT  
& AQUACULTURE

CONFERENCE • WORKSHOP

PROCEEDINGS

**PROCEEDINGS**

**POPULATION GENETICS  
& ITS APPLICATIONS TO  
FISHERIES MANAGEMENT  
& AQUACULTURE**

**ORGANISED BY THE CENTRE FOR MARINE SCIENCE  
AT THE UNIVERSITY OF NEW SOUTH WALES  
WITH THE ASSISTANCE OF  
THE FISHING INDUSTRY RESEARCH AND  
DEVELOPMENT COUNCIL**

**Held at the University of New South Wales, Kensington,  
on 25-27 June, 1990**

**Editor: Patricia I. Dixon**

**Centre for Marine Science, University of New South Wales, Kensington, N.S.W.,  
Australia.**

Published by:  
Centre for Marine Science,  
University of New South Wales,  
Kensington, N.S.W., Australia.

© Centre for Marine Science 1992

ISBN 0 7334 0277 1

This work is copyright. *The Copyright Act 1968* permits fair dealing for study, research, news reporting, criticism or review. Selected passages, tables or diagrams may be reproduced for such purposes provided acknowledgment of the source is included. Major extracts or the entire document may not be reproduced by any process without written permission of the Centre for Marine Science.

# CONTENTS

PARTICIPANTS	i
<b>SESSION 1: POPULATION STRUCTURE</b>	
Introduction	
An Overview - The Needs of Management.	3
<i>Dr. Roy Harden Jones</i>	
Chapter 1	
The Use of Electrophoresis in the Analysis of Fish Populations.	7
<i>Dr. Barry Richardson</i>	
<b>SESSION 2: POPULATION STRUCTURE (continued)</b>	
Chapter 2	
Molecular Population Structure and the Biogeographic History of a Regional Fauna : mtDNA Analyses of Marine, Coastal, and Freshwater Species in the Southeastern United States.	37
<i>Dr. John Avise</i>	
<b>SESSION 3: POPULATION STRUCTURE (continued)</b>	
Chapter 3	
The Analysis of Genetic Structure of Populations.	77
<i>Professor J.S.F. Barker</i>	
<b>SESSION 4: PANEL DISCUSSION</b>	
A Structured Discussion.	
The requirements of those concerned with stock assessments and the extent to which these can be met by geneticists.	103
<i>Dr. Murray MacDonald - Convenor</i>	

## **SESSION 5: GENE POOL MANAGEMENT**

### **Chapter 4**

Quantitative Genetics of Semi-Managed Populations. 137

*Professor John W. James*

## **SESSION 6: GENE POOL MANAGEMENT (continued)**

### **Chapter 5**

Genetics in Fish Culture. 155

*Dr. Colin E. Purdom*

Conservation of Genetic Resources. 171

*Dr. W.B. Sherwin*

## **SESSION 7: PANEL DISCUSSION**

Genetics in Aquaculture. 193

*Dr. Bert Sheridan*

## **SESSION 8: WORKSHOP - FUTURE DIRECTIONS**

### **Conclusions**

Past and Future Considerations. 221

*Dr. Bob Ward - Convenor*

## PARTICIPANTS

Dr. Jane Andrew  
Inland Fisheries Commission  
of Tasmania

Dr. John C. Avise  
The University of Georgia, USA

Professor Stuart Barker  
University of New England, Armidale

Dr. John Benzie  
Australian Institute  
of Marine Science

Tom Bergin  
Australian Fishing Confederation

Darren Brasher  
Fish Genetics Group  
University of Tasmania

Amanda Brooker  
CSIRO

Antony Brown  
Triton Trout Pty. Limited

Lindsay Brown  
Dept. of Genetics  
La Trobe University

Annadel S. Cabanban  
James Cook University of  
North Queensland

Dr. Don Colgan  
Australian Museum

Assoc. Professor Clive Crossley  
Sydney University

Professor Ross Crozier  
La Trobe University

Dr. Patricia Dixon  
Centre for Marine Science

Eric Edwards  
Fisheries Research Institute, NSW

Dr. Nick Elliott  
CSIRO (Division of Fisheries)

Elizabeth Ferdinandus  
University of New South Wales

Stewart Frusher  
Australian Institute  
of Marine Science

Hector Fuentes  
Fisheries Research Institute

Dr. Roy Harden Jones  
CSIRO (Division of Fisheries)

Dr. John Harris  
Fisheries Research Institute, NSW

Dr. Jane Hughes  
Griffith University

Professor John James  
University of New South Wales

Dr. Mike Johnson  
Zoology Dept.  
University of Western Australia

Patricia Kailola  
Bureau of Rural Resources

Dr. Bob Kearney  
Fisheries Research Institute, NSW

Dr. Clive Keenan  
QDPI, Southern Fisheries  
Research Centre

Tom Krasnicki  
University of Tasmania

Shane Lavery  
Zoology Dept.  
University of Queensland

Dr. John MacIntyre  
Centre for Marine Science

Dr. Murray MacDonald  
Marine Science Laboratories

Angus MacNeil  
Salmon Enterprises of Tasmania

Dr. Peter Mather  
Queensland University  
of Technology

Denis M'Gee  
Bayer Australia Ltd.

Andrew Miller  
Andriska Aquatics

Dr. Neil Murray  
La Trobe University

Dr. Craig Moritz  
Zoology  
University of Queensland

Dr. John Nell  
NSW Agriculture & Fisheries

Dr. Julian O'Dea  
Australian Fisheries Service

Dr. Jenny Ovenden  
University of Tasmania

Natalie Pavuk  
Inland Fisheries Commission,  
Tasmania

Dr. John Paxton  
Australian Museum

Dr. Colin E. Purdom  
Ministry of Agriculture, Fisheries  
and Food, U.K.

Pamela Reeves  
Australian Fisheries Service

Dr. Barry Richardson  
Bureau of Flora and Fauna

Harry L. Shaw  
Pacific Aquaculture

Dr. Bert Sheridan  
Fisheries Research Institute, NSW

Dr. Bill Sherwin  
Phillip Institute, Melbourne

Adam Smith  
Fisheries Research Institute, NSW

Dr. Peter Smith  
MAF Fisheries Research  
New Zealand

Adam J. Smolenski  
University of Tasmania

Peter Stevens  
Zoology Dept.  
University of Auckland, NZ

Dr. Iain Suthers  
Dept. of Zoology  
University of Sydney

Purnama Sulardi  
School of Biological Science  
University of New South Wales

H. Tanaka  
South Pacific Aquaculture Development  
Project/FAO

Stephen Thurstan  
Fisheries Research Institute, NSW

Richard Tilzey  
Bureau of Rural Resources

Dr. Michael Walker  
Fishing Industry Research  
& Development Council

Dr. Bob Ward  
CSIRO (Division of Fisheries)

Peter Ward  
Bureau of Rural Resources

Robyn Watts  
University of Western Australia

Lynette Woodburn  
Dept. of Genetics & Human Variation  
La Trobe University

Dr. R.W.G. White  
University of Tasmania

Julia Yeatman  
James Cook University  
of North Queensland



**SESSION 1**

**POPULATION STRUCTURE**

**Chairman: Dr. Patricia Dixon**

**University of New South Wales,  
Kensington, N.S.W.**

## INTRODUCTION

### An Overview - The Needs of Management

**Dr. Roy Harden Jones**

**CSIRO Marine Laboratories,  
Hobart, Tasmania**

The Fishing Industry Research and Development Council funds a significant proportion of the Fisheries R and D that is carried out in Australia for the benefit of the fishing industry, which includes the catching, culturing, transporting, processing and marketing of fish and fish products. The proposal to hold this conference was made by Pat Dixon and she argued that “the enlightened use of genetic data could, in the long term, lead to the conservation of resources in a way that was more likely than present methods to ensure sustainable yields”. It was intended that the conference should be a forum in which basic information could be provided for fisheries research workers and fisheries managers; and that there should be opportunities for discussion on the major issues relating to genetics and fisheries management, and genetics and aquaculture.

In the financial year 1990/91 FIRDC is spending about \$425K on six genetic or genetic related programmes. This is a modest sum, and represents 7% of the funds available for distribution. The Council will look to any recommendations that might arise from this conference for guidance as to the future directions of research and the need for funding.

As a member of the Research Council I am fully aware that when applicants come to the treasurer’s table they sometimes seem to promise more than they can reasonably expect to

*Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney 25-27 June, 1990.*

deliver, but I do not think that geneticists are any more optimistic than some of my colleagues who switch echosounders on or off or delve into the microstructure of the sagittal otolith.

What could fisheries biologists, who give scientific advice to managers, or Management Advisory Committees, reasonably expect from those who study genetics? Management Advisory Committees may have to decide on the magnitude of allowable catches, and in giving advice, fishery biologists must be able to assign catch and fishing effort appropriately in a particular population. For the fisheries biologist, the stock is the unit of management and has been defined as "a relatively homogeneous and self-contained population whose losses by immigration and accessions by immigration, if any, are negligible in relation to the rates of growth and mortality" (Anon, 1960). I must emphasize that this definition of a stock does not necessarily presuppose or depend on any heritable differences between stocks. The absence of genetic heterogeneity between spawning groups of the same species is not in itself a sufficient argument to manage a population as a single stock. But positive evidence of heterogeneity between spawning groups in population structure, environmentally dependent morphometric or meristic characters, natural markers or genetic variations, is a warning signal to be very careful indeed before deciding to manage them as a single stock. So here is an area where geneticists could contribute to fisheries management. And if the population is made up of several genetically different stocks, it may be important to know the mixing rates if the stocks are taken together on common fishing grounds. When conventional marking or tagging experiments are impractical, to what extent could genetics help? To be useful, mixing rates between stocks would have to be given with error limits of  $\pm 20\%$  at the 95% confidence level. Under what circumstances could the geneticist deliver; what strategic and tactical research is required; and at what cost?

The second subject to be discussed at this conference is the contribution that genetics could make to aquaculture. FIRDC is presently supporting 26 projects relating to aquaculture and these include work on disease, economics and the dissemination of information. These projects cost \$1.473 million, 22% of the Council's expenditure in 1990/91. Two areas have been identified as having priority: the first is closing the life cycle of species under laboratory or hatchery conditions, and the second the development of cheap and effective feeds. Presently there is no substantial funding for genetical studies. Council will be looking to this conference for guidance as to what role genetics might be expected to play in the development of aquaculture in Australia.

My own view is that it would be prudent to determine the extent to which important variations in performance - for example growth rates, conversion rates, and fecundity - can be attributed to environmental or genetic factors. What is the heritability of these variations? Is it not important to answer these questions before embarking on expensive selection or biotechnological studies?

I hope that the Conference will be able to give Council some guidance in these matters.

#### REFERENCES

Anon, 1960. Proc. Joint Scient. meeting of ICNAF, ICES, and FAO on fishing effort, the effect of fishing on resources and the selectivity of fishing gear. Volume 1 - reports. Spec. Pubs int. Commn NW. Atlant. Fish., (2) 45 pp. see p.8

## CHAPTER 1

### The use of electrophoresis in the analysis of fish populations\*

Dr. Barry Richardson<sup>†</sup>

Bureau of Flora and Fauna  
Canberra, A.C.T.

I understood the task today to be somewhat analogous to the challenges Roy has made. I will try and take up some of these challenges later in my talk. I thought that I would start by introducing the basic concepts of electrophoresis and the relevant genetics. I will then go on to be fairly critical about the way electrophoresis has been used and to make some suggestions as to what I think needs to happen before we can use it effectively.

#### ELECTROPHORESIS

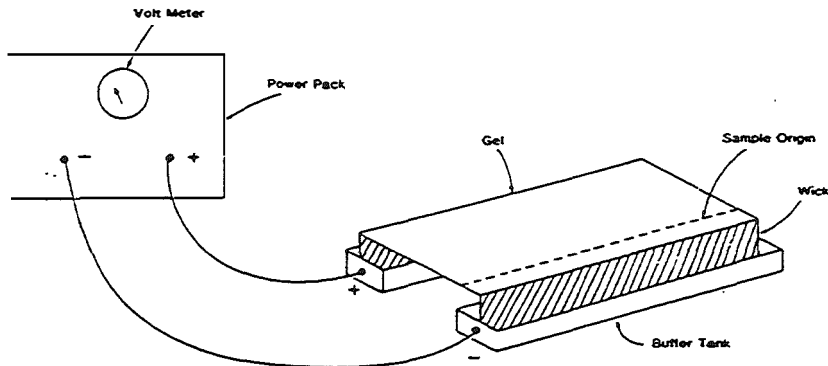
I would like to start with some acknowledgements. You will find my talk divides into two parts, the very clear and intellectually sound parts and the woolly parts. I own the woolly parts and I would like to acknowledge Peter Bavistock, Joanne Daly, Tony Lewis and Murray MacDonald who jointly can take credit for all the good parts. More detailed information on this topic can be found in Richardson *et al.* (1986).

Electrophoresis can be defined as the migration of charged particles under the influence of an electric field, so it has nothing necessarily to do with proteins. In our context it started from

*\*Paper presented at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990. †Present address: Faculty of Science & Technology, University of Western Sydney, Hawkesbury.*

observations late last century where it was found that charged particles could move in solutions under the influences of an electric field, the speed of migration varying with the size of the charge. A form of electrophoresis was developed and used in chemistry early this century, called moving boundary electrophoresis, in which different solutions were placed in either side of a U-tube separated by a removable barrier. When the barrier is removed and current applied, migration at the interface of the solutions can be studied. This technique was used for quite a lot of basic chemical research in the early days. The difficulty with a technique like this is that you only get detectable migration at the interface. There are no discrete zones, there is just an interesting effect you can follow at the boundary. Furthermore, as you run it for an extended period it begins to collapse. The osmotic pressure between the two sides is significant, especially under migration, and the boundary breaks down shortly after the current is applied. As a consequence, the technique is of limited value.

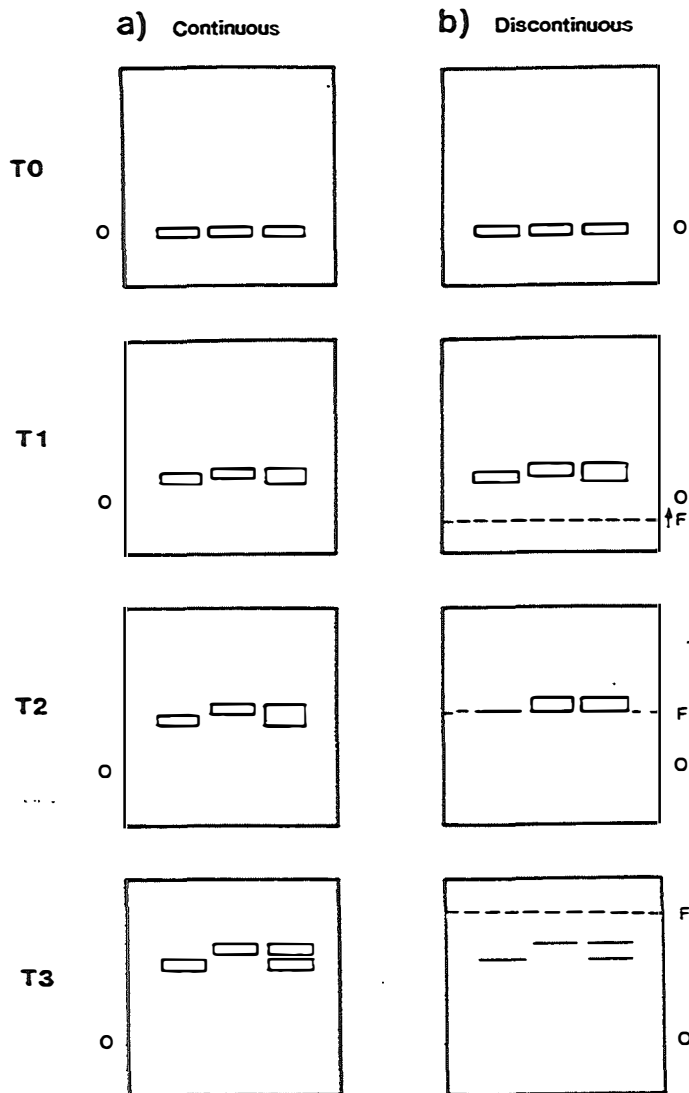
People then tried to improve the technique by filling the U-tube with glass beads and all kinds of other things to offset the osmotic effects. In the 1950's, people began to use various kinds of gels in place of the U-tube. Agar especially was used in the early days and then starch gel came along to be followed by acrylamide, cellulose acetate and a range of other media. Basically, the aim was to stabilize the system to mitigate the osmotic effects that broke down the sharp boundary between the different solutions. Having made this step, people then found that if they used a slab of gel, they could actually apply one solution to a point in the slab saturated with a second solution of buffer. The gel would allow the applied solution to be maintained as discrete zone, even under the influence of an electric current. This became known as zone electrophoresis in which the system, instead of being set up so as to produce a boundary between solutions, is set up to produce areas or zones of one solution in another solution. This has led to quite a simple system at the technical level (Figure 1). You have two baths with buffer in them, you have the gel between them, you can apply the samples at what is called the origin and you apply a current across the gel and any charged particles, for example proteins, migrate in that field. This is continuous electrophoresis with a single buffer used throughout the system.



**Figure 1.** Diagrammatic representation of a simple electrophoretic system.

A sophisticated addition to this was to introduce discontinuous electrophoresis. It is the same as continuous electrophoresis except the buffer that you put in the gel and the buffer you put in the tank are different. The significant thing is that, in the buffer that is in the tank, the proteins will move faster than they will if they are in the buffer that is in the gel (Figure 2). As you apply the current, tank buffer moves up into the gel (T1) until it catches up with the rear of the protein zone (T2). As it catches up with the rear of the zone the protein there increases speed, catching up with the front of the zone until the buffer reaches the front of the zone also. This sharpens up the bands (T3), so you get much sharper bands with a discontinuous system than with a continuous system.

Electrophoresis is used very widely. It is used quite extensively in DNA research, but I have taken my ambit to be protein electrophoresis and I will not talk about the methodology as it is used in DNA work which I assume we will be covering in later talks if relevant. It is simply one of the techniques used in processing of DNA.

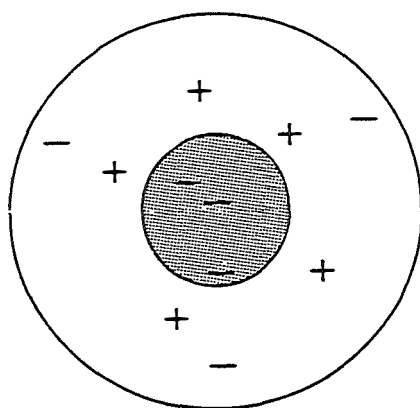


**Figure 2.** Example of electrophoresis using either continuous (a) or discontinuous (b) systems. Specimens are loaded onto the gel at the origin (o) at time  $T_0$ . During the run the front (f) between the buffers in the discontinuous system moves up the gel. The proteins, dependent on the charge they carry, move up the gel at different speeds.

A number of factors affect the rate of migration of protein molecules in a gel. The most significant factor is the surface charge on the protein and this can be changed by changing the pH. It can be that two proteins have the same charge at a particular pH, you change the pH and



one of them may change its charge and you can now separate them. Other things that can affect the migration rate include the shape and size of the molecule. Clearly a globular protein of the same molecular weight as a string shaped protein will travel through the gel in quite a different way. Consequently you can separate them on shape. Size also will affect migration rate. You can also change the way in which you make the gel so that you change the pore size in the gel and therefore increase the amount of 'friction' on the larger protein molecules as they move through the gel. As a consequence, with a skilful choice of the size of the pores you can again effect separations that are otherwise impossible.



**Figure 3.** Diagrammatic representation of the effect of charge on the functional diameter of a protein. The negative charged protein is surrounded by a cloud of positively charged ions. This cloud attracts further negative ions.

The next thing you can change is the buffer. This has a number of effects (Figure 3). Firstly, as the protein has a surface charge on it, it attracts molecules of the opposite charge out of the solution and these form a charge 'cloud' around the protein. This 'cloud' further attracts oppositely charged molecules again. The effective size of the protein as it runs through the gel is nearer size A than size B. As a consequence, if you change the nature of the buffer you use and you change the ionic strength of the buffer you can change the apparent size of the protein and thus affect the rate it runs in the gel.

Secondly, the rate and speed at which proteins run through the gel are also affected by the potential difference you apply across the gel and the ionic strength of the buffer. The weaker

the buffer the more the current has to be carried by the protein so the faster it moves. The faster it moves the more smeared the zone gets. With higher ionic strengths, the proteins move slower but the gel gets hotter denaturing the proteins and you have other kinds of trouble. So, what in fact looks intellectually simple and for which we have quite a reasonable understanding, turns out in reality, when you try and run the gels, to be a totally arcane art. You have no idea which ionic buffers are likely to work better, which is the best pH for a particular protein, whether a discontinuous or a continuous system is better, what pore size is right for a particular protein. You forget all of the theory and you just try it out. Jack mackerel like buffers with EDTA in them but gemfish do not. So this is a highly exciting and significant result, but you need to know it if you are going to run gels.

Having separated your proteins in the gel you then need to know what they are. This is the staining system and was the other key breakthrough that occurred in the late 1950's and early 1960's. In the first instance there were non-specific stains that would stain all the proteins in the gel. With these you ended up with gels with very large numbers of bands; it showed you all the proteins that were in the specimen that you applied and ran. The difficulty was that if you ran another specimen along side it, it was hard to know where the homologies were between the bands in one sample and the bands in the next sample. This was simplified by the development of specific stains which stained only particular enzymes, allowing homologies to be identified. You can now routinely stain a very large number of specific proteins and a good laboratory can identify something like 50-60 different loci.

The take home message, and this is the optimistic part I might add, is that zone electrophoresis is a quick, relatively cheap method for identifying charge or size related variation in proteins. You can detect genetic differences relatively quickly, quite cheaply and you can run very large numbers of samples in a reasonable time. That is the good side.

## GENETICS

I now need to step aside from that and introduce some genetics to relate the technique to the subject of this symposium. As I said, we are talking about electrophoresis as it relates to proteins. The simple relationship is that DNA, which is the genetic material, is translated as messenger RNA, and eventually appears in the form of polypeptide chains, one or more of which form

proteins. We detect variation in proteins using electrophoresis. The charge changes on the protein can be due to several reasons, the first of which is due to changes in the amino acid sequence of the protein. Now, changes in the amino acid sequence are due to changes in the underlying genetic material that codes for the primary structure of the protein. As a consequence, from the genetic point of view, we can treat a single polypeptide chain as the product of a single locus, or gene. If there has been a change in the DNA such as to lead to a change in the surface charge of the protein, this will lead to a different migration rate and different forms of the same protein chain will now run different distances on the gel.

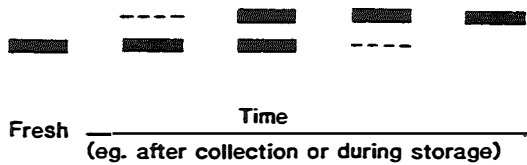
Genes may occur in more than one form or allele. Each individual has two copies of each gene, one inherited from each parent, and if they are different alleles, this individual would be said to be heterozygous for this gene. Individuals who have two copies of the same allele are said to be homozygous for the gene. So in Figure 2, from the point of view of electrophoresis, the left and central samples came from individuals homozygous for alternate alleles while the right sample is from a heterozygous individual. The particular attraction of electrophoresis is that it is easy to make such identifications. With two exceptions, the alleles identified using electrophoresis are codominant. That is, if the specimen is heterozygous you will be potentially able to see both bands on the gel. There are two cases where this doesn't happen. A reasonably common case is found in esterases and some other forms of protein in which you get alleles without enzyme activity ('null' alleles) and so cannot be detected in the heterozygous state. The normal/null heterozygotes show the same single band of activity as a normal homozygote and these two classes cannot be differentiated. So you have what you might say is a dominant condition where the visible allele is dominant over the null. One has to be careful about these because they affect the way you would interpret the results of such work. The second case, which I will just mention because everyone seems to have forgotten it, is that there are two reports in the literature of dominance for an identifiable protein that stains up (Law and Munro, 1965; Wilcox, 1966). This is in chicken phosphatase in which the two alleles are not codominant. When one is present the other is not detected. Nobody has ever followed it up. It has been in the literature twenty years and it has been conveniently forgotten. These are the only cases I know of dominance in electrophoretic typing.

The next question is, what proportion then of the genetic variation present in the DNA in the protein can you pick up using electrophoresis? In other words, is electrophoresis detecting most

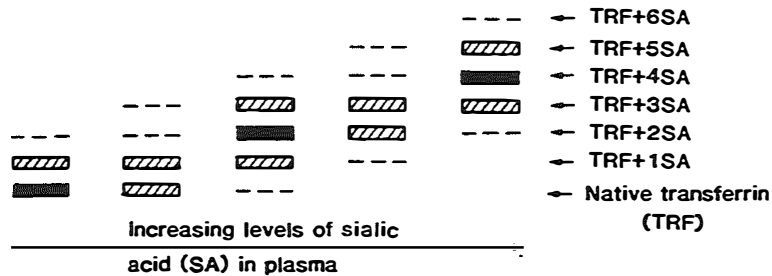
of the genetic variation in your population or only a small part of it? DNA sequencing has allowed this to be measured in the sense that people run gels from individuals and look for variation in a particular protein and at the same time sequence the DNA underlying that protein in the same individual. One such study (Kreitman, 1983) has shown of 43 alleles identifiable at the DNA level only one of them is detectable at the electrophoretic level. So, fundamentally you are looking at the situation where the vast majority of the DNA variants are not detectable using protein electrophoresis.

Furthermore, separate mutations at different points may lead to amino acid substitutions causing the same net surface charge state on the protein. It can be that a whole range of mutations have occurred each producing a protein with a single difference in charge on the surface from the original form. As a consequence, you can be misled in interpreting field data if you think a derived electrophoretic form of a protein is a single allele. You must ask the question, and it is the first biological question that I will put to you, are you looking at a single allele or not? Is it derived from a single ancestral mutation somewhere or are you looking at separately derived alleles that may have arisen in different areas independently? It makes a difference to your interpretation, clearly, because if it is the same allele and it is found in two areas that means that sometime, somebody has moved from one population to the other. If it is a parallelism, there is no logical necessity for there to have been any movement between the two populations.

The next problem, which we have to take into account when we are doing our work, is other modifications to the protein that will lead to changes in surface charge. There are two that I think we need to know about. The first one is change in protein with time (Figure 4A). When a protein is collected or when it is produced inside the organism it is in one form, as time goes on there are changes in the protein such that it changes its surface charge. In the figure there is a gradual reduction in the strength of the slower band and an increase in strength of the faster band with time. The critical thing with this is you can see that if you take the two end examples and the middle one, you are looking at something that crudely looks like the previous figure of a genetic polymorphism. So the first rule is, when you are running electrophoresis you had better look at aging effects on your samples, you had better look at the effects of how long the fish was on deck before the specimen was taken, where you kept it,  $-15^{\circ}\text{C}$  or  $-50^{\circ}\text{C}$ , how long it stood in the laboratory, whether it was frozen and then thawed and refrozen etc. You have to know what effects these actions have on the proteins if interpretation is to be valid.



**Figure 4A.** Diagrammatic representation of the effect of aging in a protein on its migration rate on a gel.



**Figure 4B.** An example of post translational modification to a protein. As more sialic acids are attached to the surface of the protein, the charge on the surface changes and the migration rate changes.

The second thing that can affect your protein is the attachment of small molecules to the surface of the protein. These are often charged, for example sialic acid, and as you attach more sialic acids so you change the net surface charge (Figure 4B). If you are not careful you can end up with the situation where you can misinterpret such data to be a genetically based polymorphism rather than an environmentally or production based polymorphism.

I would have to say these problems in interpretation are more theoretical than practical, only a very inexperienced operator would make such mistakes. They do get made though.

Take home message then: variation in proteins detected using electrophoresis constitutes a small subset of the genetic variation present in the underlying genes and is subject to interpretation problems due to parallelism and to environmental modifications of the protein.

## POPULATION GENETICS

We now have a structural difficulty if you would like to call it that, in the order of speakers as I now have to talk about the interpretation of the data. A little later in the morning, Professor Barker is going to be talking about the mathematical analysis of genetic data, so what I will do is just make the points I need to make. What it boils down to fundamentally is this, data from single genes is data of a binomial kind. You have states of data, allele 1, 2, 3 or 4, an individual has two copies of these and you can do simple analyses of the number and frequency of alleles and genotypes. In other words, if we take a simple polymorphism, you can count the number of specimens of each homozygote and each heterozygote. From these you can calculate the frequency of each genotype and then of each allele in the set of samples. In a random mating population it is possible to predict the frequencies you would expect, so you can look at the frequencies of the various genotypes you have found in your sample set and you can compare that against what you would predict if the samples were taken from a randomly mating population. This predicted distribution is called the Hardy-Weinberg equilibrium. Secondly, you do not get the frequency of genotypes expected under Hardy-Weinberg equilibrium conditions under certain circumstances. These are the ones that make the whole business work from the point of view of using electrophoresis in population studies (Figure 5). If you sample one random mating population with allele frequencies 0.8 and 0.2, and you take another sample set from one with the reverse allele frequencies and combine the samples (to reflect the effect of sampling a mixed population), you can calculate the number of each genotype and allele frequencies in the mixed group. If you calculate the expected genotype frequencies from the allele frequencies of the mixed sample set it can be seen that the predicted and the observed data do not match. You have an apparent deficiency of heterozygotes and this is called a Wahlund Effect. The implication of the Wahlund Effect is that there is inbreeding or, in our terms, mixtures of groups of organisms have been sampled that breed separately and have different allele frequencies. I will leave that for Stuart to explain further. From our point of view, you can estimate the allele frequencies, you can calculate the distribution of genotypes under the Hardy-Weinberg equilibrium conditions and you can test for a Wahlund Effect.

Population:	Genotype Frequency		
	$A_1A_1$	$A_1A_2$	$A_2A_2$
1. $p = 0.8, q = 0.2$	0.64	0.32	0.04
2. $p = 0.2, q = 0.8$	0.04	0.32	0.64
Mixed 1 and 2 in equal proportions:	0.34	0.32	0.34
Allele frequencies: mixed	$p = 0.5, q = 0.5$		
Expected genotype frequencies in mixed population assuming Hardy Weinberg Equilibrium:	0.25	0.50	0.25
Conclusion:	Deficiency of heterozygotes		

Figure 5. An example of a Wahlund Effect caused by the mixture of two populations.

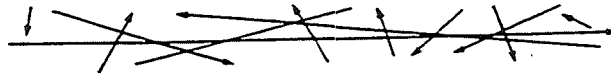
## POPULATION STRUCTURES

I next want to talk about population structures. I am going to start by saying I entirely disagree with the pattern assumed and used by fish managers and I completely disagree with Roy's proposed definition because I do not think it has anything whatsoever to do with real populations.

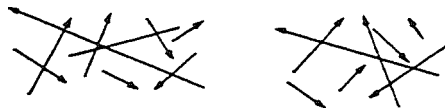
There are three possible models for population structure. The first one is called panmixia (Figure 6A). Panmixia means that over the whole range of a species, organisms move such that mating is at random with respect to the whole of the gene pool for the whole of the species. The second model I am going to call discrete subpopulations (Figure 6B). This is the model that fisheries managers prefer because it matches something that they can use when they have a real world and a real problem to deal with. In the model, the total range a species is subdivided to produce a series of geographically discrete groups. Breeding is within groups but there is very little movement, if any, between groups.

A typical example of discrete sub-populations is one Murray MacDonald (MacDonald, 1980) showed for snapper in Spencer's Gulf. These fish had quite different genetics to the ones outside of Spencer's Gulf, and because of the different water masses and so on, there was very little movement between the populations. So, irrespective of what the rest of the snapper do, there was a nice discrete sub-population secreted away in Spencer's Gulf. The third model is called an isolation by distance model (Figure 6C). In this case, there are restrictions in gene flow from one end of the range through to the other sufficient to allow genetic divergence between distant groups. The restriction is not because of barriers to movement of organisms, but simply because

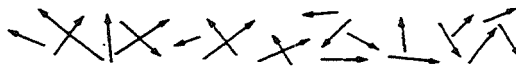
the distance moved by any given individual is small relative to the size of the population.



a) Panmixia (free interchange)



b) Discrete subpopulations (no interchange between subpopulations; free interchange within subpopulations)



c) Isolation by distance (local interchange only)

**Figure 6.** Diagrammatic representation of the distance moved between birth and breeding under the three models of population structure.

In summary, under panmixia, there is movement of different distances by different individuals but sufficient movement to mean that there is free exchange of genetic material across the full range of the species. With sub-populations there are discrete units of population, which are panmictic within themselves, but there is very little exchange between the groups, logically because of barriers, physical barriers perhaps, or behavioural barriers (e.g. time of breeding). The third model is the isolation by distance model where the movement of the individuals relative to the range of the species is small, such that individual movements are restricted to different parts of the range of species. There is the possibility of gene flow from one end of the range to the other in time, if there are sufficient generations. I don't think I want to hide what my view is. My view is that 99.9% of the world runs on an isolation by distance model.



Therefore the fundamental assumption made in fisheries stock analysis and stock management is usually in error.

## GENETIC ANALYSIS OF POPULATION STRUCTURE

How can we separate these models using genetic data? To begin with I am going to try to define a set of words, and I should warn you that other people will use the same terms in significantly different ways, there is no common vocabulary in this subject. As you can see (Figure 7), I have used stock as a management term, it just means the fish of a particular species in an area. 'Population', to me, means the whole of a population of a species over its entire range. 'Subpopulation' means a small group or subset of fish that is in some sense separated from others and is perhaps most nearly related to Roy's stock definition. A 'neighbourhood' is related to the average distance moved by individuals between birth and breeding and describes the unit of a population in terms where there are no discrete barriers, where there is a long or wide range in distribution but any given specimen or the antecedents of any given specimen come from a relatively small area relative to the entire range. You cannot really use subpopulation in this case and the proper term to use if you want to stick to the traditional literature is to talk in terms of neighbourhoods.

<i>Sample set:</i>	A group of individuals collected at one place at a particular time (e.g. one shot of the net for fish or a sample of snails from one garden on one night).
<i>Replicate sample set(s):</i>	Further sample set(s) collected in the same local area as the original sample set within a short period of time (e.g. additional shots of the net or a sample of snails from the same garden on a subsequent night).
<i>Stock:</i>	A term applied for convenience to a group of individuals in a predefined area. A stock may be represented by a sample set or a number of sample sets. In management situations, it refers to the unit of management.
<i>Population:</i>	A group of individuals scattered at various densities throughout the range of a species. The population may be panmictic or consist of subpopulations or neighbourhoods.
<i>Subpopulation:</i>	A reproductive group of animals that share a common gene pool. Panmixia occurs within subpopulations but not between subpopulations. The discrete subpopulation model applies.
<i>Neighbourhood:</i>	The region surrounding an individual within which the gametes that produced that individual can be considered to have been drawn at random. There is continuity of the population but this does not ensure complete panmixia i.e. there is isolation by distance.

**Figure 7.** Definitions of terms as used in this article.

How do you differentiate between the models? We now face the situation where we try and

become a little more practical. We have a fish population, what do you do? How can you differentiate? We start with the simplest of situations and then try and add the confusing factors as we go along. The discrete subpopulation model is relatively simple. Intellectually it is straightforward. What you are going to try and do is take some sample sets from various places in the range of the species and look to see if you can find differences in gene frequency or the presence of alleles in some places not found in others. You then begin to map your data. Certain sample sets, though collected from different places, do not differ significantly. Sample sets from elsewhere differ from the first group but are similar to each other. The geographic distributions of these groups identify the geographical regions inhabited by the different groups. If you can find this situation then you can propose that you are looking at discrete subpopulations and you can begin then to go forward with a working hypothesis of discrete subpopulations that can be managed as units. This becomes your working hypothesis and is tested by further cycles of sampling and analysis. These would aim to identify and confirm the nature and locations of discontinuities in allele frequencies.

One of the problems we have with electrophoretic studies in reality is that we tend to do one cycle of experiments and leave it at that. We go around the system once and then we assume we have an answer. There is not really the money or the time to go back. In almost every situation there should be a second detailed study after the first survey to examine in detail what happens in reality. Are there regions with intermediate gene frequencies? In your areas of intermediate allele frequencies do you get anything that looks like a Wahlund Effect? If allele frequencies at one locus change suddenly do other genes change at the same time?

How would you identify a situation where an isolation by distance model was likeliest? In the simplest form what you would find is differences in gene frequency in different places with no discontinuities. You would find changes in the gene frequencies were gradual (i.e. clines), there would be no sudden changes from one gene frequency to another.

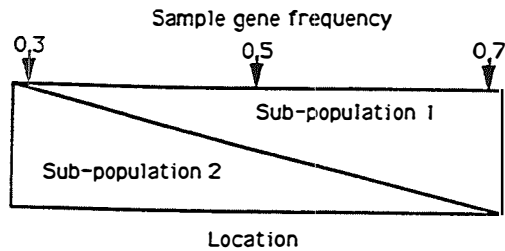
If you find clines, you are most likely looking at isolation by distance genetics. If you find no variation in gene frequency over the entire range of the species, in other words everywhere looks the same as everywhere else, your working hypothesis is that you are looking at panmixia. In other words, you've found no evidence of the other two hypotheses and therefore you presume panmixia. There are species in which the entire population is panmictic, if they cover a relatively

small area or are highly vagile.

## PRACTICAL ISSUES IN STUDYING POPULATIONS

What do we try to do in practice? We begin with surveys across areas, we are looking firstly for polymorphisms and then differences in allele frequencies. We want to know whether any differences in gene frequencies are reflected in intermediate populations as discontinuities in gene frequency or whether they are clinal in nature. If they are clinal we will want to know if there is a Wahlund Effect at intermediate gene frequencies. Can you get into trouble trying to do this? Yes, you can.

Roy introduced a key concept, from the point of view of the analysis of data, and this has to do with situations of overlap. You must remember when we are dealing with genetics we are dealing with breeding populations. We are interested in the units of breeding stock, and what happens in non-breeding times of the year is critical if that is when the samples are collected. If the fish breed in discrete areas but move to common feeding grounds and you make your electrophoresis study from sample sets from the feeding grounds, and you are trying to find out whether there are discrete stocks, you are in deep trouble straight away. You are going to find it difficult because the subpopulations are partly mixed. If the subpopulations have ranges as shown in figure 8, and you sample the ends and the middle, you might find a gene frequency change from 0.7 to 0.5 in the middle and then to 0.3. You end up with a lovely looking cline, you will think isolation by distance model when, in fact, you are looking at discrete subpopulations with overlap in the centre of the range. What you would be wanting to check then is for evidence of mixing of populations which is detectable as a Wahlund Effect. You would find that the end sample sets would be in Hardy-Weinberg equilibrium, in theory, and those in the middle would be out of equilibrium, implying a Wahlund Effect. The other point I would make here just in passing, it ought to be a truism but it is not, is that when you start your electrophoresis study you start by sampling breeding populations, you do not start by sampling the easiest place to catch them which is where the commercial trawlers catch them which is usually when they are feeding. That is the easiest thing to do but it is not the best thing to do. The best thing to do is to sample the breeding groups and work out then into the feeding populations to detect the ecological processes that occur in the population.



**Figure 8.** Diagrammatic effect of mixing two subpopulations. At the left end of the area only subpopulation 2 with a gene frequency of 0.3 is found. At the right end only subpopulation 1 is found. In the centre a mixed population is sampled with an intermediate gene frequency.

The next problem we face is deciding the nature of the sample unit relative to the sizes of the units of population you are looking at. One of the key concepts, if using electrophoresis is that the sample sets you take have to be homogeneous or you are in trouble before you start. If you have a situation like Figure 9b, you can see each sample set is taken from a discrete subpopulation. In figure 9c, the second sample is taken from the same subpopulation as the first while the third sample is taken from another subpopulation. This is intellectually, or perhaps subconsciously, the kind of thing we think we are normally doing when we are carrying out an electrophoresis study. The question is, does reality match our preconceptions? Figures 9a and 9d show cases where subpopulations overlap or the subpopulations are so small that a single shot of the net will always include more than one subpopulation.

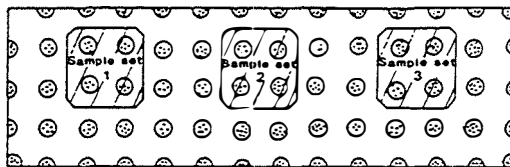
The last scenario sounds a bit crazy, it is saying that local populations are very small. I suspect that we are going to find as time goes on that this is in fact commonly the case. The first time I said this I nearly got thrown out of a manager's meeting. But the more people I talk to that do electrophoresis on fish populations, the more we find the same kinds of things. You do not find random distributions of rare alleles, their distribution is clumped. When you get, say, an allele turning up one time in a thousand specimens, you do not find it in one in a thousand specimens. One shot of your net has got six copies of it and you never find it for the rest of the study. We all have anecdotal evidence of this kind that says that something like this is happening, that there is structure in natural populations at a very fine scale and this does affect the problem of interpretation. Because if a local population is inbred, in the sense that they are relatives, you take a number of them into a single sample set and you are going to get a Wahlund effect, you are getting an average gene frequency that represents nothing. You are going to have trouble

with your interpretation.

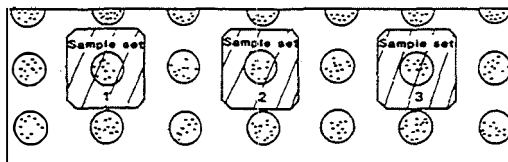
It is fundamental, absolutely fundamental, that the sample set collected must be from the most homogeneous unit of population that you can find. They will be fish of the same size taken at the same time and at the same place. If you do anything else you start with problems. Because anything you find, any variation you find by comparing between sets collected in other ways is subject to a whole range of interpretations, none of which you can discriminate and most of which are usually ignored.

The next problem I need to mention is selection. Unfortunately, reality will strike your population study in that it is possible that the gene frequency of the fish you sampled was not the gene frequency of the unit of fish produced as larvae. That is, there has been selection in the process from then through to the time when you caught your specimens. I think from our point of view at this end of the world the classic example is Peter Smith's snapper studies in New Zealand where he found quite significant differences in gene frequency between year classes caught in the same area (Smith, 1979). He went back and looked at the temperature in the year of recruitment and found that you could relate some of the differences at least in gene frequencies to water temperature for the year of recruitment. This should be taken as a warning. I would say that I think cases of selection that are so strong as to be detectable by the very low power of the kinds of techniques we use are relatively rare but you ignore them at your peril. You should take these things into account. Which is why if you are shooting on a single location you better break the sample set up by year classes and look at them separately. Do not just think you can combine them because they all came out of the one net. It is just not that simple. More commonly what you find is that selection is so low you can not detect it on a single generation basis. It does change the average gene frequencies of areas over time. This is the basis for the establishment of clines, the establishment of quite distinct gene frequencies in local populations in extreme environments like Spencer's Gulf.

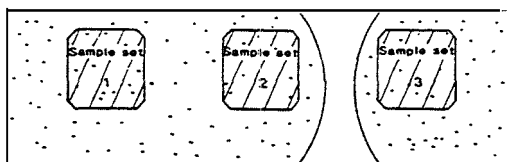
a) Sampling area includes several subpopulations each of very limited distribution.



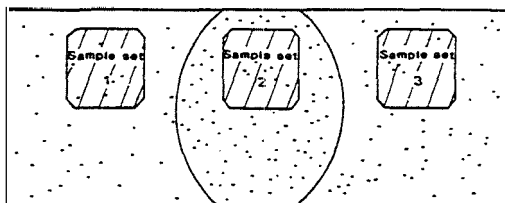
b) Each sampling area contains a single subpopulation.



c) Several sampling areas contained within the distribution of a single subpopulation.



d) Some sample sets include the distribution of overlapping subpopulations.



**Figure 9.** Diagrammatic representation of the effect of subpopulation size on the material obtained in sample sets of the same size.

The last problem is that the species you are interested in has insufficient genetic variation. You might think that you can go out there and find your polymorphisms and away you go. The truth of the matter is many species do not have sufficient genetic variation for you to do anything genetic with them at this level using electrophoresis. The study of skipjack in the Southwest Pacific (Richardson, 1983) was bedevilled by the fact that even though we ran something like 50 loci we found three polymorphisms, one of which was an esterase. This was a great pain. Tony Lewis studied scombrid species found in the Australian region. He looked at 19 species and he found the range was from no polymorphisms to 26% of loci examined polymorphic,

depending on the species. So my first comment to FIRTA would be: do not fund a big project, fund a pilot study first and make sure there is enough genetic variation to go on with before you actually support a big project.

Statistical problems are the next class we need to look at. Firstly, the power of the test. We are looking at binomial data and there are standard methods which I will not go into but on which the power of the test depends. Firstly, it depends on sample size: the bigger the sample size the more powerful the test. Secondly, it depends on the size of the difference in gene frequency between sample sets that you are willing to accept as biologically meaningful. That is not an easy question to answer. Do you want a gene frequency difference of 0.5 before you are willing to say that these are significantly different or are you willing to settle for 0.2 or 0.1 or 0.05? How different do they have to be before you are willing to say that this is not only a statistically significant, but also a biologically significant difference? There is not much discussion in the literature that I know of on how big the difference needs to be before you want to accept it. Mostly if we can find a statistical difference we are eternally grateful. Thirdly, you then need to determine the confidence you wish to place on the conclusion. Normally we set at 5% the level of obtaining the result by chance. In other words, the normal 95% confidence limit on a Type 1 error. That is, we expect the difference we found could not have happened by chance more than one time in twenty. You also need to set the Type 2 error. That is when there is a real difference between the two sample sets but you did not find it. Because the higher you make the significance level to detect a Type 1 error the lower the chances there are of detecting a Type 2 error for the same sample size. Everybody sets a Type 2 error very low mostly at 50% or even lower than that. So they are saying, we are willing to miss some differences to increase our chances of finding some differences. It is possible to calculate what sample sizes you need to meet these criteria. In figure 10, I have taken the Type 1 error to be 95%, or 5% error rate as we would normally use it. The Type 2 error can be at 50%, 20% or 10%. You can determine the difference in gene frequency for a biologically significant difference you are willing to wear. Does it have to be a 0.5 difference, 0.2 or 0.1 or are you willing to wear a 0.05 as being biologically significant? Also the power of the test depends on the gene frequency, which is printed on the top of the table. The body of the table shows the number of individuals that you have to sample, if you are to detect such a difference. It is extraordinarily chastening for anyone who does field electrophoresis. I would like to try to find 0.1 differences in gene frequency and that is a sample size of about 200 specimens per sample set. So, casual studies of small numbers

of specimens cannot tell you a great deal. The take home message is: In most field studies in the literature the sample sizes have been too small to detect the kind of biologically meaningful difference that the author set out to detect and the money spent on the study was largely wasted.

$\beta$	$\Delta p$	p				
		0.55	0.70	0.80	0.90	0.95
0.5	0.05	760	645	492	276	146
	0.10	190	162	123	69	50*
	0.20	48	40	31	25	50*
	0.50	6*	9	13	25	50*
0.2	0.05	1554	1319	1006	564	299
	0.10	389	332	252	141	76
	0.20	99	82	64	27	50*
	0.50	16	14	13	25*	50*
0.1	0.05	2081	1766	1345	756	400
	0.10	520	444	337	189	102
	0.20	132	110	85	50	50*
	0.50	22	20	14	25*	50*

\*To test for the significance of the observed difference, a homogeneity  $\chi^2$  test requires certain minimum expected frequencies in each cell. In cells marked with an \*, the minimum sample size has been set to meet this requirement, assuming a minimum expected frequency of 5 in any cell.

**Figure 10.** A table showing the number of individuals that need to be typed for predetermined levels of divergence in gene frequency to be detected.

There is a second error that we routinely make, and it is almost universally made. The problem is that we repeatedly test data sets using 95% confidence limits. If you think for a moment we have a one in twenty chance of a false significant difference. That is, if you do 20 tests then you are going to get one significant result by chance just as a background rate. Now if you take five loci at four locations that's twenty tests. So if you ran all the tests on that data set you would get one significant result by chance and you should not do it. We make it worse because we set up the matrix of loci by locations and we look at it and see that there are about three combinations where the gene frequencies are pretty different. Those would be the only three we would test because we knew the others were so similar there was no point in trying to run a statistical test. But, nonetheless we have run the full set of comparisons and we are making a mistake.

The experimental design if you are going to do this kind of work: 1. Sample sizes must be at least 100 specimens, preferably more than 200 and must be taken as a homogeneous set. These should be from one age class from a single shot, not just 200 specimens from a singleshot. That means from every age class you pull out of the water you are going to have to take one to two hundred specimens if you want it to mean anything. 2. Many polymorphic loci should be used.



You should not make a decision on the basis of a single locus. In my view if you cannot find six polymorphic loci, do not start the research. My advice to FIRTA would be unless the proponent for the study came to you showing six good polymorphisms do not give him any money. 3. Sample sets should be from the smallest, most homogeneous population possible because until you can show homogeneity in your smallest unit and there is no variation

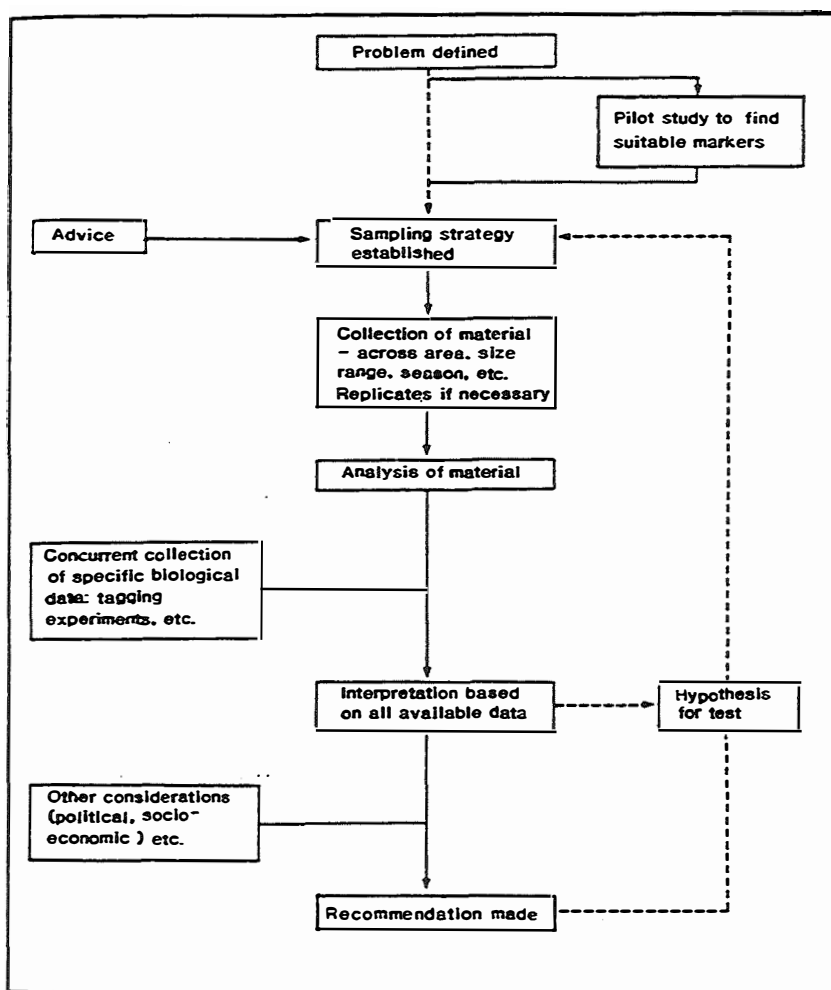


Figure 11. Diagrammatic representation of the steps to be followed in a large field study aimed at detecting population structure.

unexplained at that level you are not in any position to explain variation found at any other level in the analysis. 4. As a strategy, try and sample when possible geographical extremes and take replicate samples and sample across any potential physical discontinuities you can detect in your environment because those are the sites most likely to be informative.

