Genetic methods for the estimation of fisheries spawning stock size

Transcripts of the forum

9 May 2003

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

$$R = \frac{S}{\alpha + \beta S}.e^{\varepsilon}$$

Queensland Government
Department of **Primary Industries**

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

AFMA

FISHERIES RESEARCH & DEVELOPMENT CORPORATION

CSIRO

Genetic methods for the estimation of fisheries spawning stock size Transcripts of the forum - 9 May 2003

Southern Fisheries Centre
Agency for Food and Fibre Sciences
Department of Primary Industries

Editors: Damien Broderick
Samantha Peel
Raewyn Street
Jenny Ovenden

Queensland Government
Department of **Primary Industries**

THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

CSIRO

FISHERIES RESEARCH & DEVELOPMENT CORPORATION

AFMA Australian Fisheries Management Authority

Genetic methods for the estimation of fisheries spawning stock size: Transcripts of the forum - 9 May 2003

Damien Broderick, Samantha Peel, Raewyn Street and Jenny Ovenden

Southern Fisheries Centre
Agency for Food and Fibre Sciences
Department of Primary Industries

September 2003

Project No. FRDC 2001/018

General disclaimer

This publication has been prepared with care. DPI:

(a)   takes no responsibility for any errors, omissions or inaccuracies contained in this publication;

(b)   does not warrant that the information contained in this publication is current or that there is not more recent or more relevant information available;

(c)   does not accept any liability for any decisions or actions taken on the basis of this publication; and

(d)   does not accept any liability for any loss or damage suffered directly or indirectly from the use of the information contained in this publication.

CONTENTS

Executive Summary

Not all advances in the field of genetics are coming from sequencing entire genomes. "Genomic epidemiology" or population genetics is being used more and more in wild fisheries management. Harvesting of wild fisheries is the fifth largest export earner for the Queensland economy, and supports many regional communities. In desperation (almost!), fisheries managers have turned to this unconventional way of learning about the species so that targets for sustainable exploitation can be achieved.

Population genetics is particularly good at uncovering population structure where none is apparent. However beyond this accepted application, fisheries scientists are sceptical that population genetics can also provide estimates of the number of spawning individuals in a population. These forum transcripts represent progress towards this aim by a team headed by Jenny Ovenden and representing scientists from Department of Primary Industries, University of Queensland and CSIRO Marine Research.

The groundwork for "genetic counting" was laid in the 1930's by two mathematicians, Sir Ronald Fisher (UK), and Professor Sewall Wright (US). They made the connection between the magnitude of genetic drift (how much genotypes in a population change from generation to generation) and the number of parents that were successful in leaving offspring. A large number of successful parents means a small amount of drift and vice versa. Hence, population genetics can be used to "count" fish: measure how much drift is occurring and you have an estimate, albeit indirect, of the number of spawning fish in a population that successfully leave offspring (recruits) in the next generation. The method is now practical, as it is possible to rapidly and cheaply genotype large numbers of individuals thanks to recent advances in genetic technology.

Estimation of effective population size, or "genetic counting", may provide a unique and critically important contribution to fisheries stock assessment and management. Mathematical models are used to predict the future size of a fisheries population based on past measurements. An important, but currently impractical data source for these models is the number of spawning fish that successfully leave offspring or recruits. Because of their unique derivation, genetic counts are estimates of fisheries spawner and recruit abundance in one package. They represent not only those spawners that will be successful, but equally represent the number of recruits produced. It was suggested in this forum that genetic counts are the "glue" that connects the spawner-recruit relationship in fisheries modelling and are potentially of benefit as a new data source for fisheries population models.

Innovation is often the judicious combination of the old and new in a novel context. In this case, resource sustainability concerns of fisheries managers provided the motivation to blow the dust off some old volumes about mathematical population genetics. In combination with the latest methods of genotyping, genetic estimation of effective population size may represent a significant advance in fisheries stock assessment science.

Speakers

Greg Adcock
Department of Zoology
University of Melbourne
gjadcock@unimelb.edu.au

Tony Courtney
Southern Fisheries Centre
Department of Primary Industries, Queensland
tony.courtney@dpi.qld.gov.au

Cathy Dichmont
CSIRO Marine Laboratories

Simon Hoyle
Southern Fisheries Centre
Department of Primary Industries, Queensland
simon.hoyle@pobox.com

Jenny Ovenden
Southern Fisheries Centre
Department of Primary Industries, Queensland
Jennifer.Ovenden@dpi.qld.gov.au

David Peel
Southern Fisheries Centre
Department of Primary Industries, Queensland
David.Peel@dpi.qld.gov.au

Samantha Peel
Southern Fisheries Centre
Department of Primary Industries, Queensland
Samantha.Peel@dpi.qld.gov.au

Duncan Souter
Queensland Seafood Industry Association

Delegates

Greg Adcock
Department of Zoology
University of Melbourne
gjadcock@unimelb.edu.au

Megan Anderson
Centre for Animal Conservation Genetics
Southern Cross University
mander15@scu.edu.au

Michael Arthur
Molecular Ecology Lab
Griffith University

Crispian Ashby
Fisheries Research and Development
Corporation

Jason Bartlett
Bribie Island Aquaculture Research Centre,
Department of Primary Industries, Queensland
jason.bartlett@dpi.qld.gov.au

Kaye Basford
School of Land and Food Sciences
University of Queensland

Damien Broderick
Southern Fisheries Centre
Department of Primary Industries, Queensland
Damien.Broderick@dpi.qld.gov.au

Ben Cook
Griffith University
ben.cook@mailbox.gu.edu.au

Tony Courtney
Southern Fisheries Centre
Department of Primary Industries, Queensland
tony.courtney@dpi.qld.gov.au

Cathy Dichmont
CSIRO Marine Research

Steve Donnellan
South Australian Museum

David Gopurenko
Molecular Ecology Laboratory
Griffith University
d.gopurenko@mailbox.gu.edu.au

Peter Hale
Ecology Centre, University of Queensland
phale@ccb.uq.edu.au

Simon Hoyle
Southern Fisheries Centre
Department of Primary Industries, Queensland
simon.hoyle@pobox.com

Yutao Li
CSIRO Livestock Industries
Yutao.Li@csiro.au

Jodine McSweeny
Genetic Solutions
jodine.mcsweeny@geneticsolutions.com.au

Andy Moore
Centre for Animal Conservation Genetics
Southern Cross University
amoore@scu.edu.au

Sam Nelson
Australian Fisheries Management Authority
sam.nelson@afma.gov.au

Liz O'brien
Bribie Island Aquaculture Research Centre
Department of Primary Industries, Queensland
liz.obrien@dpi.qld.gov.au

Michael O'Neill
Southern Fisheries Centre
Department of Primary Industries, Queensland

Jenny Ovenden
Southern Fisheries Centre
Department of Primary Industries, Queensland
Jennifer.Ovenden@dpi.qld.gov.au

Tim Page
Griffith University
t.page@griffith.edu.au

David Peel
Southern Fisheries Centre
Department of Primary Industries, Queensland
David.Peel@dpi.qld.gov.au

Samantha Peel
Southern Fisheries Centre
Department of Primary Industries, Queensland
Samantha.Peel@dpi.qld.gov.au

John Salini
CSIRO Marine Research
john.salini@csiro.au

Duncan Souter
Qld Seafood Industry Association

Raewyn Street
Southern Fisheries Centre
Department of Primary Industries, Queensland
raewyn.street@dpi.qld.gov.au

Yimin Yi
CSIRO Marine Research

Genetic Methods for the Estimation of Fisheries Spawning Stock Size

Introduction

Kaye Basford:	Welcome to the forum on genetic methods for the estimation of fisheries spawning stock size. There are several people involved in this project, Jenny Ovenden, Simon Hoyle, Cathy Dichmont, Tony Courtney, David Peel, Samantha Peel, Heather Podlich, Raewyn Street and myself. People have come in at latter stages or have moved on but they still have had involvement in the project. This project has been sponsored by the Department of Primary Industries, Fisheries Research and Development Corporation, Australian Fisheries Management Authority, CSIRO, and the University of Queensland. I would like to call on Duncan Souter to provide us with the opening address. Duncan is the CEO of the Queensland Seafood Industry Association, which is the peak representative body for the state's licensed commercial fishermen, that is about 2500 people and they have been very supportive of this project. I would like to call on him now to officially open the Forum.

Opening Address

Duncan Souter:	Madam Chair, Ladies and Gentlemen,
	Why not outsmart rather than outspend? Frequent visitors to Brisbane airport will recognise that as the Sun Micro Systems slogan on the billboard on the way in. However the question is a relevant one here today. The sustainable development of our fisheries resources require answers to questions that are often complex and the answers for which aren't easy or impacts cheap to come by. Traditional fisheries research techniques have made significant progress in answering some of those questions and continue to offer significant potential into the future, however many key fisheries questions remain unanswered. The continued success of the Australian seafood industry requires, amongst other things, strong community support. That support is dependent to a large degree on the community being confident that the fisheries are being harvested both socially responsibly and ecologically sustainably. The field of genetics offers new ways for industry, fisheries researchers and fisheries managers to demonstrate to the community that fisheries are being harvested sustainably and the benefits that currently accrue from our commercial and recreational fisheries will also be available for their children. Genetics offers traditional fisheries research a reach and scope it has never had before. It offers better ways to answer old questions and at the same time offers answers to new questions once thought unanswerable by both fisheries researchers, managers and industry. Importantly it does so in a way that is both potentially more practical, more timely, more accurate and less expensive than existing techniques. Accordingly, the potential should be nurtured and its possibilities explored. This project is an important one not only because it offers an inexpensive alternative to traditional stock assessment techniques, and that it challenges for the first time long established but untested fisheries dogma, but because it breaks new ground in Australian fisheries science and lays the ground work to extend those techniques to other species. For these reasons it is highly valued by industry and its potential should be embraced by traditional fisheries scientists and fisheries management agencies alike. It's my job today to officially open the fisheries genetics forum and welcome you here today. It's a job I am particularly pleased to do because of industries long held desire to see practical and effective extension of research results. Forums such as these play an important role not only in promoting the exchange of ideas and techniques between researchers in the field but also to ensure that project results are effectively transferred to both industry and fisheries managers. So, on behalf of the project team, on behalf of the Queensland seafood industry, welcome and I wish you well in your discussions. Thankyou.

Stock assessment and genetics: a partnership approach

Cathy Dichmont

CSIRO Marine Research

Abstract

Stock assessment has been defined by Hilborn and Walters (1992) as being "the use of statistical and mathematical techniques to make quantitative predictions about the reactions of fish populations to alternative management choices". Any quantitative method that attempts to estimate the size of a population and the effects of fishing on it, is therefore seen as stock assessment.

The 'traditional' stock assessment process consists of quantifying:
a)   the basic dynamics of the animals,
b)   the degree of natural variability,
c)   model and management uncertainty,
d)   management objectives, and
e)   the degree of risk for a given reward.

In order to produce an assessment with low uncertainty, in an ideal world, one would wish to have at least the total catch of all sectors (including their biases), more than one index of abundance (i.e. from fishery dependent and independent sources), a long time series of good biological data, etc. However, in many fisheries within Australia, one has to rely almost entirely on fishery dependent data, which has many sources of bias and is not always a good indicator of biomass change. For example, the most common assumptions of catch rate data is that it is linearly proportional to biomass and is a continuous series over time. These assumptions break down in situations, for example, where the resource is highly aggregated, the fishery has changed substantially over time in its capacity to catch fish, or where a single species approach is taken in a multi-species fishery.