A suitable strategy is shown in Figure 11. First define the question to be answered. Almost all kinds of management use of electrophoresis are done by mixed groups. There is a geneticist in a lab somewhere, there is a field team somewhere, there is an agency somewhere that is paying some of the bills and probably a few other people thrown in, for example two or three State departments who are going to be involved. Surely you all ought to agree what the question is before you start and you should all agree to work towards answering that question. Next find six polymorphisms before anyone spends any more money. From the gene frequencies you get from the preliminary study you can determine using figure 10 what size your sample sets would have to be, etc. You can then go through the steps down to the interpretation at which time you might direct an answer to the agency that is supporting the work or you go back and you do another cycle of research based on hypotheses developed during the first cycle of research.

One of the problems in electrophoretic studies is that there is not enough sampling done. Most people get enough money to do one cycle of sampling and that is the end of it. The reality is the first cycle of work should be just that, the first cycle of work. You need to go back then and refine what happens between the sample sites you looked at, what happens at other times of the year. In the first cycle we might sample the breeding populations so the next time we can go back and sample feeding populations.

My take home messages to finish: 1. There have been systematic errors in effectively all electrophoretic studies done and for that reason electrophoresis has not met its potential for the use of fisheries managers. These errors are in experimental design, in sample sizes, and the lack of replicate sampling. 2. Suitable statistical methods are unavailable, I have not had time to talk about the problems with neighbourhoods. There is almost no mathematics that allows you to begin to analyse neighbourhood structuring at this stage. 3. The assumption that managers make, because it makes management easy, that populations are structured according to what I have called the subpopulation model, is in my view systematically at fault because I do not think very many populations are structured that way. 4. The work is frequently pressed on the

geneticist by agencies needing results that are quick and dirty, where you take small samples, you do not think biologically and it is done as a spin-off of a tagging program or something. The study is quick and dirty and the results are certainly quick and very dirty and you might as well not bother, except maybe as a pilot study to look for polymorphisms. Design it properly and treat it with respect and you will get good results. 5. If the isolation by distance model is true, as I believe it is, almost everything that a fisheries biologist uses in management, Von-Bertalanffy equations, the whole structure of management processes, is flawed and somebody should go back, throw that all out and start again using something like diffusion equation mathematics to reanalyze it all. I do not think it is an accident that most fisheries that have been trying to manage for maximum sustainable yield have ended up being overfished. All modelling shows you is that if you fish mixed stocks with different  $r$ 's or different growth rates you fish to extinction some subsets of that population. You do not alter the age structure, the size of the stock just gets smaller, I think that is what fundamentally goes wrong with most of the management methods we use. We need to go back to square one and redo it all over again presuming models other than discrete stock models. I hope that stirs up enough trouble for you Pat.

#### REFERENCES

- Kreitman, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature*, 304: 412-417.
- Law, G.R. and Munro, S.S. 1965. Inheritance of two alkaline phosphatase variants in fowl plasma. *Science* 149, 1518.
- MacDonald, C.M. 1980. Population structure, biochemical adaptation and systematics in temperate marine fishes of the genera *Arripis* and *Chrysophrys* (Pisces: Perciformes). Unpubl. PhD Thesis. ANU: Canberra 336 pp.
- Richardson, B.J. 1983. The distribution of protein variation in skipjack tuna (*Katsuwonus pelamis*) from the central and southwest Pacific. *Aust. J. Mar. Freshw. Res.* 34: 231-251.
- Richardson, B.J., Baverstock, P.R. and Adams, M. 1986. Allozyme Electrophoresis: A handbook for animal systematics and population studies. Sydney: Academic Press. 410 pp.
- Smith, P.J. 1979. Esterase gene frequencies and temperature relationships in the New Zealand snapper, *Chrysophrys auratus*. *Marine Biology*, 53: 305-310.
- Wilcox, F.H. 1966. A recessively inherited electrophoretic variant of alkaline phosphatase in chicken serum. *Genetics*, 53: 799.

## DISCUSSION

DR. PATRICIA DIXON: Thank you Barry. I think that between you you've probably stirred up quite a bit of trouble. I wonder if the two key speakers could come down to the front and we'll take a few questions. Further questions can be saved until our workshop this afternoon. Do I have a question for either Barry or Roy?

PROFESSOR ROSS CROZIER: You gave these diagrams of various population structures but I wonder, and I speak with the armour of relative ignorance about fisheries matters, whether in fact the fish are seeing the geography in which they are placed in the same way as the people catching them. Mainly that if you look at a current, for instance, which is moving down the coast and you go out in your boat and you collect fish every year at that spot, you say that an enormous difference is taking place, it must be selection. Whereas, in actual fact, to look from the fish's point of view you should follow that current down the coast and collect where the fish are going. So, consequently you are quite right about everything you said as far as I could see about separating the various year classes and giving the data the absolute most precision that you can. Wouldn't it be true that very many of these differences would arise purely mechanically such as we see in fact on rocky platforms with limpets and stuff like that?

DR. BARRY RICHARDSON: Yes, that certainly can happen. I don't think in Peter Smith's case that was the situation, in that snapper are fairly sedentary and are not subject to that kind of movement in this particular case. But it's certainly the case for many other species. Tuna do that kind of thing. Any highly mobile species is subject to that kind of thing and you really do have to be very careful about your interpretation for that reason.

DR. PATRICIA DIXON: Do we have another question?

DR. COLIN PURDOM: A topic which does seem to be neglected in the electrophoretic literature is the relative values of the frequencies of the two homozygotes in a two or more allele system. This is something which ought to be relevant to clines.

DR. BARRY RICHARDSON: The relative proportions of them in the populations?

DR. COLIN PURDOM: Yes, do they agree with the cline model?

DR. BARRY RICHARDSON: The test does take the numbers of homozygotes into account. The real difficulty is the power of the test. You use the gene frequencies from the sample data to calculate the expected frequencies so that it does tend to hide any effects. Unfortunately, you use the genotype frequencies to give you a gene frequency but then use that to calculate the predicted distribution of genotype frequencies. That means that it does hit the homozygotes especially hard so that if one of the homozygotes has been selected against, it tends to shift the gene frequency such that you cannot detect the deficiency in homozygotes, you see, apparently the wrong number of homozygotes.

DR. COLIN PURDOM: With a cline involving a two or more allele system it ought to be possible to demonstrate selection at each end of the cline by reference to the relative values of the two heterozygotes. It just seems never to be done.

DR. BARRY RICHARDSON: No, I didn't. I'd rather not use that method, I'd like to talk to you about it after.

DR. PATRICIA DIXON: Is there another question?

DR. JOHN AVISE: I was a little surprised at your suspicion that some 99% of species are characterized by isolation by distance. Are you thinking there of marine fishes in particular because I shouldn't think that would be the case in the freshwater fish realm where there are certainly geographical or physical barriers to dispersal, nor among terrestrial creatures which often show disjunct habitats and geographic separations of a rather substantial nature.

DR. BARRY RICHARDSON: I probably overstated it. If there are solid physical barriers to movement and I don't just mean physically moving across a barrier, but perhaps, strong changes in temperature or something, yes then discrete subpopulations become significant and that does happen in terrestrial mammals. I've done at least as much work in terrestrial environments as fisheries environments and I would have said that I was even more comfortable with isolation by distance models in the terrestrial environment than the marine environment but I think that the distances are very, very short and that's what's catching people out. They're looking for

thousand mile ranges or hundred mile ranges. I think you ought to be looking for 200 yard ranges quite frequently which is why I mentioned those very small populations. I think microheterogeneity is something we haven't come to grips with properly.

DR. CRAIG MORITZ: Barry, I agree with most of your comments but I'm a bit concerned about the implications of your sampling size. I mean, a terrestrial biologist largely has got to battle with conservation authorities to take five animals per spot. Therefore, *ipso-facto*, is all my work useless? I'm a bit concerned. I think if you ask partly from the assumption on the difference of allele frequency that you want to detect in your binomial sample equation and the implications of what you said is that we need one to two hundred animals per age class, per breeding population with replicate sampling and those samples have to be homogeneous. Now, that's plain nasty. That alone will, I suspect have a fairly large impact on the breeding population just doing the sampling to do the analysis. What I'm coming to is what's the biological basis for selecting an 0.1 difference instead of 0.5? Don't we need to understand more about what creates those differences to bring about a proper frame for sampling?

DR. BARRY RICHARDSON: I support what you're saying. I don't know of, in the literature, a rational discussion of how big a difference in gene frequency should be before it's biologically meaningful as opposed to statistically meaningful. I mean is a 0.1 difference significant or is this something maybe to do with the microheterogeneity that we were just talking about a moment ago. Is it just local drift and therefore that kind of level of variation is really not significant for the questions we are dealing with here and therefore we shouldn't try and find 0.1 differences in gene frequencies because they're not informative for the processes of management. I just don't know of anyone that's actually sat down and tried to make the debate on what is biologically a meaningful difference in gene frequency.

PROFESSOR STUART BARKER: Is there not an additional question here? The numbers you are discussing are those necessary for certain statistical parameters to detect a difference between two populations. Really, in most cases, what we are interested in is a number of populations, or a number of subpopulations, and we should recognize that the numbers are not so large where the overall pattern is of interest, as well as the specific question of whether a difference between two subpopulations in gene frequencies at a particular locus is significant. There is the additional point that increasing the number of subpopulations (rather than two) and

increasing the number of loci will provide more information.

DR. PATRICIA DIXON: One last question.

ANON: I was wondering, if I could just talk a little bit on the economic rather than the theoretical side. Early in your talk you introduced electrophoresis as being a fairly cheap and easy technique. In a lot of fisheries work, especially when you are taking the samples, you start looking at getting them from commercial catches and you are looking at getting adult fish. If you actually look for breeding populations, it means that you are probably having to use charter vessels, you might have to start using specialized gears and therefore the economics of it becomes a lot more expensive, but it appears from what you've said that the results are going to far outweigh getting mixed adult stories.

DR. BARRY RICHARDSON: I think we've gone for the cheap solution and what we've got is no answers. It is a temptation, you are going to put in a grant application. Who wants to hire a vessel if you can go out on a commercial boat and do it for a lot less money? The real difficulty is, you've got to let the experiment decide what is right and then decide whether you can actually fish off commercial vessels and do it cheaply or you have to hire boats and wear the expense. I would hope FIRTA would take this away with them when they go, it is a very cheap laboratory technique but it's paying for the boat time that breaks your heart. However, I think if we want real results then it's going to cost real money.

**SESSION 2**

**POPULATION STRUCTURE (continued)**

**Chairman: Professor Ross Crozier**

**La Trobe University,  
Melbourne, Victoria**



## CHAPTER 2

# Molecular Population Structure and the Biogeographic History of a Regional Fauna : MTDNA Analyses of Marine, Coastal, and Freshwater Species in The Southeastern United States.\*

Dr. John C. Avise

Department of Genetics, University of Georgia,  
Athens, GA.30602, USA.

### ABSTRACT

I review the mitochondrial (mt) DNA data gathered in this lab over the past decade on the comparative phylogeographic patterns of 19 species of freshwater, coastal, and marine species in the southeastern U.S. Nearly all assayed species exhibit extensive mtDNA polymorphism, although still orders-of-magnitude less than predicted under neutrality theory if evolutionary effective population sizes of females are similar to current census sizes. In both the freshwater and marine realms, deep and geographically concordant forks in intraspecific mtDNA phylogenies commonly distinguish regional populations in the Atlantic versus Gulf Coast areas. These concordant phylogeographic patterns among numerous, independently evolving species probably evidence effects of similar vicariant histories of population separation, and can be related tentatively to episodic changes in environmental conditions during the Pleistocene. However, the heterogeneity of observed genetic distances and inferred separation times are difficult to accommodate under a uniform molecular clock. Additional population genetic structure within geographic regions is evidenced by species-specific shifts in frequencies of more closely related mtDNA haplotypes, and by high frequencies of private alleles in some species. The magnitude of local population structure appears partially related to the life history pattern and dispersal

*\* Paper presented at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney 25-27 June, 1990. A substantially similar version of this paper also appears in Oikos 61: in press (1991).*

capability of a species, and is commonly higher among freshwater than marine taxa. Overall, the mtDNA results indicate that conspecific populations can be structured at a wide range of evolutionary depths. The deeper subdivisions in an intraspecific phylogeny reflect the major sources of evolutionary genetic diversity within a species, while the shallower molecular separations evidence more recent population subdivisions that may nonetheless be relevant to studies of comparative dispersal, and to the assessment of stocks subjected to exploitation on contemporary timescales.

## INTRODUCTION

Over the past several years, our laboratory has conducted geographic surveys of mitochondrial (mt) DNA variability within a number of marine and freshwater species in the southeastern United States. These studies had a variety of immediate objectives, such as characterization of the genetic status of threatened or endangered species (Atlantic Sturgeon - Bowen and Avise, 1990; Seaside Sparrow - Avise and Nelson, 1989), examination of the evolutionary genetic consequences of an unusual catadromous life cycle (American Eel - Avise *et al.*, 1986), appraisal of levels of genetic variation in "living fossils" (Horseshoe Crab - Saunders *et al.*, 1986; American Oyster - Reeb and Avise, 1990), assessment of cytonuclear associations in a hybrid zone (Bluegill Sunfish - Avise *et al.*, 1984; Asmussen *et al.*, 1987), or examination of atypical molecular features of mtDNA variation (Menhaden fish - Avise *et al.*, 1989; Diamondback Terrapin, Lamb and Avise, in prep.). To date, my students and I have assayed geographic variation in some 19 freshwater, coastal, and marine species in this geographic region.

Here I will summarize these results in the context of a comparative appraisal of geographic population structure. Despite the diverse array of species represented in these surveys, several striking and unanticipated trends in mtDNA variation and phylogeny have emerged. Perhaps our empirical experience with the evolutionary genetic structure of a regional fauna will provide some general lessons on the kinds of contributions molecular genetic data can (and cannot) make to the pragmatic concerns of "stock" assessment and population management.

Several recent reviews have thoroughly summarized the major features of mtDNA evolution in higher animals (Avise, 1986; Avise *et al.*, 1987a; Harrison, 1989; Moritz *et al.*, 1987; Shields and Helm-Bychowski, 1988; Wilson *et al.*, 1985), and two reports dealt specifically with

applications of mtDNA data to fishery management (Avice, 1987; Ferris and Berg, 1987). These works should be consulted for background to this report.

In general mtDNA evolves rapidly and exhibits extensive polymorphism within most species. MtDNA also exhibits uniparental (maternal) inheritance, without known recombination between molecules from different female lines. Thus even within interbreeding populations, mtDNA lineages are genetically isolated from one another, such that any observed homologies in structure presumably result from historical connection in a matriarchal genealogy. The phylogenies of mtDNA haplotypes within a species can normally be estimated by reasonable evolutionary criteria such as cladistic analysis or parsimony, and the results compared to expectations of theoretical demographic models to make inferences about population histories (Avice, 1989). Caution must be exercised in drawing such population inferences, since any gene tree (such as that provided by mtDNA) represents only one realization of the multi-gene process of lineage sorting through an organismal pedigree (Ball *et al.*, 1990). Nonetheless, the phylogenetic content in a mtDNA gene tree, when interpreted in conjunction with the observed geographic distributions of mtDNA clades, provides one picture of the "phylogeographic" past of a species (Avice *et al.*, 1987), and thereby adds a historical perspective to population structure and intraspecific evolutionary process.

#### THE EVOLUTIONARY THEATRE OF THE SOUTHEASTERN U.S.

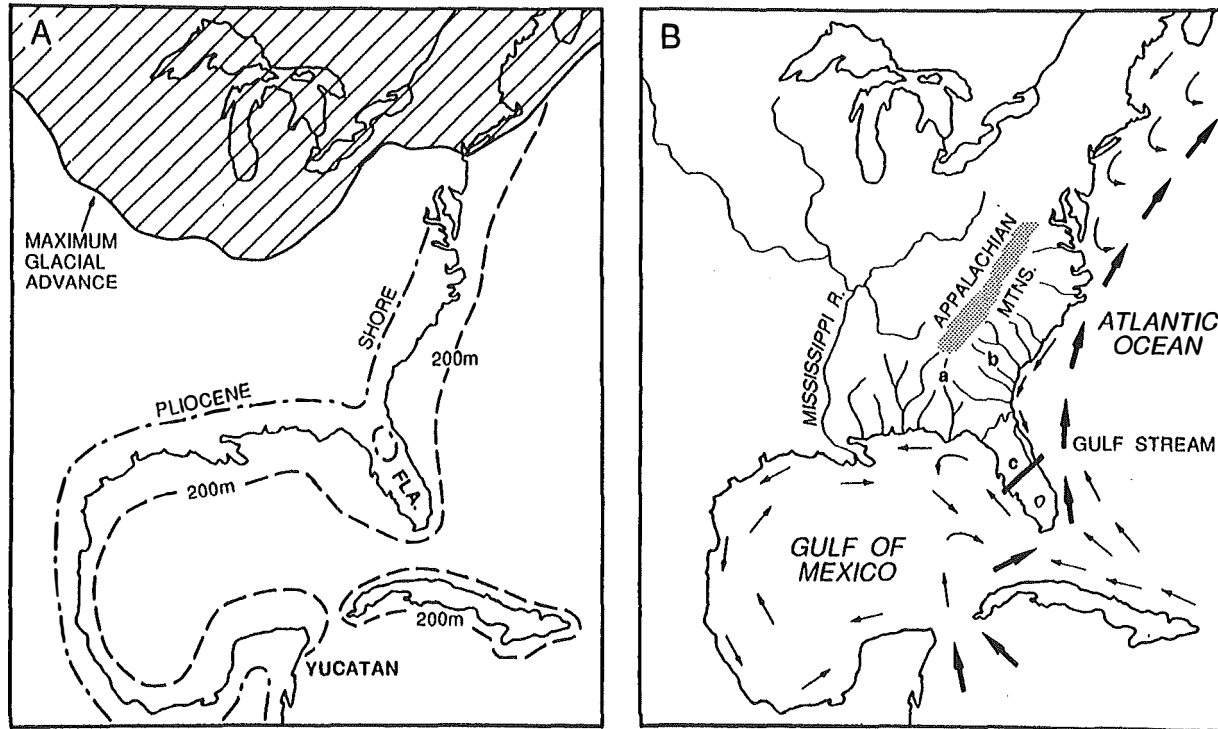
The southeastern U.S. is of special biogeographic interest because of both historical and contemporary climatic and geologic influences on its biota. Some of the major physiographic features of the region are summarized (Fig. 1). In the marine realm, the Florida peninsula currently protrudes southward into subtropical waters (25-28° north latitude), and separates temperate faunas into sometimes allopatric units on the Atlantic Coast and Gulf of Mexico. Thus the east and west coasts of Florida are well-recognized zones of transition between temperate and tropical adapted forms, with the southern ranges of many temperate species terminating in the approximate regions of Cape Canaveral and Naples, respectively (Briggs, 1974). Other "temperate" species are continuously-distributed around south Florida at the present time.

During the ten or more glacial advances and retreats of the Pleistocene, sea level fluctuations and climatic changes no doubt had great impact on the distributions of coastal and marine

species in the southeast. While the glaciers themselves never advanced beyond the north-central U.S., the associated climatic cooling pushed temperate populations southward, and may have increased the opportunity for contact of Atlantic and Gulf populations around south Florida. However, the glacial advances also caused drops in sea level (by as much as 150 metres) and exposed tremendous expanses of the Florida (and Yucatan) peninsulas (Fig. 1A). At such times, Florida was more arid than it is today, and presumably bordered by few of the intermediate-salinity estuaries and salt-marsh habitats favoured by many coastal species. Thus during glacial advances, an enlarged Floridian peninsula may actually have contributed to a separation of some Atlantic and Gulf coast populations through creation of a rather isolated pocket of estuarine habitat in the western Gulf of Mexico.

Opposing influences on species' distributions may also have been at work during interglacial periods (such as the present), when sea levels were higher and the Florida peninsula likely bordered by more extensive estuaries and salt marshes. At such times of climatic warming, some strictly temperate species may have been increasingly separated into disjunct Atlantic and Gulf populations by the tropical conditions of south Florida, while other species that were more eurythermal and estuarine-adapted may have expanded out of the putative Gulf refugium to gain increased contact with Atlantic forms around the southern tip of the peninsula. At present, swift water currents moving out of the Gulf of Mexico contribute to the "Gulf Stream", which hugs the coast of southeast Florida and may facilitate transport of Gulf-spawned pelagic larvae into the south Atlantic (Fig. 1B). The Gulf Stream tends to move offshore beyond the Cape Canaveral region of central Florida.

Cyclical changes in Pleistocene climates and landscapes must have influenced the distributions of freshwater biotas as well. At present, about a dozen major rivers and numerous smaller systems drain the southeastern coastal states from the Carolinas to Louisiana (Fig. 1B). Eastern drainages enter the south Atlantic, while western drainages enter the Gulf. During the high sea-stands of the Pliocene and the moderate sea-stands of the Pleistocene interglacial periods, smaller coastal streams were likely flooded, and freshwater faunas probably isolated in the upper reaches of the larger rivers (and perhaps in lakes or rivers of the central Florida peninsula itself). Any interdrainage transfers of fish must then have occurred via headwater or lateral stream capture, as for example between the Apalachicola and Savannah drainages which now come into close contact in the southern Appalachians. Conversely, during the glacial periods which



**FIG. 1. A)** Selected historical physiographic features of the eastern U.S. Shown are the current shoreline, the high sea stand of the Pliocene (somewhat higher than any Pleistocene interglacial shorelines), and the 200 meter depth contour (probably outside the exposed land areas of the Pleistocene glacial maxima). Also shown is the maximum of Pleistocene glacial advance.

**FIG. 1. B)** Selected contemporary physiographic features of the southeastern U.S. Shown are trends in marine currents [including the Gulf Stream (heavier arrows)], major river drainages entering the Gulf of Mexico and Atlantic [including the Apalachicola (a) and Savannah (b) drainages mentioned in the text], and the Appalachian mountains. Also indicated by the bar (c) is the approximate transitional zone between temperate and tropical faunal elements along the east and west coasts of peninsular Florida.

dominated much of the Pleistocene, the broader coastal plains associated with levels of low sea levels may have increased the opportunities for coalescence of adjacent rivers near their mouths, and hence for the lateral, interdrainage transfer of freshwater species, perhaps primarily on an Atlantic versus Gulf coast regional scale. The Apalachicola drainage (which forms the southern state line between Georgia and Alabama), now represents an important boundary region between freshwater zoogeographic provinces, as judged by the large number of species whose current ranges exhibit eastern or western termini in this general area (Swift *et al.*, 1985).

Most major rivers in the southeastern U.S. traverse several physiographic provinces. Originating as clear headwater streams in the southern Appalachian mountains, they enlarge and accumulate sediments as they traverse the hilly Piedmont region of red-clay soils, and finally emerge as warmer rivers crossing a broad and relatively flat coastal plain of sandy substrate. If present-day ecological selection pressures (rather than historical patterns of population connection and gene flow) were the primary moulders of mtDNA genotypic distributions, then clines in haplotype frequencies should probably parallel such strong ecological gradients within each river. However, as shown beyond, the intraspecific mtDNA phylogenies of assayed freshwater fishes generally orient in a regional pattern more consistent with a primary influence of historical biogeographic forces.

The discussion above provides a brief background to the major historical and contemporary features of the physical environment that may have influenced southeastern faunas. Expanded treatments can be found in Bermingham and Avise (1986), Bert (1986), Reeb and Avise (1990), and references therein. Nonetheless, it should be understood that very little is firmly known about the physiographic history of the area, and that most historical geologic scenarios are highly speculative. For example, possible opportunities for Atlantic-Gulf biotic separations have been discussed above primarily because the genetic data suggest they exist (see beyond), rather than because were necessarily predicted *a priori* on physiographic grounds alone. With the large volume of molecular data now available for this regional fauna, genetic information may actually inform some of the interpretations of historical geography, rather than the usual converse.

## THE CAST OF ASSAYED SPECIES

The species considered in this report exhibit a wide variety of population sizes, dispersal characteristics, and life history patterns (Table 1). For example, included among the marine and coastal species are the rare and anadromous Atlantic sturgeon, the common and catadromous American eel, mouthbrooding marine catfishes, demersal-spawning toadfishes, and abundant menhaden fish and oysters that produce pelagic larvae with high dispersal potential. Several of the assayed coastal species, including the menhaden, black sea bass, American oyster, horseshoe crab, seaside sparrow, and diamondback terrapin, prefer or require estuarine habitat for all or part of their life cycle. Among the assayed freshwater fishes are four common species of sunfish, live-bearing mosquitofish, and one living representative of an ancient Holostean order, the bowfin. If any shared patterns of mtDNA variability or phylogeography emerge among a significant fraction of such a heterogeneous group of species, the evolutionary forces responsible must have been of rather overriding influence.

## RESULTS AND DISCUSSION — THE EVOLUTIONARY PLAY

Most of the assayed species exhibit considerable mtDNA polymorphism, as exemplified by the values of genotypic and nucleotide diversity defined and summarized for the marine and coastal taxa in Table 2 [see Bermingham and Avise (1986) for descriptions of mtDNA variation in five of the freshwater fishes]. Genotypic diversity is an observed probability that pairs of assayed individuals from a reference population differ detectably in mtDNA haplotype (regardless of the magnitude of estimated sequence divergence); nucleotide diversity gives the mean sequence divergence between all assayed individuals in the population considered. For reasons that will become apparent, these statistics were calculated separately for the Gulf of Mexico versus Atlantic populations of nearly all species.

Figures 2-4 and 6 show phenograms relating the different mtDNA haplotypes observed within each of 12 assayed species, based on UPGMA cluster analyses (Sneath and Sokal, 1973) of genetic distances estimated from restriction fragment or site data. Phenograms are represented here not because they necessarily provide the best phylogenetic appraisals of mtDNA evolution, but rather because they facilitate simple visual comparisons of mtDNA haplotype distances and relationships across many species. The original references listed in Table 2 should be consulted for additional phylogenetic treatments applied to these data.

Table 1. Ranges and salient life history characteristics of the marine, coastal, and freshwater species in the southeastern U.S. considered in this report.

Species	Range in southeastern U.S.	Relevant life history and dispersal characteristics	Status <sup>a</sup>
<b>Marine and coastal species</b>			
Am. eel ( <i>Anquilla rostrata</i> )	continuous, Atl. and Gulf coasts	anadromous; mass spawning in tropical mid-Atl. Ocean; larvae return to coast, and juveniles reside in fresh and brackish waters	common
Hardhead catfish ( <i>Arius lates</i> )	continuous, Atl. and Gulf coasts	adults strong and active swimmers; mouth-brood eggs and fry	common
Oyster toadfish ( <i>Opsanus tau</i> )	Atl. coast to south Fla.	adults sluggish bottom dwellers; lay demersal, adhesive eggs	Intermediate-common
Gulf toadfish ( <i>Opsanus beta</i> )	Gulf coast to south Fla.	same	same
Atl. menhaden ( <i>Brevoortia tyrannus</i> )	Atl. coast to central Fla.	adults active pelagic feeders; spawn offshore; larvae move to estuarine feeding grounds	abundant
Gulf menhaden ( <i>Brevoortia patronus</i> )	Gulf coast to south Fla.	same	abundant
Atl. sturgeon ( <i>Acipenser oxyrinchus</i> )	Atl. and Gulf coasts, but perhaps absent from South Fla.	anadromous; spawn in freshwater streams; juveniles move to coastal waters after 1-8 years; adult movements poorly known	rare
Black sea bass ( <i>Centropristis striata</i> )	Atl. and Gulf coasts, but rare in extreme south Fla.	spawn near coast; larvae move to estuarine feeding grounds	common
Am. oyster ( <i>Crassostrea virginica</i> )	continuous, Atl. and Gulf coasts	adults sessile, in estuarine habitats; sequential hermaphrodites; pelagic gametes (few hrs.) and larvae (2-3 weeks)	abundant

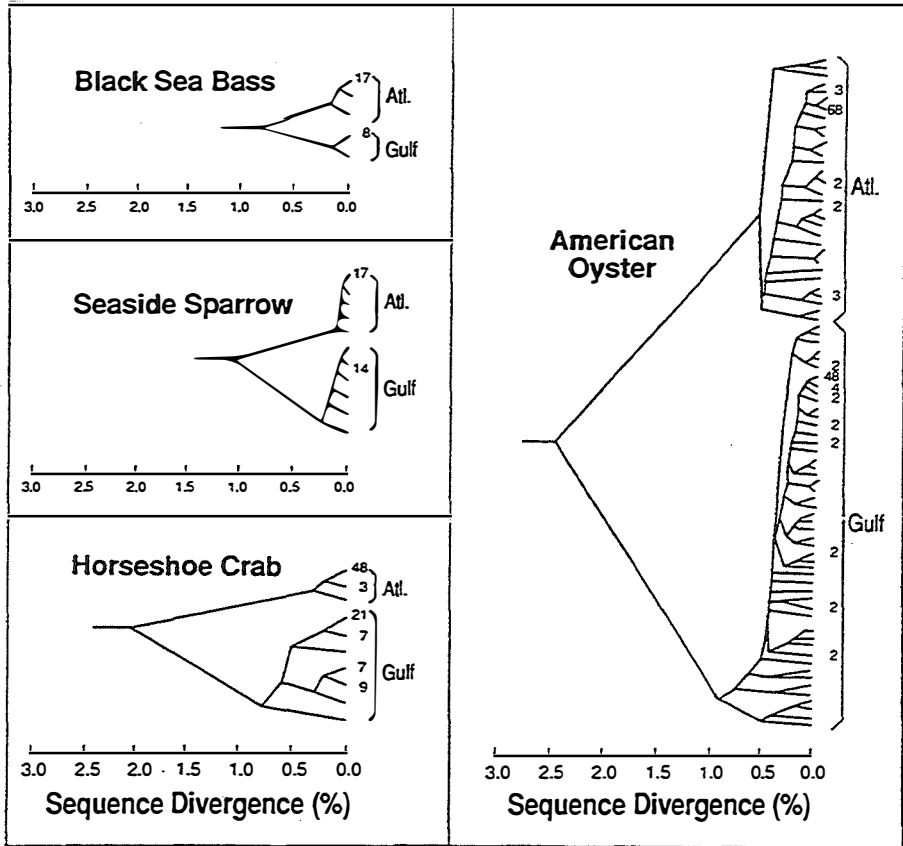
(continued)



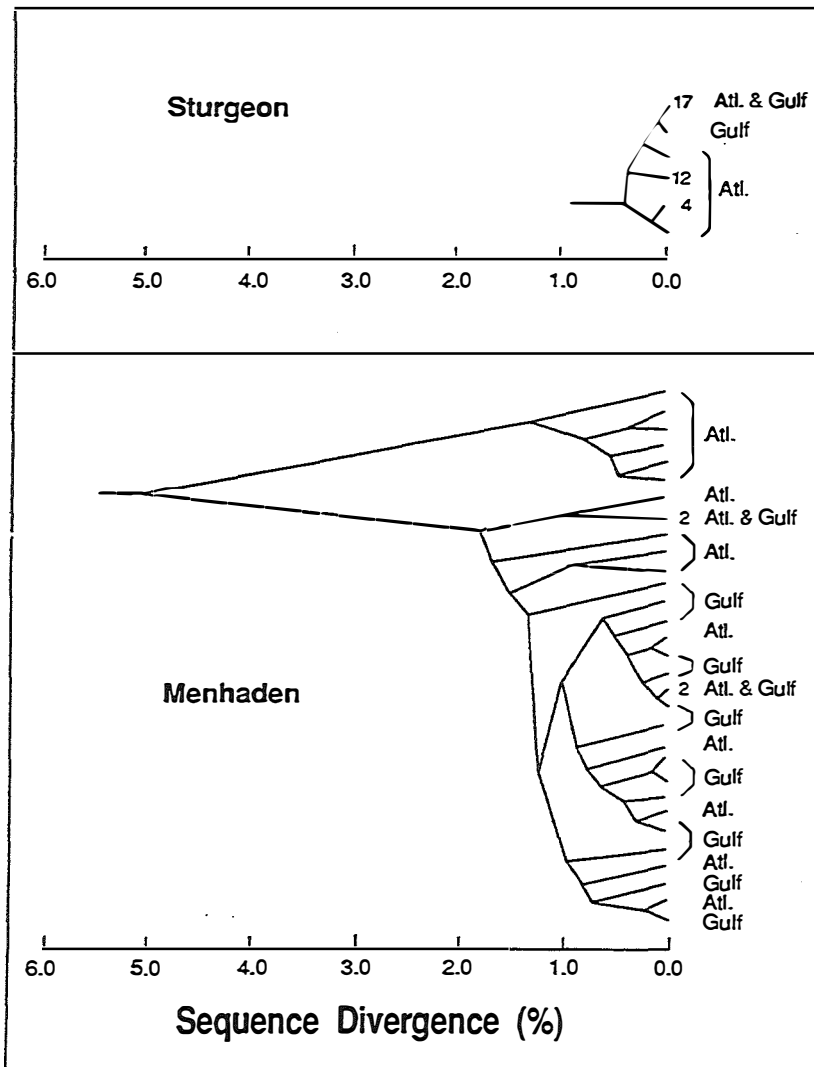
(Table 1 (continued))

Species	Range in southeastern U.S.	Relevant life history and dispersal characteristics	Status <sup>a</sup>
Horseshoe crab ( <u>Limulus polyphemus</u> )	continuous, Atl. and Gulf coasts	adults slow-moving, primarily in estuaries; eggs laid on sandy beaches; trilobite larvae stay in sand or water probably near shore	common
Seaside sparrow ( <u>Ammodramus maritimus</u> )	Atl. and Gulf coasts, but absent from southeast Fla.	confined to salt marshes; populations in the southeast U.S. non-migratory	Intermediate
Diamondback terrapin ( <u>Malaclemys terrapin</u> )	more or less continuous, Atl. and Gulf coasts	coastal marshes, estuaries, sheltered bodies of brackish or salt water	rare- Intermediate
Freshwater fishes			
Spotted sunfish ( <u>Lepomis punctatus</u> )	throughout southeastern U.S.	prefers ponds, rivers, streams with heavy vegetation;	Intermediate- common
Redear sunfish ( <u>Lepomis microlophus</u> )	throughout southeastern U.S.	similar	same
Warmouth sunfish ( <u>Lepomis gulosus</u> )	throughout southeastern U.S.	similar	same
Bluegill sunfish ( <u>Lepomis macrochirus</u> )	throughout southeastern U.S.	similar	abundant
Bowfin ( <u>Ambloplites caeruleus</u> )	throughout southeastern U.S.	similar; adults protect schooling young after hatching	Intermediate
Mosquitofish ( <u>Gambusia affinis holbrooki</u> )	<u>affinis</u> -- primarily Gulf drainages <u>holbrooki</u> -- primarily Atl. drainages	livebearer; shore-hugging in lentic waters	abundant

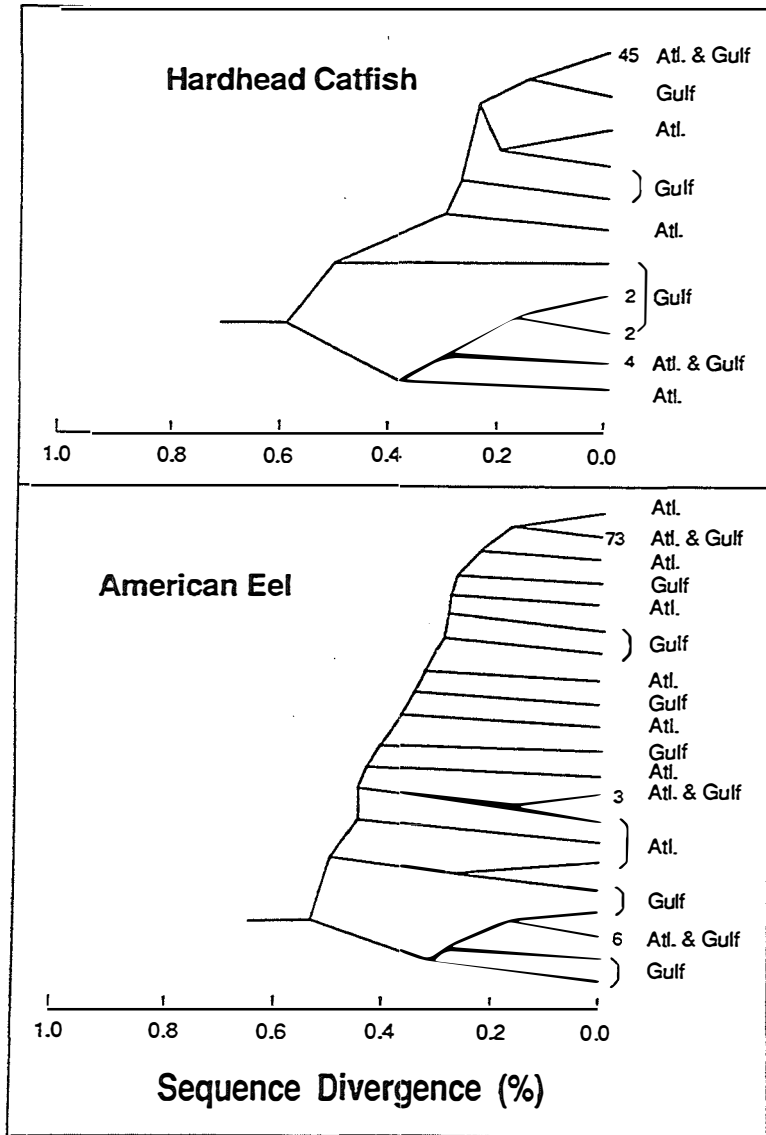
<sup>a</sup> Admittedly rough and impressionistic estimates of the sizes of adult populations at the present time: rare,  $N < 10^4$ ; Intermediate,  $N \approx 10^4$  -  $10^6$ ; common,  $N \approx 10^6$  -  $10^8$ ; and abundant,  $N > 10^8$ .



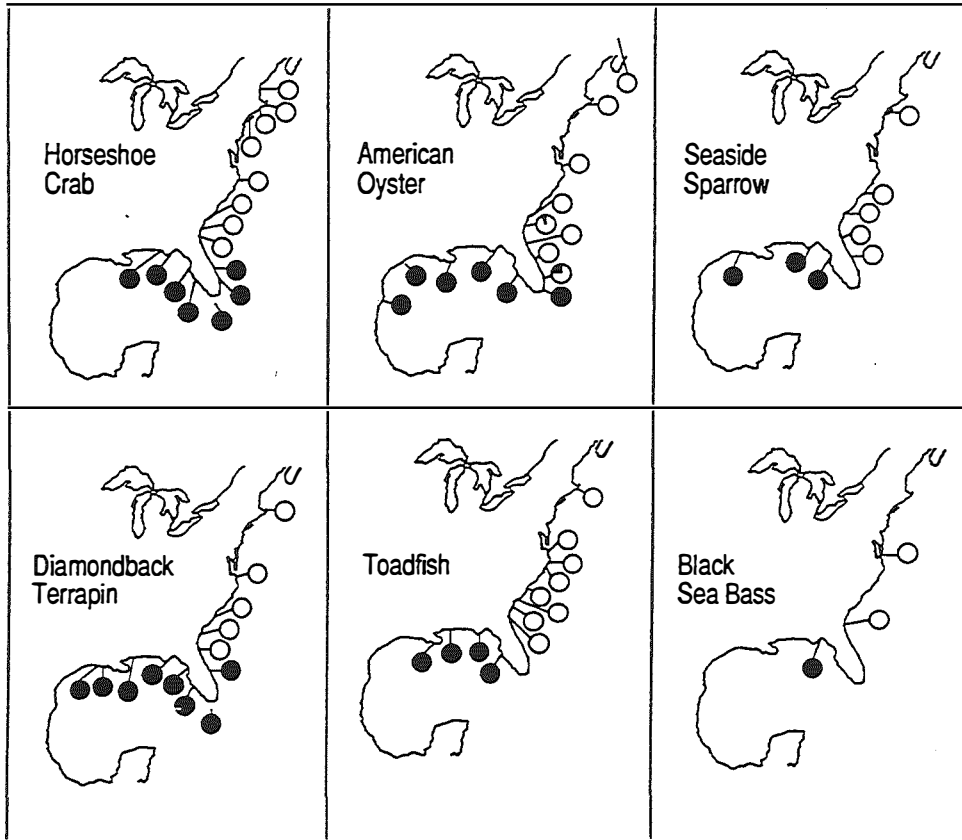
**FIG.2.** UPGMA phenograms summarizing relationships among mtDNA haplotypes observed in the black sea bass, seaside sparrow, horseshoe crab, and American oyster. Numbers of individuals of various mtDNA clones are indicated to the right; terminal branches without numbers were represented by single individuals. Note that all phenograms are plotted on the same scale of mtDNA sequence divergence.



**FIG. 3.** UPGMA phenograms summarizing relationships among mtDNA haplotypes observed in sturgeon and menhaden. Numbers of individuals of various mtDNA clones are indicated to the right; terminal branches without numbers were represented by single individuals. Note that both phenograms are plotted on the same scale of mtDNA sequence divergence.



**FIG. 4.** UPGMA phenograms summarizing relationships among mtDNA haplotypes observed in the hardhead catfish and American eel. Numbers of individuals of various mtDNA clones are indicated to the right; terminal branches without numbers were represented by single individuals. Note that both phenograms are plotted on the same scale of mtDNA sequence divergence.



**FIG. 5.** Pie diagrams summarizing geographic distributions of the two fundamental mtDNA clades in various coastal and marine taxa (see Fig. 2 and text).

The remainder of this report will address general conclusions which have emerged from these comparative estimates of mtDNA variability and differentiation.

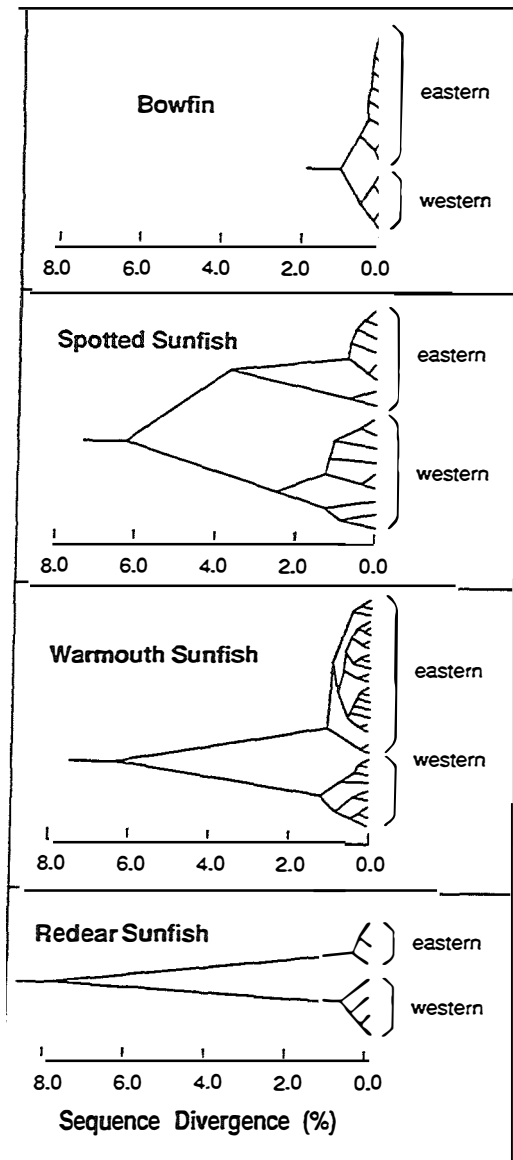
1) Major mtDNA phylogeographic patterns are shared across species.

The most important result of our comparative studies is the remarkable degree of concordance in the major mtDNA phylogeographic discontinuities across taxa. Within each of four coastal or marine species—the black sea bass, seaside sparrow, horseshoe crab, and American oyster—a fundamental split in the mtDNA gene trees clearly distinguishes individuals from the Gulf of Mexico versus those from most Atlantic coast locales (Fig. 2). In the horseshoe crab and American oyster, mtDNA genotypes normally characteristic of the Gulf clonal assemblage also

extend northward along the Atlantic coast as far as central Florida (Fig. 5). The major mtDNA subdivisions in these four species are similarly evident in other phylogenetic treatments, including Wagner parsimony analyses for which bootstrap resampling of restriction data indicates significant support for these putative clonal lineages. In the black sea bass, two taxonomic subspecies (*Centropistus striata striata* and *C.s. melana*, corresponding to Atlantic and Gulf coast locales, respectively) are conventionally recognized, but the other species had not previously been suspected of exhibiting Atlantic-Gulf distinctions.

The mtDNA data for at least two additional groups—diamondback terrapins and toadfish—can be interpreted as further provisional support for a fundamental phylogenetic distinction between Gulf and Atlantic coast populations of many species. In the terrapins, mtDNA genetic variation and divergence were atypically low, with the four variant haplotypes in the Gulf and Atlantic occurring in single individuals and usually differing from the common patterns in these respective regions by a single restriction site change. However, two mtDNA haplotypes (also differing by a single restriction site) were observed in multiple individuals such that geographic patterns could be assessed, and they exhibited an Atlantic-Gulf distribution nearly identical to those for the horseshoe crab and the American oyster (Fig. 5). Among the toadfish, two common species (*Opsanus tau* and *O. beta*) are currently recognized, and these are essentially confined to the Atlantic and Gulf coasts respectively. Since they differ dramatically in mtDNA sequence (see Avise *et al.*, 1987), an mtDNA phylogeny treating the complex as a whole reveals two major mtDNA groups, again corresponding to Atlantic versus Gulf locales (Fig. 5).

In the menhaden and surgeon, evidence supporting any putative Atlantic-Gulf subdivision is more ambiguous. Two menhaden sibling species are currently recognized, *Brevoortia tyrannus* in the Atlantic and *B. patronus* in the Gulf. However, two well defined haplotype clusters in the mtDNA phenogram (Fig. 3) do not conform exactly to these two taxon assignments or geographic regions (although their representatives do differ significantly in frequency between these areas). Thus while one cluster appears confined to the Atlantic, the other contained numerous individuals collected from Gulf and Atlantic locales. Perhaps there had indeed been an Atlantic-Gulf separation, but the two populations (or species) have not yet evolved to a status of reciprocal monophyly with respect to the mtDNA gene tree (see Avise, 1986; Neigel and Avise, 1986). However, since very closely related mtDNA haplotypes (identical or nearly identical at all assayed restriction sites) were observed in both Atlantic and Gulf locales (Fig. 8)



**FIG. 6.** UPGMA phenograms summarizing relationships among mtDNA haplotypes observed in the bowfin, spotted sunfish, warmouth sunfish, and redear sunfish. Note that all dendrograms are plotted on the same scale of mtDNA sequence divergence.

Table 2. Comparative estimates of mtDNA variability in coastal, marine, and anadromous species surveyed from locales along the Atlantic Coast and Gulf of Mexico in the southeastern U.S. Also presented are estimates of evolutionary effective population sizes for females, generated from the mtDNA data under the assumptions discussed in the text.

Species (and region)	Inds.	Number of		Genotypic diversity <sup>a</sup>	Nucleotide diversity <sup>b</sup>	Assumed generation length (yrs.)	Effective female pop. size $N_e(f)$	Reference for original data <sup>c</sup>
		different mtDNA genotypes	restriction sites or fragments per individual					
American eel ( <i>Anguilla rostrata</i> )	109	21	78	0.54	0.0011	10	5,500	1
Hardhead catfish ( <i>Arius felis</i> )	60	11	57	0.47	0.0018	2	45,000	2
Toadfish ( <i>Opsanus</i> )								
Atlantic ( <i>O. tau</i> )	43	5	50	0.58	0.0011	3	18,300	2
Gulf ( <i>O. beta</i> )	17	8	53	0.77	0.0033	3	55,000	2
Menhaden ( <i>Brevoortia</i> )								
Atlantic ( <i>B. tyrannus</i> )	17	17	55	1.00	0.0316	2	800,000	3
Gulf ( <i>B. patronus</i> )	16	16	55	1.00	0.0099	2	250,000	3
Sturgeon ( <i>Acipenser oxyrinchus</i> )								
Atlantic	21	5	68	0.64	0.0017	10	8,500	3
Gulf	15	2	68	0.13	0.0000	10	50	3
Black sea bass ( <i>Centropristis striata</i> )								
Atlantic	19	3	81	0.21	0.0003	3	5,000	3
Gulf	10	2	61	0.22	0.0003	3	5,000	3
American oyster ( <i>Crassostrea virginica</i> )								
Atlantic	104	31	65	0.57	0.0014	1	70,000	4
Gulf	108	51	65	0.80	0.0025	1	125,000	4

(continued)



Table 2 (continued)

Species (and region)	Number of			Genotypic diversity <sup>a</sup>	Nucleotide diversity <sup>b</sup>	Assumed generation length (yrs.)	Effective female pop. size $N_f(\hat{\sigma})$	Reference for original data <sup>c</sup>
	inds.	different mtDNA genotypes	restriction sites or fragments per individual					
<b>Horseshoe crab (<i>Limulus polyphemus</i>)</b>								
Atlantic	52	3	41	0.15	0.0003	3	5,000	5
Gulf	47	7	39	0.89	0.0030	3	50,000	5
<b>Seaside Sparrow (<i>Ammodramus maritimus</i>)</b>								
Atlantic	21	5	89	0.38	0.0003	3	5,000	6
Gulf	19	6	69	0.47	0.0004	3	8,700	6
<b>Diamondback terrapin (<i>Malaclemys terrapin</i>)</b>								
Atlantic	25	2	74	0.08	0.0001	5	1,000	7
Gulf	28	4	73	0.20	0.0003	5	3,000	7

<sup>a</sup>  $(n / (n - 1)) (1 - \sum f_j^2)$ , where  $f_j$  is the frequency of the  $j$ th mtDNA haplotype.

<sup>b</sup> mean  $p = (n / (n - 1)) (\sum f_j f_j p_{jj})$ , where  $f_j$  and  $f_j$  are the frequencies of the  $j$ th and  $j$ th haplotypes in a sample of size  $n$ , and  $p_{jj}$  is the estimated sequence divergence between the  $j$ th and  $j$ th sequences (Nei, 1987; p. 258).

<sup>c</sup> References: (1) Avise *et al.*, 1988; (2) Avise *et al.*, 1987b; (3) Bowen and Avise, 1990; (4) Reeb and Avise, 1990; (5) Saunders *et al.*, 1988; (6) Avise and Nelson, 1989; (7) Lamb and Avise, in prep.

3), recent gene flow between the two regions is also strongly implicated (Bowen and Avise, 1990). In the sturgeon, recognized Atlantic and Gulf coast subspecies (*Acipenser oxyrinchus oxyrinchus* and *A. o. desotoi*, respectively) again differed significantly in mtDNA haplotype frequencies. The most common genotype was observed along both coasts (Fig. 3), suggesting a recent historical connection between these populations.

Two species - the American eel and hardhead catfish - showed no evidence for an mtDNA phylogenetic subdivision between Atlantic and Gulf (Fig.4). Results might reasonably be attributable to high effective gene flow between these areas. American eels presumably spawn in the western tropical mid-Atlantic ocean (the Sargasso Sea), and larvae disperse largely by passive transport to coastal areas where maturation occurs. Conventional wisdom has been that this life history pattern should result in a nearly random distribution of genotypes along the coast (see Williams and Koehn, 1984), a suggestion with which our mtDNA data are consistent (Avise *et al.*, 1986). In the case of the hardhead catfish, adults are strong and active swimmers, and movement between the Atlantic and Gulf around south Florida may be considerable.

Overall, among the 10 coastal species or species-complexes surveyed (Table 2), at least five and as many as eight evidence a fundamental mtDNA subdivision involving Atlantic versus Gulf coast populations. Geographic distributions of the two major mtDNA lineages within each of six such taxa are summarized in Fig. 5.

A comparable degree of mtDNA phylogeographic concordance is exhibited among the surveyed freshwater fishes in the southeastern U.S. Genetic relationships among mtDNA haplotypes of bowfin and each of three sunfish species are summarized in Fig. 6. Each phenogram exhibits a fundamental split (also supported at levels greater than 95% by bootstrapping in Wagner parsimony networks - Bermingham and Avise, 1986) distinguishing mtDNA lineages from eastern versus western portions of the respective species' ranges in the area. A similar pattern of mtDNA phylogeographic divergence also occurs in the bluegill sunfish, *Lepomis macrochirus* (although a full dendrogram could not be constructed because most individuals were surveyed only at selected "marker" restriction sites - see Avise *et al.*, 1984), and in the mosquitofish (*Gambusia affinis-G. holbrooki* sibling species complex—Scribner and Avise, in prep.).

Table 3. Estimates of mtDNA sequence divergence and provisional times since separation (in millions of years) between the eastern versus western assemblages of freshwater fishes, and between the Atlantic versus Gulf assemblages of those marine and coastal-restricted species whose populations clearly differ between regions.

Species	mtDNA sequence divergence		times <sup>c</sup> of pop. separation (mva)	
	uncorr. <sup>a</sup>	corr. <sup>b</sup>	uncorr. <sup>a</sup>	corr. <sup>b</sup>
<b>Marine and coastal species (Atlantic vs. Gulf)</b>				
Horseshoe crab ( <u>Limulus polyphemus</u> )	0.020	0.016	1.00	0.80
American oyster ( <u>Crassostrea virginica</u> )	0.026	0.022	1.30	1.10
Black sea bass ( <u>Centropristis striata</u> )	0.009	0.007	0.45	0.35
Toadfish ( <u>Opsanus beta</u> and <u>O. tau</u> )	0.101	0.096	5.05	4.80
Diamondback terrapin ( <u>Malaclemys terrapin</u> )	0.001	ca. 0.00	0.05	ca. 0.00
Seaside sparrow ( <u>Ammodramus maritimus</u> )	0.011	0.010	0.55	0.50
<b>Freshwater fishes (eastern vs. western)</b>				
Spotted sunfish ( <u>Lepomis punctatus</u> )	0.062	0.044	3.10	2.20
Redear sunfish ( <u>Lepomis microlophus</u> )	0.087	0.082	4.35	4.10
Warmouth sunfish ( <u>Lepomis gulosus</u> )	0.063	0.056	3.15	2.80
Bluegill sunfish ( <u>Lepomis macrochirus</u> )	0.085	d	4.25	d
Bowfin ( <u>Amia calva</u> )	0.009	0.006	0.45	0.30

<sup>a</sup> Based on the mean genetic distance ( $D_{xy}$ ) between mtDNA haplotypes of the two regions,  $x$  and  $y$ .

<sup>b</sup> Based on a correction for within-region polymorphism:  $D_{corr.} = D_{xy} - 0.5 (D_x + D_y)$ , where  $D_x$  and  $D_y$  are the mean pairwise distances of mtDNA haplotypes within regions  $x$  and  $y$ , respectively.

<sup>c</sup> Based on a "conventional" mtDNA clock calibration reported in several vertebrate groups (1 percent sequence divergence per lineage per million years-- Brown et al., 1979; Shields and Wilson, 1987). However, caution is indicated because rates of mtDNA evolution have also been reported to differ in either direction from this estimate by several-fold (see Moritz et al., 1987).

<sup>d</sup> Corrections could not be applied because only selected restriction sites were used in the geographic survey.

Geographic distributions of the two major mtDNA clonal lineages within these species or species complexes are summarized in Fig. 7. Exact positions of the geographic boundaries between clades appear to differ somewhat among species (appearing most aberrant in *Lepomis gulosus*), but typically occur near the Florida panhandle. Consistently, a stronghold of the “eastern” forms is peninsular Florida.

Other evidence further supports a fundamental phylogenetic distinction between eastern and western populations of many freshwater taxa in the region. In the largemouth bass, *Micropterus salmoides* (Phillip *et al.*, 1981), bluegill sunfish (Awise and Smith, 1974), and mosquitofish complex (Wooten *et al.*, 1988), strong shifts in allozyme frequency at several loci reveal geographic patterns remarkably similar to those presented in Fig.7. And as already mentioned, the Apalachicola region has been recognized as an important boundary between freshwater zoogeographic provinces as evidenced by clusters of species’ distributional limits (Swift *et al.*, 1985).

Overall, among the six species or species-complexes of southeastern freshwater fishes genetically assayed to date, all exhibit a fundamental phylogeographic distinction of populations in eastern and Floridian drainages from those to the west. Such concordance in phylogeographic profiles across independent species strongly suggests a significant influence of historical biogeographic factors.

## 2) A uniform mtDNA “clock” may not exist for all species

Brown *et al.* (1979) first reported a high rate of mtDNA evolution - roughly 0.5-1.0% change in nucleotide sequence per lineage per million years - in primates. Similar rates were subsequently described for several other mammals, birds and invertebrates (Moritz *et al.*, 1987; Shields and Wilson, 1987; Vawter and Brown, 1986; Wilson *et al.*, 1985). Nonetheless, the reliability and precision of molecular dating in evolution (through use of either mtDNA or nuclear DNA “clocks”) remains highly controversial (e.g., Britten, 1986; Powell *et al.*, 1986; Sibley and Ahlquist, 1984).

Application of a “conventional” mtDNA clock calibration (2% sequence divergence per million

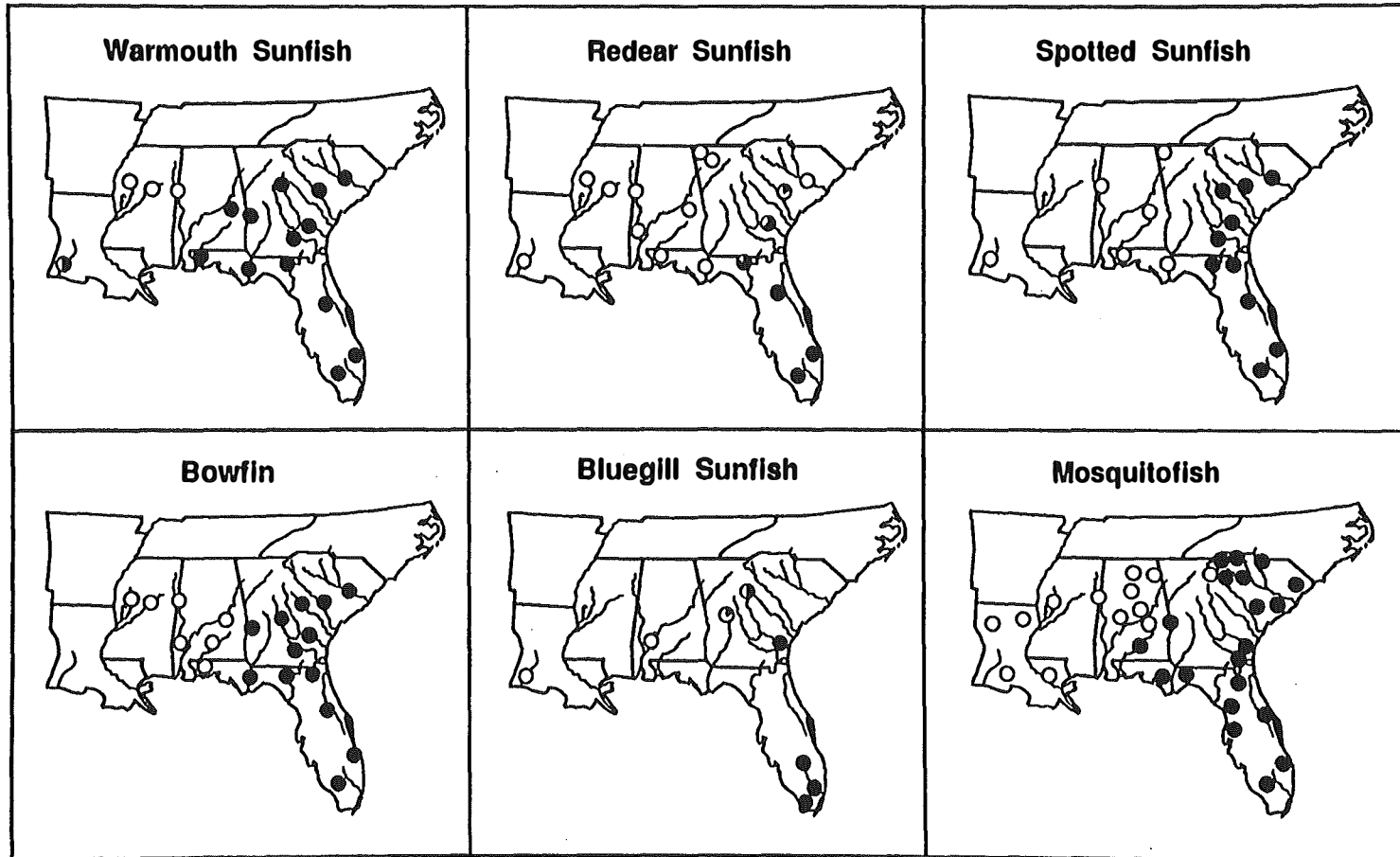


FIG.7. Pie diagrams summarizing geographic distributions of the two fundamental mtDNA clades of various freshwater fishes (see Fig. 6 and text).

years between pairs of lineages) to the species considered in this report yields the estimated divergence times presented in Table 3. Among the marine and coastal species, the Gulf-Atlantic mtDNA lineage separations range from about 50,000 to 5,000,000 years before present (bp); among the freshwater fishes, the east-west mtDNA lineage separations range from 450,000 to 4,350,000 bp. Since mtDNA lineage separations can vastly predate population separations, particularly when effective population separations are large, corrections for within-region polymorphism were also applied (Table 3, right most column). After such corrections, large differences remain in the estimated times of population divergence (100-fold and 10-fold among the various marine and freshwater species, respectively).

Thus while all of the phylogenetic separations appear to date to the late Pliocene or Pleistocene, particular times for the various species span a large range. Assuming that the sequence divergence estimates are reasonably accurate, two explanations for these discrepancies appear most likely: either (1) mtDNA evolution exhibits considerable rate heterogeneity among taxa; and/or (2) the dates of particular vicariant or dispersal events influencing population separations differed among the various species. We have some independent evidence (presented elsewhere—Avice *et al.*, in prep.) for the first possibility - that of mtDNA rate differences among certain major taxa. However, due to the cyclical nature of the climatic and geologic changes postulated to have influenced faunas in the southeastern U.S. (see earlier discussion), the latter possibility may also be a major contributor to the large range of inferred separation times. Multiple episodes of glacial advance and retreat likely provided repeated opportunities for population isolation (and perhaps later coalescence), such that the times of the regional population disjunctions may truly differ among the assayed species.

3) Shallow phylogenetic structures are also evident within regions, and may be related to species-specific gene flow regimes.

In addition to the deep mtDNA phylogenetic separations between regional populations, which are probably tied to vicariant Pleistocene events, many assayed taxa in the southeastern U.S. also show evidence of “shallower” within-region structure. These are evidenced most clearly by apparent geographic localizations of less common (and often presumably derived) mtDNA genotypes. For example, among 59 bowfin collected from 10 drainages in the eastern portion of the species’ range (Fig. 7), mtDNA haplotype “2” was observed only in 4 specimens from the Cooper River, haplotype “3” only in 3 individuals from the adjacent Savannah and

Table 4. Various comparative estimates of the geographic structure of mtDNA haplotypes within regional populations<sup>1</sup>. Freshwater and marine/coastal species are listed in rank order with respect to degree of local population structure as reflected in a "localization Index" (the proportion of mtDNA haplotypes observed in multiple individuals, yet confined to either a single collection locale or to two adjacent locales). Also presented are relative estimates of  $N_m$ , calculated (with corrections for sample size) from estimated frequencies of: (a) all "private" haplotypes; and (b) only those "private" haplotypes present in more than one specimen. See text for further explanation and qualifications.

Species	Total no. haplotypes	"Private" haplotypes, confined to...			Localization Index		$N_m$	
		1 ind.	>1 ind., 1 locale	>1 ind., 2 adj. locales	fraction	%	(a)	(b)
Warmouth Sunfish	32	17	10	2	12/15	80	0.48	0.18
Bowfin	13	3	5	3	8/10	80	0.13	0.13
Spotted Sunfish	17	9	5	1	6/8	75	0.33	0.08
Redear Sunfish	7	2	1	0	1/5	20	0.18	0.04
Total (freshwater fish):					27/38	71	0.28	0.11
Toadfish (bata and lau)	13	8	1	2	3/5	60	0.30	0.10
American oyster	82	68	4	4	8/14	57	1.88	2.14

(continued)

Table 4 (continued)

Species	Total no. haplotypes	"Private" haplotypes, confined to...			Localization Index		Nm	
		1 Ind.	>1 Ind., 1 locale	>1 Ind., 2 adj. locales	fraction	%	(a)	(b)
		Hardhead Catfish	11	7	1	1	2/4	50
Horseshoe Crab	10	4	0	2	2/6	33	0.49	0.15
Atlantic Sturgeon	6	3	1	0	1/3	33	0.42	0.17
Seaside Sparrow	11	9	0	0	0/2	00	1.15	>> 1
Menhaden ( <i>tyrannus</i> and <i>patronus</i> )	33	31	0	0	0/2	00	2.14	>> 1
American eel	21	18	0	0	0/3	00	3.04	>> 1
Total (marine/coastal taxa):					16/39	41	1.22	---

<sup>1</sup> Black sea bass are not included in this table because too few locales were sampled; diamondback terrapins because no variant haplotypes within the Atlantic or Gulf were observed in multiple individuals; and mosquitofish because the sampling is still in progress.



Ogeechee Rivers, and so on for each of 3 other mtDNA haplotypes (Bermingham and Avise, 1986). Only haplotype “1” occurred throughout nearly all eastern drainages, and by several lines of evidence it represents the ancestral condition from which the localized genotypes appear to have arisen independently by one or two assayed restriction site changes (Bermingham and Avise, 1986; Avise *et al.*, 1987).

“Private” mtDNA haplotypes, defined here as genotypes observed in only one (or two immediately adjacent) locales, were observed in most of the assayed species (Table 4). Because of the high genetic diversities within some species, these private haplotypes were further categorized as those present in a single specimen, versus those shared by two or more individuals. “Localization indices” (the observed numbers of the latter class of private haplotypes, expressed as fractions of the total numbers of haplotypes distinguished within species) ranged from 0% in American eel, Menhaden, and Seaside Sparrow, to 80% in Warmouth sunfish and Bowfin (Table 4).

Slatkin (1985) proposed a method for estimating average levels of gene flow in a subdivided population using observed mean frequencies of private alleles within samples. The approach is based on comparisons of data with results of computer simulations, and yields estimates of  $Nm$ , where  $N$  is the local population size and  $m$  is the migration rate between populations. In theory, values of  $Nm \gg 1$  indicate high gene flow between sub-populations (such that only limited genetic divergence is expected, and frequencies of private alleles are low), while values of  $Nm \ll 1$  indicate low gene flow (such that population structure is strong, and private alleles can sometimes reach appreciable frequency). Table 4 presents values of  $Nm$  [calculated by equation (3) in Slatkin, 1985, and corrected for sample size] estimated from frequencies of: (a) all private alleles; and (b) that subset of private alleles possessed by more than one individual. Values range from 0.13 in Bowfin to 3.04 in American eels (Table 4). The latter value is probably a severe underestimate, since all of the genotypes present in more than one eel were geographically widespread, and the  $Nm$  calculation was thus based solely on private alleles confined to single specimens.

Two general conclusions emerge from comparisons of these localization indices and  $Nm$  values. First, many species show evidence of local substructuring within the major phylogeographic regions identified previously. Second, the relative magnitudes of local population structure

appear plausibly related to probable gene flow regimes of several of the species. For example, in comparison to the assayed marine/coastal taxa, the freshwater fishes tended to exhibit higher localization indices (71% versus 41%;  $\chi^2 = 7.0$ ,  $P < 0.01$ ) and lower mean values of  $Nm$  (e.g., 0.28 versus 1.22 - Table 4) as might be predicted from the more isolated nature of disjunct freshwater habitats. Furthermore, among the marine/coastal species, the highest estimated  $Nm$  (and a zero localization index) occurred in the American eel, a species whose catadromous life history pattern probably involves very high effective dispersal throughout the North American coastline (Avice *et al.*, 1986). Conversely, the highest localization index occurred in the toadfish, which lay demersal eggs and are sluggish bottom dwellers.

Nonetheless, these current estimates of the magnitude of local population structure should be interpreted with considerable caution. First, the estimates apply only to female lineages, and in some species such as the American oyster, much less structure is apparent in nuclear-encoded allozymes (see Reeb and Avice, 1990). The reasons for such differences are unknown. Second, since our primary goal in most surveys was to assess phylogeographic structure on a broad scale, mtDNA sample sizes at particular sites were typically small, and actual frequencies of local haplotypes poorly assessed. Third, mtDNA genotypic diversities were so high in some species (Table 2) that few or no haplotypes were common. These factors greatly limit applicability of other conventional approaches to estimate gene flow (such as  $F_{ST}$  - Slatkin, 1987), unless the mtDNA haplotypes could be grouped reliably into more inclusive classes by the criterion of evolutionary relatedness. However, apart from the major phylogeographic disjunctions between regions, which received strong statistical support by bootstrapping in parsimony networks, most of the putative mtDNA clades within regions (Figs. 2-4, 6) remain poorly defined.

Thus overall, while the available mtDNA data suggest significant phylogeographic population substructure within regions for most species, larger sample sizes and stronger characterizations of putative mtDNA micro-clades will be required for definitive conclusions about female-mediated levels of historical and contemporary gene flow on local scales.

4) Evolutionary effective population sizes estimated from mtDNA diversities are correlated with, but much lower than, current-day census sizes. Using inbreeding theory as applied to neutral alleles inherited maternally, theoretical probability distributions of times to shared

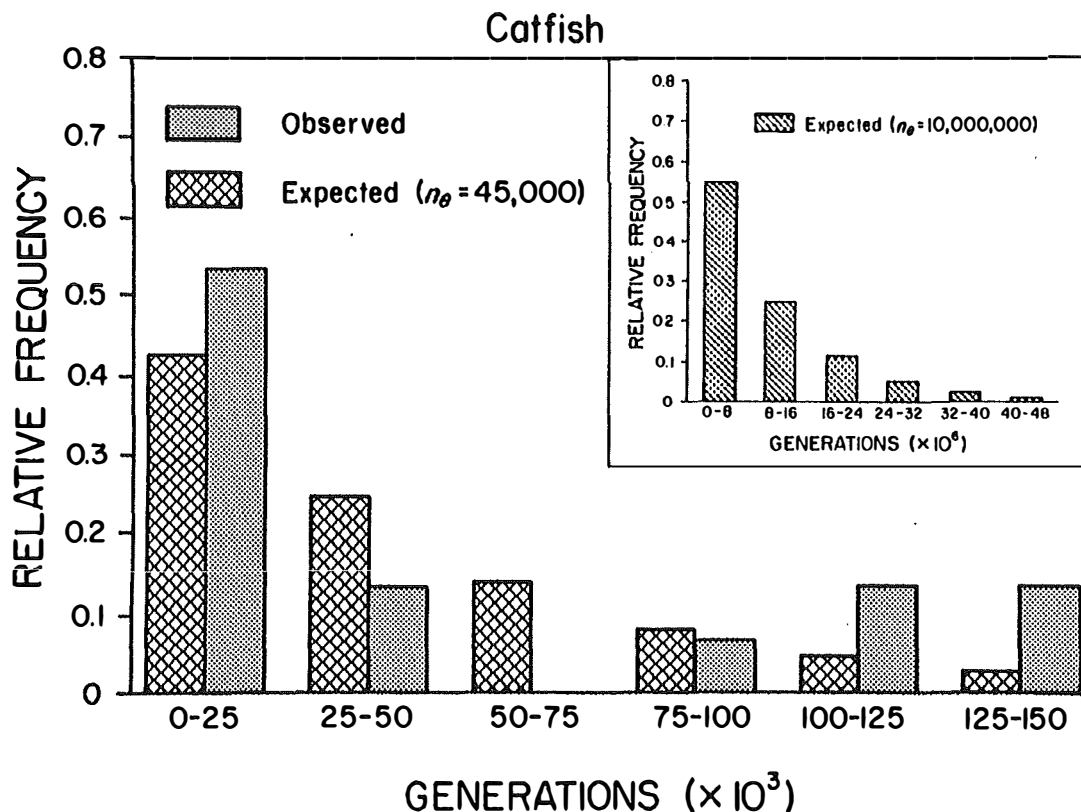
mtDNA haplotype ancestry can be generated as a function of the evolutionary effective size of a female population ( $N_{f(e)}$ ). In particular, the probability that random pairs of extant mtDNA haplotypes derive from a common ancestor that existed  $G$  generations prior is given approximately by

$$f(G) \equiv (1/N_{f(e)}) e^{-(G-1)/N_{f(e)}} \quad (1)$$

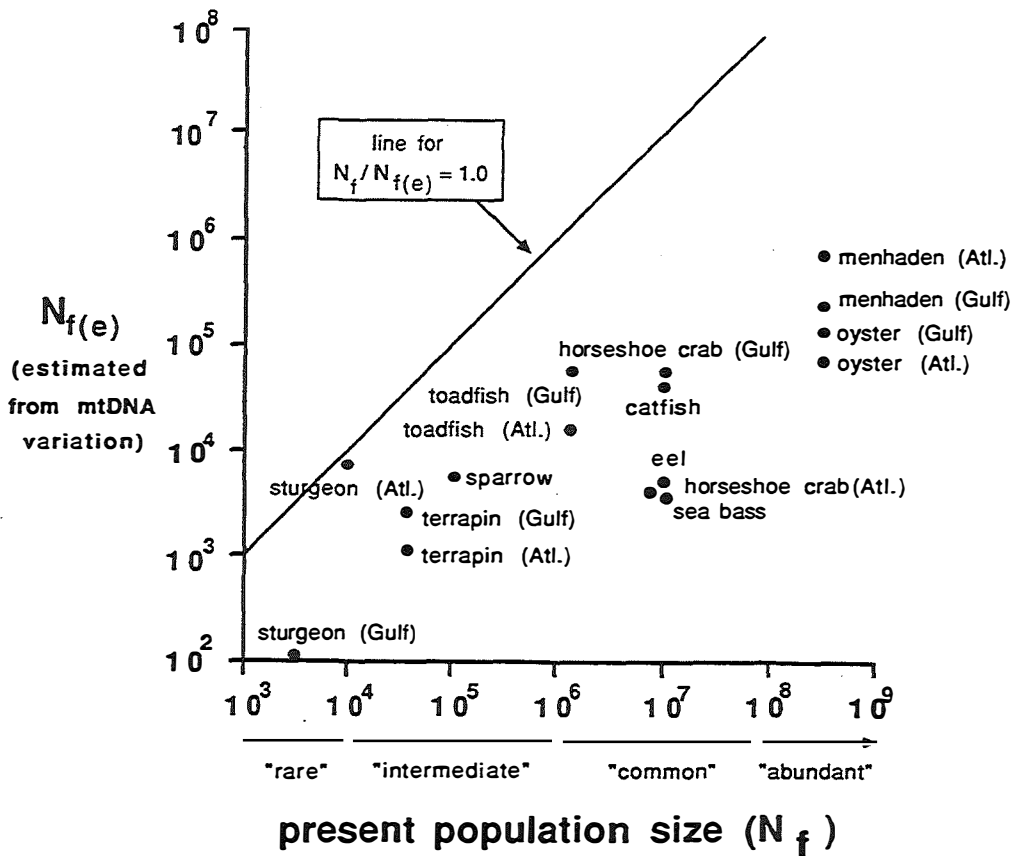
(Avice *et al.*, 1988). The geometric distributions described by equation (1) have mean  $N_{f(e)}$  [and variance  $N_{f(e)}(N_{f(e)}-1)$ ]. For any population, or entire species with high gene flow, empirical distributions of genetic distance among mtDNA haplotype pairs (converted to sidereal time using a presumed rate of mtDNA evolution and the suspected generation length of a species) may be generated and compared to such theoretical expectations (but see Ball *et al.*, 1990 for qualifications).

One example of such comparisons is presented in Fig. 8. In the hardhead catfish, mean mtDNA sequence divergence between assayed individuals was  $p = 0.0018$  (Table 2). Under a conventional mtDNA evolutionary rate of 2% sequence divergence per million years, this value translates into an estimated mean time of mtDNA haplotype separation of about 90,000 years, or 45,000 catfish generations (assuming a generation length of 2 years). Fig. 8 shows that the agreement between the observed and expected frequency distributions of times to shared haplotype ancestry is reasonably good for  $N_{f(e)} = 45,000$ . For the hardhead catfish, this estimate of evolutionary effective population size is vastly lower (by more than 200-fold— see inset to Fig. 9) than the present day size of the female population ( $N_f$ ), which conservatively might include 10,000,000 individuals (Avice *et al.*, 1988). To the extent that any local population structure exists within the species (Table 4), and may have buffered some mtDNA lineages against extinction (Avice *et al.*, 1984), the disparity between  $N_f$  and  $N_{f(e)}$  becomes even more dramatic.

In general, evolutionary effective population sizes estimated from mtDNA genetic distances are usually vastly smaller than present-day census sizes. Using the data presented in Tables 1 and 2, Fig. 9 summarizes the relationship between  $N_f$  and  $N_{f(e)}$  for the 10 marine and coastal species considered in this report. Since Gulf versus Atlantic populations of many of the assayed species exhibit sharp mtDNA phylogenetic distinctions, calculations were conducted separately for



**FIG. 8.** Frequency distribution of times to shared ancestry of mtDNA haplotypes in the hardhead catfish. Expected distributions generated from inbreeding theory (see text) are shown for each of two values of  $n_e$ : 10,000,000 (a conservative guess for the current breeding population of female catfish - inset), and 45,000 (a value which yields a mean expected divergence time equal to that inferred from the mtDNA data). The observed times were derived from the data of Avise *et al.* (1987b), using a conventional "clock" calibration (2% sequence divergence per million years) and a generation length of 2 years. Note the difference of scale along the abscissas of the inset and main graph.



**FIG. 9.** Relationship between current day census sizes (see Table 1) of assayed marine species in the southeastern U.S., and evolutionary effective sizes estimated from mtDNA haplotype diversities (Table 2). Note that both axes are in log scale. The correlation coefficient calculated between  $N_{f(e)}$  and  $N$  is  $r = 0.66$ . The methods and caveats for estimating  $N_{f(e)}$  and  $N$  are discussed in the text and Table 1.

these regional populations. Overall, among the 18 comparisons attempted,  $N_{f(e)}$  was consistently lower than  $N_f$  usually by 1 - 3 orders of magnitude. Nonetheless, a reasonably strong correlation ( $r = 0.66$ ) between  $N_{f(e)}$  and  $N_f$  across species was also evident (Fig. 9).

Possible explanations for the discrepancy in magnitude between  $N_{f(e)}$  and  $N_f$  include: (1) a much slower pace of mtDNA evolution than is generally accepted; or (2) periodic decreases in the

numbers of females through which surviving mtDNA lineages have been transmitted. The latter could be due either to: (2a) general historical demographic considerations, such as large variances in progeny survival among females, fluctuations and relative bottlenecks in female population size, and periodic extinctions (and subsequent recolonizations) of local demes; or, (2b) the occasional appearance of selectively advantageous mtDNA variants which might sweep through populations and “cleanse” the non-recombining mtDNA genome of much pre-existing genetic variability via hitchhiking of neutral markers to the selected mutations.

By hard criteria, we cannot decide between these competing possibilities on the basis of existing mtDNA variability alone. However, similar reductions of  $N_{f(e)}$  relative to  $N$  have been reported previously (in many other species) on the basis of allozyme variation (Nei and Graur, 1984). Thus whatever processes are involved (rate decelerations, positive directional selection at specific loci, and/or general demographic factors) probably relate to nuclear genes as well. The observed correlation between  $N_{f(e)}$  and  $N_f$  (Fig. 9) is not predicted under the hypothesis of a deceleration in mtDNA evolutionary rate in particular species, nor by occasional positive selection, but it is generally consistent with historical demographic influences such as proportional fluctuations in population size.

If mtDNA variabilities do indeed reflect historical demographic conditions primarily, rather than idiosyncratic episodes of selection directed at mtDNA *per se*, they may provide useful evidence on comparative levels of overall genetic variation among species, or among stocks under management. For example, among the 18 estimates summarized in Table 2 and Fig. 9, the surgeon population in the Gulf (which is currently very small) had the lowest estimated mtDNA variability and  $N_{f(e)}$ , while the abundant menhaden and oysters had the highest such values. Several authors have suggested that magnitude of genetic variability may significantly influence the probability of a population’s survival over ecological or evolutionary time, and that reliable assessments of genome-wide variation can be based on samples of various molecular genetic characters (e.g., Quattro and Vrijenhoek, 1989; Vrijenhoek *et al.*, 1985; Wildt *et al.*, 1987; but see also Lande, 1988). The correlation of mtDNA diversity with population size (Fig. 9) suggests that mtDNA variation may reflect demographic conditions whose influence should extend to nuclear loci as well.

## CONCLUSIONS

Non-recombining mtDNA polymorphisms lend themselves well to the estimation of evolutionary differentiation and historical demography within and among populations of a species. The remarkable phylogeographic concordance among numerous freshwater fishes and marine/coastal taxa in the southeastern U.S. exemplifies the utility of the mtDNA approach in identifying major evolutionary genetic stocks among the conspecific populations of a regional fauna. Additional phylogeographic substructures within regional populations indicate more recent restrictions to gene flow, and these can be related tentatively to species-specific life history patterns and dispersal characteristics. By adding a phylogenetic perspective to species demographics and zoogeographic histories, data from mtDNA can contribute to a deeper appreciation of the meaning and significance of contemporary population structure.

## REFERENCES

- Asmussen, M.A., J. Arnold, and J.C. Avise. 1987. Definition and properties of disequilibrium statistics for associations between nuclear and cytoplasmic genotypes. *Genetics* 115:755-768.
- Avise, J.C. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals. *Phil. Trans. Roy. Soc. Lond.* B312:325-342.
- Avise, J.C. 1987. Identification and interpretation of mitochondrial DNA stocks in marine species. Pp.105-136 *in*: H. Kumpf and E.L. Nakamura (eds.), *Proc. Stock Identification Workshop*, Publ. Natl. Oceanogr. Atmos. Admin., Panama City, FL.
- Avise, J.C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution* 43:1192-1208.
- Avise, J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb, J.E. Neigel, C.A. Reeb, and N.C. Saunders. 1987a. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* 18:489-522.
- Avise, J.C., R.M. Ball, Jr., and J. Arnold. 1988. Current versus historical population sizes in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and inbreeding theory for neutral mutations. *Mol. Biol. Evol.* 5: 331-344.
- Avise, J.C., E. Bermingham, L.G. Kessler, and N.C. Saunders. 1984. Characterization of mitochondrial DNA variability in a hybrid swarm between subspecies of bluegill sunfish (*Lepomis macrochirus*). *Evolution* 38:931-941.

- Avise, J.C., B.W. Bowen, and T. Lamb. 1989. DNA fingerprints from hypervariable mitochondrial genotypes. *Mol. Biol. Evol.* 6:258-269.
- Avise, J.C., G.S. Helfman, N.C. Saunders, and L.S. Hales. 1986. Mitochondrial DNA differentiation in North Atlantic eels: population genetic consequences of an unusual life history pattern. *Proc. Natl. Acad. Sci. USA* 83:4350-4354.
- Avise, J.C., J.E. Neigel, and J. Arnold. 1984. Demographic influences on mitochondrial DNA lineage survivorship in animal populations. *J. Mol. Evol.* 20:99-105.
- Avise, J.C. and W.S. Nelson. 1989. Molecular genetic relationships of the extinct Dusky Seaside Sparrow. *Science* 243:646-648.
- Avise, J.C., C.A. Reeb and N.C. Saunders. 1987b. Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). *Evolution* 41:991-1002.
- Avise, J.C. and M.H. Smith. 1974. Biochemical genetics of sunfish. I. Geographic variation and subspecific intergradation in the bluegill, *Lepomis macrochirus*. *Evolution* 28:42-56.
- Ball, R.M., Jr., J.E. Neigel, and J.C. Avise. 1990. Gene genealogies within the organismal pedigrees of random mating populations. *Evolution* 44:360-370.
- Bermingham, E. and J.C. Avise. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics* 113:939-965.
- Bert, T.M. 1986. Speciation in western Atlantic stone crabs (genus *Menippe*): the role of geological patterns and climatic events in the formation and distribution of species. *Marine Biol.* 93:157-170.
- Bowen, B.W. and J.C. Avise. 1990. The genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden and sturgeon: the influence of zoogeographic factors and life history patterns. *Mar. Biol.* 107:371-381.
- Briggs, J.C. 1974. *Marine Zoogeography*. McGraw-Hill, N.Y.
- Britten, R.J. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231:1393-1398.
- Brown, W.M., M. George, Jr., and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 76:1967-1971.
- Ferris, S.D. and W.J. Berg. 1987. The utility of mitochondrial DNA in fish genetics and fishery management. Pp. 277-299 In: F. Utter and N. Ryman (eds.), *The application of Population Genetics to Fisheries Management*. Univ. Washington Press, Seattle.



- Harrison, R.G. 1989. Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.* 4:6-11.
- Lande, R. 1988. Genetics and demography in biological conservation. *Science* 241:1455-1460.
- Moritz, C., T.E. Dowling, and W.M. Brown. 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* 18:269-292.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia Univ. Press, N.Y.
- Nei, M. and D. Graur. 1984. Extent of protein polymorphism and the neutral mutation theory. *Evol. Biol.* 17:73-118.
- Neigel, J.E. and J.C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pp. 515-534 In: E. Nevo and S. Karlin (eds.), *Evolutionary Processes and Theory*, Academic Press, N.Y.
- Phillip, D.P., W.F. Childers, and G.S. Whitt. 1981. Management implications for different stocks of largemouth bass (*Micropterus salmoides*) in the United States. *Can. J. Fish. Aquat. Sci.* 38:1715-1723.
- Powell, J.R., A. Caccone, G.D. Amato, and C. Yoon. 1986. Rates of nucleotide substitution in *Drosophila* mitochondrial DNA and nuclear DNA are similar. *Proc. Natl. Acad. Sci. USA* 83:9090-9093.
- Quattro, J.M. and R.C. Vrijenhoek. 1989. Fitness differences among remnant populations of the endangered Sonoran topminnow. *Science* 245:976-978.
- Reeb, C.A. and J.C. Avise. 1990. A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, *Crassostrea virginica*. *Genetics* 124:397-406.
- Saunders, N.C., L.G. Kessler, and J.C. Avise. 1986. Genetic variation and geographic differentiation in mitochondrial DNA of the horseshoe crab, *Limulus polyphemus*. *Genetics* 112:613-627.
- Shields, G.F. and K.M. Helm-Bychowski. 1988. Mitochondrial DNA of birds. Pp. 273-295 In: R.F. Johnston (ed.), *Current Ornithology*, Vol. 5. Plenum Press, N.Y.
- Shields, G.F. and A.C. Wilson. 1987. Calibration of mitochondrial DNA evolution in geese. *J. Mol. Evol.* 24:212-217.
- Sibley, C.G. and J.E. Ahlquist. 1984. The phylogeny of the homonoid primates as indicated by DNA-DNA hybridization. *J. Mol. Evol.* 20:2-15.

- Slatkin, M. 1985. Rare alleles as indicators of gene flow. *Evolution* 39:53-65.
- Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787-792.
- Sneath, P.H.A. and R.R. Sokal. 1973. *Numerical Taxonomy*. W.H. Freeman, San Francisco.
- Swift, C.C., C.R. Gilbert, S.A. Bortone, G.H. Burgess, and R.W. Yerger. 1985. Zoogeography of the freshwater fishes of the southeastern United States: Savannah River to Lake Ponchartrain. Pp. 213-265 In: C.H. Hocutt and E.O. Wiley (eds.), *Zoogeography of the North American Freshwater Fishes*. John Wiley and Sons, N.Y.
- Vawter, L. and W.M. Brown. 1986. Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. *Science* 234:194-196.
- Vrijenhoek, R.C., M.E. Douglas, and G.K. Meffe. 1985. Conservation genetics of endangered fish populations in Arizona. *Science* 229:400-402.
- Wildt, D.E., M. Bush, K.L. Goodrowe, C. Packer, A.E. Pusey, J.L. Brown, P. Joslin, and S.J. O'Brien. 1987. Reproductive and genetic consequences of founding isolated lion populations. *Nature* 329:328-330.
- Williams, G.C. and R.K. Koehn. 1984. Population genetics of North Atlantic catadromous eels (*Anguilla*). Pp. 529-560 In: B.J. Turner (ed.) *Evolutionary Genetics of Fishes*. Plenum press, N.Y.
- Wilson, A.C., R.L. Cann, S.M. Carr, M. George, Jr., U.B. Gyllensten, K.M. Helm-Bychowski, R.G. Higuchi, S.R. Palumbi, E.M. Prager, R.D. Sage, and M. Stoneking. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26:375-400.
- Wooten, M.C., K.T. Scribner, and M.H. Smith. 1988. Genetic variability and systematics of *Gambusia* in the southeastern United States. *Copeia* 1988:283-289.

## DISCUSSION

PROFESSOR STUART BARKER: With regard to the difference between your predicted evolutionary effective population size and the present population size, you did not comment on the fact that the latter is an estimated present census population size rather than a present effective population size. Perhaps that could account for a very large fraction of the difference.

DR. JOHN AVISE: If you look at the kind of estimates that many biologists have made for how present population size relates to effective population size, the estimates are usually that the effective population size should be on the order of 70-90% of the current day census size, given

the demography that's supposed for the species. What we find is that the reduction is vastly greater than that - the reduction is a hundred-fold to a thousand-fold in many cases, which suggests that something is going on beyond just the normal variance in reproductive successes associated with the contemporary demographics of these particular species.

PROFESSOR STUART BARKER: I have very little knowledge of fish species, but certainly in many other cases the ratio of the effective size to the census size can be very small and I have fairly grave suspicions about some of the 70-90% estimates. I think the true percentage will often be less, and perhaps very much less.

DR. JOHN AVISE: That's a very good point. You can find much wider estimates, including some that are much lower; our data are certainly most consistent with those that fall at the lower end of the range.

Another point I can make while we're on that topic is that, (and this is an old argument in molecular biology that originally came out of the allozyme data), Nei and Graur and others have pointed out that even though we talk about populations containing vast amounts of genetic variability, in fact the variation is still far less than predicted under neutrality theory given suspected population sizes and mutation rates to neutral alleles. Thus the problem really is not in understanding why so much variation exists in nature, but why so little variation exists. Nei and Graur have argued that if one looks for evolutionary explanations for this genetic variation, one therefore ought not to seek them in terms of diversity-enhancing evolutionary mechanisms, but rather in diversity-reducing evolutionary mechanisms. Our mitochondrial data are consistent with this view. There is a wealth of mtDNA polymorphism out there, but still far less than neutrality theory might predict. That's a dilemma, to account for why the variability should be low given the suspected rate of mtDNA evolution. My inclination is that historical demography may be responsible (rather than episodic waves of positive directional selection acting on favoured genetic variants). Nonetheless, it remains an open issue warranting further investigation.

ANON: I have two questions and they are interrelated. The first one is, what is more expensive to run, molecular genetic assessment or electrophoresis and secondly, in terms of cost benefit which is more beneficial?

DR. JOHN AVISE: It's certainly more expensive to run a DNA lab, but there are so many variations on how the method is employed and the distinction between capital setup and operating costs that it's hard to get an unequivocal answer to that question. In general, it's much more expensive to gather mitochondrial or other DNA data as compared to allozyme information.

I still think allozymes are immensely important sources of genetic information for all kinds of species. I became involved with mitochondrial DNA because it offers a very different perspective on intraspecific evolutionary process. In prior work with allozymes, I had seldom thought in terms of phylogeny at the intraspecific level. Allozyme data fit well into the traditional language and thinking of population genetics, where one is concerned with allele frequencies, with variances in allele frequencies across populations, and how that relates to F statistics, gene flow, etc. Phylogeny seldom enters the discussion. Thus consideration of allozyme data seldom involves talk in terms of synapomorphies or related phylogenetic concepts and terms. Mitochondrial DNA has changed my own view of intraspecific evolution considerably.

I now think of species as being phylogenetically structured in terms of their genic contents, and to me the concerns of the next decade or so are going to centre around methods and principles of establishing gene genealogies. It may be possible to gather this kind of phylogenetic data for nuclear haplotypes as well. If so, an important question will be: to what degree are the phylogenies constructed from unlinked loci in the genome concordant in their population distinctions? I would argue that only under conditions of strong genealogical concordance can one argue convincingly for historical population separations. I would predict, on the basis of our concordant patterns across species in mitochondrial DNA phylogenies, that these kinds of fundamental phylogenetic distinctions will likely characterize nuclear genomes as well, but this remains to be seen and is an important issue. So I can't unequivocally say that these mitochondrial disjunctions reflect the major sources of nuclear gene pool diversity in species. The kind of genealogical concordance I described in my talk is solely a genealogical concordance in mtDNA patterns across species. An equally important kind of genealogical concordance is that across loci within an organismal pedigree, and this form of concordance will require a good deal of formal theoretical treatment as well as empirical data analysis by those who can obtain haplotype sequences from the nuclear genome.

ANON: Can I just respond to that question? It depends on the sample sizes. I think you've got to remove the concept of the cost of the laboratory work. It is such a small part of the fisheries project together with the cost of the vessel that the only question is, do the mitochondrial techniques allow you to deal with much larger sample sizes and therefore make it relatively cheaper than electrophoresis

DR. JOHN AVISE: The sample sizes that we can deal with are typically about an order of magnitude smaller than can be handled with allozyme methods, given a comparable amount of work and effort. So we are plagued by sample size problems and I admit that freely.

One perspective that I like to take on the issue, in defense of the mitochondrial methodology, is that one can also adopt a philosophy that treats individuals as operational taxonomic units (rather than the populations). Mitochondrial DNA transmission paths follow independent matriarchal lines within sexually reproducing species, so that one can take the data from even a few individuals and ask questions about the time since they last shared a common mother (irrespective of other individuals in the population). So one can shift the focus from populations as OTU's to individuals as OTU's. There are many problems in population biology that can capitalize on this perspective, even with very small sample sizes, to answer particular kinds of questions.

For large populations or species, one can also generate, as I've said, expected frequency distributions of times to common ancestry, and then compare these with observations. It is thereby possible to address, for example, the evolutionary effective population size. I don't think such estimates depend on particularly large sample sizes.

PROFESSOR ROSS CROZIER: I think we'll leave it there. I'll just make my own comment on this too and that is that when we look at the expense of any particular research project in actual fact the greater part is probably salaries, so anything added on top is actually quite small. What we're talking about is what the available funds are. But on that final note: we've seen the importance of allele phylogenies in these studies, and I'd like to thank very much John Avise for letting us hear such a fine assemblage of pertinent case histories.

**SESSION 3**

**POPULATION STRUCTURE (continued)**

**Chairman: Professor Ross Crozier**

**La Trobe University,  
Melbourne, Victoria**

## CHAPTER 3

### The analysis of genetic structure of populations \*

Professor J.S.F. Barker

Department of Animal Science  
University of New England  
Armidale NSW 2351

#### SUMMARY

If there are differences in the average genetic composition of individuals from different parts of a population, then that population is genetically structured. Such structure is commonly assumed in fisheries biology, as evidenced by the description of stocks or races, yet detailed knowledge of the genetic and spatial differentiation of these stocks and of the impact of fishery practices on the structure is often lacking. A sound understanding of the genetic structure of exploited fish species is essential for the development of appropriate fisheries management, and for the conservation of genetic resources.

Defining the genetic structure of a particular population is the necessary first step, but management and conservation decisions will often depend on knowledge of what factors have operated to determine the structure and what factors are operating to maintain it.

Three quantitative measures of subpopulation differentiation based on allele frequency data, viz. Wright's F- statistics (or equivalently Nei's gene diversity), genetic distance and spatial auto-correlation are defined, with some brief outline of methods of estimation and statistical testing. The interpretation of results and inferences that may be drawn from them are considered.

*\* Paper presented at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990.*

## INTRODUCTION

The null hypothesis of population genetics is the ideal panmictic population - genetically homogeneous and random mating. However, if there are differences in the average genetic composition of individuals from different parts of a population, then that population is genetically structured. The analysis of genetic structure therefore depends on knowledge of the distribution of kinds of individuals (genotypes) and aims to quantify the degree and nature of heterogeneity in that distribution of genotypes. The emphasis on the distribution of genotypes is important (Barker, 1989), as the term population structure is often used in an ecological context to describe the number and distribution of individuals. This distribution of individuals in a population may range from essentially continuous to highly clumped, but populations at either of these extremes may be essentially panmictic (homogeneous genetic structure) or highly structured genetically - there is no necessary relationship between the two.

In fisheries biology, some structure of populations, or of a whole species, commonly has been assumed, as evidenced by the description of stocks or races. In some cases, clearly population structure is implied, where the stocks are subpopulations located in different places. In other cases, for example where stocks show differences in average morphology, genetic structure may be implied. But the morphological differences may be environmental and the stocks actually genetically homogeneous, so that although the population is structured, it is not genetically structured.

Population structure *per se* may well be very important for fisheries management, but knowledge of genetic structure (or lack of such knowledge) is likely to be of more profound importance. A sound understanding of the genetic structure of exploited fish species is essential for the development of appropriate fisheries management and conservation of genetic resources.

## FACTORS AFFECTING GENETIC STRUCTURE

In this paper, I will use the term population to refer to the major grouping of interest. It may be a total species population, or that part of the range of a species in a defined geographical area. The term subpopulation will be used for any grouping within the population.



Where there are distinct subpopulations because of geographical, ecological or behavioural factors, the genetic connections among them depend on the amount of genetically effective migration (gene flow) between subpopulations. With high levels of gene flow, the population may breed as a single random mating unit, while with no gene flow, each subpopulation is an independent unit. Where the population distribution is essentially continuous, it can still be genetically structured and show local genetic differentiation if the range of the population is large relative to the distances that individuals move during their lifetime.

Thus the present genetic structure of a population depends on the balance of evolutionary forces - genetic drift contributing to genetic differentiation among subpopulations, natural selection that may lead to differentiation if selective forces differ among subpopulations or, to homogeneity if selective forces are similar in all, and gene flow among subpopulations promoting genetic homogeneity.

Genetic structure implies structure at the time of reproduction, i.e. breeding structure, and this may present particular problems in defining genetic structure for many marine species. For example, consider a species that has an annual breeding season and in which reproduction occurs only in specific habitats, but individuals disperse widely during the non-breeding period. If there were any tendency for individuals to return as breeding adults to where they themselves were spawned, there could be substantial genetic differentiation among the breeding populations and significant genetic structure. If there were complete mixing of individuals from the different breeding locations during the non-breeding period, sampling at that time to assess genetic composition would give no indication of the structure. The specification of appropriate sampling strategies for the analysis of genetic structure will depend on some knowledge of the ecology, life-history and reproductive biology of the species.

## STATISTICAL TESTS OF SUBPOPULATION DIFFERENTIATION

Suppose genotype and allele frequencies for a number of loci have been obtained for each of a number of subpopulations. The primary question is whether there are genetic differences among these subpopulations. Before addressing this question, it is necessary to introduce a basic concept of population genetics - the Hardy-Weinberg genotypic proportions. Consider a large

random mating population and a locus with two alleles, A and a. If there is no selection, mutation or migration affecting allele frequencies at this locus, and p and q are the frequencies of A and a, then the expected or Hardy-Weinberg proportions of the genotypes AA, Aa and aa will be  $p^2$ ,  $2pq$  and  $q^2$ .

Returning then to the question of genetic differences among subpopulations, each locus may be tested separately by a G-test of homogeneity (Sokal and Rohlf, 1981) or the contingency  $\chi^2$  statistic. If all the subpopulations are in Hardy-Weinberg proportions, allele frequencies would be tested, while if there are deviations from Hardy-Weinberg proportions in some or all subpopulations, genotype frequencies would be tested. In either case, if the G or  $\chi^2$  is not significant, one accepts the null hypothesis that the subpopulations are identical for that locus.

This test can be repeated for each locus, and if none are significant, we can conclude that the subpopulations are not genetically differentiated. Obviously, the more loci that are tested, the more confidence we would have in that conclusion.

But suppose that some loci indicate genetic differentiation while some do not. An overall test of genetic homogeneity then could be made on the sum of the single locus statistics. However, this assumes that the loci are statistically independent, which may not be valid, and the only conclusion is that the subpopulations are apparently differentiated for some loci.

Where a number of subpopulations are being tested, there is the further problem that a conclusion of significant differentiation provides no detail of the structure, i.e. how different are the subpopulations from one another?

Statistical tests of the null hypothesis of identical allele or genotype frequencies in the subpopulations should be the first step in analysis of genetic structure, but other procedures are necessary to gain a picture of the structure.

As measures to quantify the spatial patterns of genetic variation, we will discuss Wright's F-statistics, genetic distance measures and spatial auto-correlation.

## WRIGHT'S F-STATISTICS

The theoretical analysis of genetic structure was pioneered by Sewall Wright (reviewed by Wright, 1969) as a development from his early studies of inbreeding in the history of British Shorthorn Cattle. As a measure of inbreeding, Wright (1921) defined the fixation index (designated  $F$ ) and later extended this to describe the properties of hierarchically subdivided populations. In this development, Wright (1951) introduced three parameters,  $F_{IS}$ ,  $F_{IT}$  and  $F_{ST}$  where

- $F_{IS}$  is the fixation index of individuals relative to their subpopulation, or the correlation between genes within individuals within subpopulations,
- $F_{IT}$  is the fixation index of individuals relative to the total population, or the correlation between genes within individuals,
- $F_{ST}$  is the correlation of genes between individuals within the same subpopulation, and can be interpreted as a measure of the amount of differentiation among subpopulations, relative to the limiting amount under complete fixation.

These statistics are related as:

$$(1 - F_{IT}) = (1 - F_{IS}) (1 - F_{ST})$$

When allele and heterozygote frequencies are known for subpopulations, these F-statistics may be estimated and used to describe genetic structure, including inbreeding or deviations of each subpopulation from its Hardy-Weinberg frequencies ( $F_{IS}$ ) and deviation of the entire population from Hardy-Weinberg proportions ( $F_{IT}$ ). However,  $F_{ST}$  as the measure of subpopulation differentiation is of primary interest here.

For a locus with two alleles A and a at frequencies  $p_i$  and  $q_i (= 1 - p_i)$  in subpopulation i,

$$F_{ST} = \frac{V_p}{\bar{p} \bar{q}}$$

where  $V_p$  is the variance of  $p_i$  over subpopulations, and

$\bar{p}$  is the average over subpopulations of  $p_i$ .

$F_{ST}$  may also be expressed in terms of the average of expected heterozygosity in the subpopulations, and the expected total heterozygosity (Nei 1973a) :

$$F_{ST} = 1 - H_S/H_T$$

where  $H_S = \sum H_i/n$  (n subpopulations)

$$H_i = 1 - (p_i^2 + q_i^2)$$

$$H_T = 1 - (\bar{p}^2 + \bar{q}^2)$$

These definitions of  $F_{ST}$  are strictly for a two allele locus. For a locus with more than two alleles,  $F_{ST}$  may be estimated for each allele by pooling the frequencies of all other alleles. The  $F_{ST}$  values will generally differ among alleles, but a single estimate for the locus is given as a weighted average of the  $F_{ST}$  for each allele, where the weight for allele  $i$  is  $\bar{p}_i(1 - \bar{p}_i)$  (Wright, 1978). This may be extended to average over loci, giving a single measure ( $\bar{F}_{ST}$ ) of subpopulation differentiation.