The 'traditional' methods and that of using genetic techniques to assess stock size are trying to resolve the same problems in different ways using *different* data sources. This is extremely valuable. With computing power the way it is today, ultimately all these techniques will probably be joined together into a single approach.
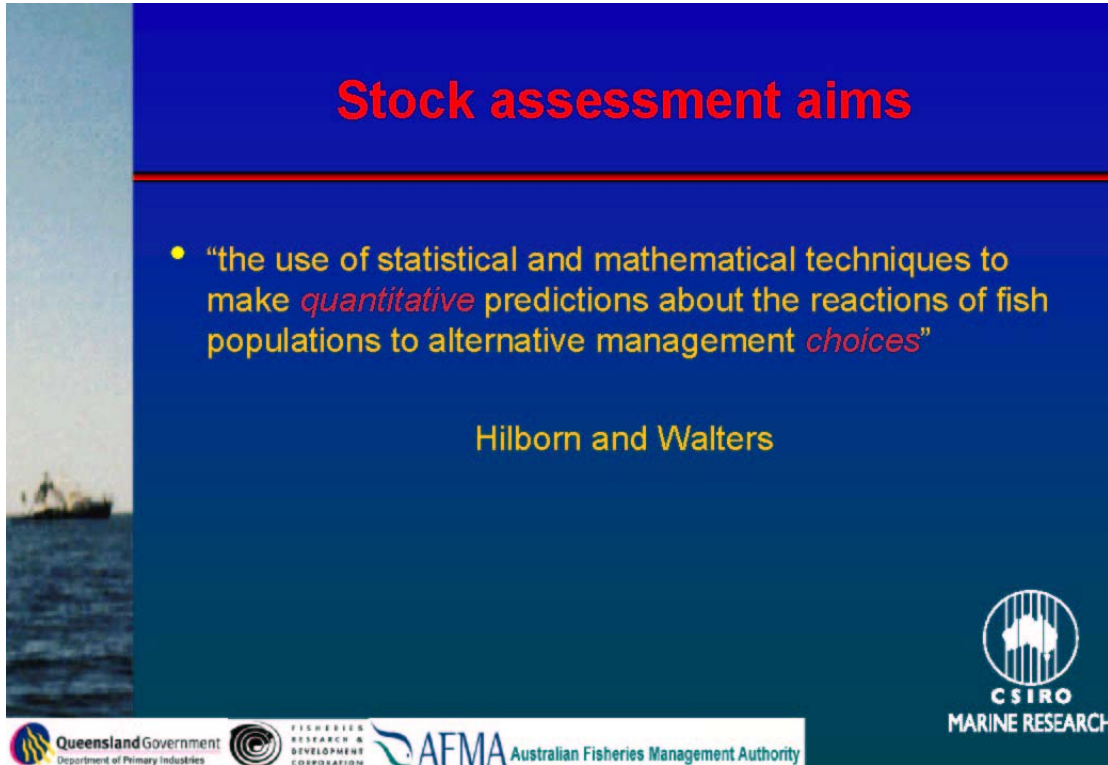
Presentation

Slide 1: Stock assessment and genetics: a partner approach.

I was asked to speak about stock assessment as well as the genetic technique that we will talk about today. I would like to propose that the two together form a very good partnership.

Slide 2: Stock assessment aims.

The definition of stock assessment by Hilborn and Walters above is very appropriate for this forum. I have highlighted *quantitative* and *choices* because they are fundamental to the management of any fishery. To manage fisheries you need to be able to come up with quantitative answers as to where you are and be able to evaluate your management options in order to be able to respond to where you are.

Slide 3: Assessment process.

In the beginning of the assessment process you are spending a lot of time looking at basic dynamics. This is really a question about how many fish are in the sea today compared to the past.

Some resources are much more variable than others so it is important to get a handle on levels of natural variability which depend a lot on the ecology of the species. For example, shorter lived species tend to have more variable population dynamics compared to longer lived ones.

Another important consideration is the level of uncertainty. It is important to understand the degree to which different levels of uncertainty affect your data and the model. It is also essential to recognise those sources of uncertainty that are important to management from those that are not important to management.

Defining management objectives is often done by the scientists themselves but I do not think that this is correct. Managers, scientists, industries and conservation groups all have a big part to play in defining management objectives to decide how best to sustain the fish stocks and the fishery. It is important to have several partners and a co-management approach because there is likely to be several areas of conflict in objectives. An obvious area where conflict often arises is trying to maximise catch whilst minimising risk of overexploitation.

Part of looking at the effects of uncertainty and its effect on your management objectives is to do a risk assessment. Risk analysis explores the various consequences of making a wrong assumption in your model or objectives.
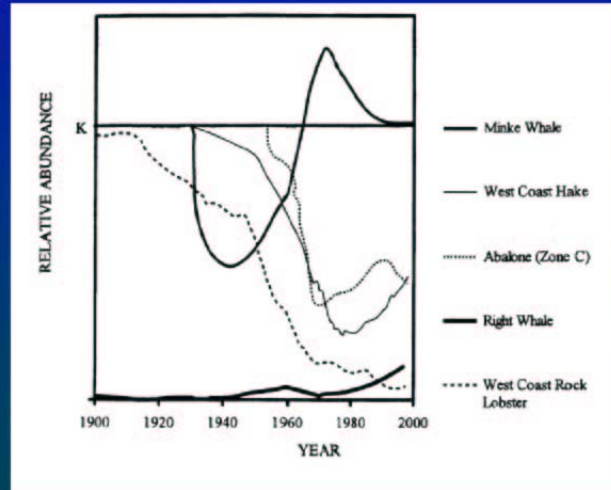
Slide 4: Some history.

These historical fisheries examples come from South Africa. Last century [20th] there was an extremely high level of exploitation on large whales, like the blue and southern right. Even around the 1900's populations of the southern right whale were clearly over exploited and critically threatened as indicated by the low values of K (stable population size/carrying capacity). A behaviour we often see in fisheries is that as populations decline, fishing effort switches to other species. In this case effort was switched from the southern right whale to the smaller Minke whale in the 1920's, from which point it declined dramatically. In the early 1940's South Africa put a moratorium on whaling within its waters. Shortly after the closure the faster growing Minke whales recovered while the slower growing southern right whale recovered slowly. Another reason for the slower recovery in southern right whales is that they are exposed to fishing pressure beyond Southern African waters whereas the Minke whales were not exposed to external fishing pressure. The point to be made here is that there was no stock assessment done on these species. It was clear from a variety of sources that they were in trouble in the 1940's and decisive management decisions were responsible for these species recovery.

In contrast to the whaling examples above, the west coast rock lobster fishery was subject to extremely good stock assessment beginning around the 1960's. As time progressed methods became more sophisticated and scientists began warning that the fishery was in a critical state. Unfortunately management action was too slow to react and only achieved very small recoveries. Here we have a situation where we have good science but poor management resulting in the collapse of this fishery.

The west coast hake and abalone also show a decline in fisheries resources after the commencement of fishing. This is normal as we expect a decline from carrying capacity but unfortunately they overshot the mark. Then stock assessment scientists stepped in, strong management decisions were made and both of them recovered. Despite the good science and management the abalone have since crashed due to the heavy international poaching.

We as scientists only play a small role in a very extensive process of fisheries management but we are integral to good management and good compliance of successfully managed fisheries.

Slide 5: Information requirements (Usual knowns).

A famous quote by Beddington about information requirements for stock assessments is "the only thing you need is catch, catch and catch". In Australia we are lucky to have good logbook statistics therefore we tend to get good catch and effort data from the industry. We often know the growth rate of these animals and we have a fishery dependent index of abundance.

Slide 6: Information requirements (Essential but usually unknown).

A major dilemma facing stock assessments is that they rarely have a fisheries independent index of abundance. Another problem is catch bias where grading occurs to select out the desirable size classes or species. The proportion of catch thrown away is not known, as it does not form part of your catch statistics. The other major problem in many fisheries is that we do not have comparable effort over time. Gear and vessels have changed dramatically over time so that it is difficult to compare one time slice with another. In general, stock assessment scientists have done a poor job at defining natural mortality. Suitable models exist to estimate it from catch data but the data needs to be very good. While it is difficult to quantify, we also need to know how variable recruitment is.

Slide 7: In the ideal world.

In an ideal world we would want to
  i)   know the total catch across commercial and recreational fishing sectors,
  ii)   have more than one index of abundance (especially valuable are fisheries independent indices of abundance),
  iii)   have long term data sets (the South African examples demonstrated earlier on showed how extremely valuable this information is), and finally
  iv)   have good biological and spatial information that are highly desirable components of stock assessments.

Slide 8: Dependent index of abundance.

There are some weaknesses in the dependent index of abundance however, which can cause them to be biased. One of the assumptions of catch per unit of effort (CPUE) is that it is proportional to abundance. While this makes inherent sense, the problem lies in our assumption that CPUE is linearly proportional to abundance. This is not always true especially in aggregated resources where we expect that the relationship between CPUE and abundance will not be linear.

Another weakness is that people tend to approach CPUE from a single species perspective when often they are taken in a multi-species fisheries context.

Depending on the particular fishery, environment and ecosystem effects, while usually not included, can have a big impact. A classic example of this is the banana prawn fishery, which is driven by strong environmental drivers like rainfall.

Spatial aspects of recruitment and growth are often excluded generally due to a lack of data.

Slide 9: Fishery dependent data 1.

The northern prawn fishery, situated in the Gulf of Carpentaria Australia, provides an example of how spatial information together with log book data can help us better understand the fishery. The yellow areas are untrawlable ground, the blue spots are vessel positions from satellite VMS (vessel monitoring systems), and the red spots are where the vessels said that they got their catch. The geographic extent of the fishery is really delimited by untrawlable grounds forcing them to trawl in amongst the coral reefs. If we only considered the logbook data we would not see this spatial information nor would you know why catch was concentrated in a particular region. It is only after the inclusion of another data source that the nature of the spatial structure becomes apparent.

Slide 10: Fishery dependent data 2.

A classic example of contraction of a resource over time is the collapse of the Canadian cod fishery and I suspect if you looked for it, you would find a similar signature in all fisheries. The fishing grounds in the northern prawn fishery were much larger in the early 1980's and fishing is now concentrated in just a few hot spots shown in magenta. The extensive contraction of the fishery since the 1980's makes comparisons of CPUE data that was collected over the entire region very difficult. We cannot compare CPUE data unless we know the reasons for contraction. For example the contractions could be due to a lack of prawns in those areas, or the fishermen may no longer have enough time to fish those areas or perhaps they have learnt more about where to go and where not to go.

Slide 11: Fishery dependent data 3.

The top picture is of a two horse trawling operation in Cape Morkim, England in 1902. The photograph on the left is of a typical trawling vessel in the northern prawn fishery during the 1970's. The picture on the right is typical of modern trawler in the northern prawn fishery today. Vessels today not only differ in terms of size and type, vessels today have onboard much more electronic information to facilitate trawling.

Slide 12: Fishery dependent data 4.

The fishing power of vessels in the northern prawn fishery has changed substantially over time. They are now 3-5 times more efficient at catching prawns compared to the 1970's. Perhaps more frightening is that, depending on the kind of data and models used, the change in fishing power over time could be interpreted as anything between a little change and a catastrophic change.   Without fisheries independent data it is impossible to know which one of these models is correct.
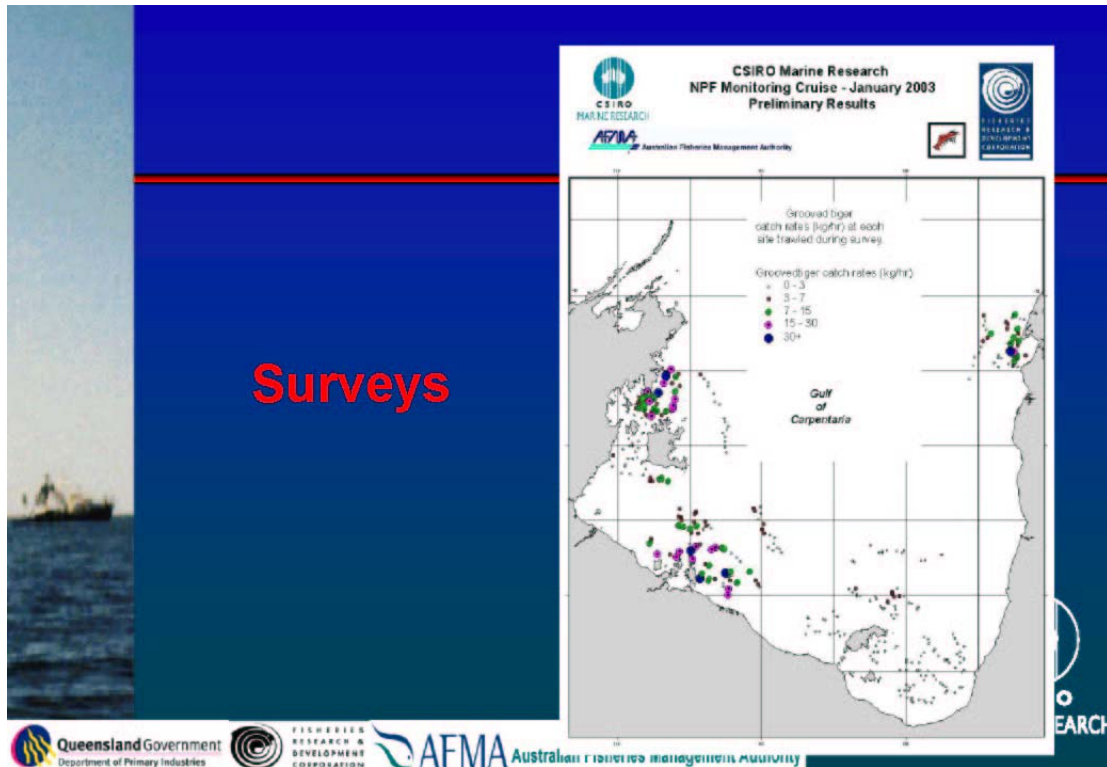
Slide 13: Surveys.

This is an extensive survey of the northern prawn fishery (NPF) conducted in January 2003. This survey covered the entire extent of the NPF grounds and also included some offshore work to look at bycatch. This survey cost around $400,000 and was fully funded by the industry. This is extremely valuable information despite its high price tag. While independent surveys are very valuable and fundamental to research on any species, we have to acknowledge that they are very expensive and time consuming. This survey involved two vessels, 16 people and took about three weeks to complete. However surveys of this kind also collect additional information on several other species, oceanographic data and bottom type data.

Slide 14: Management advice.

Fundamental to stock assessment is its capacity of forward projection. That is, stock assessments must be able to give some idea of what would happen as a result of the implementation of management actions. What is valuable about all of these techniques is that we can say this is our objective and if we do the following things we will or will not achieve those objectives. Such a system allows for objectives to have time targets associated with them. For example, management decisions in the northern prawn fisheries during 2001 were not strong enough to affect a quick enough recovery of the resource. Our objective was to have 70% of these values to go over the target by 2006. This meant that we had to make drastic changes to management.

Slide 15: Conclusion.

In conclusion, different stock assessment techniques make different assumptions. Some of these are good and others are not so good but it is important that we are up front about these assumptions. Techniques that try to resolve the same problems in different ways using different sources are extremely valuable. In the same way that having fisheries independent and dependent data enhances stock assessment, we have used several different techniques to try and find out the status of a resource. The reality is that, when we consider the computing power of today and into the future, and advances in statistical theory, it is highly likely that all of these techniques will be joined into a single process because they are all trying to do the same thing.

Questions

Ovenden:	Cathy [Dichmont] can I ask about spatial contraction of the resource in the southern gulf where you showed the yellow, the green and the pink squares. Was that a result of a real contraction of the resource? Did your surveys show that the prawns were generally absent in those pink areas?
Dichmont:	We actually designed the survey and went out in September last year. We designed it so that we would be able to compare these and these areas [west and north east Mornington]. Interestingly this contraction hasn't just happened there, it has happened in lots of places and the answers are actually, as per usual, complex. It seems that in this region [Mornington], which is where the resource is in very poor condition, it's pretty much the same but because that's if you look at our survey results for 2000 [Southern Gulf, Mornington Island region], there are a whole lot of zeros. But if you look at the contraction over there [Groote Elyandt] then a lot of that was [no prawns in the habitat]. But it makes a very big difference as to what you then do with your CPUE. Just as a further point, this is a very big problem in the southern blue fin tuna. It's probably the biggest reason why there is conflict. Other than the forward projections, there is probably quite a big conflict in the southern blue fish tuna assessment to do with the fact that they have shown this contraction and in their case it looks like a real contraction.
Donnellan:	I'm a novice in the biological side of this. What influence does the actual trawling have on the survivability of the populations in terms of disturbing habitat or irreversibly altering it?
Dichmont:	There are alternate opinions on that. If you talk to the industry, probably a lot of what they say is true in that trawling probably produces more trawling grounds. It makes it good for prawns, if you are looking at prawns. However if you look at fish for example, fish need structure. When talking about fish they are slightly structure driven like the snapper. In that case the trawling removes that structure so in that regard trawling is bad for them but there is a theory that prawn trawling may not be all that bad for prawns. It's probably bad for other things. There are experts in this room that are more knowledgeable about this question than me, if Tony [Courtney] wants to answer it.
Courtney:	I was actually going to comment on the previous question so I'll comment on both. There is *some* suggestion that trawling removes prawn predators so you can actually get an increase in prawn biomass as a result of lowering their natural mortality rates as a result of removing their predators. This is what you find from some of the ecosystem modelling work by Neil Gribble using Ecopath and Ecosym. The other thing I wanted to say about the previous question was that a spatial reduction in the areas or locations where trawling occurs can often be in response to the management measures. An example is here on the Queensland east coast where the managers have recently brought in limited nights. A few years ago the fishermen could fish as many nights as they wanted to. Now each individual fisherman has a set number of nights. This has removed a lot of the experimental trawling of areas that fishermen used to do. Every night that they work now, they try to maximise their profit because they only have a set number of nights to work under. They now go into areas where they know they can make a certain profit rather than going into more peripheral or risky areas and so they don't explore new areas as much as they used to do.
Dichmont:	That's true Tony [Courtney] and I suppose I obviously didn't answer the question properly because I think it's often many reasons why it's not happening. In the case that I have given in Mornington Island they are crossing those grounds they are actually going over. You must remember that they have got sounders that are telling them what's below them and in this fishery it is a lot of reasons why. They now have excellent databases that tell them exactly where those fish are. So they are using the GPS to go to the hot spots. So it is for all of those reasons above that there is not wider shrinkage. The point is that catch rate data is not an experimental sample. It's an extraordinarily biased data set. People tend to think well I've got

	hundreds of thousands of records it's so extensive and massively useful, and it is up until a point, but it has very strong weaknesses.
Basford:	One thing I was interested in was when you had the satellite data [VMS] and then you had the reported data, there where actually differences. Do the fishermen report to you where they are fishing and where they have caught things but this doesn't necessarily match up to where they actually are.
Dichmont:	Well I think it's pretty good. I must admit I can give you worse examples.
Basford:	So you would consider that a very good match.
Dichmont:	That's a very good match. You are asking them to give their location of their best catch of the night and for all we know there are actually records over here [areas that have rarely trawled] and people are being completely honest. I think the truth is that it is a big ask when you are giving people daily log books to give a lat and long for the whole nights catch. When you could be travelling six of these grids in a night.
Basford:	Would they also see that this could possibly be giving their competitors advantages if they knew where they fished?
Souter:	Well certainly in Queensland that data isn't given out. You know it's obviously held in commercial confidence.
Dichmont:	No that data normally doesn't get given out. It is interesting when I spoke to a very large player, someone who was a fleet manager in a very large company in the NPF. He says that when he phoned his skippers they where not telling him correctly where they where. He had to remind them that he has their VMS in his office and tell them that 'no, no I actually know where you are' and they'd go 'Oh yes I forgot.' So it does seem to be difficult to say where you are as a skipper. The location data in the logbook I do actually think that it is a good match. At least it is in the right area. They are not saying that they are in the other end of the gulf so I don't think they are lying to any great extend but what you are not picking up is the kind of structure within the data. So if you use a GIS you could really smooth between those areas whereas what seems to be really happening is that that is very aggregated fishing.
Adcock:	I was interested in the unknowns you listed for the fisheries earlier on in you talk. It seems to me that often we find the genetic data versus the fisheries model predictions that there is some sort of disparity there and that suggests that there is something missing from the models, perhaps, and this is possibly coming from those unknowns.
Dichmont:	Which models?
Adcock:	In terms of effective population size, which we will get to later in the forum, do you think therefore that you should start sorting out those unknowns as well as incorporating into these combined models that you think would be good.
Dichmont:	Yes. To be honest I can't quite see how we are really going to get out of it not combining them because the value of the whole system is the prediction and it will be very hard to predict without giving it a history. History is against the genetic technique at the moment because most fisheries won't recover a genetic answer in the last twenty years whereas we have catch rate data for the last twenty years or more depending on the fisheries. Yes I think the unknowns are the crucial things and honestly it does seem like fishery independent data the investment is well worth it even though it is expensive because it really, really reduces your risks. I have experience where I come from in South Africa. We started extensive surveys in the early 1980's. For example, in the hake stock assessment that I was involved with we had three surveys a year. There was very little doubt as to what was going on there in many respects. Just simply because catch rate data was going one way and survey data was going one way, you could understand why things were going wrong in the sense that there was quite large movements. You are much more informed with it, than without it [independent data sources]. Yes, what I was saying the unknown list should become a known list.
Adcock:	Can I suggest you put on that unknown list behaviour and reproductive biology? Both of those things are almost intractable.
Dichmont:	I must admit I see reproductive biology as a known. I can't find many fisheries that don't have that.

Adcock:	Well OK what I was thinking about in terms of behaviour is who is mating with whom and why?
Dichmont:	I imagine the genetics will really be able to answer that.
Ovenden:	Cathy [Dichmont] I think that the fisheries independent surveys that are now planned for the gulf that are industry funded will hopefully go a long way to addressing some of these uncertainties, perhaps.
Dichmont:	Sure, it is one of the reasons why we have really promoted it and successfully so and people are starting to see that. One of the ways of getting there was to be extraordinarily honest about the assessment.
Basford:	Thankyou very much. I think you have given us an excellent basis for the forum. I was very pleased to notice the interaction of the members of the audience. One of the things that the people who organised today wanted was to get feedback and discussion of the points. The aim was not just to get up and present things and sit back and not contribute. The aim is that you comment and make contributions because that is the way that this project is going to be able to gain from this forum. I would now like to introduce Jenny Ovenden. Jenny is a senior fisheries geneticist at the Southern Fisheries Centre, Agency for Food and Fibre Sciences in Queensland. Her main interests are in innovative application of population genetics in fisheries sustainability and she will talk about the theory and practice to estimating fisheries spawning stock size.

Theory and practice of population genetics to estimate fisheries spawning stock size

Jenny Ovenden

Southern Fisheries Centre, Department of Primary Industries, Queensland

Abstract

Commercial fisheries resources are renewable when managed sustainably. Resource sustainability contributes to the profitability of the commercial fishing sector with flow-on benefits to the state and national economies and to society in terms of reliable employment, particularly in regional areas. Damage to ecosystems and the environment by fishing practices are also minimized. Our work is aimed at evaluating a new data source for the process of fisheries stock assessment that is an essential precursor for sustainable resource management.

Population genetics contributes to at least three distinct areas within fisheries science; conservation biology, aquaculture and stock assessment. The main contributions to stock assessment are (1) determination of population structure to determine the most appropriate population unit for stock assessment and (2) genetic tagging for harvest rate monitoring and (3) the identification of cryptic larvae and juveniles. This research will make another connection between stock assessment and fisheries genetics – provision of estimates of spawning stock size.

Professor Sewall Wright (1889-1988) and Sir Ronald Fisher (1890-1962) independently proposed a simplified or 'idealised' set of conditions that facilitated the mathematical description of evolutionary processes including genetic drift. Along with mutation, gene flow and natural selection, genetic drift is one of the major drivers of evolutionary change at the population level. It is defined as random change in gene frequencies across generations, and results in random fixation or loss of alleles (genes). Importantly to this study, its magnitude is inversely proportional to population size. Because real populations are never 'ideal' Wright developed the concept of an 'effective population', which describes the number of individuals in an 'ideal' population consistent with an observed amount of genetic drift. In practical terms, it can be considered as the number of spawners that successfully contribute to the next generation. Effective population size can be estimated from empirical data by calculating (1) the allele frequency variance collected over a temporal interval, (2) linkage (gametic phase) disequilibrium and (3) heterozygote excess from a single genetic sample.

In a fisheries population model, spawning individuals contribute recruits to the next generation that are subsequently available to be harvested. Natural mortality on juvenile life stages and fishing mortality on adults severely limit the number of animals that are able to successfully contribute offspring to the next generation. The relationship between the number of successful spawners and subsequent recruits is critically important to stock assessment models. Genetic estimates of effective population size may assist with the more accurate, and independent measurement of these parameters.

The ratio between ecological and effective population size is determined by the variance in progeny number (or family size, reproductive success). In species with high fecundity and mortality, this variance can be large. Consequently, effective population size can be several orders of magnitude below apparent population size, inferring that the amount of genetic drift being experienced by that population is large. In an 'ideal' population, loss of alleles due to drift is balanced by input from mutation events (mutation-drift equilibrium). Other factors that counter the potentially adverse operation of genetic drift in fisheries species will be discussed.

Presentation

Slide 1: Theory and practice of population genetics to estimate fisheries spawning stock size.

This research project is a unique synthesis between the fields of population genetics and stock assessment. Until now there has only been a small amount of interchange between these fields and it is our hope that this project will bring them closer together towards the aims of sustainable management of our fisheries resources. The process is made more difficult because geneticists seldom understand the process of stock assessment and *vice versa*. Today I will introduce some of the population genetic theory that underlies this project.

Slide 2: Outline.

Firstly I will clarify the need and scope of this project in terms of what it is trying to do and what it is not trying to do. Then I will talk about how we use genetic methods to estimate spawning numbers, primarily from a genetic perspective. However I will try to integrate the genetic approach with the stock assessment scientist's approach of spawning stock size estimation. Finally, I will introduce how this project applies this genetic methodology to fisheries populations, particularly tiger prawns.

Slide 3: Need.

Unlike many of the resources that the Australian economy is based upon, we are lucky that fisheries resources are renewable. If the fisheries resources are managed correctly then they are going to be available for future generations. Several fisheries around the world are over-exploited and poorly managed. However the process of fisheries management in Australia is aimed towards sustainable goals. With sustainable resource exploitation, our economy, society and the environment is supported. This project is about the evaluation of a new data source for fisheries stock assessment. Cathy Dichmont pointed out earlier the limitations and difficulties facing stock assessment scientists; this project is one small step towards improving that situation.