Nei (1973a, 1977) introduced a concept of gene diversity and a parameter  $G_{ST}$ , based on allele frequencies at any number of multiallelic loci. This G-statistic has been widely used, and while it is sometimes argued that it extends  $F_{ST}$  to allow for multiple alleles,  $G_{ST}$  is identical to  $\bar{F}_{ST}$  as defined above. However, as an alternative formulation, it is useful to define  $G_{ST}$ . For a given locus with more than two alleles, if  $p_{ik}$  is the frequency of allele  $k$  in subpopulation  $i$ , then:

$$H_i = 1 - \sum_k p_{ik}^2$$

$$H_S = \sum H_i/n$$

Note that these are equivalent to the  $H_i$  and  $H_S$  given previously for a two allele locus. Similarly,

$$H_T = 1 - \sum_k \bar{p}_k^2$$

$$\text{where } \bar{p}_k = \sum p_{ik}/n$$

Then, defining  $\bar{H}_S$  and  $\bar{H}_T$  as the averages over loci of  $H_S$  and  $H_T$ ,

$$G_{ST} = 1 - \bar{H}_S/\bar{H}_T$$

Here we have considered just one level of structure, viz. subpopulations within a total population. Often, however, it may be desired to apportion differentiation (or diversity) in an extended hierarchy, say demes within subpopulations, and subpopulations within the total population. Wright's F-statistics would then be:

$$(1 - F_{IT}) = (1 - F_{ID}) (1 - F_{DS}) (1 - F_{ST})$$

If we now define a further statistic,  $F_{DT}$ , as the measure of differentiation among demes within the total population, then the proportions of variation are:

$$(1 - F_{DT}) \quad \text{within demes}$$

$$(F_{DT} - F_{ST}) \quad \text{among demes within subpopulations, and}$$

$$F_{ST} \quad \text{among subpopulations}$$

An alternative approach to the estimation of parameters of population structure was introduced by Cockerham (1969, 1973), and extended to multiple alleles and loci (Weir and Cockerham, 1984). This approach is based on an analysis of variance model, and estimates parameters that are formally equivalent to Wright's  $F_{ST}$ ,  $F_{IS}$  and  $F_{IT}$  (see Barker, East and Weir (1986) for an example of application of this method).

A number of other methods of quantifying differentiation among subpopulations have been proposed, but the three discussed above appear to be more generally accepted, with Nei's gene diversity measure widely used. Although these three methods are 'closely related' or 'formally equivalent' as theoretical definitions, they do involve somewhat different assumptions, and a variety of estimation procedures have been used. Further, the estimates obtained are subject to statistical uncertainty - being based on a sample of loci actually assayed, samples of individuals from each subpopulation, and samples of subpopulations. Weir and Cockerham (1984) and Chakraborty and Leimar (1987) discuss these estimation and statistical problems. There is as yet, no consensus as to the best method of analysis, but the method of Weir and Cockerham (1984) incorporates sample size and subpopulation number into the estimation, so that interpretation is independent of the sampling scheme. This is an advantage in many field studies where sample sizes are unequal and small, and in giving results which are more general and more useful in comparing results from different studies.

What then is the investigator to do? Almost certainly he will use a computer package, such as BIOSYS-1 (Swofford and Selander, 1989) which includes a Wright's F-statistic analysis, or GENESTATS (Black and Krafur, 1985) which includes the methods of Nei and of Weir and Cockerham (Note: the program GENESTATS contains an error, Black pers. comm.). Regardless of the method used, it must be emphasised that the investigator should consider the assumptions underlying the method, the statistical uncertainty (significance) of the results, and hence should interpret the analysis cautiously, particularly when based on limited data.

However, the most important consideration of the sampling strategy needs to be made long before the question of analytical methods. Sampling should include the widest possible range of subpopulations (and of subdivisions within each) - again prior knowledge of the biology, ecology and population structure of the species of interest will be advantageous. Within subpopulations (or whatever is the lowest unit of the hierarchy investigated), increasing the number of loci assayed for each individual will be more useful than increasing the number of individuals. For example, in a current study of genetic structure of swamp buffalo and native goats in southeast Asia (Tan *et al.*, 1990) our strategy is to develop electrophoretic assay systems for at least 70 loci in each species, to initially assay 25 individuals from each locality (subpopulations), and to sample these subpopulations as widely as resources permit through out the region - initially about 12 localities. The data base can easily be expanded and the genetic

structure analyses repeated as additional localities are sampled or additional assay systems are developed.

As an example of results using this method of analysis, Table 1 gives Nei's gene diversity analyses for geographically grouped samples from naturally reproducing populations and hatchery stocks of Atlantic salmon (Ståhl, 1987). More than 40% of the total gene diversity is between samples, with most of this between regions. However, the substantial component for between drainages within regions has significant management implications. This study was based on 38 electrophoretic loci, and the gene diversity results for the individual polymorphic loci are given in Table 2. Most of the diversity between regions is due to only four loci, as the other loci had very low levels of polymorphism. Had one or more of these four loci not been

**Table 1 . Gene diversity analysis based on 38 loci in geographically grouped samples from naturally reproducing populations and hatchery stocks of Atlantic salmon. Regions are: landlocked, Baltic Sea, Eastern Atlantic Ocean and Western Atlantic Ocean. (From Ståhl, 1987)**

Item	Hatchery stocks	Natural populations	Total
Number of			
Regions	4	3	4
Drainages	17	18	31
Samples	24	29	53
Fish	2410	1699	4109
Absolute gene diversity			
Total	.037	.041	.040
Standard error	(.020)	(.021)	(.021)
Relative gene diversity (in percent)			
Between regions	26.5	29.7	28.4
Between drainages within regions	14.5	4.9	9.0
Between samples within drainages	5.1	1.5	3.6
Within samples	53.9	63.9	59.0

**Table 2.** Distribution of gene diversity at variable loci among 53 geographically grouped samples of naturally reproducing populations and hatchery stocks of Atlantic salmon.  
(From Ståhl, 1987)

<u>Locus</u>	Absolute gene diversity		Relative gene diversity (in percent)			
	<u>Total</u>	<u>Within samples</u>	<u>Between regions</u>	<u>Between drainages within region</u>	<u>Between samples within drainages</u>	<u>Within samples</u>
<i>AAT-3</i>	0.410	0.203	39.7	8.3	2.4	49.6
<i>AGP-2</i>	0.0009	0.0008	0.04	1.3	-	98.7
<i>IDH-3</i>	0.006	0.005	0.5	16.0	-	83.5
<i>LDH-4</i>	0.0009	0.0009	0.3	0.5	0.9	98.3
<i>ME-2</i>	0.497	0.256	35.2	11.7	1.5	51.6
<i>MDH-1</i>	0.010	0.009	1.3	2.2	1.7	94.7
<i>MDH-3</i>	0.076	0.059	17.0	3.0	2.3	77.7
<i>PGI-1</i>	0.001	0.001	0.1	1.2	0.2	98.6
<i>PGM-1</i>	0.008	0.007	0.8	4.1	5.9	89.2
<i>SDH-1</i>	0.492	0.343	15.5	7.9	6.9	69.7
<i>SDH-2</i>	0.002	0.002	0.1	4.6	-	95.3
Average	0.040	0.023	28.4	9.0	3.6	59.0
<u>Standard error</u>	<u>0.021</u>	<u>0.012</u>	<u>5.8</u>	<u>1.1</u>	<u>1.3</u>	<u>5.2</u>



assayed, the picture of diversity between regions would have been quite different - again emphasising the need to maximize the number of loci studied.

## GENETIC DISTANCE

Several measures of genetic distance between subpopulations have been proposed, where a measure expresses as a single number the amount of difference in genetic constitution of two subpopulations. In general, if two subpopulations have the same allele frequencies at some set of loci, they will be zero distance apart. As allele frequency differences increase, so the distance becomes greater.

For just two subpopulations, a genetic distance measure is of no advantage, and more information will be gained by direct comparison of allele frequencies. As the number of subpopulations increases, a genetic distance measure will provide a better visualization of the differences and of their relative magnitude, e.g. is subpopulation A more different from subpopulation B than it is from C. Further, all of the pairwise distances may be used to reconstruct the evolutionary history of the subpopulations, i.e. a phylogeny or a two dimensional pattern of nodes and branches, where closely related subpopulations are placed on adjacent terminal branches, and distantly related ones may be separated by many nodes.

If two subpopulations are, for geographic or ecological reasons, genetically isolated from each other (i.e. no gene flow between them), they will tend to accumulate different alleles. The reason for this differentiation will be mutation, selection and random genetic drift. Thus if allele frequency data are available for only one or a few loci, the estimated genetic distances and phylogeny will not be reliable. But if a large number of loci is used (50 or more, Lewontin, 1974), effects of genetic drift or selection varying for different loci will be averaged out.

The genetic distance measures that have been proposed are mathematically rather diverse (reviewed by Wright, 1978; Nei, 1987) and for some, their biological interpretation is not clear (Nei, 1973b). If the rate of gene substitution per unit time is constant, Nei's standard genetic distance (1972, 1973b) is linearly related to the time after divergence of two populations. Further, the standard errors of Nei's distance statistics can be estimated (Nei and Roychoudhury, 1974). For these reasons, Nei's genetic distance measures have been the most widely used in

studies of population differentiation and in evolutionary biology.

Because the mathematical properties and biological bases of the various measures do differ, different distance measures could lead to different interpretations of the phylogenetic relationships among a set of subpopulations, with no way of determining the 'best' phylogeny, i.e. the one closest to the true phylogeny (see Felsenstein, 1985 for the consideration of the statistical problem of inferring phylogenies). However, correlations among various distance measures have been found to be generally very high (Hedrick, 1975; Chakraborty and Tateno, 1976), particularly for distances among local populations.

A new measure of the genetic distance, based on the coancestry coefficient, was proposed by Reynolds *et al.* (1983), particularly for short-term evolution (e.g. local populations diverging due to drift only). They argue that in this case, their weighted estimator appears to be a better measure of distance than the Nei (1973b) or other measures.

As an example of a procedure for estimating genetic distance, and simply because it has been most widely used, Nei's standard genetic distance will be used.

Suppose two subpopulations, X and Y, with a number of alleles segregating at a particular locus. If  $x_i$  and  $y_i$  are the frequencies of the  $i$ th allele in X and Y, the probability of identity of two randomly chosen genes in X is  $j_x = \sum (x_i^2)$ , and  $j_y = \sum (y_i^2)$  in Y. The probability of identity of two genes chosen at random, one from each subpopulation, is  $j_{xy} = \sum x_i y_i$ . In each case,  $\Sigma$  means summation over all alleles at the locus. Then, if  $J_X$ ,  $J_Y$  and  $J_{XY}$  are the arithmetic means of  $j_x$ ,  $j_y$  and  $j_{xy}$  over all loci, including monomorphic ones, the genetic identity between X and Y is:

$$I = \frac{J_{XY}}{\sqrt{J_X J_Y}}$$

and the standard genetic distance is:

$$D = -\ln I$$

A number of authors have used observed allele frequencies in the above equations to estimate genetic distance, but when sample sizes are small, the estimate is biased. Unbiased estimates

are obtained by replacing  $\sum x_i^2$  and  $\sum y_i^2$  by  $(2n \sum x_i^2 - 1)/(2n-1)$  and  $(2n \sum y_i^2 - 1)/(2n-1)$ , where  $x_i$  and  $y_i$  are now the observed frequencies of allele  $i$  in  $n$  individuals sampled from each population (Nei, 1978).

The exact variance of standard genetic distance is not readily obtained, but a formula for estimating the asymptotic variance is given by Nei (1987), and he also discusses tests of significance for comparing genetic distances. As noted previously, it is important to study a large number of loci (50 or more) rather than a large number of individuals per locus to reduce the variance of the genetic distance estimate (Nei and Roychoudhury, 1974; Nei, 1978). On the other hand, Archie *et al.* (1989) note that the number of individuals sampled per population should not be too small, say at least 20.

Whatever distance measure is used, the matrix of pair-wise genetic distances needs to be converted to a phylogeny. When this is derived for a set of incompletely isolated subpopulations, it will not necessarily represent the real evolutionary history in terms of divergence times, unless the rate of evolution is constant in different lineages for a given locus. But for visualizing relationships among local populations, however, this does not matter, as the pattern will represent the genetic relationships among the populations at the time the allele frequencies are determined. In this case, the diagram of relationships is generally called a dendrogram.

Again, a number of different methods are available for analysis of the matrix of pair-wise distances to produce a dendrogram. Nei *et al.* (1983) compared three methods and concluded that UPGMA (unweighted pair group method with arithmetic means) was best. Subsequently, Rogers (1984, 1986) and Swofford and Berlocher (1987) have developed new methods claimed to consider problems specific to allele frequency data and more appropriate for inferring evolutionary tree, i.e. relationships among species and higher level taxa. Swofford and Berlocher (1987) specifically note that use of their method for conspecific populations should be approached with caution.

A number of the genetic distance measures and four methods for producing a phylogeny from these distances (including UPGMA) are included in the BIOSYS-1 computer package (Swofford and Selander, 1989).

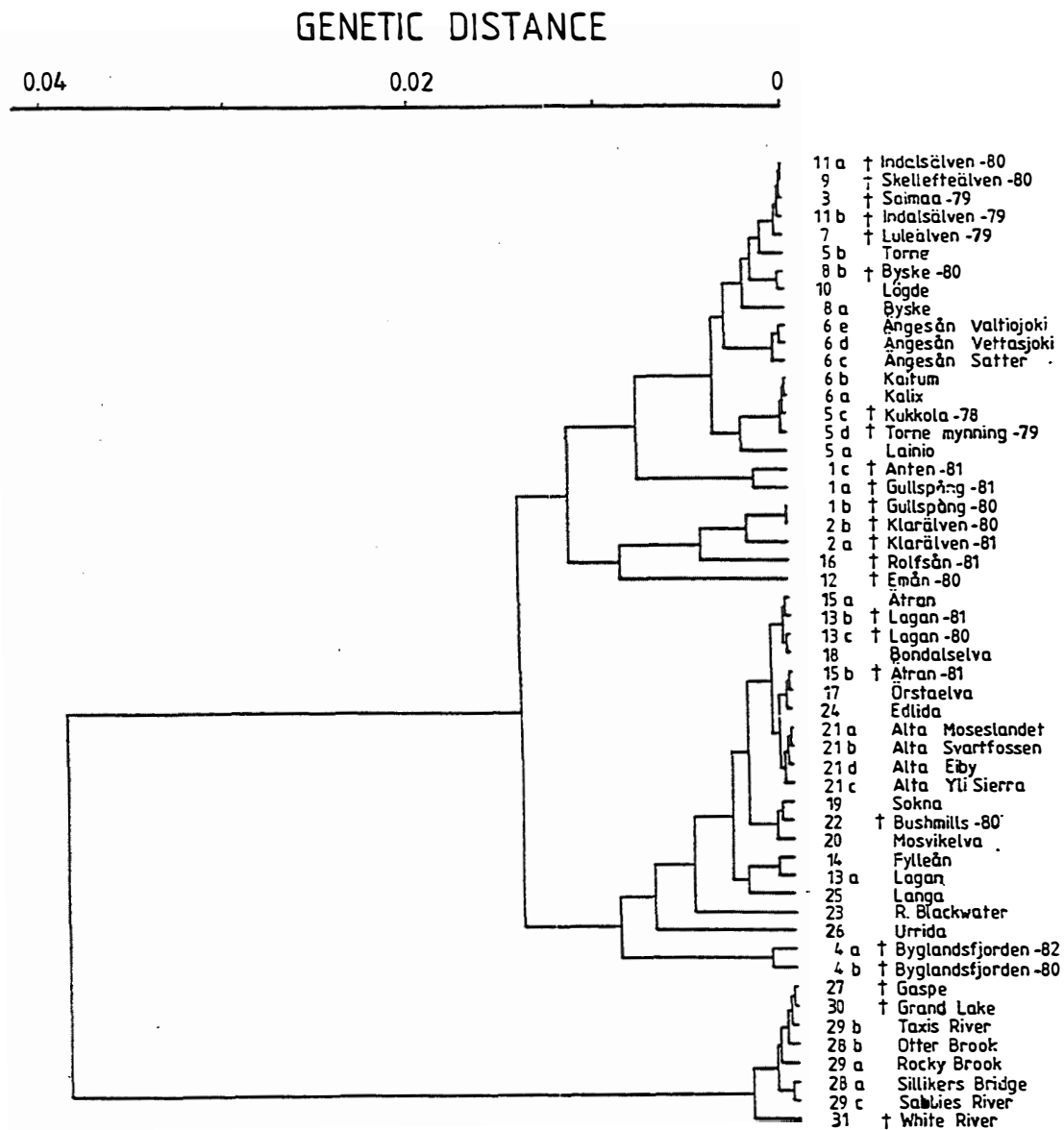


FIG. 1. Dendrogram (UPGMA) summarizing the genetic relationships among 29 samples representing naturally reproducing populations and 24 samples from hatchery derivatives (marked by †) of Atlantic salmon. The dendrogram is constructed from genetic distances between samples based on the allele frequencies at 38 loci. (From Ståhl, 1987).

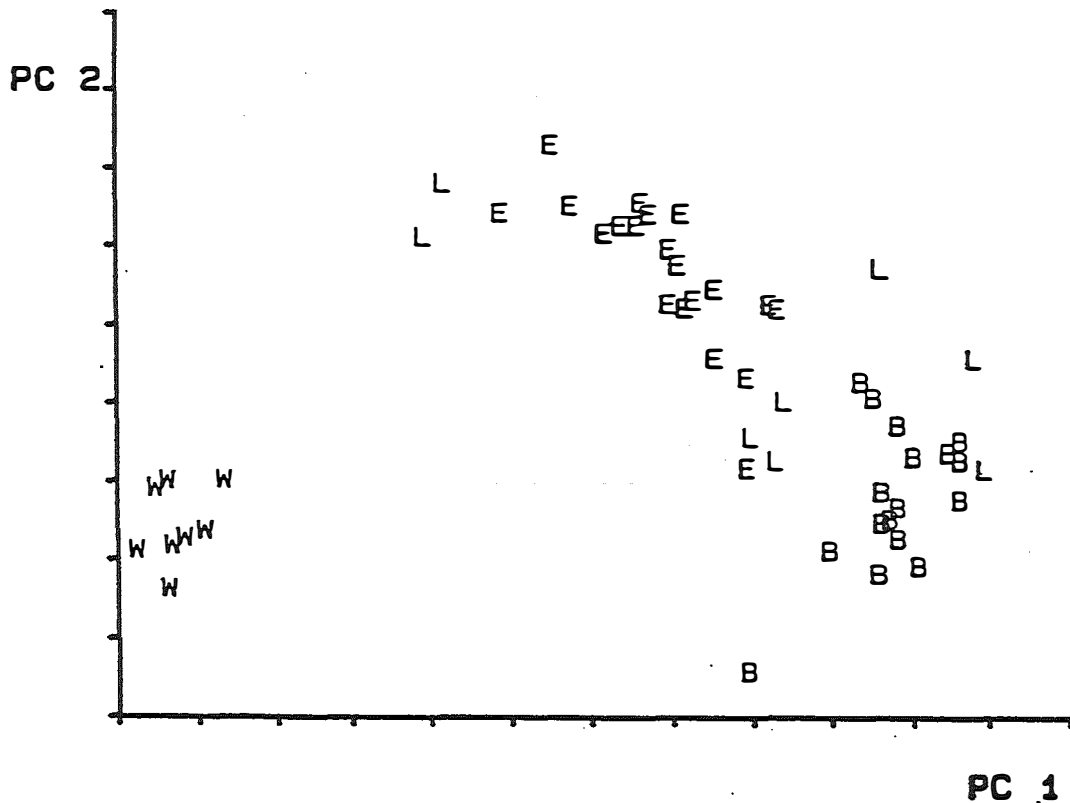


FIG. 2. Principal component scatter plot derived from analyses of allele frequency estimates at 38 loci in 29 samples from naturally reproducing populations and 24 samples from hatchery stocks of Atlantic salmon. The first principal component (PC 1) accounts for about 65% of the total variance and the second component (PC 2) for 19%. Identical scales are used for the x- and the y-axis. (From Ståhl, 1987).

B: sample from Baltic Sea,  
 E: sample from Eastern Atlantic Ocean,  
 W: sample from Western Atlantic Ocean, and  
 L: samples from landlocked populations.

The data of Ståhl (1987) for Atlantic salmon populations which were used previously to illustrate gene diversity analysis, were analyzed also for Nei's genetic distances and a dendrogram constructed using UPGMA (Figure 1). Ståhl (1987) also used a principal component analysis on these data, and his plot of the first two principal components is given as Figure 2. Both these visualizations clearly show the three distinct clusters of samples from the Baltic Sea, the Eastern Atlantic and the Western Atlantic. The dendrogram (Figure 1) has the

added advantage of illustrating the inferred relationship between samples within regions. However, little confidence can be placed on those relationships when the genetic distance is very small.

## SPATIAL AUTOCORRELATION

The  $F$ -statistics and genetic distance/phylogeny methods provide respectively a measure and a visualization of genetic structure. The present structure of a population depends on the balance of evolutionary forces - drift, gene flow and selection, and fisheries management or conservation decisions may well depend not only on a knowledge of the present genetic structure, but on an understanding of the factors that have acted to determine the structure, or are acting to maintain it.

For any number of loci in a given set of subpopulations,  $F_{ST}$  values will be expected to be homogeneous since all loci have been subject to the same structure, unless selection is acting at some loci. This expectation would appear to provide a basis for testing whether selection is acting at any loci. Tests for the heterogeneity of  $F_{ST}$  values across loci have been proposed, but are invalid for spatial variation (see Felsenstein, 1982 for comment on this and related problems in inferring history and geography of populations). For neutral loci and assuming subpopulations are at equilibrium,  $F_{ST}$  is given approximately as  $1/(1+4Nm)$ , where  $N$  is the effective size of subpopulations (assumed equal) and  $m$  the proportion of migrants entering each subpopulation each generation. Thus given an estimate of  $F_{ST}$ ,  $Nm$  can be estimated. Note that  $Nm$  is the number of migrants entering each subpopulation each generation, and if  $Nm$  is greater than one, then gene flow is strong enough to prevent substantial differentiation due to genetic drift. While one might prefer to have separate estimates of  $N$  and  $m$ , both are notoriously difficult to estimate, and Slatkin and Barton (1989) have shown the  $F_{ST}$  method for estimation of  $Nm$  to be relatively robust to selection and to variation in population structure. Thus this estimate of  $Nm$  would appear to be useful for determining the importance of gene flow in natural populations.

However, like  $F_{ST}$  itself, it is an average estimate for the population. In order to account for patterns in the allele frequency distributions over subpopulations, including any effects of selection at some loci, other approaches are necessary; spatial autocorrelation analysis and derivatives from it provide one possibility.

Statistical methods for spatial pattern analysis, specifically spatial autocorrelation, were introduced to population biology by Sokal and Oden (1978a, b). Spatial autocorrelation is defined as the association of the values of one variable with the values of the same variable at all other localities (Sokal and Oden 1978a). Thus for genetic data, each allele is separately analyzed, and for any given allele, the observations are its frequencies at each of the localities. We wish to determine whether or not these frequencies show geographic pattern, i.e. do neighbouring localities have more similar allele frequencies than those that are further apart? Neighbours are commonly defined as any pair of localities separated by no more than a given distance, say 100 km. The autocorrelation coefficient then is estimated for these pairs of localities - providing a measure of the similarity of allele frequencies when localities are separated by up to 100km. Other distance classes may be specified for locality pairs separated by 100 to 200 km, 200 to 300 km, and so on and the auto-correlation coefficient estimated for each of these distance classes. A plot of auto-correlation coefficients against distance - a spatial correlogram - summarizes the spatial relationships between populations for that allele. Methods for the computation of spatial autocorrelation coefficients, tests of significance and some examples of their application are given by Sokal and Thomson (1987).

From the patterns of variation in geographic space and from the correlogram, inferences can be made about the processes that have produced the pattern - particularly by comparison of results for different alleles and different loci. Dissimilar patterns indicate differences in the processes producing them, e.g. differential responses by different alleles to environmental patterns differing in spatial patterns, or migration at different rates and in different directions from several source populations. If genetic drift is the only process leading to population differentiation, the correlograms for different alleles will be similar (Sokal and Wartenberg, 1983).

Although spatial correlograms allow inferences about the patterns of spatial variation and the processes that produced them, they cannot show the directionality of any spatial variation pattern. For example, the correlograms for two allele frequencies may be similar, with both indicating clinal variation, yet the direction of the clines may differ. Thus Oden and Sokal (1986) have developed a method for computing directional spatial autocorrelation, which allows evaluation of spatial trends for different compass directions. In this method, pairs of localities are divided into direction/distance classes that indicate both geographic distance and

the compass bearing between them. This procedure has been applied to the study of spatial structure in *D. buzzatii* populations in eastern Australia (Sokal *et al.*, 1987) where the results are compared with spatial studies of allele frequencies in other species.

For loci where clinal variation has been detected (e.g. by spatial autocorrelation analysis), a new method for determining their direction and for comparing these directions for different loci has been developed by Barbujani (1988). This may allow a distinction to be made between effects of long-range gene flow and differential selection. Both of these processes will cause non-random patterns of allele frequencies along an axis corresponding to the direction along which migration occurred, or selection intensity varied. However, long-range gene flow should cause parallel clines at all loci, while loci subject to differential selection will not show similar directional clines, unless they are responding to the same environmental variable(s). Only in this latter case will migration and selection patterns not be distinguishable.

Commonly, spatial variation in allele frequencies may not be clinal, but subject to high rates of change in particular regions. Such areas of rapid transition may indicate steep ecological gradients and resulting selective differences or zones of contact between genetically different subpopulations. A method for detecting such regions of abrupt change, recently developed by Barbujani *et al.* (1989) may be applied to allow further inferences about factors that have acted to determine the genetic structure, or are acting to maintain it.

## CONCLUSIONS

It should be apparent that the three quantitative approaches to the analysis and interpretation of subpopulation differentiation (F statistics, genetic distance and spatial autocorrelation) each give a different perspective, so that the use of all may give a more comprehensive understanding.

Further, particularly for the first two, a variety of parameters and estimation methods have been proposed, and there are many unsolved problems in relation to statistical testing of hypotheses, and the specification of optimum methods. Clearly, this is not very satisfactory for the investigator, but we would be foolish indeed to ignore it.



## REFERENCES

- Archie, J.W., Simon, C. and Martin, A. 1989. Small sample size does decrease the stability of dendrograms calculated from allele-frequency data. *Evolution* 43, 678-683.
- Barbujani, G. 1988. Detecting and comparing the direction of gene-frequency gradients. *Journal of Genetics* 67, 129-140.
- Barbujani, G., Oden, N.L. and Sokal, R.R. 1989. Detecting regions of abrupt change in maps of biological variables. *Systematic Zoology* 38, 376-389.
- Barker, J.S.F. 1989. Population Structure. In: *Evolution and Animal Breeding* (W.G. Hill and T.F.C. Mackay, eds) , pp. 75-80. C.A.B. International, Wallingford.
- Barker, J.S.F., East, P.D. and Weir, B.S. 1986. Temporal and microgeographic variation in allozyme frequencies in a natural population of *Drosophila buzzatii*. *Genetics* 112, 577-611.
- Black, W.C. and Krafsur, E.S. 1985. A FORTRAN program for analysis of genotypic frequencies and description of the breeding structure of populations. *Theoretical and Applied Genetics* 70, 484-490.
- Chakraborty, R. and Leimar, O. 1987. Genetic variation within a subdivided population. In: *Population Genetics and Fishery Management* (N. Ryman and F. Utter, eds), pp. 89-120. University of Washington Press, Seattle.
- Chakraborty, R. and Tateno, Y. 1976. Correlations between some measures of genetic distance. *Evolution* 30, 851-853.
- Cockerham, C.C. 1969. Variance of gene frequencies. *Evolution* 23, 72-84.
- Cockerham, C.C. 1973. Analyses of gene frequencies. *Genetics* 74, 679-700.
- Felsenstein, J. 1982. How can we infer geography and history from gene frequencies? *Journal of Theoretical Biology* 96, 9-20.
- Felsenstein, J. 1985. Phylogenies from gene frequencies: a statistical problem. *Systematic Zoology* 34, 300-311.
- Hedrick, P.W. 1975. Genetic similarity and distance: comments and comparisons. *Evolution* 29, 362-366.
- Lewontin, R.C. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- Nei, M. 1972. Genetic distance between populations. *American Naturalist* 106, 283-292.

- Nei, M. 1973a. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* 70, 3321-3323.
- Nei, M. 1973b. The theory and estimation of genetic distance. In: *Genetic Structure of Populations* (N.E. Morton, ed.) pp. 45-54. University of Hawaii Press, Honolulu.
- Nei, M. 1977. F-statistics and analysis of gene diversity in subdivided populations. *Annals of Human Genetics* 41, 225-233.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583-590.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- Nei, M. and Roychoudhury, A.K. 1974. Sampling variances of heterozygosity and genetic distance. *Genetics* 76, 379-390.
- Nei, M., Tajima F. and Tateno Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *Journal of Molecular Evolution* 19, 153-170.
- Oden, N.L. and Sokal, R.R. 1986. Directional autocorrelation: an extension of spatial correlograms to two dimensions. *Systematic Zoology* 35, 608-617.
- Reynolds, J., Weir, B.S. and Cockerham, C.C. 1983. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* 105, 767-779.
- Rogers, J.S. 1984. Deriving phylogenetic trees from allele frequencies. *Systematic Zoology* 33, 52-63.
- Rogers, J.S. 1986. Deriving phylogenetic trees from allele frequencies: a comparison of nine genetic distances. *Systematic Zoology* 35, 297-310.
- Slatkin, M. and Barton, N.H. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43, 1349-1368.
- Sokal, R.R. and Oden, N.L. 1978a. Spatial autocorrelation in biology 1. Methodology. *Biological Journal of the Linnean Society* 10, 199-228.
- Sokal, R.R. and Oden, N.L. 1978b. Spatial autocorrelation in biology 2. Some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society* 10, 229-249.
- Sokal, R.R. and Rohlf, F.J. 1981. *Biometry*. W.H. Freeman, San Francisco.
- Sokal, R.R. and Thomson, J.D. 1987. Applications of spatial autocorrelation in ecology. In: *Developments in Numerical Ecology* (P. Legendre and L. Legendre, eds) pp. 431-466. NATO ASI Series, Vol. G14. Springer-Verlag, Berlin.
- Sokal, R.R. and Wartenberg, D.E. 1983. A test of spatial autocorrelation using an isolation-by-

distance model. *Genetics* 105, 219-237.

Sokal, R.R., Oden N.L. and Barker, J.S.F. 1987. Spatial structure in *Drosophila buzzatii* populations: simple and directional spatial autocorrelation. *American Naturalist* 129, 122-142.

Ståhl, G. 1987. Genetic population structure of Atlantic salmon. In: *Population Genetics and Fishery Management* (N. Ryman and F. Utter, eds) pp. 121-140. University of Washington Press, Seattle.

Swofford, D.L. and Berlocher, S.H. 1987. Inferring evolutionary trees from gene frequency data under the principle of maximum parsimony. *Systematic Zoology* 36, 293-325.

Swofford, D.L. and Selander, R.B. 1989. BIOSYS-1: A computer program for the analysis of allelic variation in population genetics and biochemical systematics (Release 1.7). Illinois Natural History Survey, Champaign, Ill.

Tan, S.G., Selvaraj, O., Sekaran, M., Mukherjee, T.K. and Barker, J.S.F. 1990. Cellulose acetate electrophoresis of native southeast Asian buffaloes and goats: some technical aspects. *Isozyme Bulletin* 23, 110.

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

Wright, S. 1921. Systems of mating. *Genetics* 6, 111-178.

Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics* 15, 323-354.

Wright, S. 1969. *Evolution and the Genetics of Populations. Vol. 2. The Theory of Gene Frequencies.* University of Chicago Press, Chicago.

Wright, S. 1978. *Evolution and the Genetics of Populations. Vol. 4. Variability Within and Among Natural Populations.* University of Chicago Press, Chicago.

## DISCUSSION

PROFESSOR ROSS CROZIER: I have a question. You didn't comment very much on the use of randomization tests and I think it was Jim Crow recently in *Genetics* who remarked that Fisher would have used randomization tests extensively if he'd had computers. Also Sokal recommended randomization tests for all the autocorrelation statistics and I wondered if you want to comment on that.

PROFESSOR STUART BARKER: You are correct - for many of these statistics, there really is no expectation of the distribution. For that reason it is essentially impossible to get good

measures of standard deviation or standard error, and therefore various randomization procedures have been utilized, or alternately, multiple computation of an expectation for a particular statistic and then testing how frequently particular observations or particular parameters would be expected as against that expectation.

DR. CRAIG MORITZ: In recent studies, particularly in connection with conservation genetics where we are often dealing with wildly fluctuating populations or populations that have been overharvested or whatever, there's been a lot of concern about the equilibrium assumptions associated with getting estimates of migration or effective population size or whatever from these statistics. It's been suggested that  $F_{ST}$  or  $G_{ST}$  respond relatively rapidly, for example, whereas distance measures probably don't or expected homozygosities don't. Do you have any comments on the sort of, general comments on the equilibrium assumption?

PROFESSOR STUART BARKER: Not really other than what you have already intimated, Craig. Much of the theory does assume that the subpopulations are in equilibrium, and I would particularly mention using  $F_{ST}$  to measure the product of effective population size and migration rate, which clearly assumes the subpopulations are in equilibrium. Certainly this is one of the other notes of caution that needs to be raised. In addition, it should be said that anybody who is looking to analyze genetic structure should use more than one method to get at least a better feel for the total population situation.

But I think it is fair to say, and basically you've answered your own question in that respect, that the distance measures are more likely less biased or less affected by non-equilibrium conditions. Remember also, that those distance measures are giving a picture, a statement of the genetic differences between the subpopulations for the particular loci that you've investigated at that point in time.

DR. JENNY OVENDEN: The calculation of  $Nm$  under equilibrium assumptions from  $F_{ST}$  or  $G_{ST}$  I found to be really of great interest to fisheries managers. But I think it's very important to emphasize that value actually estimates the number of migrants that successfully contribute genes to the next generation. There's a great amount of confusion when  $Nm$  values are presented that it's an absolute value of the numbers of individuals that are exchanged per generation. So we really need to be very careful when we talk about the estimation of  $Nm$ . That's all I've got

to say.

PROFESSOR STUART BARKER: That's a very valid point and I tend to assume that without stating it. It is essential to be clear that it is genetically effective gene flow that is relevant.

DR. MURRAY MACDONALD: Just to follow on from that in referring back to your original definitions of genetic structure and population structure. It's the population structure as defined by you which the fisheries managers are particularly interested in and the question here, particularly in relation to this estimate of  $Nm$ , is how close does your estimate of  $Nm$  reflect what's happening as far as the population structure is concerned? How much migration is occurring, is that a realistic estimate of the migration that is occurring in a contemporary population?

PROFESSOR STUART BARKER: As I said at the beginning, there is no necessary relationship at all between population structure and genetic structure. Secondly, as Jenny Ovenden has indicated,  $Nm$  is the genetically effective number of individuals moving between subpopulations, and again there may be little relationship to the actual number of individuals moving between subpopulations.

DR. NEIL MURRAY: Perhaps you could say it might be a minimum estimate.

PROFESSOR STUART BARKER: That's about as far as it goes. In other words, it does come back to the sort of feedback situation I was talking about. Genetics by itself is obviously not enough; ecology, life history, etc. by themselves are not enough.

DR. NEIL MURRAY: It may be a bit redundant now, but I wanted to add another word of caution about  $Nm$ . One thing that you haven't said, which again we all as population geneticists understand but fisheries managers get hung up about, is that the relationship of  $F_{ST}$  to  $Nm$  depends on your model of population structure and presumably the one you are talking about is the island model. There are of course stepping stone models of whatever you'd like all of which will have different relationship between  $F_{ST}$  and  $Nm$ .

I'd just like to comment that it really does come back to knowing a lot about the biology of the

organism before you interpret the genetics as well as the other way around. I wondered if you've had any experience trying to estimate gene flow on sort of neighbourhood models rather than island type models?

PROFESSOR STUART BARKER: Thank you, Neil. We've thought about it and really have not pursued it for those sorts of reasons. I don't think we know enough. If either of you know enough, it might be okay to go ahead. Apart from that, the whole concept of  $Nm$  really doesn't please me very much, I don't like it.

It would be much nicer if we could get separate estimates. There are a number of cases in this whole area of population genetics theory where you can get estimates of products of parameters easily, but you can't get the separate parameters. Estimating both  $N$  and/or  $m$  independently is extraordinarily difficult. Various approaches have been proposed for estimating migration rates. We have attempted it in *Drosophila*, where it's perhaps a little bit easier than it is out there in the sea, i.e. to estimate the actual proportion of migrants. But there was an enormous effort to do it and that was for very small scale gene flow over only a couple of hundred meters within a *Drosophila* population.

**SESSION 4**

**PANEL DISCUSSION**

**Convenor: Dr. Murray MacDonald**

**Marine Science Laboratories  
Queenscliff, Victoria**

## PANEL DISCUSSION

### A Structured Discussion - The requirements of those concerned with stock assessments and the extent to which these can be met by geneticists.\*

Convenor - Dr. Murray MacDonald

Marine Science Laboratories,  
Queenscliff, Victoria

DR. MURRAY MACDONALD: I've been given the somewhat dubious honour of leading a discussion on the usefulness of genetic information and techniques from the point of view of those who assess fish populations and fisheries; and those who make decisions about the management of those populations and the fisheries based on them. I would like to confine this discussion to the application of genetic techniques and genetic data to the assessment and the management of wild populations, that is populations which are not in some way cultured or enhanced by human activities. A discussion of cultured or enhanced populations will be conducted after relevant papers on those topics have been presented by tomorrow's speakers.

Today we've heard a number of speakers provide erudite descriptions of genetic techniques which can be used to obtain information about the genetic structure of populations of fish or invertebrates. In this case, I use Stuart Barker's definition of 'genetic structure' and I would like to emphasise the distinction between this term and 'population structure' as he also defined it. I would like this discussion to provide us with some insights into the relationship between information regarding genetic structure and what we know about population structure; in other words what does genetic data tell us about population structure? In my experience as a fisheries scientist, it is information about the structure of contemporary populations of exploited species

*\* Panel Discussion at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990.*



that fisheries managers are really interested in. Things like growth, mortality, recruitment and population sizes. What can genetic data and the genetic techniques we've heard about today tell us about those things?

As a basis for this discussion I've distributed a table of fish resource and fishery assessment topics around the room. This table was a first attempt by some people from the Australian Society for Fish Biology to summarise the application of genetic information or genetic techniques to various aspects of fish stock assessment or fisheries management. This table is a part of a paper published by Clive Keenan and myself in the proceedings of a tagging workshop which was conducted by the Australian Society for Fish Biology in association with the Bureau of Rural Resources here at the University of New South Wales two years ago. The views represented in the table are primarily those of the authors. What I'd like to do is go through each of the categories of fisheries research or assessment listed in the table and seek comment on the author's views regarding the applicability of genetic data or genetic methods to that particular aspect of stock assessment or fisheries assessment. I guess most of the discussion is going to focus on the category called 'stock identification', although if you have input in any of the other areas listed we'd love to hear it. What we'll do then, is leave the stock identification category until later and go through all the other categories to determine whether anybody here today has diverging views from those presented in the table. Let's start with the first category, initial biomass. Does anybody believe that there is an application for genetic techniques or genetic data in determining the initial biomass of wild populations which are going to be subject to exploitation? No, response? All right let's pass on to the next category, catch history. Here again the authors of the table have suggested that genetic data is of little assistance in understanding the catch history of a fishery, whether it be a new fishery, an established fishery or even a declining fishery. Any other suggestions on that matter?

DR. CRAIG MORITZ: What if you had historical bottlenecks? It seems to be something that can be detected with genetic analysis. One type of historical bottleneck is overharvesting of the fishery.

DR. MURRAY MACDONALD: The 'catch history' category in the table refers to estimating the total amount that's being harvested rather than assessing the impact of the various levels of harvesting on levels of genetic variability or other aspects of population structure.

DR. ROY HARDEN JONES: Do you mean that one might be able to get an indication, say of the original parental biomass?

DR. MURRAY MACDONALD: Yes, that was the question I was asking. Is it possible using genetic techniques or genetic data to get any estimate of the initial or original biomass of the adult stock?

DR. ROY HARDEN JONES: I think it's very easy to pass over these questions in silence and then have it handed down in history that such applications are never going to be examined again. Considering the arguments and the fears there are about the Southern Bluefin tuna, I think we should not take silences as necessarily implying that no-one thinks these applications are going to be of any use. I think more thought is required on this. It has never occurred to me. I just wonder, could we have a number of hands up, to whom has this occurred? Has anyone actually ever thought of using genetic evidence as an indicator of original parental biomass? If nobody's ever thought of it before perhaps we're passing it over a bit quickly. Perhaps we should come back to it in future discussions.

DR. MURRAY MACDONALD: That's fine. If somebody ~~thinks~~ there is a possibility that genetic ~~data~~ can be used in this way but they haven't really thought about it we would still like to hear about it. Certainly we should know about that. I mean what we want out of this discussion is some indication of areas or possible areas for future investigation or study. If somebody thinks there is a potential application then we need to know about it and we would like to discuss it in this forum.

DR. JENNY OVENDEN: The table you have given us shows plusses under the 'Initial Biomass' and 'Catch History' categories. Can you explain these plusses?

DR. MURRAY MACDONALD: The plusses and minuses at the top of the table were meant to indicate for each category whether any kind of estimate was possible, whether on the basis of genetic data or any other kind of information. It must be remembered that this table was generated during a workshop which was convened to look at various forms of tagging - including artificial tags and genetic marks - and their application to fisheries science. The plusses and

minuses indicated whether the workshop participants thought there was any way of assessing those parameters by any method. The applicability of genetic data/methods in particular is indicated further down the table.

PROFESSOR ROSS CROZIER: Could I comment on the estimation of initial biomass. Of course this is a bit tangential to what the geneticist can, it seems to me, actually estimate. He or she can, however, provide an estimate of long term effective population size, and if you multiply that by the average weight of individuals you can get some idea. I repeat that this would only be an estimate of long term effective biomass and that the variety of transitions that may have occurred during the history of the species to end up with the particular result that you get would not be known.

DR. MURRAY MACDONALD: Right, but you're suggesting then that it may be possible to back-calculate some initial estimate of biomass.

PROFESSOR ROSS CROZIER: Well you can but it's a very long term thing. You don't really know, for instance, if a particular estimate of effective population size results from there being a small population for a long time or a large population which has undergone a severe bottleneck and so forth. To some extent this can be tackled with some of the allelic diversity measures since, looking at the human population, we can be sure that there hasn't been a tiny population for a very long time cause we've got a big population now. But looking at the gene diversity values for humans and seeing that the long term effective size is supposed to be something like eight thousand people, we know that there have been big changes. If I might, I had one other comment and that is that there is a difficulty in using any method of stock identification, genetic or otherwise, in that as I understand the idea of stocks it's supposed to give you some idea, if you go and do something to one stock it will make a difference to another stock that's nearby. The trouble is that all these methods are based on the current or existing populations and assume that the individual animals are being washed around purely as particles, as it were, that are purely at the mercy of the elements in which they live. If there is any dynamic response of population movement to population changes, such as you might expect in some terrestrial populations, then unfortunately any identification of stock is not going to tell you what's going to happen if you take one stock away. The next one which shows very little migration may then move over into the vacant space. I expect that this would vary greatly in likelihood from one life history to the

next but it is an interesting and perhaps somewhat nihilistic point to raise.

DR. MURRAY MACDONALD: Sure, and in fact I was hoping that point was going to be raised a little bit later in the category relating to recruitment and movement.

DR. NEIL MURRAY: To raise Craig Moritz's point again, it seems to me that it is possible to look at historical collections of organisms or tissues and look at how much genetic diversity is in there. With DNA methods you can estimate levels of variability in museum collections and look at present stocks and see what fraction of that variability is still around and say something about the impact of catching on at least initial genetic diversity. Now, if you're not covering it under this item, I'm not sure where you're going to cover it.

DR. MURRAY MACDONALD: That's a relevant point. As I was saying before, if people have ideas how genetic techniques can directly or indirectly calculate initial biomass, whether it be numbers or total weight or whatever, then we need to know about it, we need to discuss it and flag areas for future investigation. The estimation of initial levels of genetic variability and assessment of the subsequent impact of fishing on these levels is perhaps not quite the same concept, but it is an important area of investigation in its own right.

DR. NEIL MURRAY: That's one which is certainly theoretically possible and museum collections are currently being used in that way to study conservation and phylogeny questions. I really don't see why this sort of approach wouldn't have fisheries application as well.

DR. MURRAY MACDONALD: Does anybody else have any other views on that issue?

DR. COLIN PURDOM: On the catch history it's possible that there might be management interest in whether the recovery of a depleted stock is from immigration or from regeneration of the remnants of the old population. I think there was some evidence on the recovery of the herring stocks of the North Sea. In some instances it did appear to be a regeneration and in others it was immigration. Of course, you're still left with the problem of whether the observed gene frequency array is an attribute of a population or is it the result of selection? One is back into that old uncertainty. There remains a management question. Whether you fish out entirely your stock and rely on the area being repopulated from outside or whether you only fish it half down

and let the resident population recover more quickly.

DR. MURRAY MACDONALD: Do we have any other views on either of those two issues? If there is no further comment we'll pass on to the next category in the table, estimates of population size. There are two sub-categories here; absolute population size, meaning total numbers of individuals and relative population size. The table indicates only a limited application of genetic data/methods to the estimation of relative population size. In making this assessment the authors had in mind primarily cases where subpopulations had been identified as being genetically different. Where there was some evidence of mixing between these subpopulations it would be possible to estimate the relative proportions that each breeding population or subpopulation contributed to stocks in a mixed area. Does anybody have any other ideas about any potential applications for estimating population size, either absolute or relative? The reason we suggested the application of genetic data/methods to this category was poor was because it seemed that once again we were up against the problem of determining just how realistic the genetic estimates of mixing are - particularly when genetic differences are not diagnostic. Under these circumstances, can genetic data provide a realistic estimate of different breeding populations contributing to the total stock in a given area? That's the kind of information which is of particular interest to fisheries managers.

DR. JENNY OVENDEN: With the mitochondrial genome that type of estimate is theoretically possible if you have two breeding populations that are characterized by genetically distinct mitochondrial genomes. When the two populations come together to forage or something in their non-breeding form then the mitochondrial genome of course provides an absolute marker of the degree of mixing between the two breeding populations. The only critical feature there is whether you can find absolute differences in the mitochondrial genome between breeding populations. Probably with John's examples where there were populations with different mitochondrial genomes that are very clear cut like he was showing us this morning, accurate estimates of mixing would be quite possible. If those oysters were fish and they were foraging together you could certainly tell that they were a mixed stock.

DR. MURRAY MACDONALD: Right, and I presume the same would apply to fixed differences for allozyme data as well. But where there are not fixed differences does anyone have any views about how realistic estimates of relative population size are?

DR. BOB WARD: In the United States there are very extensive programs looking at precisely this problem in salmon fisheries. Maximum likelihood techniques are used to work out the contribution of different stocks to mixed populations using allozyme data and you don't have to have fixed differences. This method provides estimates of the proportional contribution of different stocks to a mixed population, together with some standard error estimates. So it's possible with allele frequency measures.

DR. MURRAY MACDONALD: Do you know of any independent means of verifying the accuracy of these types of estimates?

DR. BOB WARD: I don't know. I'm not involved in that. There's a very big literature on this, on these sorts of estimates. They've obviously put millions of dollars into the working out of these techniques.

PROFESSOR STUART BARKER: I can't really add to that. I was simply going to make the same point, that from allele frequency differences it is possible to get estimates of degree of mixing.

DR. BARRY RICHARDSON: Another use of genetic data in this context is in identification of mixed species in cases where it is impossible or difficult to identify individuals using other methods. For instance, fish eggs could be identified by any of the genetic methods we've spoken about today or by immunological methods for surface proteins. If you are trying to identify the standing mass of eggs of a particular species in a particular year there are genetic methods that would allow you to take little round things and say which species they belonged to.

DR. MURRAY MACDONALD: But that is basically the same as what Jenny was talking about where you in fact do have fixed differences whether they be species level differences or subspecific differences.

DR. BARRY RICHARDSON: Certainly, but with the variety of techniques now available it's usually pretty easy to find a fixed difference of some kind at the species level which you can then use. Immunological methods are probably the best way for that.

PROFESSOR ROSS CROZIER: It's true that a direct estimate requires fixed differences but after all there have been very long, well established methods for using just frequency differences between populations to yield admixture estimates. So one could do that provided one does have sufficient confidence in the estimates of genetic structure in the contributing populations.

DR. MURRAY MACDONALD: That's the point I was making. Do we have any independent means of verifying those estimates?

PROFESSOR ROSS CROZIER: Well, you'd have to have that in the first place with fixed differences anyway. So I don't really see that as an extra difficulty.

DR. MURRAY MACDONALD: The point that I was referring to perhaps is that where you don't have fixed differences, how do you determine whether the individuals that you're looking at in a mixed stock were contributed from one population or from the other?

PROFESSOR ROSS CROZIER: For particular individuals it would be a difficulty. If you're trying to determine the overall contribution of a population to a mixed stock, then that estimate could be obtained with the usual associated confidence limits. But if you really want to know is this particular fish from one population or another then you do require fixed differences.

PROFESSOR STUART BARKER: Taking the question of population admixture estimates one step further. Where there are not fixed differences and one is trying to answer that question, then the major point is to use more than one locus. In other words, if you've got only one locus with different frequencies in the different populations and you analyse it to estimate mixing, you get only one estimate. But if you have a number of variable loci and repeat the estimate of mixing for each, then confidence in the estimate will increase, and increase dramatically if you're getting similar results for different loci. Unfortunately, in many cases where this has been done, the estimated admixture rates do differ and differ significantly between loci. So you might say well that doesn't answer the mixing question, but it does raise other questions as to why the loci are different. These inter-locus differences are almost certainly a reflection of selective differences among the loci, i.e. selective differences in the mixed population. So that provides additional evidence.

DR. MURRAY MACDONALD: Are there any other comments on this issue?

DR. ROY HARDEN JONES: There have been so many positive reactions to this discussion that it seems to be all signed and sealed. Tell me in what fishery has this (maximum likelihood) technique been actually applied and proportional mixture ratios given? A reference from somebody?

DR. BOB WARD: Salmon fisheries on the west coast of the United States.

DR. ROY HARDEN JONES: A little too general I think.

DR. JOHN AVISE: I think there are many examples of the sort you're questioning here. They involve hybrid zones that can be thought of as special cases of admixture, in which the parents not only occur in the same place but also interbreed. In many such hybrid studies it is possible by using multiple genes not only to assess the degree of admixture and the degree of hybridization that has occurred, but also in some cases to classify individuals as to whether they are F<sub>1</sub> hybrids, backcrosses, or later generation hybrids. Analysis of mitochondrial DNA may also allow a determination of who the mother was in the cross. Now you might claim that these are special cases involving hybrid zones, but I can think of other examples that are clearly analogous to admixture issues. For example, eels in Iceland which we have studied recently. We know of genetic differences between American and European forms of the eel, that can be used to quantify the proportional contribution of American and European genomes to the Icelandic eel population. It turns out that about 92% of the genes in Iceland are of European descent and 8% of American descent. We suspect that some of the eels are F<sub>1</sub> hybrids and others are later generation hybrids or backcrosses, and that the crosses have taken place in both directions with respect to sex. I could go on with other examples of this sort. There are quite a number of such situations, but they have more commonly arisen in the context of hybrid zone settings rather than mixed fisheries or mixed stock assessments. Nonetheless, the problems are conceptually quite similar I think.

DR. MURRAY MACDONALD: Fine. Anybody else with views?



DR. BOB WARD: Can I just mention a bit more about salmon? Salmonids are generally a good group for studies of mixing, because of their homing tendencies which tend to produce quite strong differentiation between populations - something which you perhaps won't get in many purely marine species. So it may be very difficult to do this sort of work with a marine species.

DR. CRAIG MORITZ: A student in my lab, Janet Norman, is working on marine turtles which, incidentally, are classified as fish under fisheries legislation in Queensland. Also John Avise has some students working on marine turtles and we have both found that when you sample nesting populations you find fixed, absolutely diagnostic differences in mitochondrial DNA. We are now going to feeding grounds and areas where the turtles are harvested to get absolute estimates of the proportion of turtles from different breeding areas which are mixing in those feeding and harvesting grounds. So the technology is there. You just have to find the right organism. Everyone should be working on turtles!

DR. MURRAY MACDONALD: So what criteria would you suggest we look at to determine whether an organism is suitable for this kind of application?

DR. ROY HARDEN JONES: Under the Fisheries Act a turtle in fact is a fish!

DR. MURRAY MACDONALD: I was thinking more in terms of allele distributions.

DR. NEIL MURRAY: I'd just like to also mention recent studies of humpback whales - another "pseudo fish" - where genetic data have been used to estimate the mixing of Pacific populations at least. I would also like to backtrack for a moment to the estimation of absolute population size and ask a question which arises from something Stuart mentioned in his talk. I don't think one need write off genetic data/methods altogether for measuring absolute population sizes. It's extremely hard to measure, but it is possible to get genetic estimates of local effective breeding population size, particularly if you have an isolated population and you look at the variation in allele frequencies from generation to generation. Probably there are other scenarios in which it can be done. I wonder if Stuart would like to comment on those sorts of estimates?

PROFESSOR STUART BARKER: One can get those estimates particularly if one has estimates of allele frequencies in an isolated population over time. The theory has been

developed. To my knowledge it hasn't been applied very much. I guess my feeling is one of pessimism in the sense that for the few cases where this approach has been applied the estimated effective population sizes turn out to be very small. In other words, I think what one is really measuring is the evolutionary effective size and not the present effective size. I think that's essentially the problem in relation to the question that's being asked here. It's not a minimum estimate of the population size as it exists at the present time, but a reasonable estimate of the effective size over some long period of time. Even though you're looking at changes in allele frequencies over a short term period and between recent generations, the estimate of minimum effective size is still influenced by the history of the population that you don't know.

DR. NEIL MURRAY: Is that what you found when you were talking about application?

PROFESSOR STUART BARKER: Yes. When you have a *Drosophila* population of almost certainly thousands of individuals existing in a natural state throughout most of the year, and you get an estimated effective size of two, you start to worry!

DR. BARRY RICHARDSON: Returning to estimates of mixing. A couple of years ago George Habib looked at the proportion of skipjack tuna from different breeding areas in the tropical Pacific which mixed in feeding areas off New Zealand. So attempts have been made to estimate relative population sizes in this part of the world.

DR. MURRAY MACDONALD: Any more views on the estimation of population size? Any more comments? The next category in the table is the estimation of mortality. M means natural mortality, F is fishing mortality and Z is total mortality. The feeling of the previous tagging workshop was that there is little application for genetic data or genetic techniques in estimating rates of mortality of any kind. Does anyone have any views on this, either pro or con?

PROFESSOR ROSS CROZIER: Surely relative mortalities could be obtained using genetic data. Estimating total mortality would be difficult as you mentioned, but if you are interested in seeing if there are any particular types within a cohort that were being removed or dying at a faster rate than others.

DR. MURRAY MACDONALD: You mean estimating selective mortality for a particular

genotype?

PROFESSOR ROSS CROZIER: Well, it could appear selective, either because of that same locus or because of some other locus that's affecting it, but you could get some idea about the mortality of different types of individuals within the population.

DR. MURRAY MACDONALD: So what you're suggesting is that it's possible to measure the impact of fishing itself on the genetic structure of the exploited population.

PROFESSOR ROSS CROZIER: Fishing or any other type of selective mortality.

DR. MURRAY MACDONALD: Any other views on that matter?

DR. ROY HARDEN JONES: This reminds me of a problem I've had with lots of these studies and that is looking to see whether any of the characters that form the raw material for this genetic work confer a selective advantage. Haemoglobin variants are an example that seems obvious, and I was thinking here of Sick's work in cod where perhaps at different temperatures, the different haemoglobin might have different capacities to carry out some respiratory function, such as the ability to secrete swim bladder gas quickly or something like that. Does anyone have any ideas on how the selection pressures might act on some of the characters that form the raw material of genetic studies?

DR. MURRAY MACDONALD: Anyone wish to comment on that? Well perhaps I might do so. In the context of fisheries the main way in which selective pressures might act is if animals with particular genotypes are in certain locations at a particular time so that there are differences between individuals in the vulnerability to fishing effort. In that way you end up with some kind of selective fishing mortality. Other ways I have some difficulty thinking of. Anyone else got any ideas?

DR. BARRY RICHARDSON: The surface proteins of certain bacteria mimic the antigenic attributes of the plasma protein transferrin, thereby avoiding detection by the host's immune system. As a consequence, the animal populations are under pressure to keep a wide range of alleles for transferrin to try and protect themselves. This leads to a race between selection in the

host for new alleles for transferrin and pressure on the bacteria to develop new alleles of their surface proteins to copy the new transferrin alleles of the particular host. So there's pressure on the host population to keep high levels of polymorphism for that locus because then the bacteria if it succeeds in one organism when it goes to the next one, gets the wrong transferrin type. But again this is simply one more form of selection on a protein.

DR. MURRAY MACDONALD: In this application you're talking about assessing how parasites or pathogens affect rates of natural mortality?

DR. BARRY RICHARDSON: Yes, there would be a selective advantage in maintaining a transferrin polymorphism.

PROFESSOR ROSS CROZIER: I think we could mix this particular approach with some of the previous things we've talked about and point out that when we have an admixture of populations in some common fishing area it's quite likely that it will also be desirable to estimate the admixture of different age cohorts there. If this was done it would not be surprising to find that there are fluctuations between cohorts in the relative abundance of fish from different populations of origin. Since they come from different areas and have therefore had different selective histories, they would probably have different responses to selection pressure on the fishing grounds where they're being sought. This would not involve any particular arguments about individual loci at all but would just be using them in their strictest sense as markers of what the ancestral populations' mortalities are in those new areas.

DR. MURRAY MACDONALD: Although you may have a problem there in determining whether, or not, the observed fluctuations in the relative proportions of different populations in a mixed area reflected mortality events in that area due to something like fishing or, whether it was due to natural fluctuations in cohort strength in the respective populations which are contributing to that mixed area.

PROFESSOR ROSS CROZIER: You're absolutely right. If there's continued recruitment at all stages of the life history then that of course would pose a severe problem. But once again we would return to some knowledge of the natural history of the organisms concerned.

DR. MURRAY MACDONALD: Any other comments? Moving on then from mortality to estimates of growth. Remember we're talking specifically here about wild populations rather than managed or cultured populations. Again, the feeling of the previous tagging workshop was that the application of genetic data or genetic methods to the estimation of growth rates in wild populations was minimal. Does anyone have any views about that?

DR. BARRY RICHARDSON: There have been reports in the literature that the average level of heterozygosity in electrophoretically detected loci is positively correlated with growth rates in some species. This could be used, I guess, in some ways to look at subsets of a population with high heterozygosity which may grow faster. That's about the only thing I can think of genetically.

DR. MURRAY MACDONALD: Right, there may be a correlation there, but I think the question was; is it possible to actually estimate growth rates using genetic data? Remembering always that we're talking about a fisheries assessment or fisheries management application here.

PROFESSOR STUART BARKER: Going back to the question of selective mortality from fishing, if one does have the situation where heterozygotes at some set of loci show faster growth rates surely it would be useful to have data on proportions of heterozygotes in different age classes or whatever it may be. Because, if in fact, fishing is preferentially taking the more heterozygous individuals because they are growing faster, are you in fact imposing selective mortality in that sense on the faster growing individuals. That's going to have an impact on the future of the fishery at least. So it's not so much a question of direct assessment of growth rate but recognizing the need for research perhaps to look at that possibility.

DR. MURRAY MACDONALD: To determine the impact of fishing again on the genetic structure of the exploited population?

PROFESSOR STUART BARKER: To determine what the impact would be on the quantitative genetics of the fish in that sense and therefore on the future of the fishery.

PROFESSOR ROSS CROZIER: It would seem sensible to suggest from Stuart's last point that one of the desirable goals of a well-managed fishery might be to make sure that you somehow

maximize the number of these heterozygotes in the exploited population. Therefore, if an allele is starting to become rare for some reason you would do something about it.

DR. JENNY OVENDEN: If I could redefine growth here and talk about population growth as opposed to individual fish growth, is that acceptable?

DR. MURRAY MACDONALD: Sure, although traditionally when population growth is referred to in a fisheries assessment or fisheries management context, you are either talking about growth in population numbers or growth in biomass.

DR. JENNY OVENDEN: Well, let's talk about growth in numbers. If you come across a population of fish whose mitochondrial genomes are very homogeneous, one of the explanations for that is that the population has recently passed through a founder effect or a bottleneck. If you've got large numbers of genetically homogeneous fish, then one has to assume that the population has undergone recent increase, probably a very large increase in numbers. So that might be a way of measuring the viability of the population, the ability of the population to respond to perturbations such as fishing pressure or parasite infestation, that kind of thing.

DR. MURRAY MACDONALD: It seems to me that what you've done here is to equate estimates of population growth with estimates of population size as discussed earlier, but you are not necessarily measuring growth of individuals or biomass as such.

DR. JENNY OVENDEN: No, my comments are certainly not applicable to individual growth, or population growth.

DR. MURRAY MACDONALD: Okay, any other comments? Well that brings us to the last category in the table of recruitment and movement and, on reflection, I think perhaps it might be more beneficial if we lumped that last category in with the discussion of stock identification because they are inextricably linked. As you can see, the conclusion of the previous tagging workshop was that it's in these two areas that genetic methods and genetic data are likely to have the most direct application - at least as they are being used so far. I guess the question is just what does the application of the various genetic methods and the data that's obtained from them tell us about the breeding structure of populations, in the sense that Stuart Barker defined earlier on.

Furthermore, with respect to movement and recruitment to certain areas, how does genetic data enable fisheries managers to decide whether or not they should control fishing effort and/or catch so that it does not result in undue impacts on either fish stocks in the area that's being fished or recruitment to areas other than the area in question as a result of alteration to the migration patterns or abundance of exploited populations. I think perhaps as a primer for this discussion I'll ask Neil Murray to give us a short description of some of his views on the application of gene flow concepts in this area and we'll take it from there.

DR. NEIL MURRAY: This table contains some data I've been working up lately. I want to follow on from one of the things which Stuart raised in his talk about measuring gene flow from  $F_{ST}$  and analogous measures of genetic patchiness. I think there is something to be got out of this but I don't think it's been done very well in the literature. I also think that there are some very big questions about what it all means. I've just extracted from the literature directly some published  $F_{ST}$  values, they're for various species, some which my own students have produced. I have turned these  $F_{ST}$  values into gene flow measures in the simplest way possible, assuming an island model of population structure which is what everybody in fisheries since Allendorf and Phelps' '81 paper has done. You can see from the table that generally gene flow estimates for freshwater or river-based species are lower than those for marine species. Of course most people expect that, but it's the quantitative aspects which are interesting. As Stuart said, it would be nice to know what these estimates mean biologically. What it is saying is that if you assume that population structure then, for coho salmon about three individuals move between populations which for rivers in their sample every generation. That's the number of individuals into your average size population every generation. That's a quantitative measure which some people have started to use as a criterion, at least in their minds, to define what a stock is. I think there are big problems with that. Over here is an independent estimate of  $Nm$ . It's independent only in the sense that it uses a tiny fraction of that data, in fact a fraction which contributes almost nothing to the  $F_{ST}$  data, which is one of Slatkin's gene flow estimates based on his so called private alleles. If anyone's heard of private alleles then this is what it's all about. That gives another measure. What surprises me is that the numbers are extraordinarily comparable, at least for where the data is reasonable. Those two bracketed cases there's only a couple of private alleles so you don't expect an estimate that means very much. Even though that one is actually right on that one. Something about the genetic structure of these species is being estimated in these two methods and it's telling us much the same sort of thing. My real difficulty is that I don't

know what it is that's being measured. I think that's probably shared by other people. Because if we go back to some of these numbers we have, according to the blacklip abalone data, say, eleven individuals moving between populations per generation. That's a very large gene flow measure. Marine invertebrates and vertebrates tend to produce the largest in the literature. If you compare terrestrial mammals and frogs and things, these are very large numbers. And yet what Lindsay has found is that you do get, what another definition of a stock would be distinct allele frequency differences, distinct stocks between places only a couple of kilometres apart. That's especially so where there's an isolation to abalone movement involved. Any sort of translation of these to stock recruitment relationships or average movement measures, really founders when you come back to the basic breeding biology. I think what it really says is that we've got to know the breeding biology backwards and how the local hydrology affects your population. If I could just translate that into another question that people will want to ask. This I think is almost what Professor Harden-Jones started with this morning, with the idea of stocks as being independent of immigration. What a manager would often like to know is what fraction of the breeding population are immigrants. That's not an  $Nm$  measure. That's an  $m$  measure. That's much harder to get as Stuart mentioned. These are four of those data from the last table and an  $N_e$  effective size estimated roughly by the authors in this case and by me and it's almost guesswork in these two cases for effective size. You see you are looking at  $m$  values that are actually very tiny. The biggest one is about 1% in the coho salmon. If you are looking at abalone, you are saying that about four in a thousand is your immigration rate. That probably relates more to the local stock recruitment question than the  $Nm$  value does in its simplest form. That is why I'm afraid that this  $N_e$  estimation might have to be taken more seriously before we start to term gene flow measures. However we derive them, however consistent they seem to be for some of these things before we turn them into real management parameters or variables more likely than parameters. Thank you.

DR. MURRAY MACDONALD: Anyone wish to comment on that?

PROFESSOR STUART BARKER: Not really saying anymore but perhaps just emphasizing a point which arises from Neil's data there. The thing to remember here is that the term  $Nm$  is in fact the product of two variables,  $N_e$  and  $m$ . So if in the case of sockeye salmon you assume that 3,500 is a reasonable estimate of effective population size, then you've got an  $m$  of  $1.6 \times 10^{-3}$ . Suppose, however, given the large standard errors on those estimated  $N_e$ 's, that popu-



lation size really is 350 and not 3,500, then  $m$  is an order of magnitude larger; or if the real population size is 35,000, then  $m$  is an order of magnitude smaller. So I agree with Neil that what's really important here is to obtain separate estimates of effective population size and of migration rate as a proportion of population size.

PROFESSOR ROSS CROZIER: Something that I said at the beginning of this discussion is relevant here, and that is that we're studying a current situation when we make these estimates of migration. The situation is clouded by a slightly controversial concept in evolutionary population genetics relating to coadaptation. This concept implies, for example, hybrids between different populations in many, but certainly not all, species will show some reduction in fitness. This trend will usually, of course, be mediated by the particular environment that individuals encounter, such as if they are facing competition between other individuals. So when we carry out our  $Nm$  estimates here, we are of course measuring genetically effective migrants as Jenny Ovenden stressed some time back. The number of migrants who don't make it may of course be much larger. The minute you remove one stock from an area you suddenly find the next stock actually draining into the hole left at a much greater rate than expected because previously those individuals who were immigrants were in a much more difficult competitive situation and were not actually leaving descendants. All of this suggests that in some cases - and we don't know quite which ones - other factors such as the fitnesses of hybrids could in fact greatly influence the apparent migration rates that are being estimated from genetic data, and somehow these confounding factors have to be accounted for if we are to provide realistic estimates of migration.

DR. MURRAY MACDONALD: When you refer to migrants that 'don't make it' I presume you're including both those that don't make it to where they're going and those that make it but don't contribute in a reproductive sense.

PROFESSOR ROSS CROZIER: Yes, but also those migrants that do contribute in a reproductive sense, but their progeny for some reason or other don't do well enough to boost the apparent migration rates. In other words, genes will have to successfully pass through the transitional stage of  $F_1$  hybrids before they are deemed to have made it from one population into another.

DR. CRAIG MORITZ: There is a substantial problem here. That is that we're mixing two very different styles of population genetics, historical and current, as Ross was saying. Now the  $Nm$  estimate, as Neil knows, is historical. People have looked at salamanders where they had marked a population and they know damn well that salamanders don't move from one point to another. Next you'll have  $Nm$  estimates of 1 or 2 which the authors of that paper, Larson *et al*, suggested it was due to migration of animals post-Pleistocene. He's still looking at effects from the Pleistocene. I think in all this discussion, we need to be very clear on whether we're looking at short term effects, which I guess is what fisheries managers are primarily interested in. They want to know if they overfish the current generation of a stock whether they are going to get recruitment the next year. On the other hand, a conservation geneticist might be interested in conserving overall patterns of genetic diversity, in which case - as John Avise mentioned earlier - studies of intraspecific phylogeography and long term  $Nm$  estimates should be your guide. I wonder if perhaps we're looking at the wrong thing for fisheries management in looking at  $Nm$ . What we should be doing is looking at things more related to short term gene pool dynamics, like Hardy-Weinberg Wahlund effects and things of that nature. I wonder if maybe Mike Johnson would like to comment on this?

DR. MIKE JOHNSON: Thanks, Craig. I really don't have much to add because I basically agree with the problem that you've raised and really don't see a solution to it in terms of using  $Nm$ .  $Nm$  is something which is important on an evolutionary time scale and not on a generational time scale, so from the point of view of generation to generation management of fish stocks, it's not really telling us much. The only way that I can see that you can move clearly to a short term situation is where you're fortunate enough to have genetically distinct groups so that you can estimate different levels of admixture for example. There is also a problem here, if you think about it in the long run, that if you're getting admixture of breeding populations rather than admixtures feeding in a common area, then of course that means gene flow which means that you then eliminate over time the genetic differences between the original populations. So the possibility of getting population - specific markers that you could use is going to be reduced if there are breeding connections between them. So no, I don't have an answer Craig.

DR. BOB WARD: We have focused on problems associated with the use of  $Nm$  but there are also problems with  $F_{ST}$  and particularly the standard errors associated with  $F_{ST}$ . For example, you can sample a set of populations where there's no statistically significant heterogeneity

among the populations in allele frequencies but where you still end up with a positive  $F_{ST}$  value even though the true  $F_{ST}$  value should be closer to 0. What I'm saying really is that the  $F_{ST}$  values have high standard errors associated with them, and for those which really are not significantly different from 0 the  $Nm$  number calculated could be anything from 35 to infinity. So I think that's another problem that we've got to consider.

DR. NEIL MURRAY: Can I just say something and that's actually a point which should be followed up. Stuart mentioned this morning that there were different ways of estimating  $F_{ST}$  and Weir and Cockerham is different in that it is an estimate and therefore your estimates can come out slightly negative if it's close to zero. So in that sense it's less biased than all the others. That's certainly a point though and I think that the all published fisheries  $F_{ST}$ 's are not Weir and Cockerham estimates. The other point that I could come back to is that, yes, clearly there's just an illustration of the problem Craig raised is that it is a historical thing and for the people who aren't population geneticists, thinking about it what that's really saying is that over historical, long term historical times, you've got maybe eleven individuals or twenty individuals moving between your average populations. If you've got a species which has an historically very high effective size for a long time and you suddenly chop it down by fishing, you're changing things utterly. It really does mean you've got to come back to that  $m$  measure to make any sense out of what the fishery is doing. That is really what the managers probably ought to be thinking about.

DR. COLIN PURDOM: This is almost sort of turning the problem on its head. With respect to Atlantic salmon the precision of homing seems too good. The behavioural studies that are being done in parts of England and the tagging studies in parts of Ireland suggest that it really is 100%. That there is no gene flow. The question then remains if there's no gene flow, why isn't there greater genetic differentiation of stocks? Could it be that the present populations of salmon are sparse because they've been fairly heavily fished and perhaps some of their habitat has been destroyed, but that 100-200 years ago when things were quite different the populations were so large that migration was more or less forced upon some of them. Therefore we have potentially A) a very potent effect of fishing on genetic structure of populations and B) it makes rather a nonsense of some of these  $Nm$  values.

DR. MURRAY MACDONALD: Any other points of view or comments? I'd like to ask Neil

to comment, if he would, on the circumstances under which he thinks these estimates of gene flow are potentially most useful for fisheries managers and over what time period?

DR. NEIL MURRAY: I think they're most useful when you know all the biological answers to begin with. As was said this morning, you have to know something about the breeding population structure of your species before you can go from an  $F_{ST}$  to an  $Nm$ . For example, barramundi is strung out along the northern coast in what is a classic one dimensional stepping stone model of population structure. However, nobody uses this model to analyse the data. If they did, I think the current estimate of migration in barramundi - 2.62 per generation - would probably turn into a value of about 8 or 9. So you really do need to know the biology of the population before you can get a reasonable  $Nm$ . There are some species that probably exhibit island models of population structure, but they aren't necessarily common. As Barry was saying this morning, the isolation by distance model is certainly one which should be looked at much more commonly. Under this model, though, it's likely that the number of migrating individuals are being underestimated by an  $Nm$  which is calculated the way everybody calculates  $Nm$ .

DR. MURRAY MACDONALD: I gather that in broad terms you're suggesting that gene flow analysis might be applicable, if it is at all, to situations where you don't have a clearly defined genetic population structure. Where there are quite obvious discontinuities and they persist over time, the management implications arising from that are relatively straightforward. However, I gather you're talking about situations where the genetic divergence between populations is less obvious and where there's some indication of either mixing or movement between areas. Under those circumstances what would a fisheries manager do with this kind of information if it were obtained from a genetic study or an analysis using say the isolation by distance model? What are the implications for the fisheries manager in the short term, if any? Or is this kind of analysis merely providing some insight into the overall genetic status of the exploited stock and providing some more general guidelines on what we need to do to maintain the populations genetically to avoid either extinction or some other catastrophe?

DR. NEIL MURRAY: Or loss of variation. I think that depends what the objectives for management of the stock are. It's clear from the data on coho salmon, for example, that you can have a gene flow estimate of 148 between regions and yet it's known that fish from those regions are adaptively different. In spite of some of these apparently high levels of gene flow, there can

be very big differences between populations. I think probably the only guideline a fisheries manager could operate on is that in anything that has a low  $Nm$  value was reasonably isolated historically, and if it was, it probably still is now. However, it would also be wise to expect local adaptation. On the other hand, populations which exhibit a high  $Nm$  value may have high current rates of gene flow, but the value may also reflect a historically high gene flow that's been chopped back by subsequent events including fishing, and a high  $Nm$  value doesn't preclude local adaptation anyway. So, you've got an asymmetrical result. If you get a low  $Nm$  value I think it means something fairly straightforward to a resource manager. It indicates discrete stocks and they ought to be treated as such until evidence to the contrary is obtained. But a high  $Nm$  value indicates we've got to go and do some more biological and ecological studies of the populations.

DR. MURRAY MACDONALD: Would you care to place some sort of figure on what you mean by high and low  $Nm$  values here?

DR. NEIL MURRAY: No. Well, looking at the values in the table there, on the numbers there, anything less than 10 would be suspicious. But there are problems you run into. As Bob was saying, when  $F_{ST}$ 's are low you run into essentially an asymptote on the biggest possible  $Nm$  you can measure at about 200-300. You're in a window of data where you're very subject to error when you get into the larger numbers.

DR. JENNY OVENDEN: I think some input from the fisheries biologist or fisheries manager would be really useful here because the presence of gene flow between populations doesn't necessarily guarantee that there is a large amount of mixing which will prevent a population from going to extinction. There's lots of other characteristics of a population that could cause it to crash. I'd love to know from a fisheries manager how our estimates of gene flow are used and would be treated.

DR. MURRAY MACDONALD: I suspect that they're looking for guidance on that from population geneticists, but perhaps Bob Kearney might like to provide a few comments on that.

DR. BOB KEARNEY: There's a very simple answer - they're not used. You provide fisheries managers with a format where these estimates can be used and they will consider it.

DR. MURRAY MACDONALD: I think here's your answer. We've got to try and come up with some way of determining whether or not this information is usable to the fisheries manager, particularly in the short term time frame. If not, as Bob says, it's likely that it's going to be ignored.

PROFESSOR STUART BARKER: Just another note of caution that I did make earlier. When you're talking about  $F_{ST}$  and any derivatives from that such as  $Nm$ , remember that it is a single global estimate for the total population or the total species under investigation. Using it by itself may well be very deceptive. It's a first step which can give some preliminary information, but most importantly I think you've got to look beyond that to the partitioning of the  $F_{ST}$  value. How much of the total variation is between particular regions or subpopulations or whatever they may be. That will give some additional guidance perhaps depending on the nature of the distribution of the variation among the subpopulations. The other thing is to try and get an overall picture of the population structure through genetic distance and other procedures. Putting them together is going to give you much more information at least from which to make inferences, if not something absolutely specific.

DR. MURRAY MACDONALD: I'd like to move the discussion in a slightly different direction now and ask a question of John Avise. We've been talking about attempting to estimate population structure and movement using mainly basically allozyme electrophoretic data. I was wondering whether John has any ideas on how DNA information might be used to get such estimates and whether he thinks this approach is of any value to fisheries management.

DR. JOHN AVISE: Generally, I think you have the same array of possibilities for DNA data that we've just discussed for the allozymes. Namely you can try to get estimates of  $Nm$  either from  $F_{ST}$ 's or from the private allele approach. There are two additional methods that might apply to the kind of haplotype data provided by mitochondrial DNA that we haven't yet discussed. One is an approach recently introduced by Monty Slatkin, who takes the phylogenies of alleles or haplotypes and determines in what geographic locales the different branches of this phylogeny occur. From the geographic distributions of haplotypes, he estimates the minimum number of gene flow events required to account for the super imposition of this phylogeny over the geographic locales, and converts that to another estimate of  $Nm$  which is an effective mi-

gration rate between the geographic locales. All of these three methods - the  $F_{ST}$ , the private allele, and Slatkin's phylogenetic methodology - are of somewhat limited use because you are dealing with a product of unknown parameters, and furthermore these are typically interpreted as reflecting equilibrium conditions. It's not clear to me how they're going to be of use in fisheries management, except that they give a general guideline as to whether a species is a "high-gene-flow" species or a "low-gene-flow" species.

We really also need to take into account the historical context of gene flow. One of the things I really like about the discussion today is that many people are noting the distinction between contemporary and historical population structure, and the fact that the kind of estimates of dispersal that one might get from mark-recapture studies or other direct methods may sometimes differ from the  $Nm$  values estimated from genetic data. I think that's because on the one hand the  $Nm$ 's represent the long term product of effective population size and migration rate, whereas under present dispersal conditions you're looking at the current-day conditions of movement of individuals. These can differ. Contemporary population structures can often differ dramatically from historical structures.

I want to suggest one other method that a former student of mine (Joe Neigel) is working on currently that represents a conceptually distinct approach to estimating migration. Rather than estimate  $Nm$ , he takes the historical information content that we believe is present in the mitochondrial molecule (as I discussed this morning), and utilizes that additional information to try to reconstruct per generation dispersal distances. The basic approach is to utilize a random walk model that envisions individuals moving a certain number of yards or metres per generation. If dispersal is independent of the generation, you can get an expectation through time of the accumulated distance to which individuals should have moved under this kind of a migration model. Thus, the model is of random dispersal under a diffusion-like process. Joe uses a mitochondrial phylogeny to estimate the ages of lineages, and in conjunction with their geographic distributions, works back from that to estimate single-generation dispersal distances.

We've only applied this approach so far to one data set. This involved *Peromyscus maniculatus*, which is a small mouse in North America; the mitochondrial phylogeny spans something like two million years, so we're talking about a large number of generations and a continent-wide distribution of mice. But when the random walk model is applied, the estimate of mean single-

generation dispersal distances is about two hundred metres per generation. This is a physical distance that came out of this particular application of the model. The result may be entirely fortuitous (this is the first time we tried the method), but this estimate does happen to agree almost perfectly with direct mark-recapture results which indicate that mice in their lifetime move a couple hundred metres on average. As I said, this may be a purely fortuitous outcome because we're dealing with big numbers and taking square roots to estimate dispersal. Nevertheless, I think it is an interesting conceptual development because we're trying to get away from the idea of merely estimating a composite function  $Nm$ , and are trying to generate something more concrete that can be evaluated against the kind of direct mark-recapture data that are being generated today. As I said before, this approach capitalizes on the phylogenetic information content of mitochondrial DNA. I can't think of a way to do that easily with allozyme information, since you don't know the phylogenetic orders or histories of particular alleles. In principle one might hope to get the necessary kind of information from nuclear DNA analyses as well. But that raises many other questions about how to assay particular segments of nuclear DNA such that you can view them from the kind of phylogenetic perspective that we've developed for mitochondria. I think one would likely be plagued by problems with recombination within the region under concern, such that a straightforward linear history of haplotypes may seldom be recorded in particular segments of the nuclear genome, the way they are in mitochondrial DNA.

DR. MURRAY MACDONALD: So you're suggesting that rather than look at rates of migration from point A to point B or area A to area B you try to estimate the neighbourhood size of the population, as the term was coined earlier.

DR. JOHN AVISE: Actually, we're trying to translate all of these manipulations into a concrete measure of dispersal distance expressed in absolute distance moved per generation. We are still a long way from claiming that this approach is going to be of pragmatic utility to the fisheries manager. I think it's a conceptually interesting approach - an attempt to get away from the unsatisfactory or incomplete perspective of merely addressing composite  $Nm$ .

I'd like to mention one other aspect about the utility of genetic data in a management context that follows what many other people have said here. It seems to me that one of the most straightforward applications of genetic information is its use in identifying some of the major



sources of gene pool diversity within species, particularly if there are fundamental phylogenetic subdivisions of the sort that we found for many species in the United States. These can be fairly easy to identify. They're often quite unambiguous. They are typically reflected not only in mitochondrial discontinuities but also in allozyme discontinuities, and so they're sometimes straightforward to identify. These presumptive phylogenetic partitions are supported by concordant patterns of independent genes or unlinked loci in the genome, and surprisingly they sometimes do not coincide with traditionally-recognized subspecies boundaries which have been the usual basis upon which managers have drawn conclusions about evolutionary genetic stocks.

All too often, particularly at the within-species level, current taxonomies were erected in the 1800's and early 1900's, sometimes by naturalists travelling through an area who might have seen a gopher or fish that in terms of phenotype impressed them as somewhat different, so they put a subspecies name on it. In my experience, once these subspecies names are in the literature, they inevitably assume an aura of significance that may go far beyond any empirical basis for their recognition. And yet entire management programs, including those for endangered species protection, are sometimes built around these subspecies names. We have worked with a number of taxa where endangered status was conferred simply because the population had been described as a subspecies or separate taxonomic unit sometime back in the late 1800's on the basis of very questionable and preliminary morphological evidence (that may not even hold up upon closer examination). So here's one fairly straightforward way that genetics can contribute to management of populations - in identifying the major sources of gene pool diversity within a species. There are often long term phylogenetic separations within species that lend themselves to ready identification. In terms of species that don't have these major gene pools subdivisions, you often get back into this area of isolation by distance. Basically, populations of many species are likely to be structured at a variety of levels. There may be family units staying together over microevolutionary spatial scales and microevolutionary time scales. Then there are likely to be extended families that have broader geographic distributions and are slightly older, and so on. So, under an isolation by distance model, you anticipate a whole array of hierarchical population levels that might be evidenced by significant genetic structuring. But the particular geographic structure for any one gene will tend to be idiosyncratic. There may not be a concordant buildup of genes identifying populations in particular regions. So such a finding in itself might be taken as evidence for isolation by distance, rather than long-term

vicariant separations of populations.

How much concordance exists across the population units identified by independent genes that have trickled through the organismal pedigree? Under isolation by distance, you would probably expect to see little concordance across loci in the particular population units identified, even if each gene individually reveals a significant  $F_{ST}$  and a low  $Nm$ . So I think we might be able to classify species into these kinds of categories. Is a particular species characterized by isolation by distance throughout its range, so that it has low  $Nm$  and high  $F_{ST}$ , or is it a species that is characterized by some long term evolutionary separations where we expect to see a greater development of genealogical concordance? The latter would be the kind of evolutionary stocks that can be most securely identified by molecular methods, and that fisheries managers ought to give special consideration in development of management strategies.

DR. MURRAY MACDONALD: Perhaps, I could ask Barry to make a comment about the application of the concept of single-generation dispersal distances to allozyme data, particularly the kind of analyses which result in estimates of neighbourhood size. Perhaps you'd care to comment about the relevance of that to short term decisions regarding fisheries management.

DR. BARRY RICHARDSON: I fully support what John has just said. It's the concordance or lack of concordance in distribution patterns that helps you determine which of the particular models of population structure is likely to be relevant. The next question I think is, does it really matter whether you're looking at an isolation by distance model or a discrete stock model? I think it does, because if you over-fish a discrete stock you then have difficulties with re-establishing that stock. There will be problems with recruitment from other isolated stocks into the area you've overfished. Whereas, if you have an isolation by distance model, at least in the simplest case, if you overfish a section of the species' range the chances of recruitment from other parts is likely to be higher. It may be more complex in reality, but the first assumption a fisheries manager should make is that if he's got isolated groups then he's got a serious problem in replacing them if he overfishes one or more. If he's got an isolation by distance model he's got perhaps a little more safety up his sleeve in determining optimum harvesting levels.

DR. MURRAY MACDONALD: I guess the question would be over what distance does he have safety?

DR. BARRY RICHARDSON: That depends on the size of the neighbourhood. I would think if you fished out part of a species' distribution equivalent to two or more neighbourhood sizes then you would have difficulty re-establishing stocks in that area from outside recruitment, much like you would with an isolation model. The problem is how do you measure the size of a genetic neighbourhood? The sad truth is very little work has been done on this concept since Dobzhansky and Wright in the late thirties and early forties. It is calculated using the variance of the distance moved between the place where individuals were born and the place where they breed. It's very difficult to see easy genetic methods of doing that at the moment. Almost nobody's making any concentrated effort to try and do it. The normal way to do it is from mark-recapture experiments to attempt to estimate rates of movement. But these estimates are grossly affected by the leptokurtosis of movement distances. In other words the average distance moved is not the real distance. Many specimens move very little distance and others move very, very great distances. That affects this variance quite radically and you have to take that into account in measuring it and that makes it quite difficult. At this stage I can't suggest how you might measure genetically the size of a neighbourhood. The methods we have will tell you neighbourhoods exist, but beyond that I'm not really quite sure which way you can go. Perhaps other people may have some suggestions.

DR. MURRAY MACDONALD: Any other comments?

DR. CRAIG MORITZ: Coming back to the usefulness of DNA and taking up the gauntlet Roy threw down early this morning, if you do have a discrete stock situation which depends on breeding biology and so forth then it seems to me that in many cases the amount of discussion we've had today about sampling effort and the cost effectiveness of the study are going to be determined by the type of genetic marker used. There has been a lot of talk about the need for fixed genetic markers. Where you don't have fixed differences the amount of sampling needed to detect and statistically demonstrate linkage disequilibrium or Wahlund effects is enormous. So, it seems to me, more effort needs to be put into detecting fixed differences in your pilot study, and here I echo Barry's comments about the need for pilot studies. If your pilot study doesn't reveal two or three fixed differences with allozymes then rather than put a lot of effort into sampling allozymes that's the time to shift to DNA markers. The reason DNA markers may be useful is that some of them at least have very different genomic dynamics to those of underlying

allozyme variation. For examples, effective population size of mitochondrial DNA is about a quarter that of nuclear genes because only females transmit it and females are effectively haploid. That's from the equal sex ratio. Repeated genes may be subject to concerted evolutionary pressure in which case the total distribution will be such that there is more between demes than within demes. For these types of reasons sampling a lot of DNA markers will give you a lot more statistical power in any sort of stock analysis. So I think rather than put a lot of effort into doing enormous sampling, where you don't have the right type of genetic markers, we should be putting more effort into finding the most appropriate types of genetic markers. In the long run I think it will be cheaper in boat time, lab time, all of those things.

DR. ROY HARDEN JONES: I've got a fundamental worry that perhaps some people could sort out for me once and for all. It's a paper entitled "Phenotypic variation in electromorphs previously considered to be genetic markers in *Microtus ochrogaster*"<sup>1</sup>. It's a paper in *Oecologia* which all geneticists will probably know but they don't tell biologists about it. The paper indicates that electromorphs in fact vary according to the seasons in some animals, which brings the whole concept of fixed genetic markers into question. Tell me please is this paper correct or wrong?

DR. MURRAY MACDONALD: Well, can someone perhaps comment on the possibilities with regard to that particular case.

DR. ROY HARDEN JONES: I'd be quite happy to discuss the matter after the workshop. But it did generate some correspondence in *Nature*.

DR. BARRY RICHARDSON: Can I say that the proteins used in that study were transferrin and leucine amido-peptidases both of which are renowned for exhibiting environmental effects and I'm not surprised at the result. They're not the proteins I would routinely want to use in the kind of stock identification work we are talking about.

DR. ROY HARDEN JONES: That's the sort of attack I would make.

DR. BARRY RICHARDSON: Well, I can't help it. If you want me to write on the board I can tell you which population markers you shouldn't use. Any esterase would be right at the top of

the list. Transferrin also has to be dealt with carefully because, if you remember, the electrophoretic data I displayed this morning showing environmental effects due to the addition of sialic acids. There are a subset of types of enzyme that I would not normally want to put any faith in if I was using them for the kind of work we're interested in fisheries. Esterases are at the top of the list but any functionally non-specific enzyme would be on the list.

DR. COLIN PURDOM: Can I add a worry of mine which is perhaps a little bit nearer home? Having seen the attendance list for this workshop I know the person I'm going to refer to isn't here. It's a New Zealand scientist called Gauldie, who has poured a great deal of cold water on the use of genetic markers to identify stocks, largely on the basis of the selectionist/neutralist controversy. Can anyone comment on that?

DR. MURRAY MACDONALD: We do have a person in the audience who is, or was, a colleague of Bob Gauldie's. Perhaps Peter Smith would like to make some comment on this.

DR. PETER SMITH: Bob Gauldie has left New Zealand and is working at the University of Hawaii. I think most of his concern arose from the problem of interpreting transferrin data. He looked very carefully at transferrin allele frequency differences, both within and between populations of cod. Some of the sample sizes in that data set were very small, and I believe most of the problems that he alluded to were based around small sample sizes. Barry Richardson has already covered the need for adequate sample sizes in his talk this morning.

DR. MURRAY MACDONALD: Are there any more comments on that or anything else? Alright then, bearing in mind we're already well over time perhaps we'll wrap this up. I would just like to close by giving my impression of where the discussion led. It seems to me that in general where fixed genetic differences can be detected between populations, assessments of breeding structure and mixing/movement - and the implications of these findings for short term fisheries management - are relatively straightforward. However, in situations where no fixed differences are apparent and/or there is evidence of some gene flow, assessments of contemporary population breeding structure and or rates of mixing/movement are still subject to potentially large errors because the relationship between these attributes and observed distributions of genetic variation is still poorly defined. However, this workshop has identified another potential application in fisheries management for genetic data - particularly the mitochondrial

DNA information John Avise was telling us about. That is in providing longer term guidance on the management of exploited populations from the point of view of conserving genetic diversity. I'll leave it at that and we can pick up this issue of conservation management in tomorrow's discussion. Thanks very much for your participation.

1: The paper referred to is McGovern, M. and Tracy, C.R., 1981. Phenotypic variation in electromorphs previously considered to be genetic markers in *Microtus ochrogaster*. *Oecologia* 51: 276-280. Also see McGovern, M. and Tracy, C.R., 1985. Physiological plasticity in electromorphs of blood proteins in free-ranging *Microtus ochrogaster*. *Ecology* 66: 396-403. The findings of McGovern and Tracy on transferrin and leucine aminopeptidase variation were independently investigated by Mihok and Ewing. They showed that phenotypes for the same proteins in *M. pennsylvanicus* remained the same through major seasonal changes in reproductive activity and environmental conditions, and suggested experimental errors may be the cause of McGovern and Tracy's observations (Mihok, S. and Ewing, D., 1983. Reliability of transferrin and leucine aminopeptidase phenotyping in wild meadow voles (*Microtus pennsylvanicus*). *Biochem. Genet.* 21: 969-983).

**SESSION 5**

**GENE POOL MANAGEMENT**

**Chairman: Dr. John MacIntyre**

**University of New South Wales  
Kensington, N.S.W.**

## CHAPTER 4

### Quantitative Genetics of Semi-Managed Populations\*

Professor John W. James

Department of Wool and Animal Science,  
The University of New South Wales,  
Sydney, N.S.W.

The approach which I will take is essentially to look at fish breeding from the point of view of a sheep breeder. I will deal with the genetics of quantitative characters, which are usually controlled by many gene loci, as distinct from previous discussions of essentially single locus or even single site genetics.

To deal with quantitative characters, our first task is to decide just which characters we are interested in. If we are simply looking for genetic differences between populations, any trait will do so long as it shows genetic variation, but to make genetic improvement, we must choose the quantitative characters to deal with rather carefully.

#### OBJECTIVES

The first thing to do is to decide on the objectives of the breeding program. What are the important traits whose inheritance requires study? These are the traits which determine profitability, assuming profitability is what we want to improve. Profitability depends on amount of product, quality of product, and cost of production. We can define our objective as

*\* Paper presented at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990.*



returns - costs and check which traits influence returns, and which affect cost of production.

It seems to me that in fish production the important traits would be growth rate, feed efficiency, survival, and perhaps fecundity. Because of the existing high fecundity in fish its improvement is less important than in sheep, which mostly produce one or two progeny per year, so that the cost of producing progeny is very high. The maintenance feed requirement will also be relatively higher in sheep than in fish. As a result, I believe that since the fraction of costs devoted to producing an offspring is much smaller in fish, fecundity will have much less emphasis than in sheep. Therefore, rate and efficiency of growth would be the major traits of importance.

Feed efficiency should be highly correlated with growth rate. This is the case in terrestrial animals, and given the lesser requirement for feed for maintenance in fish, the correlation should be higher still. It therefore seems likely to me that improvement in growth rate is likely to bring almost as much improvement in efficiency as would be possible if efficiency could be conveniently measured.

Since the cost of producing a single offspring is small, early survival is not likely to be very important. However, once costs have been incurred in growing animals for some time, survival would become important. Thus, as an outsider, I would guess growth rate to be of major importance, with survival, including disease resistance and adaptation to husbandry conditions, being of moderate importance, and fecundity of minor importance. Product quality would also be important, but this probably varies from one case to another, and I will not devote much attention to it.

## ASSESSMENT OF STRAINS

Having chosen the traits to be studied, the next step would be to evaluate the strains available. In this context, strains would simply be fish taken from different locations, regardless of whether we can establish that they are genetically different by comparing allozymes or other genetic markers. It may be that strains which we cannot differentiate by gene markers may differ for quantitative traits, and on the other hand, strains which may differ quite appreciably at the gene marker level may not differ in the important quantitative characters.

So the first step is to estimate the variation present among available strains. This means taking samples from a number of strains if we can obtain them. For example, Gjerde (1986) took over 30 strains of Atlantic salmon, and the mean two-year weight of these strains ranged from 4.6 to 6.3 kg, a very substantial range. Of course, most strains were close to average, but the best and worst were far different, and there were others well above average.

There are two points to be considered here. If only six or eight strains had been assessed in this example it is very likely that the range observed would have been considerably smaller. When several traits are considered, a large number of strains will be needed to give a good chance of finding a strain well above average in all traits. We should also pay attention to the sample size for each strain. To evaluate a strain we need a sufficient sample of genes from the strain to enable reliable comparisons to be made.

The second point is that fish are normally wild animals as distinct from sheep which have long been domesticated. It is a real possibility that in the early generations there would be some degree of natural adaptation to conditions, so that there would be genetic changes within the strains. While these are likely to be similar in all strains, they may be different, so that relative ranking of strains may change over the period of adaptation. There are reports of changes in behaviour of fish over the first few generations of culture, so it would be wise to monitor the performance of newly captured strains for some time.

## STRAIN USAGE

Assuming that the available strains have been evaluated, the next thing to consider is the use to be made of them. For instance, there are a large number of sheep breeds. If we want to produce fine wool for clothing then the solution is simple - use the Merino. But if we want to produce meat animals, the normal procedure is not to do this with a single breed. Meat animals are usually produced by crossbreeding. The system used in Australia typically starts with Merino ewes crossed to Border Leicester rams, and female progeny of this cross are mated to Dorset or Suffolk rams to produce the final meat animal. A similar system with different breeds is used in the United Kingdom, and such systems are also used in other species. The object of such a system is twofold.

One aim is to combine desirable traits present in different breeds. We need a fertile ewe who is a good milk producer as dam of the slaughter lamb. But that lamb needs only to grow fast and have good carcase quality. Its fertility and milking ability are irrelevant since it never expresses these traits. So we divide the production system into sections and try to identify strains to be used in each section.

In addition we commonly observe hybrid vigour or heterosis. That is, the crossbred is very often superior to the average of the two parents. This may or may not be useful. What is really important is not whether the hybrid is better than the average of the parents but whether it is better than the better of the two parents. Sometimes it is. We need to check whether some form of crossbreeding is likely to be better than production from a single breed. For production of meat sheep it is, but for production of apparel wool it is not. We must therefore investigate a range of breeding systems.

## CROSSING SYSTEMS

The example above is called fixed crossing. There are particular strains and we cross them in a fixed sequence to fit in with the production system and produce the best outcome in terms of profitability of the enterprise. But there are other options. An alternative system is rotational crossing. An example of this using three strains, A, B and C would be as follows. Strain A females are mated to strain B males, the crossbred female progeny being then mated to strain C males. Then the female progeny are mated to strain A males, and so on. This has the advantage that only small numbers of pure breeds are required, even for a very large number of crossbreds. A large fraction of the potential heterosis can be permanently maintained. The system may be recommended for convenience even if it does not produce better animals than a fixed crossing system.

Another possibility is formation of a synthetic breed. Here the object would be to combine the advantageous characters from different strains. We would also expect the genetic variation in a synthetic strain to be greater than in an individual strain because genetic variation between strains would contribute to the variation in the synthetic line. The main difficulty would be if heterosis were very important. Only about half of the heterosis can be maintained in a synthetic line, so half must be sacrificed if a synthetic is formed. However, the additional genetic variance

should allow more rapid subsequent improvement in a synthetic than in a pure strain.

So the first stage of setting up a breeding program is to look at purebreds and crosses. Gall (1975) found in rainbow trout that heterosis in two-year body weight was +19%, in egg number was +9%, and in egg size was -8%. Egg volume thus remained much the same. In sheep what we usually find is that growth rates show heterosis of perhaps 5 to 10%, but fertility often shows a great deal of heterosis. The example indicates that heterosis may be appreciable for one trait but negligible for another trait in a given cross. On the other hand, Gjerde (1981) found no significant heterosis for slaughter weight among crosses of five strains of Atlantic salmon, so weight does not always show heterosis. Thus one cannot assume that a trait will or will not necessarily show heterosis in a given cross. This is fairly general, in that only after a great deal of experimental evidence has been collected is it reasonable to make a prediction of what will happen in any cross. But while heterosis may not always occur, there may be a considerable amount in some cases.

On the basis of examining the performance of available strains and crossing systems we can decide on the breeding system to use. This will be the best available system using existing animals. The next step is to improve production above this level. It may be that further strains become available over time, and these should of course be evaluated. But if the original examination has been thorough, further improvement from this source is likely to be slow.

## ESTIMATION OF GENETIC PARAMETERS

Assuming we have chosen the best available system with existing animals, the next step is produce animals better than currently existing ones. That means setting up a selective breeding program. For this we need to know how much variation there is in important characters, to what extent that variation is inherited, and what correlations there are among these traits.

Another point to be considered is the possible existence of indicator traits. Quite often in practice traits which we would like to improve cannot be observed and selected for directly. So it would be useful for traits which are difficult or expensive to measure if other characters can be found which give reasonable indications of breeding values for traits we want to improve.

It is not very difficult to estimate how much variation there is in a character. It is more difficult to estimate the heritability of a character (the fraction of variation which is heritable) and harder still to estimate accurately the correlation between breeding values for different characters. To estimate heritability, essentially what we need to know is how similar related animals are compared with unrelated animals. It is easy enough to estimate the correlation between relatives provided we can identify the relatives. It is relatively easy to tag sheep to enable identity to be established, but rather more difficult in fish, especially when they are very small.

However, it is possible to group animals together as family members and keep them that way until they can be identified. But this introduces a problem in that it is likely to introduce non-genetic differences between families and lead to an overestimate of the heritability. In other words, under these conditions differences between families are partly genetic and partly environmental. So we tend to attribute environmental differences between families to genetic effects. The genetic differences between families can be transferred to progeny in the next generation, but the environmental differences cannot.

This problem can be overcome, but at a cost. One can split families and replicate them across tanks. Then one can estimate how much variation there is between replicates within families, remove that from the variance between families and get a clean estimate of genetic variation. This is a straight forward matter of design, but does mean that at least twice as many tanks as families would be required for a balanced experiment.

One of the difficulties of estimating heritability is that a rather large number of families are needed for a reasonably accurate estimate. I would think that 100 families would be an absolute minimum for reasonable precision. Many published estimates of heritability in fish are not very accurate. Of course, this could be said about estimates in sheep too, especially in the older literature. But standard errors of about 0.15 to 0.2 are very common for heritability estimates in fish. Given that the heritability might be estimated as 0.2 or 0.3, such standard errors do not give much confidence in the value reported. One would expect, therefore, to observe considerable variation between estimates obtained in different studies. This turns out to be true, and raises a problem of interpretation. There are many studies and heritability estimates vary widely. Given the standard errors, this variation could be simply due to sampling, but it could be that heritability does differ from one population to another. Unless heritabilities are

accurately estimated these possibilities cannot be distinguished.

Table 1 shows some heritabilities taken from Gjedrem (1983). These are averages of one to four estimates.

TABLE 1. Heritability estimates in different species, using half sib correlations

TRAIT	RAINBOW TROUT	ATLANTIC SALMON	CARP	CHANNEL CATFISH
Juvenile body weight	0.12	0.08	0.15	0.42
Adult body weight	0.17	0.36	0.36	0.49
Mortality resistance	0.14	0.11	-	-
Fat %	0.47	-	0.14	0.23
Age at maturity	0.18	0.71	-	-

The estimates sometimes differ considerably between species. For example, heritability of age at maturity is nearly 4 times as large in salmon as in rainbow trout. This difference is probably real, and would indicate very different consequences of selection in the two species. For adult body weight, estimates are more similar except in rainbow trout. The estimates indicate that selection for adult body weight could be quite successful. Mortality resistance has a fairly low heritability. This is not too surprising since mortality may be caused by many different factors.

Also, it is an all-or-none character, and being alive or dead is a rather crude classification, even if an important one. So there would be a lot of variation due to crudeness of classification.

Genetic correlations are even more difficult to estimate accurately than are heritabilities. 200 or 300 families would be the lower limit for useful estimates of genetic correlations. However, though it is difficult, it is important to estimate genetic correlations so that correlated responses to selection can be predicted.

## RESPONSE TO SELECTION

Given that heritabilities and other parameters have been estimated, the next step is to predict the consequences of selection. The first thing we require is the standardised selection differential, the difference between the mean of selected parents and the population average, expressed in standard deviation units. Provided a character is more or less normally distributed these can be found from tables, given the proportion selected, as in Falconer (1981). For instance, if 50% are selected, the standardised selection differential is about 0.8, if the best 20% are selected it is about 1.4, if the best 10% are selected it is 1.75 and so on. The response to one generation of selection is: (standardised selection differential) times (heritability) times (phenotypic standard deviation). This response per generation is converted to response per time unit, usually a year, by dividing by the generation length, which is the average age of parents when offspring are born. This may be quite important, because sometimes a procedure for getting maximum response per generation does not give maximum response per year. If it increases the generation length, it may be better to use a method which gives less response per generation but turns generations over quickly. In fact, getting generations turned over quickly is often a key factor in animal breeding programs.

Selection responses can be predicted in this way, but it is also possible to run a selection experiment to check what happens when the selection is actually done. Kinghorn (1983) gave examples of one generation of selection for slaughter weight in Atlantic salmon, reporting work by Gjerde. Two separate experiments were conducted, in which progeny of selected parents were compared with progeny of control unselected parents. The responses were 14.4% and 10.7%, with a 4 year generation interval, giving an annual gain of 3.6% and 2.7%. Sheep breeders would be very pleased with such gains, since responses of about 1% per year are

expected in well designed selection programs. Of course, it is not possible to get such a high selection intensity in sheep as in fish. These results give clear evidence that response to selection can be quite rapid.

So far I have dealt with responses in the trait selected for. But one must also look at responses in characters other than that actually selected for. If we decide to run a simple selection program, selecting for, say, two-year weight because we believe this will genetically improve the population, we need to know what responses to expect in other traits. Two-year weight will not be the only trait to respond to selection; traits genetically correlated with it will also change, so we need to predict changes in other characters of interest.