Slide 4: Scope.

The field of population genetics and fisheries science have come together in three main areas; conservation biology, aquaculture and stock assessment.

In the area of conservation biology we can use population genetics to look at conservation units enabling us to make decisions about which part of the resource should be conserved for the future. Population genetics can also provide information about stock enhancement through translocation (breeding and release) programs, assist with the design of marine reserves, and provide information about the genetic diversity in order to assess the 'health' of particular populations in the fisheries resource.

In the area of aquaculture, population genetics can be used in breeding programs to improve profitability. Quantitative genetic markers are particularly useful for the selective breeding of desirable traits.

The area of stock assessment, where I mostly work, comprises of at least these three major streams. Firstly spatial heterogeneity and patterns of temporal stability of population genetic data can be used to define population structure and identify demographically independent populations. This spatial information can then be feed back to stock assessment scientists so that their models operate on the appropriate scale. Genetic tagging is a new field of population genetics that can assist stock assessment with mark and recapture data, among other aspects. Myself and many other population geneticists have been using genetic methods to identify larvae of several species that are difficult to identify through other means. The main emphasis of this study is to use genetics to estimate the effective population size. There are important links between the use of effective population size and conservation biology, but for the purposes of this forum we need to look upon these new genetic data as contributing to the process of stock assessment. Thus we will interpret the consequences of differing effective population sizes in the context of how it contributes to stock assessment.

'Idealised' population for studying genetic drift

- Diploidy
- Sexual reproduction
- Non-overlapping generations*
- Random mating
- No migration
- No mutation
- No selection
- (distribution of number of progeny per parent approaches the Poisson distribution)

Professor Sewall Wright (1889 -1988) &

Sir Ronald Fisher (1890 -1962)

Slide 5: Idealised population for studying genetic drift.

Professor Sewall Wright and Sir Ronald Fisher are the founding fathers of population genetics and invented the concept of an ideal population to facilitate the development of mathematical models to study evolutionary processes. There is a long list of assumptions that define an ideal population; the list is perhaps so overwhelming that it is difficult to see how any natural population could possibly conform. While this may be the case, the ideal population provides us with a starting point from which evolutionary models can be built and progressed. The population genetic literature is full of mathematical and empirical descriptions of the effects of departure from these assumptions.

For example, this project has already investigated the departure of the assumption of non-overlapping generations even though we are currently dealing with tiger prawns in Moreton Bay that have non-overlapping generations; they are an annual crop, they breed and then they die. However tiger prawns in the Gulf of Carpentaria have overlapping generations and we have been looking at computer simulations to see how departures from this assumption bias our estimates of effective population size. In this case, if we do not apply a correction factor for overlapping generations our estimates will have a downwards bias of 10%. The point being that all of these assumptions about an ideal population can be addressed, have been addressed and will continue to be addressed by population geneticists.

An additional assumption to estimate effective population size, that was not part of the original ideal population of Fisher and Wright, is that the variance in progeny number should approach a Poisson distribution. The effects of departure from this assumption will be discussed later.

Genetic drift

- Drives evolution, along with mutation, gene flow and natural selection.
- Equal to random change of allelic frequencies across generations.
- Magnitude inversely proportional to population size.
- Results in fixation or loss of alleles.

Slide 6: Genetic drift.

The measurement of effective population size relies heavily on the concept of genetic drift. Genetic drift is one of the factors that drives evolution and operates in conjunction with mutation, gene flow and natural selection to cause genetic changes in populations that eventually result in the formation of new species. The definition of genetic drift is the change in allele frequencies from generation to generation and it is a random process due to the sampling of gametes. Most importantly for this project is that the magnitude of genetic drift is proportional to breeding population size. Genetic drift ultimately results in both the change in allele frequencies and their eventual loss or fixation in the population.

Effective population size (N_e)

- Is the number of breeding individuals consistent with an observed amount of genetic drift
- Equivalent to the number of spawners that successfully contribute to the next generation
- Is not census size, number of breeding individuals, or "long-term" effective population size
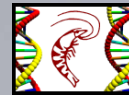
Slide 7: Effective population size.

Genetic drift occurs in all natural populations and the effective population size is the size of the breeding population that is consistent with the measurable amount of genetic drift. In a practical sense, the effective population size is the number of spawners in a fisheries population that effectively contribute genes to the next generation. A lot of animals may be involved in spawning but only a few of them, due to variable recruitment processes and environment, are actually successful at leaving behind offspring that make it to the next generation. It is important to remember that the effective population size is not the same as the census size of the population, nor is it the total number of breeding animals, nor is it the long-term effective population size. This project is concerned with the short-term measurements from generation to generation of extant genetic drift not the long-term estimates of genetic drift that are extrapolated from the number of breeding individuals over evolutionary time scales.

Estimation methods

- **Temporal**
  - *Allele frequency variance* across generations
- **Point**
  - Deviation of expected compared to observed degree of *linkage (gametic) disequilibrium*
  - *Excess of heterozygotes* compared to proportion predicted under Hardy-Weinberg equilibrium

Slide 8: Estimation methods.

There are two methods available to estimate effective population size and both depend on being able to detect genetic drift.

The temporal method detects change in allele frequency through time. Samples are taken from the population at two different time intervals (one generation or more) and used to determine the amount of drift that has occurred and convert that into effective population size estimates.

There are several other ways to estimate effective population size that do not require samples to be taken over time. We have used computer simulations to trial two of these point methods, linkage (gametic) disequilibrium and heterozygote excess. The latter method is more suitable for estimation of effective population size in small captive populations rather than large populations like Moreton Bay tiger prawns.
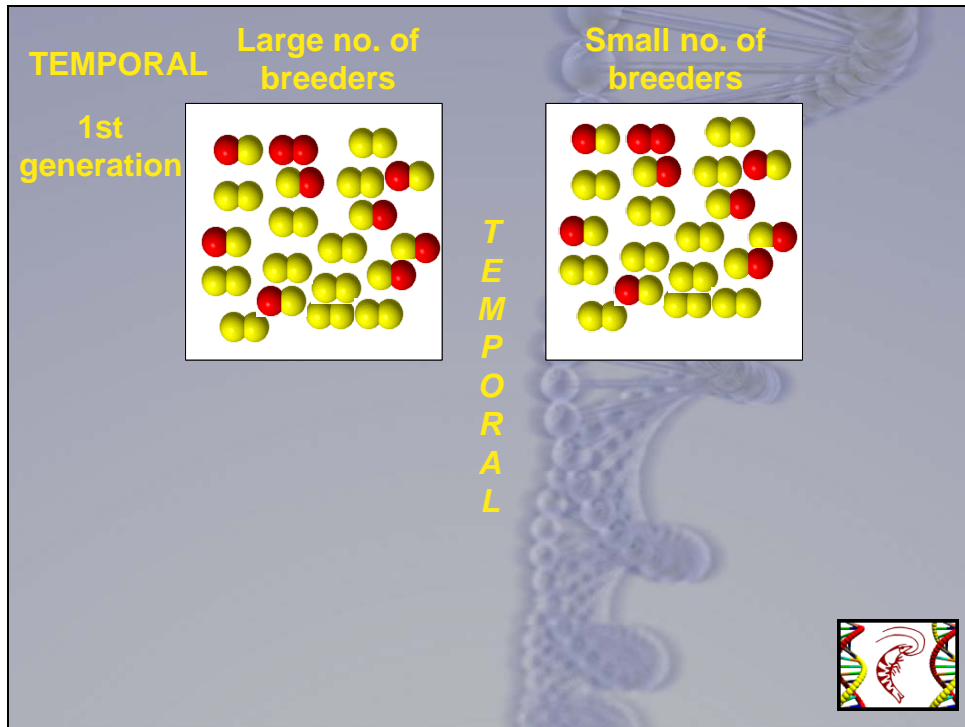
Slide 9: Temporal method demonstration – 1.

This example shows two identical populations each with the same number of individuals. Each animal is represented by a pair of balls depicting their alleles at each locus. Some individuals are homozygotes having identical alleles (both red or both yellow) at a particular locus and some individuals are heterozygotes having different alleles (one red and one yellow) at a particular locus. The population on the left has a larger proportion of successful breeders than the population on the right.
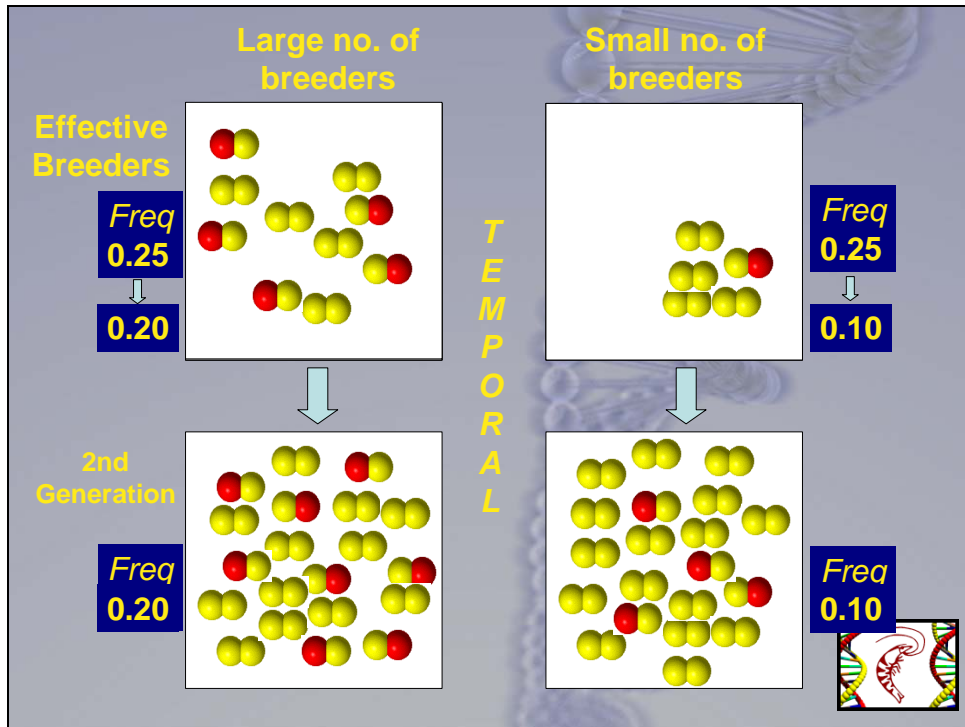
Slide 10: Temporal method demonstration – 2.

The starting frequency of the red allele is 0.25 in both populations. If we randomly sample the individuals that will be successful at transmitting their genes to the next generation the frequency of the red allele changes just by chance. The frequency of red alleles changes from 0.25 to 0.20 in the larger breeding population and from 0.25 to 0.1 in the smaller breeding population. While the direction of change is random, the magnitude of change is likely to be bigger in the population with the smaller number of breeders because of sampling effects associated with choosing fewer individuals.

The second generation have the same frequency of the red allele in the population as was observed in the population of successful breeders. The magnitude of the variance of change in allele frequency is proportional to the size of the breeding population. This is essentially what we are measuring when we measure genetic drift. Measurements from a large number of individuals are done over a number of independent loci to obtain more accurate estimates of genetic drift. For example, we measured the change in frequencies of 133 separate alleles from 1340 tiger prawns in Moreton Bay to get accurate measures of genetic drift over one generation.
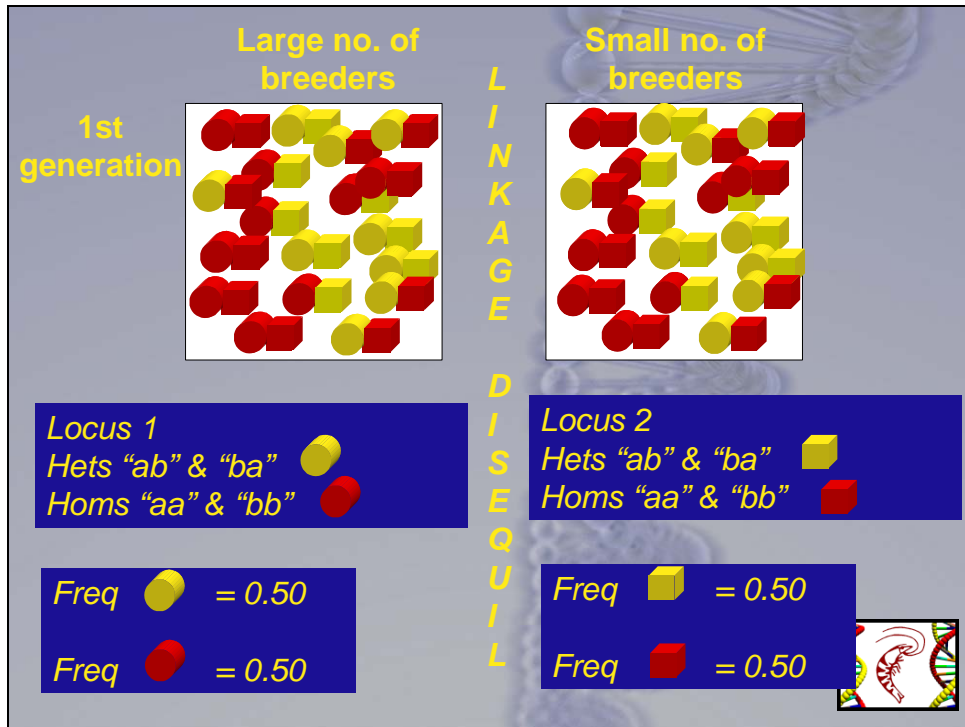
Slide 11: Linkage disequilibrium method demonstration – 1.

The linkage disequilibrium method does not need temporally spaced samples to estimate effective population size. Estimates can be generated from a single sample.

This example shows two identical populations each with the same number of individuals. Each animal is represented by two loci (a circle and a square) each with two alleles at each locus. Some individuals are homozygotes having identical alleles (red) at a particular locus and some individuals are heterozygotes (yellow) having different alleles at a particular locus. The population on the left has a larger proportion of successful breeders than the population on the right.
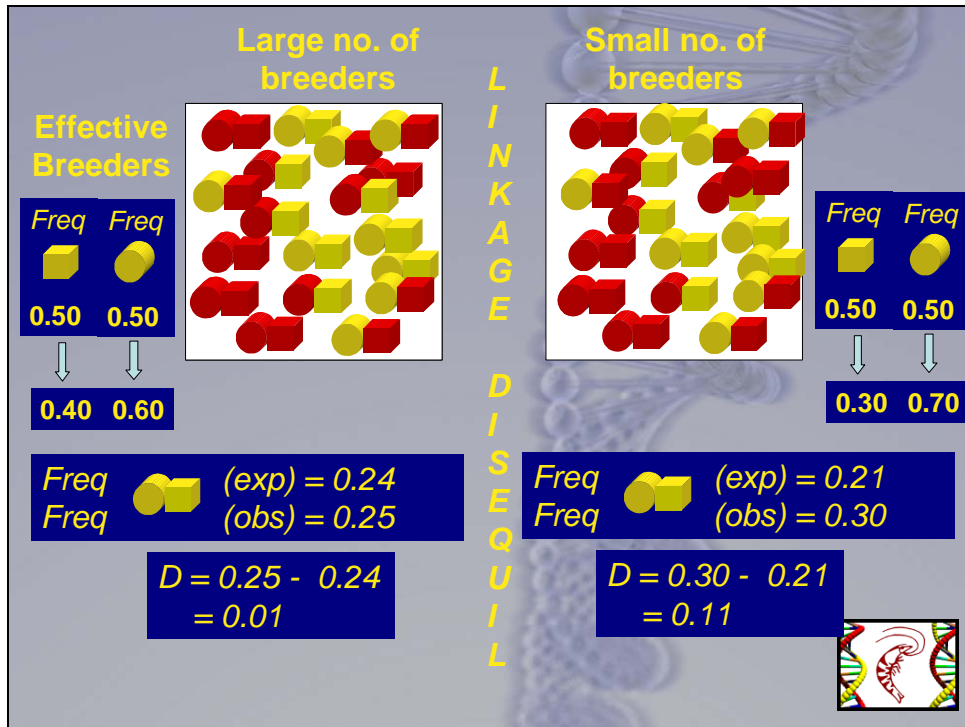
Slide 12: Linkage disequilibrium method demonstration – 2.

The proportion of heterozygotes is 0.5 for both the circle and square loci in both populations. If we randomly sample the individuals representing those that successfully transmit their genes to the next generation, the proportion of heterozygotes at the circle and square loci change just by chance.

The proportion of heterozygotes has changed from 0.5 to 0.40 at the square locus and from 0.5 to 0.60 at the circle locus in the population with a large number of breeders. In the population with the smaller number of breeders the proportion of heterozygotes has changed from 0.5 to 0.30 at the square locus and from 0.5 to 0.70 at the circle locus. While the direction of change is random, the magnitude of change is likely to be bigger in the population with the smaller number of breeders because of sampling effects associated with choosing fewer individuals.

From our data we are able to predict the expected number of animals that should be heterozygous at each loci and compare that to the number of heterozygotes we observe. The observed proportion of heterozygotes in the population with the larger number of effective breeders is 0.25 and the expected proportion is 0.24. The observed proportion of heterozygotes in the population with the smaller number of effective breeders is 0.30 and the expected proportion is 0.21. Parameter D is the difference between the observed and expected number of heterozygotes and is inversely proportional to the effective population size. The population with the smaller number of effective breeders experiences more genetic drift resulting in larger values of parameter D.
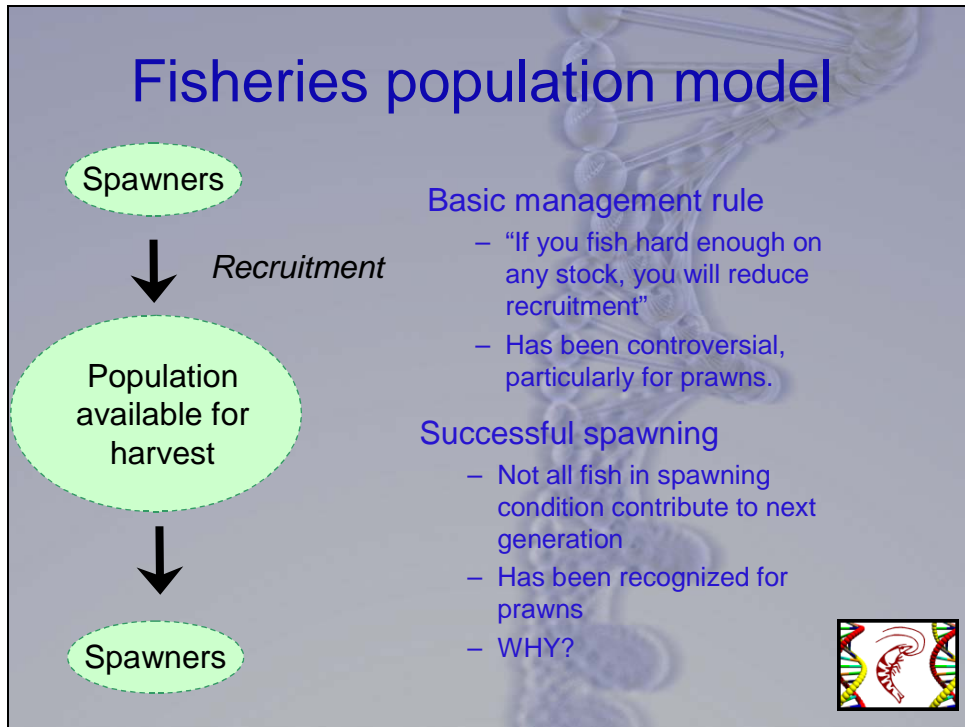
Slide 13: Fisheries population model.

Having introduced you to effective population size and how to calculate it, I am now going to demonstrate how the concept of effective population size can be integrated into fisheries population models. Fisheries stock assessment scientists use models of populations through time to investigate population dynamics and respond to different management regimes. In these models offspring are produced that appear in the next generation through processes of recruitment to become available for harvest. From the population that is harvested a certain number of animals are randomly selected to become the successful spawners for the next generation. While this may appear basic it is important to understand the structure of this model in a fisheries context. When you look at the fisheries population in this way you realise that if the spawning population is fished hard enough then the number of animals available in the next generation for harvest will definitely be reduced. The connection between the numbers of spawners and the amount of recruitment is very controversial, especially in highly fecund species like fish and prawns. It was not thought that harvesting spawning populations could have an effect on the size of the population available for harvest in the next generation.
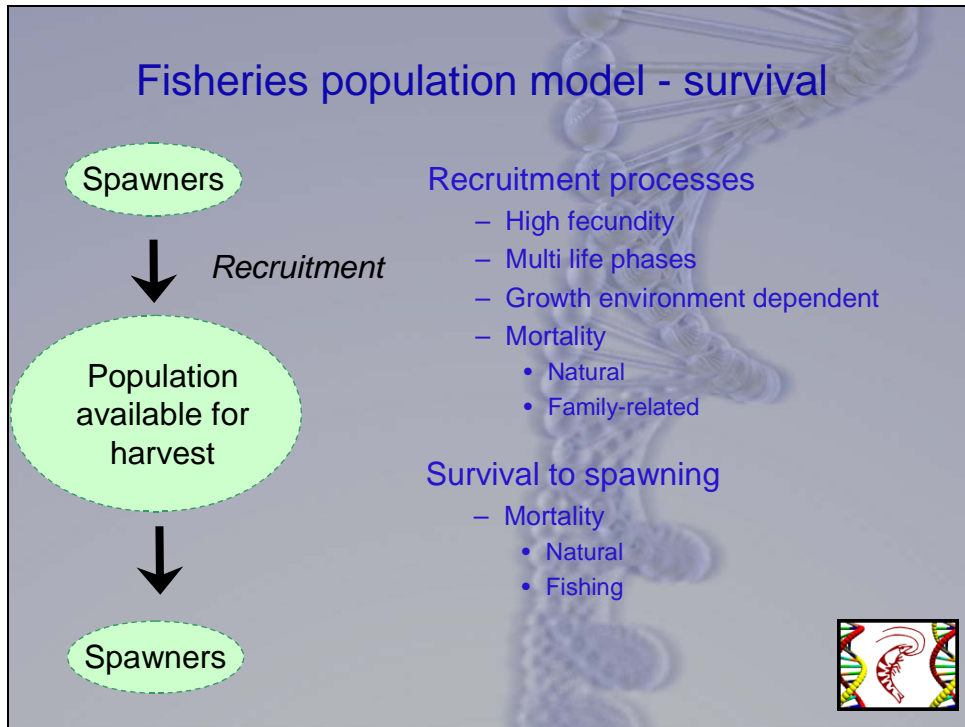
Slide 14: Fisheries population model – survival.

It is important to emphasise that not all individuals in the spawning population will contribute to recruitment because of high fecundity and low survivorship. Tiger prawns, for example, produce 500,000 eggs per female (approximately) and not all of them are going to survive. Many fisheries have different life phases that require different environmental conditions for growth and survival. There are many aspects of mortality that can affect recruitment. Family related mortality provides a particularly interesting case where whole batches of eggs are wiped out due to predation or cold shock, for example. Mortality can also affect the spawning individuals; both natural and fishing mortality can have significant effects.

Fisheries population model -measurement

Spawners

Recruitment

Population available for harvest

Spawners

Spawners
- Number of eggs
- Number of females times fecundity
- Number alive times fecundity at age
- Biomass at reproduction
- Index of abundance at reproduction

Recruits
- Number of individuals alive at a given time after egg stage.

Spawners and Recruits
- CAN BE INFERRED EMPIRICALLY BY ESTIMATING N_e

Slide 15: Fisheries population model – measurement.

Fisheries models need good measurements of the numbers of spawners and the number of recruits. Finding a relationship between the number of spawners and number of recruits is not as simple as it sounds. How do you measure the number of spawners? Do you count the number of eggs that float to the surface after spawning? Do you look at a number of females and work out how many eggs they are carrying? Do you look at the number of individuals that were caught during their reproductive period? All of these methods have been documented in the literature but they are less than satisfactory.

The estimation of the number of recruits is more straightforward because it is a measure of the number of individuals alive after a given stage and can be estimated through CPUE and fisheries independent data.

Incorporating effective population size estimates into fisheries models can give an index of both the spawning and recruitment numbers and therefore offers great potential for fisheries stock assessment.

Species with type III survivorship

- High fecundity and high mortality
- Characteristic of many fisheries species
- Have non-Poisson distribution of progeny, so

$$N_e = \frac{4N - 2}{V_k + 2}$$

- V_k is variance in progeny number (=variance in family size/reproductive success)

Slide 16: Species with type III survivorship.

Animals with type III survivorship like prawns and many other fisheries species have very high fecundity and mortality. These species do not have a Poisson distributed expectation of the number of progeny where the mean and variance of the number of offspring produced are the same. In species with type III survivorship the reproductive variance is enormous and certainly non-Poisson; for example many prawns in a population may not produce any offspring at all. The equation presented here relates effective population size (N_e) to the census number (N) by being inversely proportional to the size of the reproductive variance. Census number is the total population size, regardless of the life history phase of the individual. In our empirical measurements of the Moreton Bay population of tiger prawns we are comparing N_e to that proportion of the census population size that are adults in spawning condition (N_a).

Slide 17: N_e and type III survivorship.