If  $C$  denotes the criterion selected on,  $T$  is the trait whose response we wish to predict, the response per generation is

$$i r_G h_C h_T s_T$$

where  $i$  is the standardised selection differential,  $r_G$  the genetic correlation,  $h$  the square root of the heritability and  $s$  is the phenotypic standard deviation. Selection on  $T$  would give a response  $i h_T^2 s_T$

so the ratio of correlated to direct response is

$$r_G h_C / h_T.$$

One reason why we are often interested in looking at correlated responses is that we try to avoid selecting on the character that we actually want to improve, and select on another trait instead. Perhaps we want to reduce generation length, or what we want to improve is difficult or expensive to measure. A good example of this in many animal breeding situations is feed efficiency. Efficiency of conversion of food to final product is something everyone wants to improve, but it is often extremely difficult to measure individual feed consumption - it can be done, but is very expensive. So breeders try not to do that but select on something else which they hope will improve feed conversion. Or else, we may try to measure something early in life so selection decisions can be made before the trait of real interest becomes available. So we need



to be able to predict correlated responses.

An example of prediction of correlated responses is given by Huang and Gall (1990). They made estimates of heritabilities and genetic correlations and phenotypic standard deviations in rainbow trout. Then they used the equation for correlated response given above, assuming that the best 10% of available fish were chosen on the basis of yearling weight. From tables the corresponding standardised selection differential was found, and correlated responses were calculated. One generation of such selection would be expected to increase mean yearling weight by 12 g. The correlated response expected in 25 month weight is 72.5 g. Egg volume would be expected to increase by 12 ml, egg size to change by - 2.39 eggs per 30 ml, and egg number to increase by 139. Thus the expected outcome of such a breeding plan could be evaluated in terms of all these responses.

## INBREEDING

The high reproductive rate in fish means that selection can be very intense, with a resulting large -selection response. But if we breed from only a small number of parents we will soon run into inbreeding problems because the rate of inbreeding is inversely proportional to the number of parents. If there are S sires and D dams used each generation, the rate of inbreeding is

$$\frac{1}{8S} + \frac{1}{8D}$$

There are two consequences of inbreeding. The amount of genetic variation in the population is reduced, so that future selective gains are reduced. The second consequence is a depression of production due to inbreeding. As an example, Kincaid (1975) presented data on the effect of one generation of brother-sister mating in rainbow trout, where the inbreeding coefficient is 0.25.

TABLE 2. Effect of one generation of brother-sister mating in rainbow trout.

<u>Trait</u>	<u>Depression</u>
Crippled fry	37.6%
Food conversion	5.6%
147 day survival	14.6%
147 day growth	23.2%
Number of fish/year	17.4%
Weight of fish/year	36.6%

The final figure of a decline of 36.6% in weight of fish per year with 25% inbreeding is a clear demonstration that inbreeding is undesirable. So, as well as predicting selection responses as discussed earlier, we should also check on how much inbreeding is expected for a given breeding plan, and we should devise programs which do not produce significant amounts of inbreeding. In virtually all species which do not naturally inbreed, the effect of inbreeding on productivity is quite serious.

## CONCLUSION

As well as the things I have discussed so far, animal breeders seek to develop ideal selection criteria. That is, they look at a range of possible characters which can be used to make selection decisions, and put them together into an overall score or selection index which will maximise response in the breeding objective. If this is to be done properly, it is necessary to have reasonably good estimates of the heritabilities and genetic correlations involved. Reasonable estimates of the relative economic importance of various characters are also needed. The need for such information is not a limitation of selection index procedures. In any breeding program which will work effectively it is necessary to have this information, as otherwise it is not possible to make a rational evaluation. One advantage of selection index procedures is that the breeder

is forced to consider just what information is required. Procedures for doing this are standard and discussed in many texts.

From the literature it seems that fish breeders are becoming quite active in such areas as estimating heritabilities and planning breeding programs. They have advantages over sheep breeders, particularly because of high fecundity, and it seems that selection responses are potentially greater than a sheep breeder would expect. There are some practical difficulties in implementing more complex breeding systems involving use of information from relatives, involving collection of pedigrees, but it seems that simple mass selection programs should be effective

### REFERENCES

- Falconer, D.S. 1981. Introduction to Quantitative Genetics. 2nd ed. Longman, London.
- Gall, G.A.E. 1975. Genetics of reproduction in domesticated rainbow trout. J. Anim. Sci. 40: 19-28.
- Gjedrem, T. 1983. Genetic variation in quantitative traits and selective breeding in fish and shellfish. Aquaculture 33: 51-72.
- Gjerde, B. 1981. Growth rate and mortality of crosses between strains of Atlantic salmon. 32nd EAAP Meeting, Zagreb. Paper IV-16.
- Gjerde, B. 1986. Growth and reproduction in fish and shellfish. Aquaculture 57: 37-55.
- Huang, N. and Gall, G.A.E. 1990. Correlation of body weight and reproductive characters in rainbow trout. Aquaculture 86: 191-200.
- Kincaid, H.L. 1976. Inbreeding in rainbow trout (*Salmo gairdneri*). J. Fish. Res. Board Can. 33: 2420-2426.
- Kinghorn, B. 1983. A review of quantitative genetics in fish breeding. Aquaculture 31: 283-304.

## DISCUSSION

DR. JOHN MACINTYRE: Some of the advantages are no bagging, no footrotting, no muelsing. Do we have questions?

ANON: You pointed out that in selective breeding of fish the percentage gained is very much higher than that stated in sheep. What would you anticipate that you are really dealing with to grade your sheep that have been subject to selective breeding for a long time? Most of the gains have been made, whereas with fish an animal in which very little selective breeding has taken place and therefore more relative gains can be made, and I suggest that the percentage gained would decrease as more breeding takes place. Shouldn't it be at this stage that we be looking beyond the immediate gains and maybe looking at gene manipulation as a way of really advancing production. We don't often get quite a clear image on what we achieve by selective breeding.

PROFESSOR JOHN JAMES: So far, experience is that limits to what can be achieved by selective breeding seem to recede as you approach them. One can still make progress at the rate of say 1% per year in production in sheep and the same can be achieved in cattle. These are things which have been going for quite some time. There are good theoretical reasons why, as you select, you would expect that eventually the rate of response to selection would decline. But even 1% a year is actually really very worthwhile. While I don't want to suggest that things other than selective breeding shouldn't be looked at, because I think they should, I don't think you want to feel any sort of despair about selective breeding. I say 1% a year improvement doesn't sound much but if you look at it over a period of time it really is well worth having.

DR. COLIN PURDOM: In anticipation of what I am going to say later, I would like to pour a fairly large douse of cold water over the concept of selection for growth. The difference between growth rate in sheep and in trout is that sheep have access to food all the time and for trout, feed is given in short bursts because it deteriorates rapidly once in water. A decision has to be made on how much to feed. Normally this is defined as a percentage of the body weight of the fish. Now this immediately introduces a very interesting bias. If a batch of fish gets marginally longer than another it gets more feed; they grow more and they have got to have yet more feed, so I think it's extremely difficult to design proper experiments to describe the sorts

of growth rate for example that Gall has produced.

PROFESSOR JOHN JAMES: I think that this is actually something that goes back a fair way in animal breeding because perhaps about 30 to 40 years ago, there was a lot of concern about exactly this in relation to production of dairy cattle where the question of feeding to production arose. Because the cows that give more milk get fed more. The question then is what are you defining as the environment within which production takes place. If you define that as an environment in which if you grow faster you get fed more, then the feeding to production is part of the environment. I think it is a legitimate way of looking at things to say that although this does introduce a correlation in one sense between genotype and environment, and it is possible to redefine the environment so that this is actually a genetic effect in that the reason they get more feed is that they are genetically superior. Perhaps this is not something that everyone would agree with, but I think it is a legitimate way of approaching things.

ANON: Just to comment on the exploitation of heterosis. You've impressed there is heterosis important enough to exploit and if one is going to use them, how would you reverse the cross? It's important, I think, to check out what actual level of heterosis does occur. Can it vary from 50% and be as low as 4% and can it in fact be considerably higher?

PROFESSOR JOHN JAMES: Yes, naturally, the figure of 50% is a simple prediction and it does depend essentially on the assumption that dominance is the major factor causing heterosis. If epistasis is important anything can happen and does.

ANON: I was interested in the problems that you described in attempting to estimate heritability with the traits. It seems to me that, as you pointed out, huge standard errors equate with heritabilities and one would assume correlations. The fact that it's nice to have some sort of estimates to represent the heritability irrespective of a response to selection which would completely correlate that one literally on a case by case basis.

PROFESSOR JOHN JAMES: I think that there a couple of things to consider. Selection experiments to estimate heritability are much more expensive because you need a separate selection experiment for every character whereas you can estimate heritabilities for a large number of characters just in the one set of data. Similarly, if you are trying to estimate

correlations the same sort of thing applies you must have a selection line for all of those things. Of course the other thing you have to bear in mind is that estimates derived from selection experiments also have standard errors. Particularly if the heritability is low, the standard error of an estimate from a selection experiment can also be quite high. I know they are real responses in a sense, but also they are subject to sampling error. The problem is that, essentially, it's quite a bit of work to get all these estimates. Without wanting to be discouraging, we know that the more characters you try to handle at once the very much worse it gets. In fact if you try to estimate, say, heritabilities and genetic correlations for something like 15-20 characters you are virtually certain to get a set of estimates which couldn't possibly be correct in the sense that the heritabilities and genetic correlations could not simultaneously be true. This is a very high probability of getting that. Well, that's not as bad as it sounds in practice because you usually don't want to make use of all of those things simultaneously anyway.

**SESSION 6**

**GENE POOL MANAGEMENT**

**Chairman: Dr. Craig Moritz**

**University of Queensland  
St. Lucia, Queensland**

## CHAPTER 5

### Genetics in fish culture\*

**Dr. Colin E. Purdom**

**Ministry of Agriculture, Fisheries and Food,  
Lowestoft, U.K.**

DR. COLIN PURDOM: This paper is about practical fish genetics. Basically the material can be divided into two different genetic approaches. First, the conventional approach which covers selection and cross breeding already covered in some detail at this conference, and secondly, unconventional approaches which include induced polyploidy and sex control and biotechnology. Cross breeding usually involves inbreeding in one way or another, and that provides a suitable bridge to go from the conventional to the unconventional via gynogenesis, more of which later. A further important aspect of fish genetics is the environmental threat that aquaculture is seen in some parts of the world to be posing.

Any scrutiny of fisheries statistics will show that catches have been in decline for the last two decades whilst values have been going up. This is what has driven the search for more effective ways of farming fish. There was at one stage a feeling that fish farming was about producing cheap protein to supply the underdeveloped nations of the world; it almost never is the case. Fish farming is about making good quality food and making fair profits.

Twenty years ago when fish genetics in aquaculture took off, fisheries biologists expected a lot

*\* Paper presented at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990.*



of geneticists, and on some of the theoretical grounds covered at this symposium they still seem to! Selection is still the most popular notion, and selection for improved growth remains high in farmers' esteem. A great deal of controversy is aroused over the potential for selectively improving the growth rate of fish, but that isn't to say that selection itself is not, or has not been, a highly successful development in other aspects of fish culture. Huge levels of success have been achieved in terms of changing the colour patterns of fish or changing the shapes of their bodies and fins. Any visit to a pet fish shop anywhere in the world will provide examples of how successful this sort of selection has been. Some of the fancier breeds of goldfish may present doubtful improvement, but some are more decorous, particularly those with highly decorous fin shapes, and these are determined genetically in a fairly straightforward and simple way even though they involve quite a lot of gene loci and perhaps a lot of complex interactions.

The colour patterns of goldfish and nishikiqoi are not determined genetically, although in other fish they are. Thus xanthic flatfish lacking black pigment are periodically caught by fishermen, and although no precise data are available it is probably true to say that one in a million fish is of this sort. The colour is almost certainly due to a simple recessive, so the gene frequency would be something like one in a thousand or thereabouts. The significant point here is that in any breeding plan with fish, sooner or later these pigment deficient recessive mutants turn up, and there are strains of most fish that are bred artificially that lack melanin. Some have curious uses; for example, trout of a golden hue have been used in put and take fisheries, and one reason is that anglers who catch nothing and state that the fishery had no fish in it could be disabused by the sight of one of these 'goldfish' which are visible at 50 to 60 yards! This may be regarded, perhaps, as a trivial use of genetics, but nonetheless it satisfied the demands of a particular management problem.

The question of growth rate does need more discussion, however, because it is the characteristic of fish that has most interested the fish farmer for the last 25 years. Despite all this interest, however, there is no evidence that fish selectively improved for growth rate have ever been used in commercial practice. The most definitive work in this area was that by the Israeli group working with carp and tilapia. Although they were able to demonstrate some heritability of growth rate they were never able to achieve selection progress for this particular characteristic. They did argue that it was possibly due to the fact that the carp had been cultivated for a hundred years or so and perhaps had plateaued. But it seems to me that this is too naive an explanation,

and that the real explanation lies in the fact that variation for growth rate in fish is almost overwhelmingly determined by environmental factors. It is possible to manipulate the environmental factors, for example, to measure repeatability - a crude estimate of heritability. Work done in England some twenty years ago using plaice and sole (marine flatfish rated at the time as possible farm fish) produced figures in the range of 0.1 to 0.01 - very low levels even for repeatability. A recent experiment in Japan has demonstrated the problem quite elegantly using a triploid fish *Carassius auratus gibelio*. This is a naturally occurring all-female species found widely in Asia and Russia. These triploids form natural clones, they reproduce without meiosis, they are heterozygous but are all identical to each other. When growth was measured in these fish the usual normal-type distribution, skewed at the high end, was found - the rapid growers are called shooters or jumpers, and the point is that they arose even though all fish were genetically alike. This sort of experiment is not of course 100% conclusive, but it does demonstrate the enormous environmental variance that typifies growth rate in fish. The corollary of this is that the potential of the fish really is expressed by what happens with the fast growers. There have been suggestions that one should select these aggressive, dominantly hierarchical fish, but that of course cannot work when the source of the hierarchy is environmental. To select for lack of aggression might make more sense.

Turning now to other selection programmes. In the work in the UK on rainbow trout genetics, the first important job was to try to spread the spawning time of the fish - at that time some spread of egg availability was achieved simply by buying eggs from Australia or New Zealand. Eggs were imported to the UK in mid summer from New Zealand. The normal spawning period for trout in the northern hemisphere is, of course, in the winter. People who imported these New Zealand stocks were of the opinion that they would spawn again the following July or August, but of course they did not because their spawning time was determined by day length and in their new home in the UK they spawned in the winter just like other rainbow trout. However, by searching the northern hemisphere it was possible to acquire a range of stocks which varied in spawning time roughly from September through the turn of the year into April. So a government programme of choice widened spawning potential to cover half the year. Later, a Japanese strain was obtained which was reputed to spawn in mid summer, in fact it spawned twice a year, but its growth characteristics were so poor that it is probably not going to be worth maintaining. Nevertheless it is with a commercial farmer and may be subjected to further trials.

The question arises - where does this variation for spawning time come from? There is no really solid answer. Some think it derives from selection over the past 50 or 60 years. Alternatively it could be that the wide distribution of rainbow trout in America, which is from Alaska in the north down to California in the south, does embrace a range of natural spawning times and that these "domesticated" fish simply represent stocks derived from that geographic range.

The stocks collected in the UK programme differed in other aspects of performance including growth rate. The worst grower was the New Zealand strain and in terms of length it had only about 75% or 70% of the performance of the best of the other strains, the Winthrop, a strain widely used in North America. A further interesting feature of these strains was that the enzyme characteristics, the gene frequency arrays, were significantly different but not substantially so. There was a high level of heterozygosity with one exception, which was a stock (Washington) that was selected at the University of Seattle for some 30 or 40 years. Selected for fast growth it in fact did not grow any faster than the better of the other stocks, but its great benefit was that it was the most inbred of the rainbow trout stocks available in the UK, although it did have one allele segregating at a frequency of about 50%.

The variable growth rate in these strains was genetic in the sense that crosses between different strains generated growth rates in the  $F_1$  intermediate between those of the two parent types. None of the crosses showed evidence of heterosis. The further question was whether one could use the variance generated in an  $F_2$  from a hybrid for a selection experiment. However, variances in  $F_2$ 's and back crosses for length measurements were no greater than the variances in the parental types. The  $F_2$ 's, some of them were complex, involving 3 or 4 different strains, and the back crosses had significantly lower levels of variance than the parental genotypes. Once again the growth rate differences in these strains do not appear to be attributable to additive genetic variance.

It is necessary to emphasise the negative aspects of selecting for growth rate because selection experiments or procedures are very expensive and very time consuming. They have been performed for 20 years or more seemingly without any progress at all. It makes a lot more sense to seek a faster growing strain of natural origin, and fish farmers in the southern hemisphere could obtain much faster growing stock from Europe and the USA if they wanted to.

Cross breeding is almost standard practice in plant genetics. Cross breeding involving different species is a very popular pastime with those who keep fish. Many thousands of hybrids have been produced but the result is almost always unsatisfactory. Most hybrids between different species or between different genera (and even different classes have been crossed) are defective and not worth much in fish farming terms. But there have been exceptions and perhaps the best known is that of the sturgeon hybrid between the large marine species *Huso huso* and the small freshwater sterlet *Acipenser ruthenus*. The objective of the cross undertaken by Soviet scientists was to produce a fish that was suitable for pond farming but grew a little faster than the sterlet. The hybrid, called bester, did just that. It was adaptable to freshwater all its life and grew at a satisfactory rate for economic purposes. More importantly, its chromosome complement was balanced so that it was fertile and from it has been derived an artificial species which now breeds more or less true and provides a fish, a sturgeon, which is suitable for cultivation. The other common sort of hybridisation is the within-species cross, i.e. the crossing of strains within a species. This can also have the objective to combine different characteristics into an economically viable form.

One cross in the UK that has had some success with anglers involves steelhead trout and its normal cousin. In its native environment the steelhead is an anadromous species and it typically has a long and sleek silvery appearance. The normal trout are shorter and fatter and somewhat differently coloured. The hybrid is almost exactly intermediate but it has the advantage of looking almost salmon-like. It also has a growth rate which is intermediate between that of the fast growing fat trout and that of the slim anadromous form. So the fish turn out to be a commercially sensible product.

Much the more popular form of cross breeding involves inbred lines to produce genetically uniform heterotic animals or plants. Inbreeding is normally very long term, but rapid non conventional methods have been developed, which brings us to the subject of gynogenesis, the production of an embryo using only maternal genetic material. It is achieved by fertilizing an egg with spermatozoan containing inert genetic material. Ionising radiation or UV can be used to inactivate sperm whilst leaving its mobility unimpaired. Following fertilisation the second phase of meiosis is completed in the egg, the second polar body is lost and the resultant embryo is a haploid - these rarely survive long but they do sometimes hatch. Amongst the haploids generated by gynogenesis there is always a very low frequency of diploids and this can be

increased by applying a physical shock to the egg immediately after fertilisation. The physical shock can be either high temperature or low temperature or pressure, or the same effect can be achieved by chemicals such as cytochalasin. Different fish species respond differently to these various approaches. In the early work with marine flat fish, low temperature (0°C) was very effective but not so with salmonids. French scientists achieved success, however, using heat as the shock and later on many scientists around the world refined these methods plus the use of pressure.

The effect of the physical shock was to suppress the second phase of meiosis and the concomitant loss of the polar body to produce diploid gynogenetic individuals and the original hope was that these would be inbred, and that in this manner we would be able to make hundreds of inbred lines in fish in little more than one or two generations. In the event it did not turn out as easily as that, because of the problem of crossing-over which can generate heterozygosity within chromosome pairs. In actual fact the first generation of gynogenesis produced by this suppression of meiotic metaphase leads to fixed heterozygosity for the terminal bits of chromosomes. Current emphasis is now more on producing the diploid state by doubling up chromosomes at first mitosis in the embryo. If this can be achieved, fish homozygous at all loci are produced, following which another generation of meiotic gynogenesis produces clones of homozygous fish. This has in fact been done in the United States with experimental fish (*Zebra danio*), but unfortunately the full methodology has never been published. Repetition of the method as far as it could be understood has produced large numbers of diploid homozygous trout alevins, but they have never survived more than a few months so the cycle has not been completed. There are reports that confirm the *Zebra danio* work in trout but evidence of the existence of clones has yet to be produced. In rainbow trout it would take six years to produce that evidence.

An important point about gynogenesis is that a very simple modification of the cycle can lead to the production of polyploids. If ordinary fertilisation is followed by physical shock the resultant embryo has two sets of maternal chromosomes and one set of paternal chromosomes - it is a triploid. One reason for producing these is to generate sterile fish. Other methods of achieving sterility have been explored, but none seemed as promising as induced triploidy. The reason why triploids would be sterile is simple; the process of meiosis which leads to the formation of the gametes requires close pairing of analogous chromosomes and if three chromosome sets are present it is impossible to achieve this. By analogy again with gynogenesis,

if the shock is delayed until the time of first mitosis, the embryo would not be a triploid, of course, but a tetraploid.

The objective of producing tetraploids was again to produce triploids in another fashion. If a tetraploid stock is crossed with a diploid stock (such things are possible in potatoes and other plants) the  $F_1$  is triploid. Unfortunately, the delayed application of shocks is difficult to achieve successfully, but there are reports that it has been done in rainbow trout and in tilapia.

Once triploids could be produced, the question was how to improve the situation for commercial production, and that means defining the conditions under which the physical shocks are given. Taking heat, for example, the parameters to decide on are the extent of the heat shock - what temperature, the time of the start and the time of the finish. The temperature itself is critical. If too high a temperature is used the eggs are killed, if too low the yield of triploids is low. Other than that, the timing is not so important and any period during the first 10, 20 or 30 minutes after fertilisation seems to be good for survival and for yield of triploids. Similarly, the duration of shock is not too critical, from 7 to 10 minutes is reasonable but longer produces problems. For cold shock, hours of exposure can be given, duration is not critical; low temperatures are probably less harmful physiologically to an egg than high temperatures. Similarly with pressure, as it increases, the frequency of triploids increases and almost 100% triploids can be achieved with appropriate pressure treatments. It was felt that pressure would be physiologically less harmful than heat or cold for producing triploids but there is some disagreement here. What does make a lot of difference to survival is the quality of the eggs. Poor quality eggs do not seem able to survive any of these treatments very successfully. One thing to note is that it is very rare to get 100% triploidy, there are always a few diploids, sometimes a few haploids as well, but it is possible now with refined treatment, to get quite close to 100%.

The next question is how to recognise that the fish are triploid. One way would be to count chromosomes, but fish have got a lot of chromosomes and counting them is extremely tedious. The methods used initially were based on cell and nuclear size in a chosen tissue. It is important to use a tissue which can be identified of course, and the obvious tissues were red blood cells and cartilage cells. The disadvantage of this method is that the fish must be reared to a size to get blood out of them. That is not too difficult with rainbow trout, but it does present problems with marine fish which are often very small at hatching. Another way of detecting triploids is

to use hybrids where differences in pigmentation can reveal the genotypes. This has been quite effective in marine flatfish. Sundry expensive methods exist such as flow cytometry, and a new and elegant way of detecting triploidy may be to score nucleolus number histologically.

Much work has been done on the performance of triploids and the messages are mixed. The expectation was that the triploids would be sterile, and this was confirmed genetically, but for the males the gonads still develop albeit with nonviable spermatozoa. The triploid testis is not quite as well developed as the diploid counterpart and not quite as advanced in the spermatogenic cycle, but nonetheless it is large and sufficient to generate all of the undesirable, secondary sexual characteristics of the fish that sterile techniques seek to avoid. On the female side, however the situation is quite the reverse. The ovary of triploid rainbow trout is just a strip of tissue with occasional oocytes but not enough to generate steroid production and sexual maturity. These females thus remain forever juvenile in appearance!

What of growth performance? Much research has shown that diploid and triploid, males and females, respectively, all grow more or less at the same rate up to the spawning season but then males virtually cease growing. The diploid female has a period where she actually declines in weight simply because of the production and release of eggs, but then starts to grow again. The benefit of the triploid female is that it does not waste its energies producing eggs and continues to grow at all times. The benefits are not just that the triploid continues to grow throughout the year, it also maintains its condition. The female trout after spawning or egg production is a much less saleable commodity than a fish that has not produced eggs at all.

The ideal is therefore the triploid female, which means that it is desirable to devise a technique for producing only females. In actual fact, such work started independent of polyploidy, and such a technique was developed in parallel with the triploidy methods.

The motivation for this sex-ratio control work simply lies in the fact that males are poor growers, poor converters of food and have poor flesh quality when they are sexually mature. They are also very aggressive, ugly, and ill adapted to life in sea water.

It is possible to produce female fish just by feeding them at a very early age with appropriate hormones, but this is not really the answer to the problem because it is unreliable (sometimes

hermaphrodites are produced, sometimes there is no effect at all) the estrogens are often toxic, feminisation has to be done every generation, but most of all use of hormones meets with very considerable consumer resistance. People do not like to think that the food they are going to eat has been contaminated with hormones. Another way had to be devised, and basically it was achieved by extending the Japanese studies of sex control in carp to trout and other species. The Japanese demonstrated that sex determination in fish was basically chromosomal but that the actual determination of sex itself could be manipulated during the early part of the fish's life, immediately after hatching, i.e. during the sexually indeterminate phase. By feeding either male sex hormones, e.g. methyl testosterone or female sex hormones such as estradiol, one can produce males or females at will; this was also found to work with salmonids. Fortunately, rainbow trout and salmon have a chromosomal sex determining mechanism based on the concept that male is XY and female is XX. So the technique to produce all females required first the feeding of ordinary fry with methyl testosterone to produce 100% males, a paradoxical start! Half of those fish, however, would have been males anyway and they are identified by progeny testing and discarded. The other half would have been females and therefore, have the genetic constitution XX and when used for subsequent breeding purposes with ordinary female (XX) produce offspring which are all female. The method is simple and works very successfully. The only problem is that, in order to keep the all female stock going, it is necessary to produce some males artificially; but of course just a few hundred instead of the 50% production that would happen by normal breeding systems.

There was one drawback to this process of sex reversal, and that is that the gonadal ducts of the normal male do not develop properly in a sex reversed female. This was useful initially for actually getting the technique off the ground, but later on it presented problems because farmers were unable to strip milt to determine whether or not it was ripe and ready for use. Fish had to be killed in order to assess the milt, and an improvement was needed. Success was achieved by refining downwards the dose of methyl testosterone. The early work used a dose of 3 milligrams of methyl testosterone per kilogram of food fed for 600 to 700 degree-days. Perfectly adequate results can now be achieved using 0.5 parts per million over the same time period. The important final point is, of course, that once a stock has been produced which has no Y chromosome, it is not necessary to worry thereafter about males of the wrong genotype getting into the system. So even if masculination is not 100% it doesn't really matter, and currently the use of these very low doses of estrogens generates sufficient numbers of reversed male to enable farmers to



control effectively their breeding programmes.

The final stage was to take this technique and the triploid technique to the farmers to convince them of their usefulness. This was very hard work! The most difficult part of the case was to convince the farmer that the females were not peculiar, just ordinary females, and all we had done was get rid of the males. The females were just the same as females by any other breeding system. Farmers were more intrigued by the triploids, and after two or three years of persuasion a farm survey of England and Wales was undertaken. The encouraging result was that only 28% of UK producers followed conventional breeding methods, female-only methods were used by 72%. Triploidy was not quite as popular as the all-female method, but was still reasonably well received by the fish farming community with about 50% of farmers using it. The problem with triploidy remains that the eggs have to be of excellent quality to get good results, and that there is often a loss due to mortality of eggs during the early hatchery phase. Some 83% of egg production in England and Wales is either female-only or triploid female.

Biotechnology is now widely studied in fish. In the UK, pioneering work has been done at Southampton University where metallothionein genes and growth hormone genes have been incorporated into rainbow trout embryos with a success rate of about 15%. The snag has been that the expression of the gene has not yet been fully demonstrated. This sort of result has been achieved in several laboratories around the world, and better results will possibly come when the growth hormone and other genes currently of mammalian origin are available of salmonid origin.

Environmental problems are generating a great deal of heat in the UK because in any large scale fish farming exercise it is inevitable that fish escape. Rainbow trout, in particular, manage to find their way into all sorts of seemingly impossible positions, but they do not represent a threat in the UK because they do not breed in British conditions. However, the problem does exist with salmon, and so there has been a great controversy about whether salmon farming is going to diminish the vitality of the wondrous sport fish *Salmo salar*, and all sorts of proposals have been made for curtailing commercial production so as to “save the British salmon”.

It seems overwhelmingly obvious that introductions of new fish species, or new species of any animal or plant, can have and have had disastrous consequences for native flora and fauna. No

one disputes that, but it is very difficult to discover any example of the genetic integrity of a species being affected by the introduction of others of that same species. The argument from those who think that such hazards are real comes from two concepts. The first is that of the specific gene pool, the adapted gene pool - that it is something very special and fragile and must not be disturbed. The other is that the domesticated salmon is very different in its genetic structure from those of natural populations. To take the second point first, salmon have only been cultivated for a few years. They have a generation time of something like 4 or 5 years, depending upon where they are reared, so there have been very few generations of domestication. On top of that, until quite recently, salmon farmers very often went back to the wild for their stock. So there has been no really consistent domestication (for want of a better word) in salmon farming. The salmon that are farmed today are less than a step away from their wild ancestors - they are not that different. The question though is if they come from a specific geographic region, will they have specific genetic arrays, and could these constitute genic pollution to another stock? The answer to that is - probably not, for the simple reason that any detrimental effect that is likely to arise by an admixture of alien genotypes is going to be subject to severe selection, and the prospect that this would, in itself, have an overall effect on the fitness of a population seems to me minimal where fecundity and natural mortality are both very high. The death of individuals constitutes a genetic risk to individuals, yet guarantees fitness of the population. Populations are robust; even though individuals are genetically damaged a population has the property to recover; it is not a static entity. There are probably far more important things to worry about than the possibility of genetic pollution from farmed fish - habitat contamination or loss of habitat are the real problems.

## DISCUSSION

DR. JOHN BENZIE: Could you expand a little bit on other taxa? Your comments were related particularly to fish but there will be people here interested in other organisms and to what extent do you feel that what has been found for fish is also the case in molluscs and crustaceans? Also, and here I would be interested to hear Professor James' comments - to what extent do either of you feel that the lack of success in selection programmes in fish has been due in one sense to an inappropriate selection regime which has made assumptions, which are not tenable, largely

because the bulk of the variation present is under a large degree of environmental control and therefore the sorts of practice which have been undertaken are perhaps inappropriate.

DR. COLIN PURDOM: In answer to the first question, I think the methods are generally applicable. In fact they began in the early 20th century when German embryologists looked at the effects of ionising radiations on frogs and continue up to the present, with a variety of animals including frogs. Triploids have been produced in molluscs - in the UK the emphasis has been more on environmental protection. Triploidy has been produced in manilla clams which creates one further barrier to their successful reproduction and spread in UK waters. The methods are generally applicable simply because they depend basically upon meiosis and mitosis and these are pretty well standard across the animal and plant kingdoms. The consequences vary though; the pattern of the production of triploid gonads in teleosts is probably going to be repeated in most vertebrates, but in invertebrates there is some production of both testes and ovaries in some triploids. They are, of course, genetically sterile in the sense that balanced gametes cannot be formed. On the question of selection, it is necessary for people to define very precisely what they want to do. If they want better growth rate, then the best thing may be to choose a strain or a geographic variety that grows better - this is just common sense. On the other hand if, for example, it is desired to produce a rainbow trout that spawns in mid July and having discovered that the only option is a poor grower that nobody wants, then the best stocks should be taken and subjected to a heritability and selection programme for that purpose. And that seems to make sense if it is important to do it, but defining the objectives is vital. Growth rate genetics is a very difficult subject in fish and is probably best avoided.

ANON: I do think that there is a possibility that the best selection methods haven't been used yet and I would quote another example. This time, not from animal breeding, but from plant breeding. For a long time it was believed impossible to make any great genetic improvement in maize because selection programmes had been unsuccessful. But it turned out really that this was because of poor experimental design and that with a good experimental design it was possible to make selective improvement in maize. I think the same thing may be true in fish if this problem, say of extreme individuals because of environmental effects is such a serious problem, there may be ways of getting out of that by for instance selecting families on the basis of average family performance. This is a question that has arisen in pigs, for instance, where competition in a pen arises. It would be a question of finding some practical way of actually

running the selection programme.

DR. PETER STEVENS: I'd like to address my question to Bert Sheridan. While acknowledging that there are problems with severe inbreeding, I was wondering whether you could comment on whether genetic variation is an intuitively attractive criterion to use for management and conservation, such as we saw by geography. It seems intuitively right. Or whether you think there's a necessary correlation between genetic variation and the success or survival of a species bearing in mind that there are a lot of apparently successful species that have very little if any detectable variation.

DR. BERT SHERIDAN: It is certainly true that in the mammals, the cheetah is often given as an example of a species which is very widespread but they accept skin grafts from one another as if they are the same individual. However, despite being so widespread they are very vulnerable to episodic feline enteritis virus in a zoo which affects most cat species very mildly. The cheetah population was reduced to one-tenth of its size through virus. All the other cats showed minor symptoms. So one wonders how long the cheetahs can persist like that. Certainly there are lots of populations which have been through some bottleneck and have virtually no variation but we need to know which of course we can't know at the present slice of time. We need to know what the rate of formation of such population is and what the rate of extinction of them is. Does that answer the question?

DR. PETER STEVENS: I was also meaning isogenic species and things like that. Also some other vertebrates have very high levels of variation, but vertebrates in general have very low levels of variation. Can it be necessarily used as an indication of how successful or vulnerable a species is?

DR. BERT SHERIDAN: No. I don't like the idea of comparing, even between genera, I don't like making such comparisons let alone with vertebrates to mammals. What research has been aimed at is - I don't know if you know the biology of the species I am working on, but they are not actually endangered species - so we can look at individual populations and have a reference population which has what we hope is the normal level of variation for that species and look at a population which has been perturbed in some way and had its variation pushed to some low

level and see whether we can relate that to some problem such as the sperm disorders I mentioned. So we are actually, like you, not convinced by this association and we are trying to get better data to back it up.

COLIN PURDOM: There is an English aphorism that hard cases make bad law and I think perhaps that we should reflect that sorting out very special circumstances is not necessarily providing very good advice. *Rivulus marmoratus*, for example, is a self fertilising hermaphrodite and in consequence is highly inbred, but it seems to be very successful. But that doesn't mean that inbreeding can go on indefinitely. My views on the resilience of population genetic structures apply to commercially important fish species, which by definition are abundant ones. In the cases of a rare species of fish, or of fish in some decline, these considerations may not apply.

**SESSION 6**

**GENE POOL MANAGEMENT (continued)**

**Chairman: Dr. Craig Moritz**

**University of Queensland  
St. Lucia, Queensland**

## CHAPTER 6

### Conservation of Genetic Resources\*

**Dr. W.B. Sherwin<sup>†</sup>**

Department of Natural Sciences,  
Phillip Institute, Bundoora, Victoria

and

Department of Genetics and Human Variation,  
La Trobe University, Bundoora, Victoria

#### ABSTRACT

Strategies of gene pool management vary, depending on the biology of the species concerned and the overall aim of the management program, but one theme common to all strategies is that preventing genetic problems is usually far easier than fixing them once they have occurred.

Possible aims of a fisheries management program, classified according to their genetic requirements, include:

- (1) long-term conservation of an endangered wild population as either
  - (a) part of conservation of the community to which it belongs, or
  - (b) a genetic resource for the future
- (2) rational exploitation of a relatively abundant wild population
- (3) long-term aquaculture
- (4) aquaculture for release to replenish wild stocks
- (5) aquaculture for release as part of put-take angling.

\* Paper presented at *Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990*. <sup>†</sup>Present address: School of Biological Science, University of NSW.

Factors which must be considered in managing the genetics of a particular species are:

- (1) the level of genetic variation present in the wild population
- (2) the way in which this genetic variation is partitioned within and between populations
- (3) inbreeding and its effects on reproductive success (inbreeding depression)
- (4) the possible adverse effects of mating between individuals who are too distantly related (outbreeding depression)
- (5) conversely to (4), in some conditions, crossing between distantly related stocks may produce superior individuals, at least for one generation
- (6) the recognition and management of artificial selection imposed by the management program, and natural selection
- (7) other options such as creating triploid individuals and cryopreservation of sperm.

To create a management program, we must identify (i) manageable processes, and (ii) measurable outcomes by which we can monitor the progress of the program. Examples of manageable processes are the numbers of individuals used at each stage (i.e., in the natural population, or as founders or broodstock in aquaculture), and the timing of events (e.g. fishing or releases) relative to the reproductive cycle. Monitoring of the genetic management program requires (a) traditional genetic data such as protein and DNA markers, as well as (b) ecological and physiological information, such as recruitment levels, homing success, temperature tolerance and growth rates. Collection of the necessary ecological data can often be aided by genetic analysis.

## INTRODUCTION

I have worked mostly on conservation genetics of marsupials, but many of the broad principles are similar for fish. For all species, genes are the basis of their adaptation to natural conditions, or to artificial conditions, if we want to try to breed them. Therefore, genes are seen as a resource, both to enhance the chances that particular lineages will be able to survive in a wild population that is facing a changing environment, or for human use in breeding. Human intervention has caused various documented genetic changes to wild populations, both deliberate and accidental, and the aim of conservation genetics is, in many cases, to minimise artificial genetic change so



that normal evolutionary processes can occur (adaptation to changing conditions). In other situations we may actually want to cause particular genetic changes, in a breeding program. There are two particular factors about fish that make their conservation genetics somewhat different compared to other wild populations. One is mass removal - harvest of very large numbers; and the other is mass release - hatchery releases to restock areas. Both of these can have considerable genetic consequences that are not found in many other species.

Having been told that about half of you are not geneticists at all, I am going to try to concentrate on the things we can measure when we are worried about a population, and explain a little of the importance of each of these variables. A lot of that has been done for me this morning, and as we go along you will see that a lot of these variables are often measured by non-geneticists. After I have gone through all that, I will deal with details of particular management programs for particular objectives and stress the importance of these particular measurables.

## MEASURING GENETIC VARIATION

First of all, how do we measure genetic variation itself? There are basically two methods which I think are getting pretty familiar to you now. One method starts with the individual loci: generally we look at either the DNA or the protein or some fairly simple direct effect of one of those. We may want to look at particular genes like the major histocompatibility loci (MHC). These genes determine tissue typing (they are what you have typed if you get a new kidney) but they are also of great interest in mammalian conservation: variation at MHC loci affects reproductive success and disease resistance which are obviously things that we care about. Other types of variation include mini-satellite variants, sex linked variants, and mitochondrial variants. We may not be interested in the direct effects of these, but they are markers for particular changes that we might care about: they are very good for identifying relatives of particular individuals and so monitoring levels of inbreeding and various other things we will come across during the course of this talk. Mitochondrial DNA is also particularly useful for identifying genetic stocks (Ovenden, 1990).

When analysing data from individual loci, we often try to summarise many loci, partly for statistical power, and partly in the faint hope of getting a picture of the genome as a whole. I am going to talk of two methods of summarising: average heterozygosity, which I think you

heard about yesterday, and  $n_a$ , the actual number of alleles. Average heterozygosity is used a lot as a measure of genetic variation. The actual number of alleles is not used so much but it is valuable because it is very sensitive to the number of rare alleles; if you are aiming to breed for some characteristic such as disease resistance, rare alleles can be very important. Perhaps the reason  $n_a$  is unpopular is that conservation of the total number of alleles requires a huge population size. Another way of making use of this data on variation at individual loci is to estimate gene flow; I gather you had a lively workshop yesterday on estimating gene flow from  $F_{ST}$ , and other methods.

A second method of measuring genetic variation is to start with the characteristics which we consider to be important (growth rate, disease resistance, etc.), and work back to the genetic basis of these (if any). The ways of doing this include breeding experiments (not always possible), relocation experiments, and searching for association between variation in the characteristic and genetic markers - molecular biology is opening up a flood of new markers for the latter approach. The difficulty of this work is offset by the value of the data; many traits, when analysed this way, have been shown to have some genetic component (ie: they are "heritable"). In fish, traits which are heritable and relevant to the management of a wild or captive population include: adult body weight, adult body length, and "meatiness", Allendorf *et al.* (1987).

How heritable these traits are I will leave to the experts, but they all have some genetic component in some species.

While discussing variation of measurable traits, I must mention meristics and fluctuating asymmetry. Meristics are the things that you can count, like the number of gill rakers on each side or the number of fin rays. Meristic variants may not be something that we are terribly interested in as production characters, but they are laid down very early in the life, so they are very strongly genetically determined; therefore they are very useful for monitoring genetic changes in a population. One way of analysing these data is to quantify fluctuating asymmetry (FAS). To search for fluctuating asymmetry, you measure something or count it on each side of each animal and take the difference e.g. left arm length minus right arm length. Fluctuating asymmetry is when these differences show a normal curve distributed about zero. The simplest measure of FAS is the size of the standard deviation; an increase of that standard deviation can

result from environmental or genetic stresses which may upset developmental processes. It probably would be reasonable to say that environmental and genetic stresses are the cause of all extinctions, so it's nice to be able to measure something (FAS of meristics) that responds to these stresses.

Leary and Allendorf (1989) recently summarised studies of FAS; examples included three-spine sticklebacks *Gasterosteus aculeatus* which showed increased fluctuating asymmetry in lateral plate counts when subjected to industrial waste (an environmental stress). In these sticklebacks, and other species, increased FAS was associated with genetic stresses of various types: hybridization between marine and freshwater form, decreased variation, or strong selection (Clarke *et al.*, 1986, Palmer and Strobeck, 1986, Leary and Allendorf 1989). We'll mention FAS quite a bit later on.

#### EFFECTS OF CHANGES IN GENETIC VARIATION

Now we want to look at what are the effects of increased or decreased amounts of variation. Allendorf and Leary (1986) give many examples supporting the notion that heterozygotes at single or multiple loci have superior fitness relative to homozygotes; many fish and marine invertebrates are included in their tabulation. Of fourteen cases reported, twelve indicated heterozygote superiority in some component of fitness. Later, I will discuss a few examples where high heterozygosity actually has deleterious effects. Beneficial effects do seem to be more common though: is this just selective reporting of data? To answer this, we must consider the mechanism by which high heterozygosity at a number of loci might be beneficial: Koehn *et al.* (1988) explained that we are beginning to understand theoretical and empirical reasons why heterozygosity at a number of loci may make more efficient individuals.

Closely related to changes of heterozygosity is changes of levels of inbreeding. Inbreeding, as you probably all know, is mating between relatives; with small population size, individuals are more likely to mate with a relative by chance. Aside from causing a loss of variation in the population as a whole, this can produce increased inbreeding: an increase in the frequency of homozygous individuals (who have two copies of the same allele). Inbreeding can produce undesirable effects, called inbreeding depression. In virtually every case out of fifteen lines of ungulates (mammals) the lines which were more inbred had lower fitness (as measured by

fertility or viability of offspring) than outbred lines of the same species (Ralls *et al* 1978). In a second example of inbreeding depression (Gall 1987) five teleosts and one oyster were tabulated; various characters which are related to growth and reproductive success were scored, and in 23 out of the 27 different individual investigations listed, performance decreased with an increased level of inbreeding.

## MANAGING GENETIC VARIATION -

### I. EFFECTIVE SIZE AND ITS MEASUREMENT

To avoid a decrease of genetic variation we need to have a high effective population size or  $N_e$  as it is sometimes called (Sherwin and Murray, 1990). A rough definition of  $N_e$  is: the number of individuals making approximately equal contributions to subsequent generations. If  $N_e$  is kept high you will never have to worry about inbreeding and loss of variation problems.  $N_e$  can particularly be a problem in very fecund species, like fish, because there may seem to be many, many individuals around but they may be the progeny of very, very few individuals and so the number of actual breeding individuals passing on their genes is very low. Therefore, the genes that are in the other potential parents are simply not being passed on.

To estimate  $N_e$ , you need to analyse three sets of factors. (i) Lifetime dispersal allows you to define the area in which there is apparent random mating, so the individuals are passing on those genes together. (ii) Population density and its fluctuations give an estimate of the number of individuals able to pass genes in each season. (iii) You need to look at the demography in detail: (a) generation overlap, (b) the lifetime production of young by individuals and particularly its variance, so you identify whether all parents make equal genetic contributions, and (c) the sex ratio - each sex will be equally represented in the next generation, half the genes came from males, half from females. So if there is one breeding male and a whole lot of females you will have very low representation of each female.

I'll give an example from my own work, of the way that you determine the effective size (Sherwin and Brown, 1990). As part of a larger study, I measured  $N_e$  in the eastern barred bandicoot *Perameles gunnii*, which is an endangered marsupial in Victoria (Table 1). The census size at the time was about 633 individuals, but there was considerable overlap of generations so when you look at the input each generation of new individuals (we counted them at the age of weaning) there is a much smaller number of individuals. Those individuals did not

contribute equally to the next generation; the production of young by different mothers is shown in Fig. 1. We followed a number of mothers from weaning to what we believed was the end of their life. One female who had a particularly good territory produced 18 young in her lifetime and chased away all comers. As far as we know, a lot of the females showed absolutely no sign of ever producing young (nipple enlargement in the pouch or evidence of the young themselves in regular trapping), so these females do not contribute at all to the effective population size. Other females have low contribution and one has a disproportionately large contribution; the result is that when we correct for variation in reproductive output we have much lower numbers than the raw number of weaners entering (Table 1). Then the correction for sex ratio which I mentioned before gives an effective size which is only one tenth of the census size. Effective size is usually lower than census, but the difference in these bandicoots is one of the most dramatic differences known for a wild population.

TABLE 1

---

CALCULATION OF  $N_e$  IN THE EASTERN BARRED BANDICOOT  
(HAMILTON, VIC) (from Sherwin and Brown, 1990)

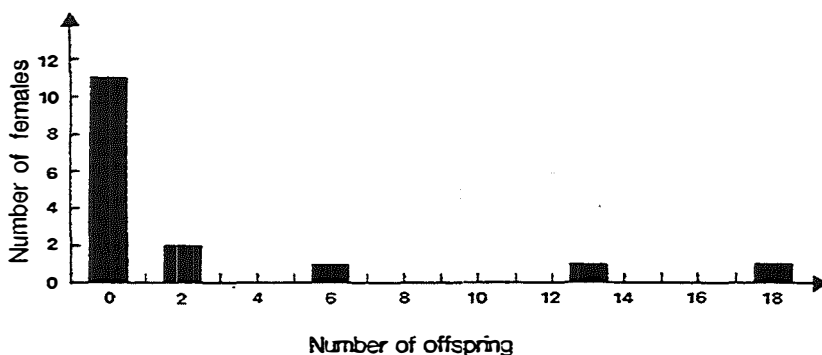
---

CENSUS SIZE		633±24
NUMBER OF NEW WEANLINGS PER GENERATION	male	153
	female	340
CORRECTED FOR VARIATION BETWEEN THE LIFETIME		
REPRODUCTIVE OUTPUTS OF INDIVIDUALS	male	24
	female	54
EFFECTIVE SIZE		67

---

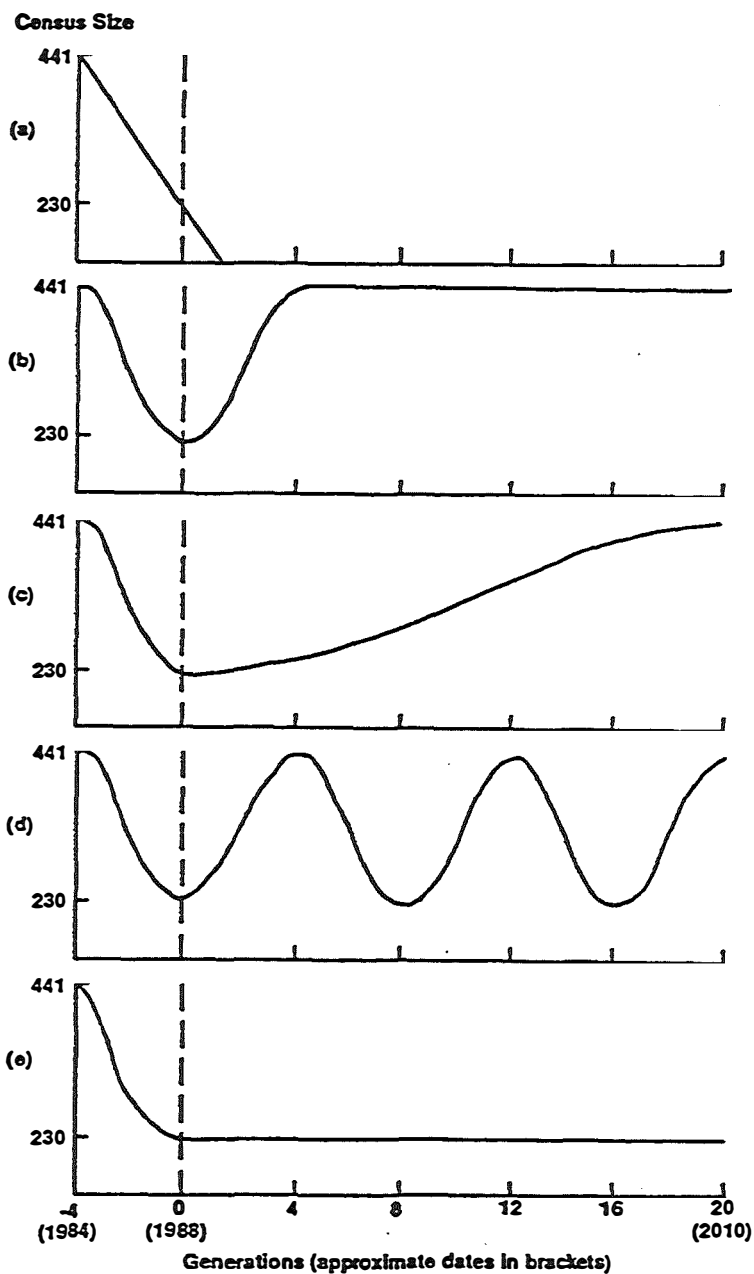
What will be the effect of this lowered effective size? The answer depends on how long the effective size is low, and how frequently this occurs. Figure 2 shows a few different possibilities projected for the bandicoot population. Figure 2 (a) is the worst one - the population is just about

extinct by next year on that projection. Figures 2 (b) to 2 (e) show different patterns of crashes and recovery, arranged in order of decreasing desirability, for genetic purposes. Figures 2 (b) and (c) show different rates of recovery. Figure 2 (d) shows repeated dips, such as might be attained by deliberately fishing an area very intensively then stopping until it recovers; these repeated crashes, exactly the same curve as 2 (b) but repeated, result in much greater loss of genetic variation. Figure 2 (e), no recovery, obviously causes a very big loss of genetic variation. Even 2 (c), one with a slow recovery over the projected time span, would cause a drop of heterozygosity or increase in inbreeding of 33%; if these animals respond to inbreeding in the same way as the ungulates as I mentioned before, there would be a 47% reduction in juvenile survival as a result of inbreeding depression.



**FIG. 1.** Lifetime production of pouch young by individual female eastern barred bandicoots at Hamilton, Victoria. From Sherwin and Murray (1989).

You can see there can be a nasty cycle: artificial changes can depress the census size, which automatically depresses the effective size (in the bandicoots' case  $N_e$  is one tenth of census). This leads to inbreeding which can therefore depress the census size because of lowered recruitment; there may be other effects related to the loss of the heterozygosity (Fig.3). The same thing can happen in a fish population. It has been suggested that what you need to avoid this cycle of inbreeding depression (Fig.3) is 50 individuals, or so, per generation in the effective population. To avoid a cycle due to loss of heterozygosity, the correct  $N_e$  is somewhat more arguable; it has been suggested that effective size should be in the hundreds. I have found one example of effective size for a fish: in the sockeyesalmon (*Onchorynchus nerka*) the local effective size is 200 and that is within a range that is probably safe from these genetic problems (Nelson



**FIG. 2.** Models of trends in the census size of eastern barred bandicoots at Hamilton. The maximum (1983-1985) census size is corrected from 633 to 441, using Dufty's (1988) estimate of the relative size of the expectation (230) and upper 95% confidence limit (330) of census size.

and Soule, 1987). Given our uncertainty of what size is needed, it is best to keep the effective size very big (several hundred), so you won't have to decide if genetic problems are likely to occur or not.