How might the variance in offspring number effect our genetic models? Large variances in progeny number can result in very small effective populations. A review of the relatively limited literature has shown that the effective population size can be orders of magnitude below the number of spawning adults.

Slide 18: Mutation-drift equilibrium (MDE) 1.

When effective population size is small the implication is that the population is experiencing a large amount of genetic drift. The general question is how can a population retain genetic diversity in the face of such large amounts of genetic drift. To understand this process we need to understand the concept of mutation-drift equilibrium. Alleles that are lost in a population due to drift are generally replaced by either mutation or migration. At equilibrium then, despite the loss of alleles due to drift, genetic diversity is generally maintained.

This equilibrium can be perturbed by a severe reduction in population size (bottleneck). Extreme fishing pressure can disturb this equilibrium, as can other catastrophic events. Admixed populations from translocation events can also perturb the mutation-drift equilibrium. However after 4Ne generations populations will generally return to equilibrium. This process occurs more rapidly in smaller populations because of stronger genetic drift.

Mutation-drift equilibrium (MDE) 2

- Some type III marine species (eg. oysters Hedgecock et al 1992)
  - Have N_e / N ratios of 10^{-6}, and thus
  - High rates of genetic drift
- MDE predicts type III species are losing, have lost, or are unable maintain 'adequate' genetic diversity due to small N_e
- Can this be reconciled with their biology?

Slide 19: Mutation-drift equilibrium (MDE) 2.

Species with type III survivorship typically have small N_e/N ratios and are subject to very high levels of genetic drift. How can these species maintain levels of genetic diversity with such small effective population sizes? This is perhaps beyond the scope of this talk and not directly relevant to the use of effective population size estimates as a new data source in fisheries stock assessment. The issue is important, however, and is being addressed in the literature.

Conclusion

- Goal is to contribute genetic effective population size estimates to fisheries stock assessments.
- Measuring genetic drift in an 'ideal' tiger prawn population.
- Converting genetic drift into estimates of effective population size (=number of successful spawners)
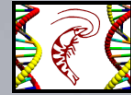
Slide 20: Conclusion.

In summary, this project uses effective population size to contribute to fisheries stock assessment models that estimate spawning and recruitment numbers. To do this empirically we are measuring genetic drift at microsatellite loci in Moreton Bay tiger prawns (*P. esculentus*). In addition to having non-overlapping generations, as discussed earlier, the population is unlikely to be exchanging genes via migration with adjacent populations because they are too distant. These, and other features of this population, make it an ideal study population for this project.

THE
END

prawn to be wild

Questions

Donnellan:	You have enormous power these days to use large numbers of loci and generate large numbers of alleles. How do you estimate the confidence of your estimate N_e?
Ovenden:	There are several methods for doing this under the temporal framework. Robin Waples proposed that the distribution of confidence intervals around the true estimates conforms to the chi-squared distribution. He gives formulas to calculate confidence intervals. You can also do bootstrap resampling on your data to estimate confidence intervals. The theory behind the linkage disequilibrium, which was one of the point estimates, incorporates equations to estimate the confidence intervals. It has been of paramount importance to us also, because we need to know the statistical power of the estimates that we are producing. We have addressed this through simulation experiments that David [Peel] will report on later in the day.
Donnellan:	What sorts of numbers of loci are you looking at to get reasonable confidence intervals?
Ovenden:	It depends a lot on what you think the effective population size of your population is. If you are dealing with a naturally occurring population then your effective population size estimates are going to be relatively large, probably larger than have been attempted to measure in captive or experimental populations. When we approached this study we thought we would get as many loci as we can and as it turned out we got eight highly variable microsatellite loci. I think the answer to your question is that for every study you would have to do a power analysis beforehand to access the number of loci that you would need.
Moore:	With the techniques you have been using, certainly the temporal method has been used previously; one of the problems has always been infinite error due to very little change in large populations. Is there any way of getting around that? I mean the old model suggests that if you add overlapping generations and a large population size then error would basically be infinite.
Ovenden:	Yes that's very true and I take on board all of those comments. What we have been trying to do is look at ways of getting around that. When you estimate the amount of genetic drift in a population the formula subtracts from the genetic drift estimate the amount of sampling error and the sampling error is inversely proportional to the number of samples you have analysed. One way of increasing the power is to increase the sample size. In the simulation experiments we've tested increasing sample size up to several thousand individuals to assess how this increases the statistical power of the estimates.
Moore:	Also increasing loci?
Ovenden:	Increasing the loci is also very important as well. David [Peel] will talk about the simulated results there as well. We've been able to address this not only under the temporal framework but also under the linkage disequilibrium framework and there are some quite interesting contrasts between the two methods as you increase the resources available to your study.
Ye:	How many samples and loci do you need to obtain reliable estimates for a fishery having hundreds of millions of spawners like the Northern Prawn Fishery?
Ovenden:	Firstly the number of samples you need is proportional to the true effective population size, which of course is unknown to start with. The practical way of doing that would be to ask your self the question what the ratio between effective population size and census size is in this species. In the literature the effective population size is several orders of magnitude below the actual number of spawning individuals. You may have one hundred million prawns in the population but the number of effective prawns is going to be many times less than that. Once you have estimated that you can do some power analysis to work out what the sample size is going to be necessary to estimate effective population size. Later in the forum you will hear some more detailed presentations about how sample size and number of

loci need to be increased to estimate effective population size when it is a large value.

Ye: Normally a population varies very much in size from year to year. How can you reliably estimate N_e over time? Is the ratio of N_e/N_a constant over time?

Ovenden Well this information is not available. This field is very new and there have been no instances where people have gone back a couple of years later and re-estimated effective population size and made these comparisons again to census population size. So the answer to your question is that we do not know.

Souter: Yes my question may be the same thing. Given the variability in recruitment success and other variables, how easy would it be to get a consistent relationship between N_e and actual population size. How easy would it be to estimate actual population size from N_e?

Ovenden: It would be great to say that given an effective population size we could instantly convert it to population size by multiplying it by this magic factor, but we know that that's not true. There are many environmental factors driving recruitment; we have to look at the different processes of mortality, we've got natural mortality, family related mortality, and fishing mortality that affects this. The application of this method to fish stock assessment has to go hand in hand with the estimation of these parameters as well. The flip side to your question is can we use effective population size to measure recruitment as well? Again I'd love to say yes by a factor of X. There is the potential there to do that but we just don't have enough information available yet to be able to do that.

Donnellan: Yes I was wondering, not so much with prawns, but in teleosts where you might get cohorts of fish that are hanging together, you may then have population structure on both a temporal and spatial scale. This means that sampling, especially on a temporal scale, without assessing the age or cohort of a fish, might be confounded.

Ovenden: What Steve Donnellan is asking about is for species, which do have overlapping generations and also have spatial population structure as well. It's very important when applying effective population size methodology that you have already understood the genetic structure of the population. That would be a precursor to the implementation of any of these studies because you can understand that the influx of new alleles through migration or the loss of those alleles by migration simulates the effects of genetic drift. Therefore in the presence of migration you overestimate genetic drift, which means that you underestimate effective population size. This is a very important area. The other question about overlapping generations is that you might for example have a fin fish and the three year old individuals that you sampled this year will actually be the four year old individuals you sample next year. When you are dealing with species that have overlapping generations you must also be able to age those individuals quite accurately to make sure that you are not just selecting the same cohort from year to year; or have some other way of guaranteeing that.

Basford: What is a type 3 species?

Ovenden: A type three species is one that has high fecundity, in other words produces lots and lots and lots of offspring, like the prawns, which on average produce 500,000 eggs per female and consequently experiences a high degree of mortality as well.

Gopurenko: Just a pretty general question for you. Assuming sometime in the future after all of the work you have done that you do come up with a good concrete and persistent estimate of N_e, and that it is reliable, it is going to have to be an ongoing process, isn't it? And not just ongoing over the years but also ongoing within the season itself just to breakdown the compartments of error. Also you are going to have to look at transferring it to other species and fisheries. Do you think fisheries industries are going to fund this kind of analysis over great periods of time? Granted that technology is going to get cheaper.

Ovenden: We've been fortunate in that we've been able to obtain funding for this very preliminary step towards the implementation of this new method in stock assessment, so that's excellent. Whether it continues or not very much depends on the support that we get from our colleagues and from the

industry. Yes, there are many areas that need to be investigated. It's a bit like the use of growth rings in otoliths in stock assessment, it's another tool you can add to stock assessments to make them more accurate. Yes, it would have to be measured on an annual basis, so if you were dealing with prawns which are an annual crop to make it most useful you would need estimates every year. So I suppose it's an ongoing monitoring tool. We haven't actually done the sums but Cathy [Dichmont] talked about the enormous cost of doing fisheries independent surveys, which involves boats, huge numbers of people and lots of time. Compared to that kind of cost genetic monitoring of population size is more straight forward, it all happens in the lab, after the samples have been obtained of course, and it's very cost competitive with that kind of extra data.

Gopurenko: Certainly the initial outlay might not be high but once the system is set up and provided that it is going to provide the consumer with data and obviously you've got the price of the technology dropping, yeah it will become much more cheaper. Gee it would be great to get some of those $400000 grants. You could do a lot of species for that.

Salini: Jenny [Ovenden] there are some implications for general population genetics.

Ovenden: Yes lots.

Salini: Most studies in the literature deal with snap shot estimates and that's published in detail for the stock structure of many species. If you were able to get historical samples, for example from Moreton Bay from 10-20 years ago and then compare it to current generation allele frequencies and if you got no differences what would you infer.

Ovenden: The implication would be that sampling error has been too high or there has been a low amount of genetic drift occurring, which implies an infinitely large effective population size.

Salini: Couldn't you also infer that populations are stable because the allele frequencies are stable?

Ovenden: Yes. The thing is that evolutionary theory predicts that genetic drift is occurring constantly. The question is do we have the power to detect it?

Salini: That's what I was getting at. I always used to think that those sorts of differences were over thousands of generations. It's the timescale that is a big factor here.

Ovenden: The effects of genetic drift are cumulative so that if you measure it over several generations you should have a much larger amount therefore your estimates of effective population size will be correspondingly smaller. We have certainly toyed with this concept within this project instead of looking at a temporal interval separated by only one generation, which is what we are doing with tiger prawns. For example, what would happen if we looked at drift over ten generations and converted that to effective population size? Not only that, we would have a sliding window approach to that [measuring drift].

Salini: Am I assuming something wrong here? What would be wrong to assume that the populations, the allele frequencies were stable over time, that there was no difference over that period of time, is that a fair result?

Ovenden: I think where you are coming from here John [Salini] is that in tests for population structure population geneticists often ask the question is this population stable over time, is there significant differences in allele frequencies through time? In many cases our population structure tests show that there is no significant difference in allele frequencies. What this methodology does is that it doesn't just rely on whether the change is significant or not, what is the total magnitude of change.

Dichmont: I was just wanting to comment about survey information and say something in the defence of surveys in there. I don't think it is one or the other. To be honest, where you make savings is perhaps not to do it as often. The reason why I say that going with one or the other is not a good thing is that when you look at the survey we did we have data from eight prawn species and all byproduct, so that's seven other species, and all the bycatch we also collected including temperature and depth. To replace that with a genetic technique would really be quite difficult, would be very expensive, in the

	order of $400000. I think that what I am trying to say all along is that these things are complimentary.
Ovenden:	I was never advocating the replacement of surveys with genetic data. I was just doing a comparison of the research money that was available for these studies.
Hale:	Wouldn't this method give you retrospective information about stock biomass and what you are after is predictive information about biomass.
Ovenden:	Yes. It's the role of the stock assessment models to make predictions into the future. The data that we are collecting is allowing us to strengthen the biological data that the models are based on currently. In other words, refine the models and then it is the purpose of the stock assessment models to make predictions into the future. There is no way that I can predict what the effective population size is gong to be in the future, but I can validate the current models.

Software demonstration (NeEstimator)

Samantha Peel

Southern Fisheries Centre, Department of Primary Industries, Queensland

Abstract

The NeEstimator software estimates the effective population sizes (N_e) from allele frequency data. The user can estimate N_e using any of the three internal methods or three third party programs. Genotypes from a sample of the population are used as input. The user provides this data in GENEPOP, ARLEQUIN or simple column (eg. saved as a tab delimited text file from Microsoft Excel) format.

The three internal methods available to calculate N_e are:
- A point estimation method using linkage/gametic disequilibrium, (Hill, 1981).
- A point estimation method using heterozygote excess (Pudovkin, Zaykin and Hedgecock, 1996).
- A temporal method using moments based F-statistics (Krimbas and Tsakas, 1971; Nei and Tajima, 1981; Pollock, 1983 or Waples, 1989).

NeEstimator is also able to utilise the following third party programs:
- A temporal method using a Bayesian based approach called TM3 (Beaumont, 2002).
- A temporal method using a maximum likelihood based approach called MCLEEPS (Anderson and Williamson, 2000).
- A temporal method using a pseudo likelihood approach called MLNE (Wang, 2003).