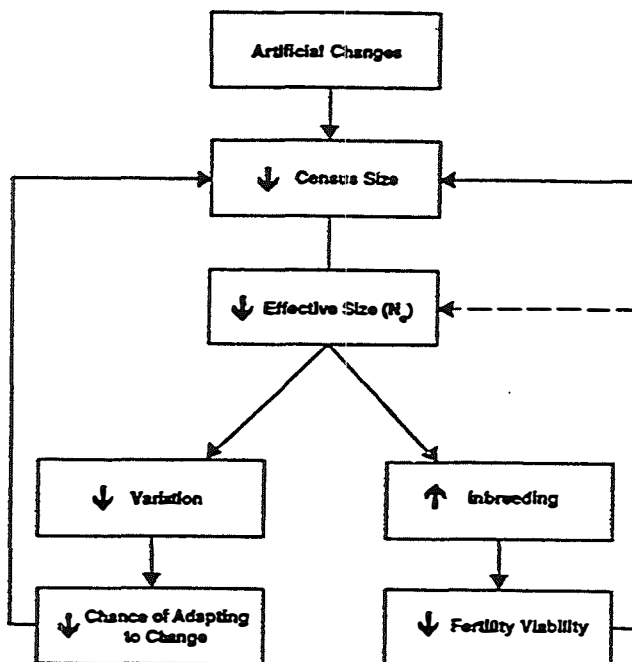


FIG. 3. Summary of conservation genetics. Artificial changes which reduce the size of a wild population can trigger one or both of the negative feedback loops shown, which may lead to a continued reduction of census size, even if the initial changes progress no further.

Getting back to the theme of measurable things, you will notice that all of the factors involved in estimation of  $N_e$  are routinely measured in ecological studies of managed populations; dispersal, density, fluctuations in number, reproductive output, sex ratio, etc. Measurement of these contributes to estimating the effective size, and on the other hand, genetic work can contribute to estimating these factors. There has already been a lot of discussion of the genetic estimation of dispersal at this conference. Genetic analysis can also help us estimate individual contributions to reproduction, for life table contribution. Useful analyses include fingerprint DNA, mitochondrial DNA, and sometimes protein analysis. There are also genetic ways of estimating the effective size, circumventing demographic work altogether (Pollak, 1983, Ball *et al.*, 1990). Whichever way we measure it, when we find out what the effective size is we may



wish to enlarge the population to avoid possible genetic problems - the connection between measurables and management is quite direct.

## MANAGING GENETIC VARIATION -

### II. SUBDIVISION

So far I have talked about the population as if it is a single population with no subdivision. What if there is genetic variation between subpopulations (stocks) as well as variation within subpopulations? There are basically three management options (a) manage subpopulations individually, (b) manage them as a single unit, or (c) deliberately mix subpopulations.

#### a. Separate management of subpopulations

If there is much genetic variation between the subpopulations, separate management will maintain whatever adaptation there is to local conditions. Differentiation of stocks may be spatially or temporally such as different spawning times (Fairbairn, 1981).

#### b. Management of multiple stocks as a single unit

Secondly, we can manage a subdivided population as a single unit even though there is high variation between subpopulations; this may be very appealing in particular cases. For example, Pacific salmon stocks are only separate at spawning time, when they have low catchability and low food value (Allendorf *et al.*, 1987), so to try and manage them as separate stocks would be very difficult. One option is to manage the species as a single unit and make sure that the intensity of fishing is such that the most vulnerable of the different stocks (e.g. the one with the lowest rate of recruitment) is conserved. This approach would mean that the other stocks were not being exploited to their fullest extent. Whenever managers aim to manage a species as a single unit, they should check first for variation between the subpopulations, looking at both genetic markers and any evidence of strong local adaptation, such as life history differences which were very suitable for particular spawning grounds. There is an example of the witch flounder *Glyptocephalus cynoglossus* of Canada where allozymes were used to define six parapatric stocks; and there was more than one stock in several of the previously defined management areas (Fairbairn, 1981). These stocks differed in: the proportion of mature

individuals, the spawning time, the growth rate, and temperature and depth of capture. Management as a single stock could eliminate one or more stocks through size selectivity of capture methods. Therefore, there would be a loss of fish adapted to a particular temperature and depth, and therefore a reduced trophic basis for the fishery as a whole, and a loss of useful genotypes for future management.

### c. Deliberately mixing subpopulations

The third possibility for management of differentiated subpopulations is deliberate mixing of the stocks; this happens quite frequently. The first problem is that the founders may not establish; Altukhov and Selmenkova (1987) transplanted eggs of chum salmon (*Onchorynchus keta*) in north western Asia. After their migration, the rate of return of fish from transplanted eggs was inversely correlated with the genetic distance between the eggs and the local population. The fish apparently were not returning anywhere else either, so the transplant just didn't work at all. A second problem is that when a transplant does work, if only a small number of individuals were taken, then not all the variation of the transplanted stock may be represented, which would reduce their chances of adapting to the new locality. A third problem is that if the transplanted stock happens to out-compete the local stock, we will lose whatever unique genotypes there were in the local population. A fourth, and final, undesirable result of mixing stocks is outbreeding depression which I will discuss in greater detail later on.

As an example of the problems of mixing subpopulations, I will present my own work on koala relocations. Koalas in south eastern Australia went to very low numbers in the first 100 years after European settlement and suffered a number of crashes; on islands in Western Port Bay, they not only had crashes of numbers, but also had relocations of very small numbers of individuals (as few as two) onto various islands and then further onto other islands. If you regard this as a hatchery there is very bad hatchery management, with repeated bottlenecks of very, very small effective size. Subsequently, because of freedom from a particular parasite koalas reached very large numbers on some of these islands; they overgrazed their food source and they are now still being used in relocations throughout south eastern Australia. They are going to areas where there were no koalas before the relocations, to areas where we think there were some, and to areas where we are reasonably sure there were some koalas before. We looked at 35 blood-protein coding loci and used allele frequency distributions to work out the number of individuals

transported and successfully breeding. We estimated that about seven to twenty individuals per generation successfully exchanged genes between south eastern Australian populations. (Ramus *et al.*, in prep.). That is a very high exchange rate for marsupials, or any mammals, so this particular management, with the bottlenecks in “hatchery” and then the relocation without regard to local differences, has produced an unusually low level of variation between these koala populations. We are trying to get samples from less perturbed populations in Queensland to compare with this result. What is more, these koala populations show various important differences in characters that matter for management of the koalas, such as the reproduction of females, the sperm morphology of males and disease tolerance; we are hoping with the complete dataset to be able to correlate these management factors with the genetic changes that we are quantifying now.

Returning to deliberate mixing of subpopulations, this will produce hybrids (if you allow products of interspecific crosses to be called hybrids), so I am going to look briefly at the characteristics of hybrids. Hybrids are not always intermediate (Neff and Smith, 1979): compared to the parental lines or the mean of the parental lines, they may have much higher or much lower scores of whatever character you are studying. For example, crossing rainbow trout (*Salmo gairdneri*) and westslope cutthroat trout (*Salmo clarki lewisi*) leads to a reduced growth rate relative to the parents, while the latter species crossed with yellowstone cutthroat trout (*Salmo clarki bouvieri*) gives hybrids with increased developmental rate (Allendorf and Leary 1988). When hybrids show a decrease in some characteristic relative to parental lines, this is known as outbreeding depression. This depression can result from (i) chromosomal problems producing low fertility of the hybrids, (ii) hybrid dysgenesis, which is an increased tendency of transposable elements to move around and disrupt the genome as they move, (iii) a loss of local adaptation, or (iv) breakup of coadapted gene complexes. A mammal called the ibex provides an example of loss of local adaptation (Templeton 1983). The Czechoslovakian ibex went extinct, and the Austrian stock was used to replace it. The Austrians bred nicely in Czechoslovakia. When the managers added some Nubian and Turkish ibex for good measure, they also bred, and the hybrids were fertile. However the hybrid rutted (went into breeding season) far too early in the year and the kids froze to death; this resulted not just in the extinction of the hybrids or the Nubians, but the entire population went extinct again (Templeton, 1987). This is an extreme example, showing how it is possible to disrupt local adaptation by introducing individuals with differently adapted genotypes. How frequent this problem is in fish I will have

to leave to the fish biologists.

Another reason which has been suggested for outbreeding depression is breakup of coadapted gene complexes. Whether or not these exist has been somewhat of a raging argument but the idea is that there are allelic combinations at different loci which happen to work well together in that species, and different combinations in another species or population. In the offspring of hybrids, these combinations start to be broken up, so the alleles don't work together as well. There is an example in a *Drosophila* species, an experiment of Templeton (1986), which shows that the percentage of hybridity of the genome is inversely proportional to the fertility of the females.

Hybrids are not always inferior: the alternative is hybrid superiority, or heterosis as it was called earlier today. This is quite well documented in domestic plants and animals, but if you've ever tried to use the seeds from a hybrid tomato on your dinner plate to grow a new crop of tomatoes you'll know that the superiority may only last for one generation, perhaps because of breakdown of coadapted gene complexes. There are examples of hybrid superiority in natural populations as well; two freshwater fish in South America, *Poeciliopsis occidentalis*, and *Poeciliopsis monarcha* have ranges which abut (i.e. parapatric) (Moore, 1977). There is a hybrid all-female form produced in the contact zone. The reproductive output per hybrid female is greater than the parents in two restricted localities near the contact zone. Note that it is only in restricted areas that the hybrids are superior; to rely on hybrids being superior everywhere is not a good idea.

With what frequency do hybrids occur in natural populations and how do we detect them, particularly in fish? Morphological methods alone are suspect because of the possibility of non-intermediacy of hybrids, although there are statistical ways you can get around this. Also, we can combine morphological information with data from (i) nuclear genes, measured by DNA or protein methods and (ii) mitochondrial DNA data, which sometimes show sharp discordance with the other two methods, giving evidence of past hybridization. A combination of all those methods is best for trying to analyze the frequency of hybrids and backcrosses. In analysis, it is best not to assume intermediacy of the backcross and hybrid. The results of this type of work are that hybrids seem to be quite frequent in fish, so perhaps problems with hybrids are not too serious in fish. Alternatively, external fertilization plus competition for limited spawning sites

may mean that mistakes happen more often in fish than in species with internal fertilisation.

### MANAGING GENETIC VARIATION - III SPECIFIC MANAGEMENT GOALS

There are three broad aims, to conserve an endangered wild population, to harvest a wild population, and aquaculture (Table 2). When managing an endangered wild population, the aims are essentially as I have stated already: keep a large effective size, try to maintain the natural structure, avoid hybridizations and avoid inbreeding. Managing a harvested wild population is very similar, but there is also the possibility of very strong artificial selection. This may be deliberate selection, or it may be inadvertent selection such as pollution, which has been shown to exert selective pressure on barnacles (Nevo, 1977). For exploited populations the harvesting itself can create large selection pressures. If a high proportion of individuals are removed at certain stages of the life-cycle, and if the removal is before the end of reproductive life, and nonrandom with respect to some kind of heritable variation, then there will be artificial selection. For example, if spawning time is heritable, and harvesting is mostly early in the season then there will be strong selective pressure favouring late spawning.

TABLE 2.

#### AIMS OF FISHERY MANAGEMENT

---

1. CONSERVE AN ENDANGERED WILD POPULATION
    - FOR ITS INTRINSIC VALUE
    - AS A GENETIC RESOURCE
  2. HARVEST A WILD POPULATION
  3. AQUACULTURE
    - LONG-TERM
    - RESTOCKING THE WILD POPULATION
    - PUT-TAKE FISHERY
- 

In the Atlantic salmon (*Salmo salar*), which is harvested in the ocean, there appears to have been selection for an increased proportion of males which become sexually mature very early, either

after one year at sea or without even leaving the freshwater (Allendorf *et al.*, 1987). This is presumably because of the selective pressure from the harvesting in the ocean, and it is a nuisance because the “average fish” spends less time in the ocean, so they become more difficult to catch. How do we avoid this? Nelson and Soule (1987) have come up with one suggestion which might horrify many fisheries biologists: they suggest deliberately avoiding size selective fishing, to the extent of opening the markets to undersize fish, but controlling the total level of fishing to a low enough effort so that the demographic characters of the population are not disturbed. Whenever there may be heritable variation which affects catchability, there will have to be a trade-off between avoidance of selection and other management goals.

There are three rather different aims in aquaculture (Table 2): long term aquaculture for food production, aquaculture for restocking of a wild population (perhaps an endangered one), and aquaculture for restocking in put-take angling. The genetic considerations for these three are rather different.

I will deal first with restocking of wild populations. The first consideration is the choice of founders. If there is only one remaining population, this is quite straightforward; take founders from the remaining population. If there is significant subdivision between different populations, then founder choice depends on the planned end-use of the hatchery stock. If you want to restock exactly the same localities, it makes sense to keep separate hatchery stocks for each of the localities that you are going to restock. Each stock would be managed essentially as described above for wild populations; I'll go into more detail of the management in a moment. The alternative is to restock vacant habitat, so you don't know what genotypes would have been there before. Presumably it's a good idea to mix genotypes from areas with habitat that is similar to the vacant habitat. These areas should not be too dissimilar to one another, or you may get outbreeding depression.

Having decided where to obtain your stock, the next step is to maintain the lineages in the hatchery; care must be taken to maintain the genetic variation that you may have sampled and this is not so easy. Allendorf (1987) has tabulated many examples of genetic changes in hatchery stocks, in some cases showing association with poor survival or reproduction. How do we avoid genetic changes in the hatchery populations? Even in the founding generation there can be much loss of variation, for example in lines of Atlantic salmon with an effective size (not census) of

6-46, 26% of alleles were lost in the first generation (Verspoor, 1988). A quarter of the alleles vanished in the first generation! Moreover, lines with lower effective sizes had worse losses. The advice that we've been given on the basis of farm animals is to try to keep the effective size greater than 50 at this stage. Next, we must consider subsequent generations; there can be further losses, and they will compound. It has been suggested that  $N_e$  should be much larger than 50 for the brood stock each generation. It is also a good idea to have regular input from the wild if possible, and maybe to keep sperm and ova in banks if it's technically feasible. These latter options may be cheaper than trying to maintain a very large hatchery line.

The next thing to consider in aquaculture is selection. How important this is depends on the aim of the hatchery management (Table 2). For restocking wild populations, we must avoid selecting for "hatchery types", such as individuals which are very docile, or spawn at a time which fits in with the schedule of the institution. Selection in these directions may minimize the chances of a successful release, so it should be avoided. However, there should also be active selection to eliminate obvious deformities.

It is not enough to plan the numbers of founders and broodstock, then hope for the best; loss of genetic variation is a stochastic process, so it must be monitored. The time scale of monitoring will depend on the biology of the species; generation length, iteroparity etc. Sample size (the number of individuals that are scored to monitor any changes) has to be considered very carefully, with advice from a statistician or population geneticist. Various traits discussed above can be measured; obviously fluctuating asymmetry is a good indicator of genetically induced problems. Protein, nuclear DNA and mitochondrial DNA markers should also be used, including (i) markers which vary in the wild population, so we can tell whether that variation has been lost, and (ii) markers which are monomorphic in the wild population, because if some other variant crops up in your brood stock then you know there has been accidental mixing from one of the other stocks or species in the hatchery. Therefore, it is important to thoroughly characterize the wild population, the founders, and also the brood stock each generation. It is useful to store tissues so that as new techniques become available we can use them (Sherwin, 1991).

The two other types of aquaculture that I have not mentioned in detail are long term aquaculture and put-take aquaculture. Compared to aquaculture for restocking, artificial selection and

manipulations like gynogenesis are less of a worry, but it will still be important to maximise the genetic base of the stock, and maintain this base, unless repeated input from the wild is cheap and easy. Notice that conservation of the wild populations will take on high priority if there is continued reliance on them for genetic input. The example given previously, of increased developmental rate in hybrids between fish from different localities, highlights the importance of wild stocks as resources for aquaculture.

## CONCLUSIONS

My final messages are to emphasise that (i) there are a lot of factors which are relatively easily measured to monitor genetic problems, and (ii) there are a lot of steps which can be taken relatively easily to avoid genetic problems. Therefore, it is best to start management and monitoring early to minimise these problems. It will be much cheaper and more efficient to avoid genetic problems when the population is very large, secure and unperturbed than to wait until the problem occurs after which it may be non-fixable or incredibly expensive to fix. Thank you.

## REFERENCES

Allendorf, F.W. 1987. Genetic management of hatchery stocks. In: Population genetics and fishery management (N. Ryman and F.M. Utter, Eds.), pp. 141-160.

Allendorf, F.W. and Leary, R. F. 1986. Heterozygosity and fitness in natural populations of animals. In: Conservation Biology: The Science of Scarcity and Diversity, (Soule, M.E., Ed.), pp.77-104, Sinauer, Sunderland, Massachusetts, U.S.A.

Allendorf, F.W. and Leary, R.F. 1988. Conservation and distribution of genetic variation in a polytypic species, the cutthroat trout. *Conservation Biology*, 2:170-184.

Allendorf, F.W., Ryman, N. and Utter, F.M. 1987. Genetics and fishery management: past, present and future. In: Population genetics and fishery management (N. Ryman and F.M. Utter, Eds), pp. 1-20.

Altyukhov, Y.P. and Almenkova, E.A. 1987. Stock transfer relative to natural organization, management, and conservation of fish populations. In: Population genetics and fishery management. (N. Ryman and F.M. Utter, Eds), pp. 333-343.

Ball, R.M., Neigel, J.E. and Avise, J.C. 1990. Gene genealogies within the organismal pedigrees of random-mating populations. *Evolution*, 44: 360-370.



- Clarke, G.M., Brand, G.W. and Whitten, M.J. 1986. Fluctuating asymmetry: a technique from measuring developmental stress caused by inbreeding. *Australian Journal of Biological Science*, 39: 145-153.
- Fairbairn, D.J. 1981. Which witch is which? A study of the stock structure of witch flounder (*Glyptocephalus cynoglossus*) in the Newfoundland region. *Canadian Journal of Fisheries and Aquatic Science*, 38: 782-794.
- Gall, G.A.E. 1987. Inbreeding. In: *Population genetics and fishery management* (N. Ryman and F.M. Utter, Eds.), pp. 47-87
- Koehn, R., Diehl, W.J. and Scott, T.M. 1988. The differential contribution by individual enzymes of glycolysis and protein catabolism to the relationship between heterozygosity and growth rates in the coot clam, *Mulina lateralis*. *Genetics* 118, 112-130.
- Nevo, E., Shimony, T., and Libni, M. 1977. Thermal Selection of allozyme polymorphisms in barnacles. *Nature*, 267: 699-701.
- Nelson, K., and F.M. Utter. 1987. Genetical conservation of exploited fishes. In: *Population genetics and fishery management*, (N. Ryman and Soule, M. Eds.), pp. 345-368.
- Leary, R.F. and Allendorf, F.W. 1989. Fluctuating asymmetry as an indicator of stress: implications for conservation biology. *Trends in Ecology and Evolution*, 4: 214-217.
- Moore, W.S. 1977. An Evaluation of Narrow Hybrid Zones in Vertebrates. *Quarterly Review of Biology*, 52: 263-278.
- Neff, N.A. and Smith, G.R. 1979. Multivariate analysis of hybrid fishes. *Systematic Zoology*, 28: 176-196.
- Ovenden, J.R. 1990. Mitochondrial DNA and Marine stock assessment: a review. *Australian Journal of Marine and Freshwater Research*, 41: 835-853.
- Palmer, A.R. and Strobeck, C. 1986. Fluctuating asymmetry: measurement, analysis, patterns. *Annual Review of Ecology and Systematics*, 17: 391-421.
- Pollak, E. 1983. A new method for estimating the effective population size from allele frequency changes. *Genetics*, 104: 531-548.
- Ralls, K., Ballou, J.D. and Templeton, A. 1988. Estimates of lethal equivalents and the costs of inbreeding in mammals. *Conservation Biology*, 2: 185-193.
- Robinson, N.A. and Sherwin, W.B. 1990. Applications of conservation genetics to the eastern barred bandicoot, *Perameles gunnii*. In: *Management of small populations* (Clark, T. and Seebeck, J.M., Murray, N.D. and Graves, J.A.M., Eds.) Chicago Zoological Society, Chicago.

Ryman, N. and Utter, F.M. 1987. Population genetics and fishery management, University of Washington Press, Seattle.

Sherwin, W.B., and Brown, P.R. 1990. Problems in the estimation of the effective size of a population of the eastern barred bandicoot *Perameles gunni*. in Hamilton, Victoria. In: Bandicoots and Bilbies, (J.H. Seebeck, P.R. Brown, R.L. Wallis, and C.M. Kemper (Eds.), pp.367-374, Surrey Beatty and Sons, Sydney.

Sherwin, W.B. and Murray, N.D. 1990. Population and conservation genetics of marsupials. Australian Journal of Zoology, 37: 161-180.

Sherwin, W. B. 1991. Collecting mammalian tissue and data for genetic studies. Mammal Review 21 (in press).

Templeton, A.R. 1986. Coadaptation and outbreeding depression. In: Conservation biology: the science of scarcity and diversity, (M.E. Soule, Ed.), pp. 105-116, Sinauer, Sunderland, Mass.

Verspoor, E. Reduced genetic variability in first-generation hatchery populations of atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Science, 45: 1686-1690.

**SESSION 7**

**PANEL DISCUSSION**

**Convenor: Dr. Bert Sheridan**

**Fisheries Research Institute  
Cronulla, N.S.W.**

## PANEL DISCUSSION

### Genetics in Aquaculture\*

**Dr. Bert Sheridan**

**Fisheries Research Institute,  
Cronulla, N.S.W.**

**DR. BERT SHERIDAN:** Good afternoon and welcome to this afternoon's panel discussion on genetics in aquaculture. Most of you don't know anything about me so, briefly I'm an expatriate poultry and pig geneticist with a bit of *Drosophila* genetics thrown in who's spent about a year now being associated with fish. So when I say I don't know very much about fish genetics I am being honest rather than modest. Anyway, we have here three speakers from earlier today and without further ado I'll throw this session open for your comments and your questions.

**PROFESSOR STUART BARKER:** Can I come back to this question of selection for growth rate? John James produced one set of data indicating that substantial selection progress had been made. Colin Purdom said he believed that very little progress had been made in many cases. Colin referred to the Israeli selection experiment on carp. If I remember that paper correctly, they found reasonable estimates of base population heritability but the realized heritabilities, while not zero, were smaller than would have been predicted from the base population estimates. Now if that's the case, it is not really all that unusual. Therefore, can we have some further discussion or can we get any other examples, or any other information on results of selection experiments for growth rate in fish species.

*\* Panel Discussion at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990.*

DR. COLIN PURDOM: Can I answer that? The realized heritability for improved growth rate was, at best, zero. There was some progress for slower growth, no progress for improved growth. That was in carp, I think they got no progress at all in *Tilapia*. But certainly the work took place a long time ago and I think it is significant that they gave up their endeavours to improve carp through selection and turned to cross-breeding. This is really my theme; if there's been progress, nobody's making any use of it and that I find surprising.

With the present crop of commercial salmon farmers, you don't expect them anyway to write papers about the results they produce. Certainly the Norwegian academics talk a great deal about improved strains but my understanding is that they are not used by industry. Neither do the academics use their joint expertise; there are two or three different groups involved, and they seem to work in opposition to each other rather than in collaboration. At a recent ICES working group on fish genetics in aquaculture, Noel Wilkins stood up and asked after twenty years of earnest endeavour where is the evidence of improved growth rate in salmon and there was a deafening silence.

Can I add one tiny bit to the discussion? If you plot the growth in length of fish against time you get a straight line with slope  $x$ . If you divide the population into two, the biggest half and the smallest half, and you continue to plot length you get two parallel lines but  $x$  remains the same, i.e. no differences in growth rate.

In all fish farms, the manager has to look at how tightly he can commercially crowd his fish and use his space efficiently. Stress therefore becomes a very basic factor. So genetic selection for stress tolerance seems important. I would propose at this stage that it's one major factor that should be looked at.

PROFESSOR STUART BARKER: I wouldn't disagree with that. However, I find the conclusions on growth rate unacceptable. It may well be a question of definition - what is growth rates and what is being selected? Some traits have responded well to selection and it's quite likely that others such as stress response would also do, but I would expect growth rate also to respond.

PROFESSOR JOHN JAMES: How do you explain the results of the selection experiments, then?

DR. COLIN PURDOM: To start with there are very few selection experiments. The best of them was that by Rom Moav and his team with carp, which has an advantage over the salmonids in that it's not cultivated intensively but extensively rather like your sheep and therefore the question of adjusting ration size is not so acute. That work did not support the concept of realizable heritability for improved growth rate.

The other studies with salmonids have been much smaller and scrappier and have this, I think, built-in bias that you have to deliberately decide how much to feed a tankful of fish. If you decide to feed it on the basis of body weight then that immediately determines the growth rate of the fish in the tank. The possible way round that is to feed to satiation and that is more a measure of the patience of the person doing the feeding than the appetite of the fishes. The rainbow trout's ability to feed is legendary, they really are quite voracious. If you want really to grow rainbow trout fast you can, at reasonable temperatures say 15-18 degrees centigrade, double their weight each month and individual specimens can attain weights of 7 and 8 kilograms within 2 years of birth. But that is way beyond anything that is produced in farms where the intention is to maximize the total output of a tank or a pond in relation to the water that's available to it. That's the limiting factor in that sort of intensive fish cultivation. It is not the individual growth rate of the fish, *per se*.

PROFESSOR STUART BARKER: John James, have you got any more on the Norwegian selection experiments?

PROFESSOR JOHN JAMES: No, I haven't but I believe they're not the only ones who have had success in response to growth rate selection. I'm not familiar with the actual details of the selection experiments. It may be that in fact all of the experiments except Moav's are defective but I don't know enough about the details of them to say that. I think that there are perhaps some 10 or 12 selection experiments which have shown success in response to selection for growth rate as against one or two that haven't and why there is the discrepancy I don't know.

PROFESSOR ROSS CROZIER: Would you care to comment on the agreement between

predicted responses to selection and the actual responses to selection, because if the theory wasn't any good it really would seem amazing that you should come by it fortuitously.

PROFESSOR JOHN JAMES: Well it wouldn't be so amazing in a sense, given that both the heritability estimates that people have are subject to appreciable standard errors and the responses to selection are also subject to standard errors of some reasonable magnitude. So that what would appear to be reasonable agreement could, in fact in the end, be good agreement or poor agreement involving the true parameters.

But, in at least some of these experiments, the apparent agreement, the agreement between what's estimated and what's observed, is quite reasonable. In that sense, it seems to be quite reasonable to say that it appears that the agreement is quite good and therefore the predictions work. What the causes of this are, and we do have the question of what actually is happening on the ground, is something that I don't know. If it is in fact the question of feeding to growth rate, then I would say this, that if that is the system then the fact that the larger animals have been fed more is not necessarily a biasing factor because to be fed more they have to be bigger in the first place. If they are not bigger in the first place they are not showing a response to selection. If they are bigger in the first place they are showing a response to selection and whether you are then spreading that response by the feeding regime is simply a question of what the definition of the environment is. I would not regard that as invalidating the selection experiment because if the response is there it's there. If there is no response there then feeding to growth, feeding to size, will produce no difference between the selected and the control groups.

DR. COLIN PURDOM: The difficulty there is that the diets that are given to salmon and to trout in commercial practice are way below the optimum level and there is always scope for giving a bit more food and getting a bit more growth. If they were feeding at the maximum level, what you just said John, would, I think be correct. But that is not the case. The diets are always substantially below what they're capable of eating.

DR. JANE ANDREW: We've been trying to raise rainbow trout which have been held for a number of generations. For the last, say four or five years, we've been raising them in the sea. Now, quite understandably, they didn't respond terribly well, so far, to salinity. But we've been trying just in the last few years, to pond brood stock in the sea in the hope that there would be

some genetic response, and that the progeny would then grow better and survive better in the sea. Do you know of any similar situations?

DR. COLIN PURDOM: We do grow rainbow trout in seawater and most of our strains were tested for ability to transfer to seawater with minimum mortality and they all perform quite well if the transfer is progressive; for example, it takes maybe two or three days to go from fifteen parts per thousand to thirty parts per thousand salinity. The brood stock don't like it in saline water, the males in particular tend to have osmosis problems and die. I don't think that has ever been viewed in a genetic context. But certainly we don't have any problems putting rainbow trout in salt water. It's easier, in fact to put rainbow trout into salt water than to put salmon. Salmon have to go through the physiological process of smolting before they can be transferred. Rainbow trout needn't be.

DR. JANE ANDREW: Yes, they do acclimatize eventually, but what we found was that the progeny would actually perform better in the sea water compared with the brood stock which had been held and raised in sea water before maturation.

DR. COLIN PURDOM: No, I don't think it would make any difference. Unless you believe in Lamarckian theory!

DR. JANE ANDREW: With triploid hybrids, has there been any work done on the reasons for increased survival in triploid hybrids over diploid hybrids?

DR. COLIN PURDOM: I think the general finding is that the viability of triploid hybrids is intermediate between that of the hybrids and the parental group that provides the female. I know that in a lot of hybrid work you don't get neat intermediacy but in a lot of others you do. Most of the hybrids that I've seen are precisely intermediate and very often the triploid is again halfway between that and the parent just on a sort of straight gene count, two thirds of plaice and one third of flounder, or two thirds of brown trout and one third of arctic char, or whatever. That's my finding anyway.

DR. BILL SHERWIN: This intermediacy is presumably a count of the allele frequencies rather than performance in growth, breeding time etc.



DR. COLIN PURDOM: Quantitative gene expression is found for pigment patterns which, although not simply inherited, certainly are inherited. But for growth rate, using the plaice/flounder hybrids we found intermediacy. Growth rate depended upon the amount of plaice genetic material.

DR. CRAIG MORITZ: There certainly are some cases in natural unisexual species of fish and also in frogs. In Elinson's work on frogs the diploid hybrids were largely inviable. He demonstrated true triploid rescue effect, I think it is called, where you add one of the parental genomes back in resulting in two balanced sets of chromosomes. It seemed to make the thing work whereas the diploid hybrid didn't. I may as well follow on with the question and we'll come back to that.

DR. COLIN PURDOM: I don't think it's a question of having a balanced pair of chromosomes. In the triploid there is no pairing or anything of that sort, it is I think just straight quantitative effect. But I think most of the French work on allotriploids, as they call them, is supportive of what I've just said that you get an improvement in the viability in the general direction of the female part of the hybrid combination. As well as this, we have tried to produce allotetraploids for exactly the reasons you've advanced. Because one would then hope to make what otherwise was a sterile hybrid fertile and capable of use in breeding programs. But it still didn't work.

DR. BERT SHERIDAN: Perhaps the problem in the lack of progress in some of the selection programs, comes back to an earlier comment about inappropriate experimental design in terms of trying to control the positive effect of the environment on the growth rate of the slightly larger fish. We seem to have the situation where the faster growth rate of the larger fish in a group can be due to their superior competitive ability rather than due to a superior genetic potential. Would anyone else care to comment?

DR. COLIN PURDOM: It doesn't happen in sheep?

PROFESSOR JOHN JAMES: It happens in a number of species, and if there is a serious problem through competition then probably the most direct way out is through selecting groups of individuals, to form a group, allow the competition to take place within that group and then

select the best group. In particular, there would then be no advantage to the best competitor within a family group over the worst competitor. There would not, as a consequence of this, be any automatic selection in favour of increased aggressiveness in the population and in fact to the extent that a particular family, let's say, had a lower level of aggression than another and this reduced stress contributed to an increased average growth rate in that family that could lead to a reduced level of aggressiveness. But I don't know whether this would work. It's something you would have to try before you could be confident of it.

DR. COLIN PURDOM: I think this plan would face a problem of common environment effect. Most families of fish would show a maternal effect for a start. Since you'd have to raise them to a certain size before you could mark them in order to mix them together, you would have a maternal effect and a common environment effect which would probably give you a greatly enhanced estimate of heritability. However, I think what would happen again is that you would get no selection progress.

PROFESSOR JOHN JAMES: Well there are a couple of simple ways out of that. One is to use sire families instead of full sib families by which we'd get rid of the maternal effect and the other one is to replicate tanks so that you spread the sire family over a number of tanks. This, of course, calls for more facilities and it depends on how keen you are on getting the results. There are methods available which can overcome these problems. It's a question of whether it's worthwhile putting the facilities to work to overcome the problem or whether you're better putting your facilities to work to do something else.

DR. COLIN PURDOM: I was delighted to hear you give that simple solution to the problem because that's exactly what I had proposed at the last ICES working group on fish culture and genetics. If we are to produce reliable heritabilities that we base the estimates on sire groups only and not use maternal groups. But, you see what then happens is that you get negative results and nobody likes publishing negative results. This may be another reason why you see more examples of significant heritabilities. No one likes publishing negative results.

PROFESSOR STUART BARKER: There are actually quite a lot of sire component estimates of heritability in the literature.

DR. COLIN PURDOM : Not in my experience. They're mostly midparent or female.

DR. BERT SHERIDAN: The mention that you made of aggression being involved as a behavioural trait reminds me of some genotype by environment interactions observed with poultry. There was a comparison done some years ago with layers housed on litter, where the relationship between aggression and laying performance was examined. Within any one pen, the more aggressive hens tended to have superior production. However, the more productive pens tended to contain less aggressive groups of layers. Thus, when one selected for increased egg production from hens housed on litter, one tended to select the more aggressive hens rather than those with the best genetic potential. When cages came in, this problem was overcome by testing the birds in individual cages. Thus, one can inadvertently run into behavioural problems in a breeding program in that the more aggressive and more successful animal isn't necessarily the best animal to have in a group production type environment. Do we have any other comments or questions?

DR. NEIL MURRAY: Would you mind if I resurrect a question that was dropped a while ago, and that was the question of whether we were just talking about freshwater fish here or, aquaculture in general? We could talk about selective responses in growth rate in other organisms that aquaculturists might be interested in. Certainly I know one case in abalone, Japanese aquaculture attempts succeeded in producing a very rapidly growing strain of *Haliotis discus*. I'm very interested in any information on what the experience is with selection in crustaceans and other molluscs.

DR. COLIN PURDOM: There has been a big program to improve *Macrobrachium*. I don't know what the result is. When I last heard about it, which was probably three or four years ago, the plan was to bring in *Macrobrachium* individuals from as wide a geographic range as possible and by a sequential series of hybridizations to produce a really mixed up pool and then to start with that. It is the opposite of the coadaptive gene pool concept but the programme was just beginning and I've no idea how far things have gone.

On molluscs, I'm surprised about the *Haliotis* result. I'd like to see the paper. We did some attempts at parent offspring correlations with *Crassostrea gigas* twenty years ago. It was a program that lasted about ten years and that was completely negative and not published.

DR. BERT SHERIDAN: Any more comments or questions?

DR. PETER SMITH: There's concern in some parts of the world about the genetic impact of introduced stocks on the native gene pool. I wonder if you'd like to comment.

DR. COLIN PURDOM: As I said this morning, there does appear to be some concern in the UK and elsewhere about salmon. It seems to me that this was a very abundant, very tough species. It survives, despite considerable damage to its environment by man in various ways and despite extraordinarily high levels of fishing pressure, 85% in some cases. So it's not an easy pushover by any means. It has also successfully colonized a very large area of the globe since the recent ice age, ten thousand years ago. So the likelihood that it has developed very highly specialized genotypic arrays to support it in its very brief stay in freshwater seems to me to be not very believable. Even if you do accept that individual rivers have some genetic adaptation, that certain genotypes are preferred by selection pressure, then introducing some alien genotypic arrays will increase the genetic selection and simply return the thing to where it started. So from whichever way you look at it, it doesn't seem to me as if the noble salmon is about to become destabilized by the introduction of a few genes from so called domesticated species. Now, that's for salmon.

This summer in Lancaster, the Fisheries Society of the British Isles have a symposium on rare fish and there will, I'm sure, be papers that refer to the possible problems arising from introductions of alien genotypes into populations of very rare fish, but that is a different circumstance. Does that answer the question?

DR. PETER SMITH: Yes.

DR. BILL SHERWIN: Could I perhaps make that question a little bit more specific? In your work in various species, you obviously use crosses between different races, and I'm wondering what kind of value you'd place on those different races for your work. How many different races per species is it worth looking after for your work? It's obviously very expensive looking after them individually.

DR. COLIN PURDOM: Yes, the work that we did with the different strains of rainbow trout was fairly expensive and that's one reason why my government has decided it can no longer continue to support it. We aren't continuing this work and what we've had to do is to send the stocks to enlightened trout farmers to use, hoping that they try to keep them pure. That's not easy for a commercial trout farmer. Yes, it is expensive maintaining individual strains. It's even more expensive trying to maintain individual selected groups of individual strains. It's even more difficult in a commercial farm keeping them that way.

DR. BILL SHERWIN: Thank you for that answer, that was useful. But it's not exactly what I was asking. I'm thinking of the wild populations as a resource for your work. Would you agree that they are a resource for your work and if so, how many different populations, genetically differentiated populations, of one species are needed for your work?

DR. COLIN PURDOM: I don't think I can answer that simply. However, if one examines salmon for the electrophoretic variance it is not possible to tell whether it belongs in North America, in the Baltic or in the North Atlantic area. None of these populations is distinguished by any particular single allele. In addition, within the United Kingdom, there is only a very low level of variation in salmon over the three or four loci that are polymorphic, but when we did investigate more closely we argued if there's going to be a racial difference it will be between the freshwater environments because the marine environment is common for all of these salmon from wherever they come from. So if there's going to be a selective force it will be in the freshwater life.

Within England we had the sharp discontinuity of the chalk stream compared to the northern spate rivers. The chalk stream is highly eutrophic, it's alkaline of course, very productive, whereas the spate rivers in the north are very oligotrophic and variable and when we looked at salmon from these two groups, that was where we found the difference, the major difference in the salmon stocks of England. But it was still a question of the allele frequency for AAT being something like 0.7 in the spate rivers and 0.1 in the chalk streams. So what I'm saying is the allelic resource is there in one strain, it is just that it varies in proportions.

DR. BILL SHERWIN: If I could try one more step of this question. The allelic differentiation may not reflect the differentiation at genes which are of significance to your work in selecting

for a particular character. This is what I'm really trying to get at.

DR. COLIN PURDOM: Yes, I think I know what you're trying to get at and I don't believe it. In a completely different context there is the argument that the Brazilian rainforests have to be saved because there may be a fungus that has a compound in its hyphae that is vital to medical research. I don't find that kind of hypothetical justification very convincing. I think there ought to be much better reasons for preserving the Brazilian rainforests than that speculation. Similarly for the genetic resource for something like fish, we are after all talking about a species which is very, very widely distributed and very, very successful.

DR. BILL SHERWIN: Perhaps it's not the appropriate species to ask about, but what I am interested in is how useful you've found it to cross two different populations, and as a result of that cross something useful appears which wasn't present in either of the two parental strains.

DR. COLIN PURDOM: I don't think I've ever experienced that sort of serendipity.

DR. BILL SHERWIN: But you haven't noticed the signs of it in crosses between different hatchery stocks.

DR. COLIN PURDOM: Different hatchery stocks perform differently, they look different, but none of that, I think, was consciously extracted from a known natural variation.

Could I enter another concept into the argument? Salmon have got 56 pairs of chromosomes. If we assume that each chromosome has one gene locus of which there is an alternate allele, the number of possible genotypes is simply  $3^{56}$ . I think that's, give or take a bit,  $10^{27}$  and there aren't that many animals on the earth. It therefore follows that the vast majority of the potential variation within genotypes is untested by natural selection or anything else. So should we be worrying about specific genotypes that may never even have faced the test of time?

DR. BILL SHERWIN: Well there are examples from other people's work, where crossing strains from different drainages results in an increased developmental rate, or something like that. But you have either never tried that or have never found it.

DR. COLIN PURDOM: We've never found it but hybrid vigour is often sought.

PROFESSOR ROSS CROZIER: Would this be an argument then for impossibility of selection if there's that much variation? No, I don't really see that knowledge for the question that he was asking. I'm not really quite sure of how relevant that is. If we look at other species which have also arisen and become highly modified during that same general period of time, like maize, people are desperately trying to conserve all sorts of stocks of maize. Although I wouldn't be too surprised if you carried out an electrophoretic survey it would become difficult to separate them, the same as you find separating your salmon. It's the loci that are affecting the characters of interest to you that are important, not the marker loci, which may not be reflecting what happens at the loci you want to know about.

DR. COLIN PURDOM: I was responding to the question, 'do you think there are particularly favourable genotypes?'. I'm saying that within this enormous variety there must be fabulous genotypes that have never been tested by evolution.

PROFESSOR ROSS CROZIER: I think that is absolutely true.

DR. COLIN PURDOM: That was the context of that argument. But, can I say for maize it must be rather easy, you just keep a bag of seeds tucked away in a locker but for fish you've got bigger problems.

PROFESSOR ROSS CROZIER: There is indeed a significant factorial problem although even with maize you find occasionally that the seeds don't last forever.

DR. CRAIG MORITZ: I think there is a place where this becomes relevant and that is what I alluded to at the end of the last section. It comes in where you have demonstrated significant spatial heterogeneity. What should our fallback position be? Do we try to maintain those stocks? If there is such a thing as local adaptation, the local stock could probably be the unit over which that adaptation occurred. Maybe our fallback position should be, if we don't know there's local adaptation, we should assume it's there. Therefore, try not to muck around too much with the overall map or the geographic pattern of variation in the course of our aquaculture or release of transplants or whatever we're doing.

DR. COLIN PURDOM: I think it is going to be a question of the cost of management options. Do you wish to conserve this array of adapted types or do you wish to conserve the fishery as a whole; they are not necessarily the same question.

DR. JOHN AVISE: I think there's another issue here that some might find trivial, but I believe is of considerable importance. Every time we conduct a transplantation or artificially move genes from one locale to another within a species' range, we're likely blurring or destroying the natural historical record of that species. An analogy might be the unauthorized movement of artifacts between archaeological sites, before a physical anthropologist had a chance to study those sites. A more blunt analogy is the burning of history books in a library. From genetic studies, we now know that in many species a very significant fraction of overall genetic variance is distributed among populations, and that this geographic population structure has a strong historical component. My concern is that transplantations are a powerful force for homogenizing previous genetic structure, as well as for accelerating the loss or extinction of overall gene pool diversity within a species. In my opinion, the burden of proof underlying the rationale for any contemplated transplantation program should fall squarely on the proponents rather than opponents of this strategy. Unfortunately, just the reverse is normally the case at present.

DR. BERT SHERIDAN: This may be the exception rather than the rule, but some months ago we had Walt Courtney from the United States. He is a freshwater biologist and he gave an example of what occurred some years ago when it was decided that largemouth basses were dangerously low in one of their freshwater river streams. So the biologists in their wisdom translocated some largemouth bass from another river. It appeared the ones they'd brought in normally spawned at a different time of the year from the ones that were there. The hybrids that were formed were very competitive and eliminated both of the parent species from that river system. However, the hybrids spawned at a different time of the year, again at a time when survival of the spat wasn't successful. So they ended up with no bass at all in that particular river system. Thus the indiscriminate translocation of fish between river systems can sometimes have disastrous effects although this could be the exception to the rule. Anyway, we'll move on.

DR. JOHN BENZIE: In relation to the comments for restocking and preserving natural populations, one consideration is the huge expense of maintaining stocks artificially. It is pretty



cheap to maintain them in their natural environment. You then also preserve a record of their history and other pieces of information that could be of value.

The point I wanted to take up this time is selection. Really there's one problem which hasn't been discussed and that is that in taking something into culture, and reproducing from a small closed population, some changes will occur. There have already been examples like, I think it may be *Tilapia*, where the management practices resulted in a loss or a reduction in the size of the fish with culture. In choosing the ones to breed for the next generation they chose the fast growers, and the fast growers were the ones who became reproductive first and at a smaller size. So the farmers were producing these nice little fish and those that they were harvesting became smaller. So obviously there was a response to selection there, an unintended one. But that would also lead one to believe that perhaps there should be responses that could be more positively dealt with. Most of aquaculture is concerned with primary domestication of organisms which are recognised as entirely wild. The aquaculture organism longest in culture has still been kept for relatively short periods, when compared with wheat, cattle or sheep, and are there any lessons to be learned about the history of sheep and cattle production? Was the first huge leap in production obtained from them, simply a fact that you reproduced the colonies that were productive in culture. Is this the sort of thing that we're seeing in aquaculture situations at the moment and, well, maybe I should get discussion on that. Are there any lessons to be learned from history? What sort of data is around for those early situations in agriculture?

PROFESSOR JOHN JAMES: Essentially none, because it all happened before anybody made any records of these things. So we really don't know about that. We do know that there are certain problems of adaptation which do arise when, for instance, you transport animals to a different country. If you take Australian animals to Southeast Asia where there's a whole range of different diseases that produces some difficulties which have to be overcome. I would normally expect to find that if you shift animals into a completely different environment they're going to spend a few generations adapting to that environment. This will happen regardless of whether you do anything about it or not. They'll look after that for themselves. This probably is very important but the problem is that really we don't know what it is that they would like to change so probably the best thing is for us just to leave it to them.

DR. COLIN PURDOM: The thing that bothers me about that, is that there must be a lot of

mortality somewhere to drive this adaptation. If it takes only one or two generations there must be enormous mortalities.

PROFESSOR JOHN JAMES: Not necessarily mortality, selection acts also through differences in reproduction.

DR. JOHN BENZIE: The use of polyploidization techniques, which result in immediate increase in yield and other obvious advantages in the short term, would be extremely attractive. The use of those now essentially locks things as they are currently. I think it is important to look at the sorts of management strategies that ought to be pursued now in order to prevent a huge genetic loss by the time, say in a hundred years, when naturally occurring resources have disappeared. Is it worthwhile putting a large investment, in some cases, to look in great detail at what you can do with selection, and once you've got those selected organisms, you use polyploidization techniques on those.

DR. COLIN PURDOM: The question is, what domesticated animal is facing problems from a deficit of genetic variance? That's the first question, the second question is how long is it before we get to that stage with fish? I think it is going to take a very long time. Farmed fish are still wild animals. But perhaps my colleagues can tell me what domesticated animals are now suffering from a deficit of genetic variance.

DR. JOHN BENZIE: Well perhaps the amount and quality varies. How many crops are now currently being looked at worldwide, to get those wild races for example found on the edges of fields, looking for genes to introduce into stocks of rice or fruit that have lost important characteristics relating to disease resistance etc.

DR. COLIN PURDOM: I was steering clear of plants.

PROFESSOR STUART BARKER: I think there is a very important point to take into account here in relation to what John Benzie just said. There are two aspects, one is the quantitative genetic variation for growth rate or such like traits that are continuously variable, the other relates to traits that are essentially due to just one or a few loci. Particularly in the plant world, this sort of exploration work is being done, primarily looking for single genes for disease

resistance, or other adaptive traits. There is a very important distinction to be made between maintaining or introducing new quantitative variation, and introducing new qualitative variation, i.e. specific alleles.

DR. JANE ANDREW: I'd just like to introduce another interesting idea which relates to the use of polyploids to maintain any line, whatever technique you use, you still have to keep the animal going. So it's not really a question of one or the other. In Tasmania, we're facing problems with conservation groups who don't want trout as an exotic species introduced to native waters, to our waters, and the use of sterile fish is of great value in this way so it's possible that doing, creating polyploids and having domestic strains which are sterile and stocking them into natural waters will have no effect at all. So we can have both things at once.

We've also found that hatchery strains, which in rainbow trout in Tasmania have quite a different behaviour, don't survive very well in the wild. This is what Dr. Purdom was also, I think, hinting at this morning, that there is a natural process which means that these hatchery or domestic strains don't necessarily pose a threat. We've actually found that in practice that in netting surveys, where domestic trout were stocked a few years ago, there is no sign of them now. Because they're voracious feeders, because they don't have the behavioural adaptation to the wild environment, they've actually died out very quickly.

DR. BERT SHERIDAN: Can anyone comment on that?

DR. MURRAY MACDONALD: I wonder if you could change the discussion slightly and get in to an area which doesn't seem to have been discussed much, if at all, by any of the speakers today. That's in relation to the use of genetic techniques or genetic markers as a tool for tracking the progress of hatchery bred fish which are then released into the wild and comparing their performance in a variety of ways to those of the existing wild populations. I know of some work that's being done particularly in Scandinavia in this area. I wonder if any of the main speakers or anyone else has any comments on that.

DR. COLIN PURDOM: I do not know about the Scandinavian work. We do, however, have an interesting situation on the river Itchen which is one of the prime brown trout rivers of the world. It also has a good salmon fishery, but because it is heavily exploited the managers have

been restocking with parr and smolts. We've begun to take an interest in this practice and are now tagging the released smolts and checking them electrophoretically. We have discovered to our delight that the smolts that are being introduced are very different to the natural smolts that are already there; they come from a hatchery in Scotland where the genotypes are characteristic of a spate river as I said earlier. The Water Authority is putting as many smolts into the river as exist naturally. So there's a very substantial addition to the gene pool.

It is a splendid experiment which we now have control of. Within the next few years we will be getting a lot of information on the survival of the introduced fish, the introgressive hybridization, if any, and so on. What is particularly pleasing is that if I had asked permission to do such an experiment it probably would have been refused!

PROFESSOR STUART BARKER: While we're on this point, can I throw the question back to Murray MacDonald and ask if he is asking the question as to whether this would be a desirable research or general management initiative? In other words, should we be getting this sort of data for the various fisheries of importance in this country?

DR. MURRAY MACDONALD: It's certainly my own impression that we should where we're considering the prospect of breeding stocks of particular species and releasing them into the wild and wanting to measure their performance against any existing natural populations. I was more interested in raising this particular application of genetics to aquaculture so that it could be discussed and perhaps gaps identified and future directions identified which perhaps will be discussed more in the next session.

MR. TOM BERGIN: Just to follow on this avenue of thought. What we are looking at is the possibility in a few years of restocking recreational or commercial fish from hatcheries back into the wild. To maintain such operations it would be essential to demonstrate that the experiment has contributed substantially to fish stocks. Now the only way we can see this happening is either by taking a genetic sample out of that stock, retaining it in some preserved form and, using the genetic fingerprints of that release, find out whether in fact you will recover an economically viable proportion of that release. Doing the whole exercise would either involve a microchip and you can't put them in very tiny fish, and they work out at around \$13/\$14 per head anyway. The only other alternative is tagging, which means you have got to wait until the fish is

sufficiently large to take a tag. The economics of holding for that extra time is probably also prohibitive for sizeable releases and so we believe that really at this stage the only alternative we can see is to develop some sort of genetic marking.

DR. COLIN PURDOM: It is possible to genetically mark a stock by selecting a very rare electrophoretic variant for example and make the stock homozygous for it. The passage of that variant can then be followed but it will become integrated with natural stock if the fish breed. The environmentalists, or the conservationists, will say that is a bad thing to start with anyway but it is very expensive. It would take perhaps two generations to produce the stock of homozygotes. It then has to be enlarged, so I don't think the genetic tag, the artificial genetic tag is the cheap option. I think the cheap option is the microtag. I don't know what species you are referring to, presumably not salmonids because they will take a micro tag within a few weeks of yolk sack absorption.

MR. TONY BROWN: I was not suggesting for a minute that they should look at my trout, however, after Tom's point we have for our own purposes selected a strain which has its own marker, namely that it has no spots and this can be included quite easily. I see no reason why other spot patterns which are inherent could not be used for this purpose. Some may not be quite so readily visible at close examination. Maybe even the number of scale counts between certain fins could be honed in on as a selection factor.

DR. BERT SHERIDAN: Maybe it would be worthwhile to discuss setting up a breeding program for fish and try to consider the various specific problems which one would have to take into account if we were going to adapt a program from sheep, or from poultry, or pigs or some other animal and then try to apply it to fish as such. I do not want to lead the discussion here but it seems to me that one of the problems lies with the testing environment in trying to eliminate competition. We have already discussed this to some extent. We have of course got the domestication of the wild species as such - the problem of tagging which we were just talking about - we have got what I could call the double edged sword, of the enormous fecundity rate in fish. You can mate a group - get some three million progeny and find that the three million progeny are full brothers and sisters. There seem to be some specific problems that perhaps it could be worth our while to try and consider that are perhaps unique to fish and not to other animals when it comes to designing a breeding program.