The minimum computing requirements for running NeEstimator are a PC with 95, 98, ME, 2000 or NT (version 4.0 with service pack 3 or later), approximately 9Mb free hard disk space, and a mouse or compatible pointing device. Certain options require Microsoft Excel 97 or later.

Presentation

NeEstimator 1.0

Slide 1: NeEstimator 1.0.

The NeEstimator software estimates the effective population size (N_e) from allele frequency data. It was developed as part of a current research project using the Moreton Bay tiger prawn as a case study.

A total of 6 methods are available to estimate N_e using NeEstimator, two point estimation and four temporal estimation methods. Three of these methods are internal to the program and three are third party programs that can be executed from within NeEstimator.

New Analysis
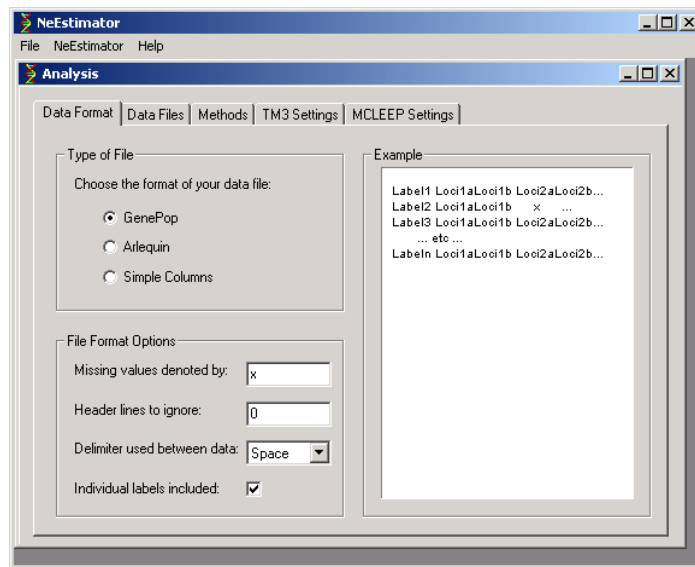
Slide 2: New Analysis.

NeEstimator uses genotype frequency data in GENEPOP, ARLEQUIN or simple column (eg. tab delimited text) format.   The program allows you to set various input file format options such as missing values. The program assumes that all data files are in the same specified format.

Loaded Data Files
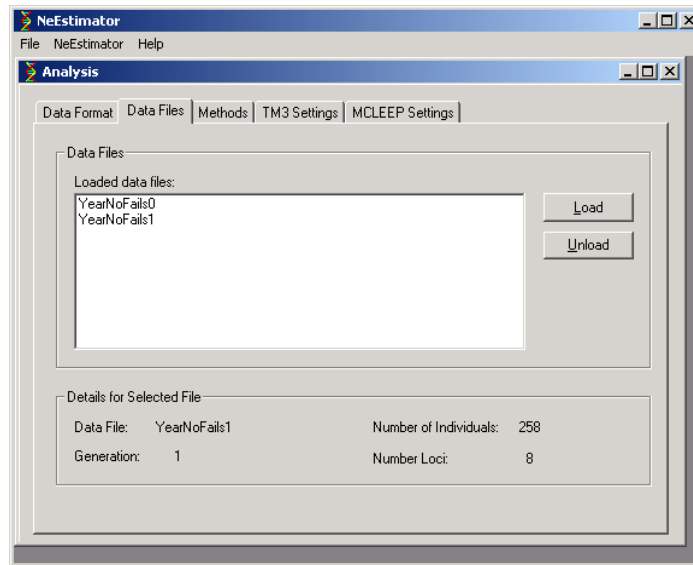
Slide 3: Loaded Data Files.

Each generation is loaded as a separate data file. The first generation is generation zero, the second is generation one and so forth. Generation summary information is displayed for each data file.

Configuring Third Party Methods

Slide 4: Configuring Third Party Methods.

Software for the three third party methods does not come bundled with NeEstimator. Links to download this software and installation instructions are provided in the online help. Once the third party software has been downloaded, installed, and configured it can then be executed via the NeEstimator interface.

MLNE Configured and Checked
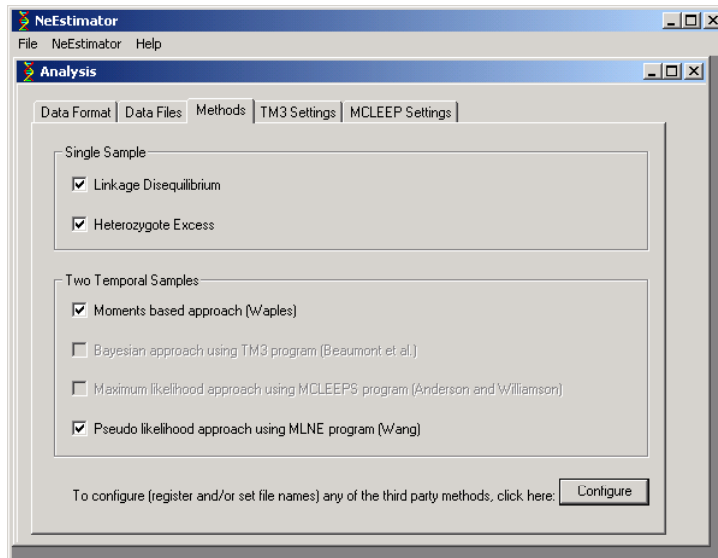
Slide 5: MLNE Configured and Checked.

The user can select any combination of the available methods to estimate N_e by clicking the corresponding check boxes. The third party methods are available once they have been configured (i.e. NeEstimator has been informed where to find the method's application file).
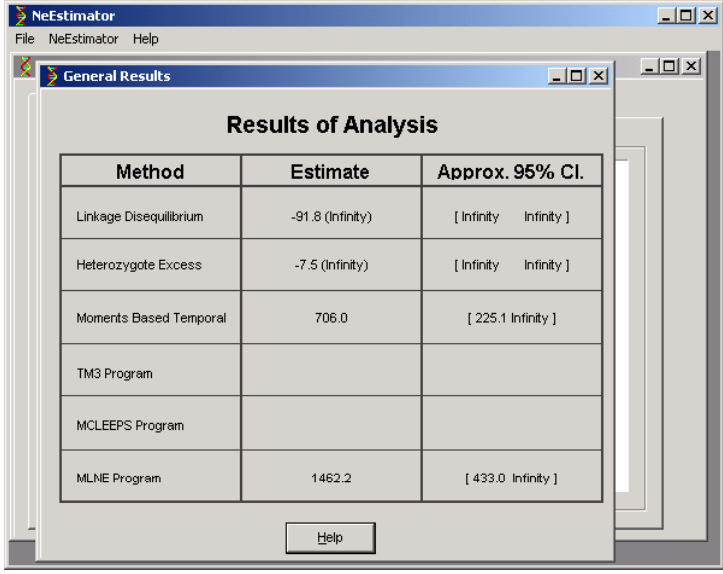
Running NeEstimator
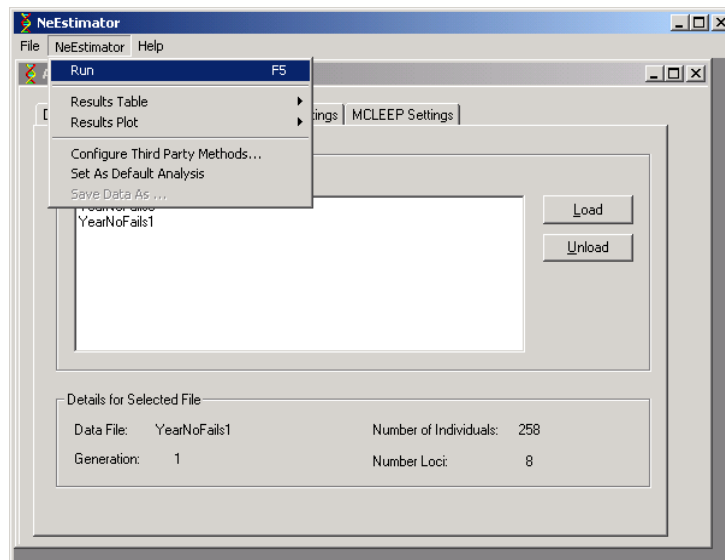
Slide 6: Running NeEstimator.

Some of the methods are very quick to run such as the heterozygote excess method while others are very computationally intensive such as TM3. Running times will vary depending on the size of the data set and the computing platform.

General Results

Slide 7: General Results.

Summary results are presented in a tabulated format for easy comparison among implemented methods. Detailed results for each implemented method are also available.
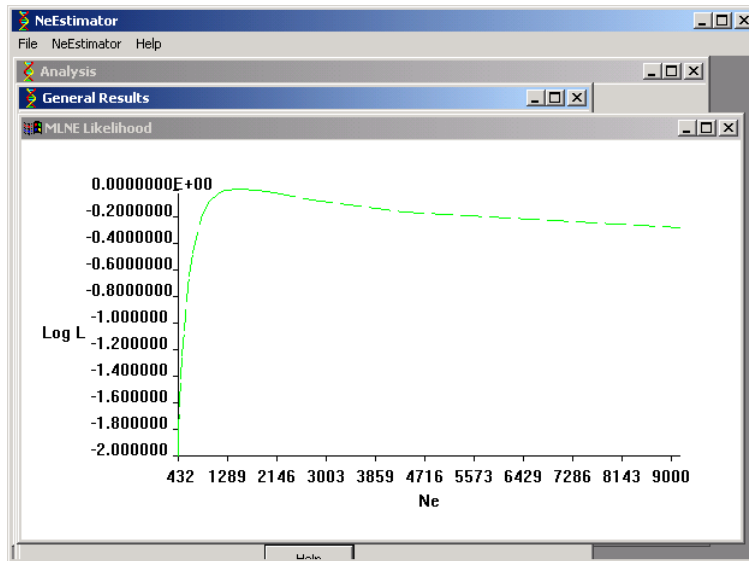
Slide 8: MLNE Results Plot.

Plots of the likelihood curve are provided for the MCLEEPS, TM3 and MLNE programs.

Questions

Basford: [Jenny Ovenden] did you want to comment about the output in the sense that often you were getting estimates with infinite confidence limits.

Ovenden: Part of the rational behind developing this software was to allow investigators who collect this data to easily and simultaneously implement many methods of calculating effective population size and then to quickly an easily compare results between methods. Sure, all of these methods exist in the literature but as far as we are aware this is the only software that allows a user to implement them quickly and easily. I can't really comment on the size of the confidence intervals other than to say they are presented for all of the methods and they can be used to assess the accuracy and efficiency of methods that were used. The three bottom methods TM3, MCLEEPS and MLNE are external to the program. They are based on executables written by the authors who wrote those papers. The way the program works is that you install the software from your CD and then you go to the web sites of the authors for these other programs and you download their executables onto your hard drive and you tell the software where those executables reside on your hard drive and the software will go there and call them and perform those analyses. I believe that MCLEEPS and TM3 take several hours to make their estimates as there is a tremendous amount of iteration involved in those likelihood estimates.

Basford: From a user point of view of course the concern would be that some methods are giving you estimates of zero and some methods were giving you estimates in the order of 1000.

D. Peel: We are essentially getting infinity as an answer for all the methods for the test data. The answer of 1000 has an upper confidence interval of infinity. The zero answer is a minor mistake, it should actually be reported as infinity.

Basford: There is still a little bit of work to be done explaining the output.

Ovenden: Yes. It is entirely up to the user to interpret the output from the program. It's not a panacea for analysis of the data. The user has to understand the methods. What we are doing is just providing a platform for the implementation of all of these methods in a very flexible user interface.

Basford: The ability to have the various methods in one software package is enormous but it doesn't take away the scientists involvement in having to understand what the different methods are doing and why they give you different estimates. But I think that is also a good point that David [Peel] made in that you have to look at those confidence intervals.

D. Peel: To be honest, when I first saw the results for Moreton Bay I thought the answers were way too small. I wondered what was going on and it's only when you look further into it and see the infinite upper confidence intervals that you really understand what is going on.

Basford: You really do have to understand very clearly the methods being used in the software to properly interpret the results.

Ovenden: Yes.

Moore: Is this software only for a PC interface or will it be available for Macs?

D. Peel: Its only PC based. It would be quite expensive time-wise to make a Mac version. We thought about that at the beginning, as it would broaden the user audience.

Moore: Is there an instruction manual that comes with it?

D. Peel: I was just about to say there is actually a full online help

S. Peel: Our web site is printed on the disk that is in your folders. It is also in the help file as well. So that could be where you go for updates, bug fixes, contact information, that sort of thing

D. Peel: Help also exists as separate documentation.

Moore: Is it going to be a program that you can run the analysis over the web or will it be one that you install on the site?