DR. ROY HARDEN JONES: We have let go quite unchallenged Bill Sherwin's first aims of fisheries management, which I have written down: Is to conserve endangered wild populations for their intrinsic value and as a genetic resource. Would you accept that as a reasonable paraphrase?

DR. BILL SHERWIN: I count that as one possible aim.

DR. ROY HARDEN JONES: Apart from two prepositions, it was what you had on your overhead.

DR. BILL SHERWIN: It was one of three aims on the overhead.

DR. ROY HARDEN JONES: I don't know anything about the aims of management for the states or territory, but they certainly are not the aims of the Commonwealth and in fact you are confusing ends and means.

DR. BILL SHERWIN: Would you like to state the aims of the Commonwealth?

DR. ROY HARDEN JONES: They are in the Fisheries Act. I haven't got a copy with me, but I will tell you what the main thing is and that is the rational exploitation of the resources, that is the objective. It may be that preserving the genetic resources is a means to that end but it certainly is not the identified one. If, in fact, you launch a crusade with the wrong objective you are not going to get very far.

DR. BILL SHERWIN: Well, I am trying, but I don't know a lot about fish. Is there no legislative mechanism for looking after any fish unless it can be exploited?

DR. ROY HARDEN JONES: Well, the Fisheries Act deals with commercial fisheries.

MR. TOM BERGIN: I go along with Dr. Harden Jones to the extent that fish covered under the Fisheries Act, the conservation is not the main feature they are talking about, sensible resource usage. But, fish are also covered under quite a number of other Commonwealth Acts; National

Parks and Wildlife Act being another one where they do specifically take on the role of conserving endangered species.

DR. JULIAN O'DEA: As one of only two people from AFS I think I ought to comment. As I remember it, there are two aims in the Fisheries Act; one is to conserve the resource - I think that should be made clear. The other is to achieve optimum utilisation, which is usually taken to mean economics, so what we look at is - protecting the biology of the fish, making sure that it continues to be there and protecting the economics of the fishery.

DR. MURRAY MACDONALD: As one of the state managers referred to by Roy, I would like to confirm that there are in fact both legislation and policy statements by Federal and State governments with respect to conservation of resources, including marine and fisheries resources and that if you are going to have rational management for a resource, whether it be for fisheries purposes or any other purpose, it must include conservation. By conservation, I mean wise use and there must be some genetic component to that conservation according to lots of the discussion that we have had here already. To lead on from that, I would like to ask a specific question of Bill Sherwin - if we might digress back to wild populations again from aquaculture or enhanced populations. Can you comment or do you have any ideas, Bill, on how we can measure the effects of fishing, whether it be commercial or recreational, on the genetic structure of wild populations, in other words, how can we measure selective fishing mortality? Presuming we can measure it, what kind of use can be made of that information by fisheries managers or could be made, and what kind of tools could they bring to bear on that problem to alleviate or to eliminate it?

DR. BILL SHERWIN: The simplest way is to score the characteristic which interests you (i) in the commercial catch and (ii) in fish caught in some other way, which you believe is not selective, e.g. a wide range of different mesh sizes for nets compared to the one mesh size used in the normal operation. Then you simply tally up the numbers caught by each method and you will find out if there is selection. The next question is: how much of the variation which you studied is genetic? If not much of it is genetic, then you can have all the selective fishing you like, and nothing will happen. So you also need to follow a time series to see if there is a change in the direction predicted by the selectivity of the fishing gear or other manipulations. Now we seem to have the legal aspects cleared up, I think I should repeat that I stated three aims for

genetic conservation of fisheries; the second aim (looking after an exploited population) has most of the same considerations as the objectionable first aim, except that we do also have to consider selection as we have just discussed.

MR. ANTONY BROWN: I would like to ask Colin Purdom his reaction to some research that occurred some fifteen or sixteen years ago. There was a gentleman who went on a Churchill Foundation tour of the United States looking at the salmon fisheries and there was a lot of concern about the limited genetic pool expressed there and their small rivers where some would run up, spawn and come back down again but the spawning beds were very limited in area. The effect that was that at the end of the day the last spawner was the successful spawner and that the subsequent generations that ran up the stream had a very high incidence of brother sister mating because they arose from the last fish that spawned. From my experience in Argyll, similar spate streams there might present that problem. Do you have experience of this at all or any knowledge of this?

DR. COLIN PURDOM: No, most of the UK spawning areas are a bit bigger than that. There is a certain amount of overcutting but nothing quite as extreme as that you describe and I would be very surprised if the spawning grounds in any viable salmon river was that limited. Can I add to that though, that I think the extent of inbreeding depression is very often overstated and my thoughts go back to the days when I worked with *Drosophila* and I believe the figure there for the beginning of perceptible inbreeding depression was about 75% which I think corresponds to six or seven generations of strict brother sister mating. We have for the last two years been running a similar program with fish, with guppies (*Poecilia reticulata*) and I am now up to generation 7 or 8 of brother sister mating. At generation 6, we conducted a whole series of matings between inbred, within inbred, and with the outbred base population and so on and from that array at generation 6 there is no obvious effect of inbreeding on the size of the parents, nor on the size of the broods and so, I was very surprised to see huge effects being claimed for inbreeding of one generation with an F of 0.25.

PROFESSOR JOHN JAMES: I think your experience with *Drosophila* must be rather unusual because there are I think a lot of experimental results in the literature which show that you don't have to wait anything like six generations of brother sister mating to get quite drastic reductions for instance in competitive ability in *Drosophila*. If you look at Latter and Robertson's work



you will find that the competitive ability dropped off extremely sharply and in fact the situation was that the competitive ability dropped off very sharply with inbreeding in the initial generations and then, it tended to drop off less sharply later on. The same thing is true - I think if you look at, say, work in Japanese quail, that shows a lot of drastic inbreeding depression. On the other hand, what you do find is that on occasions, you get an inbred line or a few inbred lines which show virtually no inbreeding depression and that can happen from time to time. But the usual story is that, certainly in *Drosophila*, you get for competitive index, reproduction rate and so on, quite drastic inbreeding effects. However, if you look at things like abdominal bristle number, you don't see any corresponding inbreeding effect.

DR. COLIN PURDOM: Would you like to put a figure for F for these circumstances? Is it 0.25 do you think? One generation of brother sister mating?

PROFESSOR JOHN JAMES: I think one generation of brother sister mating would certainly (from an outbred population) reduce fitness quite drastically in things that I have seen anyway.

DR. NEIL MURRAY: The whole question of whether inbreeding depression is a real problem has arisen a lot in the conservation of zoo mammals and there has been a lot of work put together looking at the effect from an outbred gene pool. This has been summarised by Ralls, Ballou and Templeton in *Conservation Biology* a couple of years ago. I think for forty species there is a significant inbreeding depression in juvenile survival in 38 out of the 40. There was wide variation. However, for an F at 0.25, it corresponded to about a 30% drop in juvenile survival on average. That ranged from almost nothing in tigers to very large inbreeding effects in some rodents. So, there is variation as John said, but I think there is concordance from wild studies of bird populations and mammals that inbreeding of the order of 0.25 is pretty serious.

DR. JOHN HARRIS: The subject of inbreeding depression is a particularly interesting one for people studying carp in Australian waters. The situation was that we had presumably a very small founder group of a new strain imported illegally into the country in the early 1960s and we probably have had seven or eight generations since. The early history of the spread of carp was usually described as being explosive and it was a fairly graphic description because the increase in biomass was phenomenal and their distribution increased rapidly. But over the past decade, there have been many widespread reports that the population numbers of carp are

declining, and declining in a way which seems to have followed the pattern of their spread. I would be very interested to know, whether that would be likely to be attributable to inbreeding depression? Secondly, if the authorities responsible for managing the importation of fish into Australia were to permit the widespread importation and use of the highly selected strain called Koi Carp, through aquaculture industries and the ornamental trade - and the almost certain result of that would be the escape of numbers of Koi into waters already containing inbred carp - would it be reasonable to predict a level of heterosis and some repeat of the previous explosion?

DR. BILL SHERWIN: Could you clarify what you mean by the level of reproductive problems following the introduction - do you mean that in the peripheral areas there is quite a problem with reproduction?

DR. JOHN HARRIS: There has been very little work to quantify any of those aspects. It simply seems that the populations of carp have changed dramatically, but in a rather consistent way, as they spread. There were extremely large populations which radiated out from their introduction area and subsequently there have been declines which seem to follow that radiating pattern. The declines were in certainly fish numbers, perhaps also in fish size.

DR. BILL SHERWIN: So, the central populations are the ones which are suffering the declines are they? Is it the margins which have the healthy populations?

DR. JOHN HARRIS: That has been the pattern suggested from a number of sources for many years.

DR. BILL SHERWIN: I would be surprised if there was much inbreeding or loss of variation in something that has an explosive recovery from a small bottleneck.

PROFESSOR ROSS CROZIER: Wouldn't it be more likely to be pathogens or adaptation of other local organisms to exploit this new source. It doesn't however, negate some thoughts along the lines that you mentioned since, after all if the pathogens have become adapted to the carp that were previously released and you then release yet another, you might get some further spread. But I don't think it would be inbreeding because, as you described it, after the initial bottleneck there was a great increase in size - populations should have maintained the variation

that was sufficient for their initial spread.

PROFESSOR STUART BARKER: I would like to raise a subject that has not been commented on. For the beginning of an aquaculture program, John James referred to testing available strains to determine which one or more of these strains might be most appropriate for initiating a program. He also referred to testing various crosses among strains to determine whether there is any significant heterosis or hybrid vigour that could be utilised in commercial production. In these aquaculture situations, is the manager or the aquaculturist likely to have many more strains available than he can readily imagine doing a complete comparative evaluation of all at the same time? If this were the case, then allozyme markers and genetic distance estimates could be used to divide the set of strains into subsets. Defining those subsets that are similar within, but different between, would be one way of assisting an initial rational decision among all strains to determine which should be tested. This is one possibility, the other is in terms of considering strains for crossing. Again, one does not have to have very many strains available before the number of crosses that are possible increases dramatically. Here, there is the possibility of using genetic distance as a predictor of potential heterosis. Taking the simplest hypothesis of heterosis, the further apart the strains are, the more different they are in gene frequencies, the more important is heterosis likely to be. Thus, allozyme markers can provide at least a potential basis for selection among strains for evaluation and for selection of strains for heterosis testing.

DR. COLIN PURDOM: I have not observed any heterosis in the crosses I have made between rainbow trout strains. Rom Moav did detect heterosis in crosses of, I think Chinese and European carps, but I don't think you can lay down any hard and fast rules. I don't think it follows that those that look more genetically distant because of electrophoretic differences are in fact likely to produce heterosis. It seems to me more likely that heterosis would arise if there has been some inbreeding in the past. Some geneticists distinguish between heterosis, which is the opposite to inbreeding depression, and hybrid vigour or luxuriance, which is what you very occasionally get when you cross two related species. But, as I say, in my experience it is not normal to get heterosis.

DR. STUART BARKER: I was not asking the question whether heterosis was important, but simply pointing out this would be a way of reducing the number of strains, or taking from a set of strains that were available, those for testing to determine the importance of heterosis. This

idea derives from some evidence that is available from cattle where there are a couple of cases where distances between breeds are in fact correlated with the magnitude of heterosis. So there is some empirical evidence supporting this idea, as well as the simple dominance theory of heterosis where the bigger difference in gene frequency, the bigger the expected heterosis.

PROFESSOR JOHN JAMES: I would agree that if I had some strains which were otherwise indistinguishable and I wanted to reduce the number of crosses that I was looking at, I might then use electrophoretic markers as a basis for it, although that does make the assumption that the differences in the markers correspond with the differences in the loci concerned with the characters we are interested in. But on the other hand, if there were differences in important characters between them, I would be picking the strains that were best anyway and only in the case where they were more or less interchangeable would I use the markers as a basis. But in that case, yes I would.

PROFESSOR STUART BARKER: I was making the case for a new program - starting with unknown strains, unknown materials and having to make some choices on the basis of no other information, then it would be better than nothing.

DR. BERT SHERIDAN: We have run out of time. I thank you all for your interest and I would like you to join me in thanking our panel for their contribution.

**SESSION 8**

**WORKSHOP - FUTURE DIRECTIONS**

**Convenor: Dr. Bob Ward**

**C.S.I.R.O. Marine Laboratories  
Hobart, Tasmania**

## WORKSHOP - FUTURE DIRECTIONS

### Conclusions - Past and Future Considerations\*

Convenor: Dr. Bob Ward

C S I R O Marine Laboratories,  
Hobart, Tasmania

DR. PATRICIA DIXON: Before we get this final workshop underway, I'd like to ask the person in the audience that's organized the funding for probably the largest, most extensive and most expensive genetic study of fish populations, whether he'd like to come to the front and just make a few comments that we might take into account later on when we get around to discussing things. Bob Kearney?

DR. BOB KEARNEY: The reason why I thought I was going to be asked to comment was in relation to the question that was asked yesterday, and that was, "what is it that managers really want to know in relation to genetic work?" I gave a rather cryptic comment at that time. I said that on this subject managers didn't really want to know anything. It was that statement that I thought I was going to be asked to address, and that is what I shall try and do.

While it is important not to be too pedantic about the choice of words, it is equally necessary to be sure that we differentiate between what managers want to know, and what they need to know. But before considering even these fundamental questions, it is necessary to define what type of fisheries management we are considering. The following comments are relevant to fisheries managers who are dealing with capture fisheries operating on wild stocks. The comments are

*Panel Discussion at Population Genetics & Its Applications to Fisheries Management & Aquaculture Conference • Workshop, Sydney, 25-27 June, 1990.*

not necessarily relevant to those who manage aquaculture or stocks in impoundments, or stocks which are significantly artificially enhanced.

When considering what managers should know I would like to refer to a statement from Stuart Barker's paper to this meeting. Stuart said that a sound understanding of genetic structure of exploited fish species is essential for development of appropriate fisheries management. This is a statement that many fisheries researchers and managers have made many times; in fact I have used it myself on numerous occasions. There is only one real problem with the statement: it is wrong. In theory, some aspects of fisheries management require a detailed understanding of the genetic structure of exploited stocks, and even the degree of mixing of separate stocks. In practice, this is very seldom known and yet appropriate fisheries management strategies are developed. Often fisheries can be managed successfully without any detailed knowledge of the stock structure. This does not mean that a knowledge of the stock structure is not desirable, but it does mean that it is not always essential for appropriate management. In reality, managers don't even consider stock structure unless scientists can convince them that it critically influences their management decision.

Unfortunately, in most cases fisheries managers' objectives are short-term solutions to existing crises. Often the crisis has to be solved without the knowledge of the stock structure; often the knowledge of the stock structure would have no impact on how this crisis was solved. My contention is therefore, that managers don't want to know anything about the stock structure of resources and they don't usually need to know. One would hope that somebody giving the managers advice would know something about the stock structure, or at least understand whether a knowledge of the stock structure would influence the final management decision.

Let us consider the type of information that scientists do need to give managers in the hope that this will lead to the correct management decisions. As an example, I refer to some data available on the southern bluefin tuna fishery which is widely accepted as one of Australia's most mismanaged fisheries; specifically the plot of catch against effort in the Japanese longline fishery for southern bluefin. In this fishery, over the period up to 1987, total catch and catch per unit of effort fell alarmingly and yet no serious management action was taken to restore the total catch to the maximum sustainable level, or significantly improve the efficiency of the fishery as indexed by catch per unit of effort. The stock structure of southern bluefin was considered

by scientists to be important, but as management would not take action over dramatic declines in the total productivity and efficiency of the fishery, it was clearly not going to make any difference what scientists said about the stock structure. In summary then, management did not want to know anything. Scientists even concluded that the two most common alternative stock structure hypotheses (one or two stocks) would have made no difference to the recommended management action, which was to dramatically curtail fishing effort in all sectors of the fishery. Therefore, it could be argued that management did not even need to know what the stock structure was.

I believe we must accept that management does not want to know anything about stock structure except in those few cases where scientists can convince managers that they have to know what the stock structure is to solve their short-term management problems. Therefore researchers can be assured that management is not going to come to you and ask you to tell them how you can assist with management unless they have a crisis. Fisheries managers have a tendency to stay with the status quo and only to change management practices when forced to do so. If scientists don't indicate to them that there is a real problem that must be solved, then you can be certain the managers won't come to you looking for problems.

Research managers, on the other hand, have a different problem, and therefore normally a different approach to that of fisheries managers. Research managers normally have a relatively finite amount of resources that can be allocated depending upon certain priorities or policy guidelines. They are far more likely to see the end for genetic research than are fisheries managers, and again more likely to be influential in giving such research the priority it requires to receive funding. I have assumed that most of the audience here are researchers and that most of you would benefit from increased funding, or security in funding, to enable you to continue to carry out the work so dear to your heart. Those of you who are waiting for managers to come to you offering money have a very slim chance of success. If you really wish to increase your chances for funding, then you have to convince research managers that you have a scientific solution to a problem that they have, and hope that they have a means, with your help of course, to convince managers that they need you.

Even as a research manager, I find it difficult to accept that a knowledge of stock structure is always critical to management advice. Let me use as an example the South Pacific skipjack



situation. Pat Dixon referred, obliquely, to this program in her introduction when she stated that I had in fact funded a great deal of genetic research work in fisheries. One of the things we did in this area was to contract Barry Richardson to carry out analyses of skipjack in an attempt to determine the precise stock structure of skipjack across the Pacific Ocean. This work was prompted because Kazuo Fujino of Japan and Gary Sharp of the United States had both hypothesised that discrete subpopulations of skipjack exist throughout the Pacific and that these sub-populations greatly influenced the way in which skipjack fisheries should be managed.

In considering the results of Barry's work, and our own tagging results, we came to the conclusions that the studies had been extremely well carried out and that Barry had fairly clearly demonstrated that the Pacific skipjack resource was not comprised of a single panmictic population. There was obviously population structure but it could not be decided whether this was due to the presence of two large separate populations at either side of the Pacific or some clinal structure across the whole of the Pacific. One could therefore argue that the original hypothesis on stock structure remained unanswered. However, the results did negate Fujino and Sharp's arguments that management should be based on the hypotheses of separate stocks.

I was particularly interested to hear Barry Richardson's comments yesterday when he said that there were three major issues which under-pinned the reasons for genetic work on skipjack: firstly, did we have a panmictic population; secondly, did we have a clinal situation; or thirdly, did we have separate stocks. Surprisingly enough, alternative answers to these three issues would presently have no impact whatsoever on the management of skipjack fisheries in the Pacific Ocean. The Pacific skipjack fishery is presently one of the world's largest with annual catches exceeding one million tonnes, and the South Pacific Commission's work indicates that it has considerable scope for development with potential yields of at least three million tonnes a year. Therefore, we are faced with the realisation that the management of one of the world's largest fisheries would not be in any way influenced by further knowledge of the stock structure. Perhaps what is even more surprising is that there is at present no significant research program anywhere in the Pacific targeted on skipjack. I do not mean genetic research but fisheries research of any type. The reason for this is that the resource is stable and management has no major problems or conflicts, therefore managers perceive that they don't need to know anything more about skipjack and therefore there is no need to fund skipjack research.

In concluding, I repeat my conviction that fisheries managers believe that they don't need to know anything about stock structure. Research managers want to know a great deal about stock structure and are more prone to be supportive of requests for funding in this area. The question of what managers need to know was raised because people wanted to identify ways of obtaining funding for continuing their research. My answer is that if you wait for managers to ask you for advice your chances of funding your research are very close to zero. If you really wish to obtain support then you must convince research managers that your particular line of research is relevant to the broader understanding of fisheries science. If there is any urgency in the funding request then you must demonstrate the relevance to a particular management problem that the research director has been told is a priority. Certainly the standard of excellence of your research will influence the decision. Those of you who wait to be asked are doomed. Those of you who believe that stock structure information is always essential for management are misguided.

DR. BOB WARD: Thanks very much Bob. I'm sure it's absolutely true that we have to convince managers that we have some technique, something to offer them, which maybe will help them manage a fishery. Clearly, managers have all sorts of things to take into account in managing fisheries and perhaps, as Bob was implying, genetic stock structure is one of the least of their worries. There are all sorts of socio-economic and socio-political factors that are taken into account in looking at where boundaries should be drawn between stocks and what sorts of catch yields should be employed.

What I wish to do now is to consider where biochemical genetics has been and perhaps where it's going, and give some conclusions about how it can help us look at population structure.

I'll start by briefly considering the past history of population genetics. Really molecular population genetics dates back a long time, back to 1900 in fact which is when Landsteiner started studying blood types in humans. So molecular population studies really pre-dated the re-discovery of Mendel's laws. As Barry Richardson outlined yesterday, in the 50's and 60's people started using electrophoresis to look at proteins in fish populations, mostly studying transferrin and albumen variation. Also, at that time, there was some work done on blood group variation in fishes but as soon as Lewontin and Hubby, working on the fruitfly *Drosophila pseudoobscura*, and Harris, working on man, as soon as they produced multienzyme studies of these two species showing that there was lots of genetically determined variation present then

fisheries biologists, fisheries geneticists, very quickly hopped on to the bandwagon. In the couple of years after 1966 there was an explosion of interest in fish population genetics, and in 1969, just three years after these seminal discoveries of Lewontin, Hubby and Harris, Wilhelmina Deligny wrote a review of serological and biochemical studies in fishes which ran to more than a hundred pages. So by 1969, a lot of electrophoretic work had already been done.

The next major technological step forward in the delineation of fish populations came with studies of mitochondrial DNA diversity, with John Avise's group leading the way in terms of population studies. That had the same sort of effect on population genetics as Lewontin and Hubby's earlier studies of protein variation. A large number of people then saw the advantages of looking at variation in mitochondrial DNA in analyses of population structure and evolutionary history.

I think perhaps the next technical advance, which hasn't been used a lot yet in population studies but perhaps will be in the future, is the technique of DNA fingerprinting that Alec Jeffreys pioneered around 1985. I was interested to see the poster here on DNA fingerprinting in barramundi. So far there has been little study of DNA fingerprinting in populations because in most species every individual turns out to be different from every other and it's impossible generally to work out allele homologies. However, such studies of hypervariable multiple or single loci are likely to offer further advances in analyses of population structure.

What I want to do now, is to ask John Avise to outline where the next advances in techniques may come from for further refining our studies on population structure or for extending those studies to look at species that it's hitherto been difficult or impossible to study in any great detail, for example endangered species. So, John, perhaps you could spend a couple minutes talking about future prospects here?

DR. JOHN AVISE: Since my comments are of an impromptu nature, I can only hope that something of interest may come out. I'm not a fishery manager, not a practising manager at all. I'm an evolutionary biologist interested in basic research questions. Thus, I feel a bit uncomfortable trying to convince managers that they ought to be listening to our field, but we really do have something to offer. We're trying to develop molecular methods and principles, and see where they may lead. I hope that someday such approaches will have many concrete

and routine practical applications, but the immediate rationale for development of the molecular population genetics field has normally been in the interest of pure science.

We are in a very exciting time now, where molecular techniques are being introduced at an incredible pace. Often, by the time a lab is converted to the latest new technique, it may already be dated, and another switch is required. So I'm not really sure where things are going to lead. There is a bewildering diversity of approaches to DNA and protein analysis and the laboratory methods seem to be accelerating in terms of the pace of introduction. One way I like to organize my own thoughts about this bewildering diversity of methodologies is to realize that almost all techniques in molecular evolutionary biology, when used in a descriptive context as opposed to the manipulative context of gene transfer, can be viewed as addressing issues of phylogeny and genetic relationship. In the broadest possible sense, these topics range from issues of identity versus non-identity to macro-evolutionary phylogeny. Let me list these (writing on board):

- a) genetic identity vs. non-identity
- b) parentage
- c) pedigrees within a population
- d) geographic population structure within a species
- e) species and higher level phylogenies

There are some genetic markers available that allow us to distinguish one individual from another, or to distinguish self from non-self in a tissue sample. Methods such as DNA fingerprinting can address that kind of micro-phylogeny problem. The next level involves questions of parentage, by which I mean establishing maternity and paternity. There are certainly many situations in population biology where it is of interest to identify the parents of particular offspring. Can we accomplish paternity and maternity exclusions? Ultimately, can we attempt paternity and maternity inclusions? Can we identify the parents of particular individuals when matings are not observed or fertilizations are questionable? These can be thought of as micro-phylogeny issues, where we are dealing with a single generation of transmission.

The next level up in this hierarchy involves assessment of genetic relationships among

individuals within a local population, or in other words to establish pedigrees on a micro-evolutionary scale. Will molecular methods become available that could establish whether individuals are first cousins, second cousins, and so on?

The next level of phylogeny is that of geographical population structure. I find it interesting that this topic is where almost all discussion at these meetings has centered - how is genetic variation distributed within and among geographic populations of a species? The final layers in the hierarchy are those of species phylogeny, and ultimately the macro-phylogenies of higher taxa. The only point I'm trying to make here is that there exist several levels at which one can talk about phylogeny (broadly defined), ranging from identity versus non-identity to macro-evolutionary relationships.

For this meeting, as I said, the level of intraspecific population structure is where most attention has been focused (appropriately so for concern with managing populations within species). I also happen to think that this is the conceptually most exciting area in this entire hierarchy right now, and may be so for the next few years. I say this because in some of the other levels we're probably in better shape conceptually and technologically. Thus for many years we have had numerous molecular methods for address macro-evolutionary phylogenies and species phylogenies - these include DNA-DNA hybridization, various immunological techniques such as microcomplement fixation, protein electrophoresis, and direct sequencing analyses either of proteins or of DNA. Researchers have been conducting these kinds of phylogeny assessments in evolutionary biology for quite some time, and there are relatively few novel conceptual difficulties in such applications. This is not to say that the tasks are easy, but in principle at least, one need simply choose methods of laboratory and data analysis that are appropriate for the evolutionary timescales under investigation.

At the other end of the scale, with Jeffreys' development of DNA fingerprinting methods and their extension to other kinds of DNA fingerprinting probes, for many species we're in a good position to establish genetic identity versus non-identity. There is sufficient genetic variation in most species that when appropriate assays are used (for example, Jeffreys' hypervariable minisatellite probes), one can often get unique genetic profiles that distinguish one individual from any other. This approach is also very useful in assessing paternity and maternity — that is, in examining parentage and transmission across a single animal generation. Much attention

in the literature has been devoted to these kinds of applications, and we're in increasingly better shape in this arena.

I think we're in worse shape at the remaining two levels in the hierarchy (estimating relationships within a population pedigree, and examining geographic population structure within species). I think that's mainly for a conceptual reason that may ultimately prove impossible to circumvent. In reality, as someone pointed out in these meetings, among the variety of population structures possible, many species are probably characterized by isolation by distance. In such cases, there may not be clear phylogenetic relationships revealed with assays of a particular locus. Let me frame the problem in another way. What we would now like to have in molecular population biology is the same kind of information from the nuclear genome that we currently have for mitochondrial DNA. In other words, as I indicated in my talk the other day, recorded in the mitochondrial molecule is a linear history of allelic changes, such that we can establish an allelic genealogy for that particular small piece of DNA. In principle, it would now be desirable to delve (even at random) into the nuclear genome and assay small segments of DNA, one at a time, for purposes of establishing haplotype relationships among the alleles of a locus. This approach might be attempted for each of a number of such nuclear genes scattered around the chromosome set. Armed with such information, haplotype genealogies for each of a number of unlinked loci in the genome would provide a fascinating class of information that one might wish to bring to bear on the questions of population structure.

The approach would be analogous to what we attempt with mitochondrial DNA, but would involve several unlinked nuclear genes that are independent in transmission through an organismal pedigree. If we had that ultimate kind of sequence information for many individuals, I can't see how we could ever do much better in terms of having an empirical data base for describing population structure.

That leads us to the question of what kind of inferences about population structure might be made with such extensive information in hand. Let us even suppose complete nucleotide sequences were available for each of a series of haplotypes at unlinked loci taken from the nuclear genome (as well as from mitochondrial DNA). How would we employ that information to make inferences about population structure? I suggest that under an isolation by distance model, you might well see different loci exhibiting distinct geographic haplotype distributions within the

species range. These could either refer to particular haplotypes, or to branches in a phylogeny of haplotypes for a particular locus. For each locus, an entirely different geographic picture in terms of particular areas recognized might emerge. There might be significant structure evidenced in the gene genealogies for each locus, but unless there were concordance in the particular branching patterns across loci, I would claim that you had little or no evidence for anything beyond isolation by distance. In other words, in the absence of concordance in the particular geographic partitions revealed by independent loci, there are no demonstrable long term vicariant population separations.

From a conceptual point of view, in the next decade much attention will have to be devoted to principles of genealogical concordance. With "the ultimate" sequence information from many genes available, how will such data translate into conclusions about population structure? How much gene flow and historical connectedness will prove to characterize different populations within a species? My current impression is that only under situations where there is a high degree of genealogical concordance can firm conclusions about strong historical population subdivisions be made. In the absence of such concordance, you may well have population structure but it will have been of short term duration. Different loci, although each exhibiting population subdivision, will show idiosyncratic structures, and under isolation by distance there will be very little genealogical concordance across loci. I think this is going to be an important area for development of formal theory in the next decade.

On the empirical side, I nonetheless have strong reservations about whether we can ever get this kind of idealized information. There are two categories of concern. One is the technical concern of dealing with diploid, sexually reproducing creatures (including most fishes). If one wishes to establish haplotype genealogies at particular loci, one has to assay particular haplotypes for sequence or restriction sites. From a diploid creature that may prove to be a difficult prospect. The few cases where it has been attempted have thus far been in *Drosophila*, where it is possible to make chromosomes identical by descent in a controlled breeding program. One can then examine the haplotypes directly (e.g., generate restriction maps) and estimate a haplotype phylogeny for each locus. Little attention has been devoted to attempts to analyse particular haplotypes (that is, establish the cis vs. trans phase of variants) in creatures other than *Drosophila*.

There are some theoretical ways one might go about it. One could perhaps amplify particular gene products from haploid gametes. Or one could attempt dilution studies to isolate a single haplotype for amplification. Even if these technical difficulties could be overcome, however, there remains the fundamental biological problem of intragenic recombination in nuclear genomes. This is a complication we may never be able to overcome, I'm afraid. If there has been significant recombination among the haplotypes over the evolutionary time scale of interest, this will have blurred the linear history of mutational differences that otherwise characterize haplotype differences, and would make it difficult or impossible to get the kind of clear phylogenetic reconstructions that are routinely possible for mitochondrial DNA (because of its uniparental transmission and non-recombining inheritance). This is a fundamental biological reality that may continue to plague the field, even if the technical difficulties of haplotype characterization can be overcome. However, until we genuinely attempt some of these approaches, we're not going to really know the true limitations of "gene genealogical" approaches.

Thus there is great room for both conceptual and technical advance, and I think much of the exciting effort will come in the population structure portion of the hierarchy of phylogeny analysis. It is precisely at this level where one can most meaningfully address questions of genealogical concordance across loci. Once reproductive isolation is achieved, a species phylogeny evidenced by any one gene is expected to be mirrored faithfully by the allelic phylogeny at other loci. For populations or species long isolated, genealogical concordance is expected to be well established. But when populations are connected by some intermediate or low level of gene flow, it becomes of special interest to ask just how much genealogical concordance would be exhibited under various kinds of population structure. There is much conceptual and empirical work to be done in this area.

All of this groundwork will be, in a sense, prerequisite to the full utilization and interpretation of molecular information in a management context. Much remains to be accomplished in the realm of pure (as well as applied) science. One long-term hope is that we as conservation biologists, interested in conserving genetic diversity, may employ molecular methods to provide a better window on diversity - an opportunity to describe its distribution within and among species. That should be an important development in genetic diversity management. In many species, we need to have a better description of genetic diversity. I'm afraid that our



existing subspecies taxonomies are too often grossly inadequate guides to how genetic variation is partitioned. That has been quite clear in our own work in the United States, where existing subspecies taxonomies have too often been woefully misleading indicators of biotic diversity. Now that we have begun to peek into the DNA's of a few species, we're gaining a better handle on phylogenetic diversity, but there is a very long way to go.

DR. BOB WARD: Thanks very much John. Coming back to the possible impacts of biochemical genetics on management, I'd like Peter Smith from the New Zealand Ministry of Agriculture and Fisheries to spend a few minutes talking to us about his experiences. In New Zealand, the research side is very much more closely linked with the management side than it appears to be in Australia. I think this closer linkage will give us an interesting perspective on the interrelationships from this New Zealand point of view.

DR. PETER SMITH: My talk is based on a New Zealand perspective. Given the catching capacity of the modern fishing fleets, in terms of the size and mobility of the boats, then the major problem facing fisheries biologists is accurate biomass estimation and the generation of a figure for a total allowable catch which is allocated to each major species to ensure both the long term exploitation and the survival of the resource. In this respect stock separation is a minor component of fisheries research. I say that as one who has spent ten years working with allozymes and more recently mitochondrial DNA markers.

Fishing fleets are large enough to eliminate stocks within a short time period unless we have tight management controls on what is caught and where and when. I believe that the long term value of population genetic studies is not going to be in producing a figure to punch into some fisheries stock assessment model, but in fundamental knowledge of fish populations.

Already with the results of allozyme studies work we are starting to see a change in our conceptual framework of the structure of fish stocks. The idea that fish are subdivided into discrete subpopulations, or stocks, goes back eighty or ninety years right to the beginning of fisheries science in the northern hemisphere around the turn of the century. John Avise has already alluded to the problems of the subspecies concept and the difficulties we face when dealing with something that has become established in the literature. I believe we are facing that problem with the stock concept. Allozyme studies on fish stocks are changing our way of

thinking about stock discreteness. It may seem pretty obvious to this group of eminent geneticists, but one of the surprising results of the early electrophoretic studies on fish stocks was the high amount of genetic variation found in fish species, and perhaps more importantly, that the majority of that genetic variation could be found in a sample of fish, say fifty fish, taken anywhere in the range of the species. There was very little genetic divergence between the traditionally recognized stocks and most of the genetic variation appeared to be contained within stocks. Where genetic breaks occur between fish stocks the genetic stock units tend to be larger geographical groups than unit stocks recognized by traditional methods. Examples can be seen in coastal species found in the Tasman Sea. Several studies have shown significant genetic differences between Australian and New Zealand coastal stocks but within countries there are minimal genetic differences between samples taken anywhere along the respective coasts.

Following on from these observations, there has been a conceptual change from a discrete stock model to a population concept or, using Barry Richardson's terms, a change from a subpopulation to an isolation by distance model. Related conceptual changes have had great impact in taxonomy. Perhaps I can remind you of Dobzhansky writing, more than twenty years ago, about what he called the greatest conceptual revolution: a change from a typological concept to a population concept in which species were recognised as polytypic units and not unique types. This conceptual change has been accompanied by a major reduction in the number of species that are recognized by taxonomists, not because those species have become extinct but because the species have been reclassified.

I believe we're facing a similar situation with stock discreteness and that the future application of genetic techniques will shift the emphasis away from routine stock identification studies to understanding some of the genetic processes that are occurring within populations. There are areas where I think genetic techniques are going to be useful.

The first of these is in the genetic effects of fishing. Several speakers have touched on this area but it has been assumed that wild populations of commercially important species are extremely large in number and are not likely to suffer genetic changes due to fishing activities. However, there are long term data sets for Atlantic cod, Pacific salmon and to some extent *Tilapia* in east Africa, that show a reduction in mean size at age and a reduction in age at first reproduction in these fishes associated with heavy fishing pressure. Although there could be an environmental

component to these changes the characters do have a low heritability suggesting that fishing has changed the genetic structure of the resources. After all, fishing is not a random process: one only has to look at the length frequencies of a sample of snapper taken by a longliner and a trawler from the same area to get two very different pictures of population structure. Thus fishing pressure could be a selective agent on the populations that we are attempting to manage. Out of interest, in the last couple of years we have shown a decrease in heterozygosity in the orange roughly caught around New Zealand. The orange roughly fishery developed rapidly in the past ten years and in that period the virgin biomass has been reduced by about two thirds. In the early stages of the fishery we carried out a routine stock separation survey using electrophoretic markers and showed a high level of genetic variation, but only minor differences between the spawning groups. We repeated the exercise six to eight years later, to look at the distribution of rare alleles, and much to our surprise the major finding was a reduction in average heterozygosity over three spawning sites.

The other area where I believe that genetics will make a contribution to our fundamental knowledge is the study of genetic changes that are occurring within fish populations, or what we might call microstructure. Barry Richardson has briefly touched on this in his presentation. In molluscs in particular an excess of homozygotes has been observed in many species: an excess much greater than would be expected by simple population mixing. Furthermore, the excess is greater in juveniles than in adults suggesting that selection may favour heterozygotes over the growth period and high mortality period. The remaining problem is what produces the homozygous excess in the first place. In this respect Mike Johnson has been doing some exciting work on this problem with a variety of species in Western Australia. If I was asked to fund future research in genetic studies I would like to shift the emphasis away from routine stock separation to examine the genetic problems which I have outlined. These comments apply specifically to population genetics of natural populations and do not include the field of aquaculture genetics.

DR. BOB WARD: Time is pressing now. If we do have a few minutes at the end maybe we could come back to some questions for Peter or John. Perhaps what I could do now is just, in a way of summarizing these proceedings, to put up a list of uses of mitochondrial DNA and allozyme techniques which are of interest to fisheries managers and scientists. They are:

1. Stock discrimination,

2. Solving systematic and taxonomic problems,
3. Identification of larval and juvenile stages,
4. Monitoring genetic variation in aquaculture.

I think that, despite what Peter's just said, perhaps managers at the moment, if they're interested at all in genetics, are most interested in it as a stock discrimination tool. The null hypothesis is that all subpopulations are taken from a single panmictic population. When you test your null hypothesis there are two outcomes. You can reject it, in other words you can say that there is some sort of population structure and you can advise the managers of this. It's up to them to decide whether they want to incorporate that knowledge in their management strategies. They may be wise to because it may help conserve their resource.

Alternatively, of course, when you test your null hypothesis you may fail to reject it. It doesn't mean to say though that your null hypothesis is correct and that you do have a single panmictic population. It just means that you can't reject that hypothesis. It may be there's a single panmictic population there or it may be that you do have population subdivision but you haven't detected it because maybe you haven't looked at adequate sample sizes or the right isozymes or the right restriction enzymes. Maybe migration among these subpopulations is very limited but is sufficient to prevent genetically detectable subdivision. Then the managers may decide to manage the resource as a single entity or, as several subpopulations.

It is probably in this area of stock discrimination that we can be of most use to fish managers, at present anyway. When it comes to stock discrimination, certainly for the foreseeable future, I'm sure that allozyme techniques and mitochondrial DNA techniques will continue to be used. There's nothing on the horizon that really threatens to supersede the preeminence of these sorts of techniques.

An area that we haven't touched on at all in this meeting is that of systematic and taxonomic problems. These are especially evident in Australian marine fauna, certainly when compared with the marine fauna of northern European waters. Although these taxonomic problems may not be of great interest at the moment to managers because they mostly involve fish of limited commercial value and by-catch species, in the future these species may have more importance in commercial terms. Certainly allozyme and mitochondrial DNA techniques are extremely

powerful in solving systematic or taxonomic problems.

A third area where our genetic studies can be useful in fisheries research is in the identification of larval and juvenile stages. Again this is something which we haven't really touched on in this conference but it's often important to be able to identify the larvae of commercially important fishes or shellfish, prawns or lobsters. It's very difficult to do this in many cases using standard techniques. If you have to rear up these larvae or juvenile stages in aquaria to adult size to see what the species actually turn into when they metamorphose or when they mature, this can take a long time and can be very expensive. Using allozyme and mitochondrial DNA techniques, providing you've typed the adults and you know what the genotypes of the adults are, you can screen the larvae and juveniles and assign them to a particular species very readily.

Then of course we've got various uses in aquaculture, that we've heard quite a bit of today, monitoring changes in genetic variation using allozyme or mitochondrial DNA markers. The populations you are interested in monitoring may be hatchery stocks or introduced stocks where the native populations may be in danger of disappearing or you may want to introduce stocks into drainages or areas previously lacking the species.

These are probably the four major areas where biochemical genetic techniques do have some important applications to fisheries research and fisheries management.

What sort of recommendations have come out of this meeting, which we can use in the future to improve the value of the genetic advice that we can provide to managers? There are perhaps five such recommendations:

1. Improved experimental design
2. Improved statistical analysis
3. Combining allozyme and mitochondrial DNA approaches
4. Combining genetic analyses with morphological and age analysis
5. Don't consider genetic data in isolation.

Barry Richardson spent some time yesterday talking about improved experimental design. He was recommending that ideally you should try and characterize genetically 100-200 fish from

each cohort, each age class, within a particular collection or shot of a commercial species. That's certainly an ideal situation. It would be nice to be able to do that. There are some problems with doing that in some species. For example, if we take the roughy, the orange roughy that we've heard of already today, this species can live to maybe 70 or more years of age. If you wanted to try and take a hundred fish from each of those 70 age classes then you've got a lot of fish on your plate, so to speak. Fortunately, actually we can't do that for roughy because although we know it lives to maybe 70 or 100 years of age, we can't accurately characterize individuals to particular cohorts. So, although it would be nice to carry out this ideal experiment, it just isn't feasible in some species. It probably is feasible in shorter lived species where you can readily age the animals.

But certainly we need improved experimental designs. We need longer term studies, we need repeated sampling. We need to get a better idea of population structure than we've done in the past. We need to know their changes from year to year and we do need to know the differences in different cohorts where possible. We need to improve the statistical analysis of the data. I'm thinking here perhaps of more routine use of unbiased estimators. We heard something about this yesterday. Examples are estimates of genetic distance and  $F_{ST}$  values which incorporate sample size. We need to be able to put confidence limits to such estimates of genetic variation within and between populations, especially if we are attempting to use the genetic data to estimate migration rates among populations. We need to be able to give figures with associated standard errors. We have to be wise in applying some of these techniques and not go at them because the computer programs are there. We have to understand the basis of the techniques and make sure that the data we've got are actually usable and we can get really meaningful results out of this statistical analysis. So we've got to improve our analysis but we've got to do so carefully.

It would add substantially to the power of these biochemical genetic surveys, if you don't just do allozymes or mitochondrial DNA. Wherever possible you should combine the use of both techniques because they are looking at variation in different parts of the genome and they have different sorts of resolving powers and they can answer different sorts of questions. But where you combine the two approaches I think you stand a better chance of picking up interesting and valuable data. Allozyme data from a particular population can give you indication of Wahlund effects and possible population mixing within that area. Mitochondrial DNA surveys can't by

themselves tell you anything about possible population mixing. It doesn't have, in a sense, the power of allozyme data where you can look at Wahlund effects and deficiency of heterozygotes. Yet, of course in other respects mitochondrial DNA surveys are more powerful than allozyme analysis. Mitochondrial DNA evolves more rapidly than nuclear DNA, you are able to study variability at synonymous codons and non-coding sites, mtDNA has a smaller effective population size than nuclear DNA and as a consequence of these factors you expect greater mtDNA differentiation among populations than allozyme differentiation. So I think both of these techniques have their particular advantages and combining the techniques is a more powerful approach than looking at either one by itself.

The genetic approach becomes yet more powerful when combined with studies of morphometrics, sizes and shapes of individuals, growth rates, cohort analyses, meristic counts, asymmetry and so on. Of course what this means is you've got to get the whole fish into the lab. You can't just get tissue samples sent to you. You can do that for the allozyme and mitochondrial DNA surveys but you need to get the whole fish into the lab to carry out a complete analysis of morphometric and meristic measurements. Doing this means that you can also answer interesting genetical questions such as are there relationships between genotype and phenotype, or between genotype and growth rate. It's not going to be possible to carry out a complete examination with some species, such as tuna for example. You can't very well bring in a 100 southern bluefin tuna into the lab and expect to do this. Unless you've got a lot of money of course. Generally speaking we don't quite have that sort of money.

An associated point is you shouldn't consider genetic data in isolation. I'm sure that actually no-one does do that. There are all sorts of ways of looking at stock structure if this is your main aim, including, for example, parasites and population dynamics. Genetic data should always be discussed in conjunction with data from other sources because this gives you a much better understanding of the value of the data and its usefulness to management.

Those are just some possible recommendations that come out of this meeting. It is now five o'clock and this meeting is supposed to finish at five but maybe we can spend a couple of minutes talking about some of these uses or recommendations or perhaps people would like to ask John Avise or Peter Smith some questions. If we perhaps stick with these recommendations, do people feel that these sorts of recommendations are worthwhile? Is this the sort of thing that we

want to try and get out of this conference? Are there other recommendations that people would like to add to this list that may be of use to managers? I'd certainly appreciate any views any of you may have on this from a management side or a research side.

DR. JOHN BENZIE: I wondered if in a sense one could be more explicit in some of the suggestions. For example, Barry indicated in his talk that there had been very little work done in terms of isolation by distance, to get effective methods of analysis of these models. He's been talking about that for eight years. If that is at all important in being able to interpret the data once they're collected then that perhaps is something that should get some sort of priority. Another thing again which Barry brought up is the method of funding - how to appraise a project. Whether you do a pilot first, and whether the funding can be organized in such a fashion as to allow more easily that sort of cycling, working on a problem, better defining it, reappraising it, reappraising the hypotheses, and going perhaps in a slightly different direction and then better being able then to come up with answers to the more general problems in mind. I'd welcome other discussion on those points.

DR. COLIN PURDOM: Nobody's mentioned the polymerase chain reaction techniques during the last two days. Is that because they've fallen out of fashion?

PROFESSOR ROSS CROZIER: I hope it does stay in fashion! I'd say that the reason you probably raised this was that for instance we may be looking at very young stages of fish. It would obviously be very highly useful to be able to look at very, tiny amounts of DNA. Seeing as, in some hands at least, one can get amplification of two different genes from the same human sperm, obviously a fish egg should give some reasonable opportunity for looking at mitochondrial or other DNA variants in it. So, if it hasn't been mentioned, I think it's been subsumed by the people that perhaps should have mentioned it in the overall set of things that they had been talking about.

DR. COLIN PURDOM: John has talked about the crossover recombination problem with nuclear DNA. Isn't it the case that the repeat, the highly repeated sections of DNA, can be examined by this PCR technique? Isn't it possible that they are not subject to crossover diminution of specificity?



PROFESSOR ROSS CROZIER: I'm a little uncertain about some of the things that you are getting at there because if you're looking at nuclear DNA one problem with the PCR approach is that you have potentially two different alleles. During the polymerase chain reaction procedure you can get effective recombination taking place as you go. So in fact, it is a little difficult looking at nuclear genes to always know if you've got the same alleles back at the end of the amplification as there were actually at the start because if there are two that differ by more than one base pair difference you've got a good chance of picking up scrambled new alleles. In fact, the way this happens is that you get incomplete production, new copies of each round, and these copies then compete at the next round but they don't necessarily pair up with the same primer DNA each time. This is thought to be the explanation, for example, why in that 7,000 year old human brain from a Florida locality, they were actually able to amplify significant stretches of DNA even though all the little pieces of DNA turn out to be quite small, smaller than the final products of amplification. That was of course mitochondrial DNA and by standard beliefs, hopes and dogma that individual of course would be the same all the way through. You wouldn't get any scrambling because there's only one type there in the first place. So I think there are problems with looking at nuclear genes with the PCR reaction because you would get new combinations back that were never there in the first place if there was that much variation.

DR. CRAIG MORITZ: Maybe I can make a slightly more general comment there. I think these recommendations are certainly endorsed but it seems to me a lot of them could have been made ten years ago. We still haven't done a whole lot about it although certainly in item 3 now there's a lot of data sets coming out along that line.

I'm glad you raised the PCR issue because I think there are some new advances now in molecular population genetics that we should be taking further note of and trying to incorporate into more management-oriented studies. There are recent studies of highly repeated genes that show concerted evolution which tends to homogenize sequences within demes and which will accentuate the differences between stocks. We have, I guess, two or three studies where people have looked at allozymes, mitochondrial DNA, repeated genes, maybe using PCR, maybe using something else. We don't know yet enough to be able to make predictions about what type of genetic marker is going to be best for which particular purpose. So I think we need a lot more basic studies just looking at different types of genetic systems, just patterns of natural variation. Once we have that data we should be able to make some fairly strong predictions.

Back to the PCR, I think one of the big advantages there is that you can work with partially degraded DNA. Often, for example, if you want to work on deep sea fish, there are problems in getting good high molecular weight DNA out a lot of the time. With PCR at least there you have an opportunity to work with that or with alcohol preserved material. I know Ross has been doing some work with, I think, dried material. Ross?

PROFESSOR ROSS CROZIER: Yes.

DR. CRAIG MORITZ: People are exploiting a whole new range of tissue types and types of specimens. I think this is going to allow us to ask some different types of questions than we were able to ask before. I think it's a very exciting time.

DR. BOB WARD: I know with respect to PCR that Jenny Ovenden is hoping to do some work on Euphausiids where it may be difficult to get sufficient mitochondrial DNA out without PCR type reactions.

DR. JOHN PAXTON: I'd like to add a recommendation to the geneticists and fisheries biologists, which stems from your comment that the systematics and taxonomy of Australian fishes are not perfectly known. They're certainly not, even for such commercial fishes as gemfish and ocean perch. The recommendation I would make is that you take a few voucher specimens from your studies and place them in your local museum. In ten years time then maybe somebody can look back on your studies and know which species or which population you were dealing with.

DR. BOB WARD: Well if there are no further points then, there is just a brief summarizing statement that I would like to make here. It involves this booklet. If some of you haven't seen this booklet and you're involved in fisheries research you should read this. This is the government policy statement released in December, 1989, entitled New Directions for Commonwealth Fisheries Management in the 1990's. It is a very clearly presented account of where the government sees the future management of Commonwealth fisheries and outlines the three main objectives of sustaining fish stocks, maximizing economic efficiency, and providing a payment to the community for the use of resources. It has a chapter on research, research

orientations, and it says specifically that "Research expenditures should be directed towards areas likely to provide the highest level of benefit and the beneficiaries should contribute to costs in proportion to the benefits received". That's one thing. It also says that "Fisheries managers will be responsible for setting the priorities and administering the resultant research projects for management related research." Clearly as geneticists we need to ensure that both industry and managers are aware of the benefits that genetic analysis can bring to areas such as stock discrimination.

Also I think we have to make sure that they are fully acquainted with the cost effectiveness of these approaches. I believe biochemical genetics approaches are cost effective. They can be carried out in a reasonable span of time certainly compared with say tagging studies that may be horrendously expensive to establish and may sometimes yield very small returns. So we have to ensure that managers are fully acquainted with genetic research and the cost benefits involved in funding genetic type research. I think that this conference hopefully, in its eventual publication, will go some way towards achieving these goals. I hope that the managers do get round to reading the publication and take some note of it.

Just to conclude, I should like to thank Pat Dixon for organizing this conference and bringing us all together in what I think has been a very stimulating and worthwhile two days.

DR. PATRICIA DIXON: Thank you, Bob. Thank you all for coming. If you hadn't come we couldn't have had the conference. I think it's been quite worthwhile. I think that contacts have been made here that are going to be continued. I think we will have continuing exchange between workers in this field to a much greater extent than we've ever had before. I'd particularly like to thank the speakers, all of you, but especially our two overseas speakers, Colin Purdom and John Avise. Thank you very much for coming and sharing your expertise with us and injecting a bit more enthusiasm into the conference. To John in particular, I found your paper particularly inspiring, and I think that there are a lot of other people in the audience who are quite keen to get on with things since we've heard of your outstanding achievements. I'd also like to thank my students and staff who have helped over the last few weeks, especially Annette my secretary who isn't here to hear this at the moment but she will hear it when she plays the tape. She has in fact borne the brunt of the work and she still has a lot of work to do yet as we prepare the proceedings for publication. I've injected nervous energy, I think, into this conference but

Annette has done most of the real work. I'd also like to thank FIRDC for their generosity and also foresight in funding a conference of this sort. I think that in the future the fishing industry will gain great benefit from having got us all together to discuss some of the things which are of great importance at the moment to us but we hope to you also in the future.