D. Peel: No it's not like GenePop, it's actually stand alone. Accept for the implementation of some of the methods, they require an external executable.

S. Peel: It's something that you download and install on your computer.

Basford: That's why you have to download the other software so that you can put them all together and make a comparison.

Arthur:	Can I just make sure I understand what that's estimating in the results table? You were saying it's basically a negative N_e from negative infinity to infinity. It is basically zero but you have no confidence in N_e?
D. Peel:	The negative is actually a repercussion of the way the equation was used and when you get an infinite answer it gets recorded as a negative answer. As far as the methods are concerned they cannot distinguish between an infinity population and an extremely large population.

Empirical results for Moreton Bay; comparison of swept area trawl survey and genetic estimates of spawning stock size of tiger prawns (*Penaeus esculentus*)

Tony Courtney and Jenny Ovenden

Southern Fisheries Centre, Department of Primary Industries, Queensland

Abstract

This presentation compares two estimates of the spawning population size of brown tiger prawns (*Penaeus esculentus*) in Moreton Bay. One estimate was based on swept area surveys undertaken in 2001 and 2002 to directly measure the number of adults in the spawning population (N_a). The other estimate used genetic methods on the same individuals to estimate the number of spawners that effectively contributed to recruitment (N_e). The Moreton Bay tiger prawn population was used because it is relatively small and closed. During the spawning season (October-November) the population of adults (>=30 mm carapace length) is at a minimum and was estimated from the swept area surveys to be about 650,000 in 2001 and 465,000 in 2002. Samples were genotyped for eight microsatellite loci, consisting of six tri-nucleotide repeat type, one di-nucleotide repeat, and one tetra-nucleotide repeat for a total of 133 alleles. As expected, F_{ST} was non-significant between years and grids. No significant linkage disequilibrium among pairs of loci was detected, as expected from a species with a haploid chromosome number of 44. Allele numbers and homozygosity per loci were in accordance with the infinite allele model for selective neutrality. Null alleles were detected at two loci, but assuming their frequencies did not significantly alter among years, the loci were included in estimates of effective population size. The moments-based temporal method yielded effective population size of 1077 with a non-infinite 80% upper confidence limit based on the chi-squared distribution. Maximum likelihood and Bayesian alternative methods gave similar results. The linkage disequilibrium method returned an infinite result. The ratio between the survey estimates (N_a) and effective spawning population size (N_e) was 0.002. This is equivalent to other reports for marine invertebrates (0.000001) and vertebrates (0.001 to 0.00001). Spatial separation between deep water spawning sites in the Bay and juvenile seagrass habitat may be contributing to the variance of reproductive success that is a factor in the relationship between survey estimates and effective spawning population size.

Presentation

Empirical results from Moreton Bay;
comparison of swept area trawl survey and
genetic estimates of spawning stock size of
tiger prawns (*Penaeus esculentus*)

FRDC Research project
2001/018

Tony Courtney and Jenny Ovenden

Thanks to

James Haddy, Keith Chilcott, Matt Campbell, Shane Gaddes, Raewyn Street,
Sally O'Connor, Darren Roy, Brett Davidson (QDPI Southern Fisheries
Centre)

Kaye Basford and Heather Podlich, University of Queensland

Ivan Rogalevs, Qld Commercial fisherman

Slide 1: Empirical results from Moreton Bay.

This presentation compares swept area trawl survey estimates and genetic estimates of effective spawning stock size for brown tiger prawns *Penaeus esculentus* in Moreton Bay. I would like to thank the several people who have assisted with the design, development, execution and data analysis of the trawl survey.

Slide 2: Affiliates and sponsors.

Several research bodies and funding agencies are involved in this project. FRDC is the main funding body, but we also acknowledge and appreciate the funding from AFMA.

Slide 3: Objective.

The objective for this part of the project is to compare the swept area survey estimates of tiger prawn spawning population size (N_a) with the genetic estimate of effective population size (N_e). I will present the first part that deals with the swept area survey estimates of tiger prawn spawning population size (N_a) in Moreton Bay. Jenny Ovenden will present the second part that uses molecular methods to estimate effective population size.

Slide 4: Biology and distribution of brown tiger prawns.

The brown tiger prawn is of major commercial value in Australia. There is a similar species of tiger prawn (the grooved tiger prawn, *Penaeus semisulcatus*) caught throughout northern Australia, but we are focused only on the population dynamics of the brown tiger prawn *P.esculentus* in Moreton Bay.
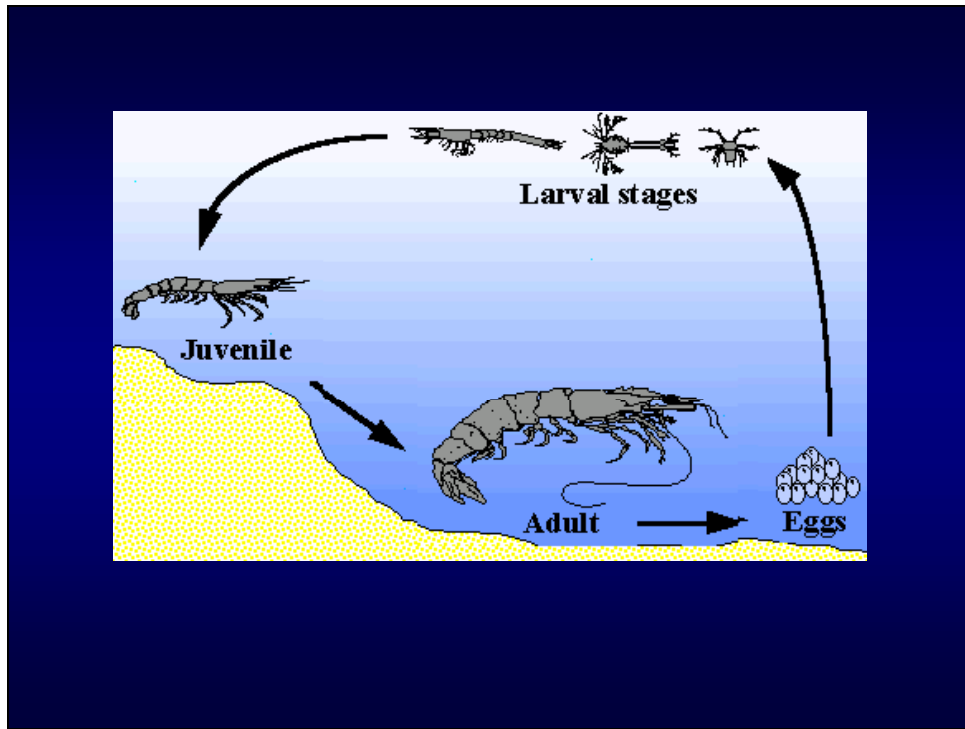
Slide 5: Prawn life history.

Female penaeid prawns, such as the brown tiger prawn, produce large numbers of eggs (i.e. 100,000-450,000). The eggs hatch within about 24 hours of being spawned and the young prawns spend two to three weeks in the water column during which time they undergo a number of larval phases. After about three weeks they move towards the shallow inshore nursery areas. Juvenile tiger prawns have a strong association with seagrass areas. They remain in the shallow water seagrass nursery habitats for about three months before moving offshore to deeper water where they continue to grow, mature and reproduce. Their life cycle takes about 12-15 months and mortality rates are high, especially during the larval stages.

Why choose the Moreton Bay tiger prawn population?

1. Relatively small, isolated population.

2. Close to southern geographical limit and therefore comparatively short spawning period.

P. esculentus endemic to Australia's tropical and sub-tropical waters from mid-New South Wales to Shark Bay, Western Australia (Grey *et al.* 1983)

Slide 6: Why choose the Moreton Bay tiger prawn population?

Tiger prawns are endemic to Australia and are found throughout northern Australia from northern NSW to Shark Bay in Western Australia. There are major fisheries throughout this region. The Moreton Bay population is a small, well-defined and relatively isolated population close to the southern limit of this region. Moreton Bay has extensive seagrass beds but to its immediate north and south there is little seagrass habitat. Brown tiger prawns do not undertake major migrations and the limited distribution of seagrass along the southeast Queensland coast indicates that the Moreton Bay population is relatively isolated, making it an ideal population to study. The spawning period in Moreton Bay is shorter than that of more northern populations making it easier to design sampling surveys to measure the spawning stock population size. About 90 tonnes of tiger prawns are caught annually in Moreton Bay, but at the time of spawning the population is actually quite small. Spawning is from October to November each year with a large pulse of recruits entering the fishery and being fished from February to May. The population experiences high fishing and natural mortality rates and as a result there are very few individuals alive to spawn in October to November.
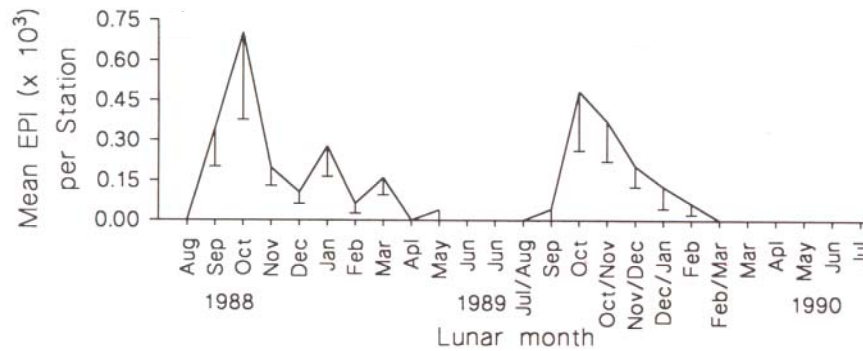
Egg production index (EPI) is a function of the size, abundance and proportion mature of adult females (from Courtney and Masel 1997)

Slide 7: Egg production index over time.

This graph of the egg production index (EPI; a compound index of relative egg production that incorporates the effects of the size of adult females in the population, their abundance and the proportion of females that were classified as ripe) demonstrates why we sample prawns in October and November. Most of the egg production occurs in these months with virtually no egg production in the cooler winter months. Each year (2001 and 2002) we undertook a stratified trawl survey during these months and estimated the spawning population size based on the catch rates of adults.

Survey design and population estimates

- Design considered the distribution and abundance from previous research sampling in the bay (1988-90), a large prawn trawl survey in southeast Queensland (1999) and logbook catch and effort data (1988-99) for Moreton Bay.

- Surveys undertaken in October/November 2001 and 2002 during peak period of egg production (Courtney and Masel 1997).

- Sampling undertaken in five 6'x6' logbook grids (out of about 30 grids). Grids treated as strata with sampling sites randomly allocated within each.

- The five grids accounted for 0.78 of all reported landings of tiger prawns from Moreton Bay. Final population estimates derived by extrapolating stratified mean upwards (1/0.78).

Slide 8: Survey design and population estimates – 1.

The survey design considered the distribution of tiger prawns from previous research and logbook catch and effort data. Tiger prawns are only a small component of the Moreton Bay trawl fishery with bay prawns and eastern king prawns being much more abundant and economically important. Catch rates for tiger prawns are relatively low.

The surveys were conducted in October and November 2001 and 2002 corresponding to times of peak egg production. Sampling occurred in 5 out of the 30 6' x 6' [6 arcminutes x 6 arcminutes] nautical mile logbook grids. Each grid was considered as a different stratum and we randomly allocated sampling sites within these strata. We were unable to sample the entire population so we focused on those areas that we knew contained the majority of the tiger prawn population. About 78% of the population occurs within the five sampled grids. The final population estimates were therefore extrapolated upwards by a factor of 1/0.78.

71

Survey design and population estimates (con't)

- Adult spawning population defined as "all male and female tiger prawns >= 30 mm CL" (Does not consider the proportion mature or ripe at present).

- 2001 survey – the five grids/strata had approx. equal sampling effort. 2002 survey design considered the strata variance from 2001 – strata with higher variance received higher sampling effort (Haddon 1997).

- Resources available for about 6 nights of vessel hire each year sampling approximately 13 one nautical mile transects (beam trawl 5 m wide x 1,852 m long transect). Resulting in 70-80 sites sampled each year.

- Population size estimates assumed 0.5 of those adults in front of the net were captured and retained (efficiency+selectivity = 0.5). (Joll and Penn 1990, Sparre and Venema 1992).

Slide 9: Survey design and population estimates – 2.

We needed to explicitly define what it was we were measuring in the trawl surveys. Rather than measuring the abundance of all size and age classes we were only interested in quantifying the adult population size. Male and female tiger prawns were classed as adults if they were equal to or larger than 30 mm carapace length (Courtney and Masel 1997). Although we macroscopically noted the development phase of the females' ovaries after capture, we did not consider the proportion of females that were mature or ripe in the final estimates. Instead we assumed that all adults contribute to spawning and recruitment and for this reason our estimates of the adult spawning stock size (N_a) may be slightly high.

There was equal sampling effort in each of the five strata for the 2001 survey. Based on the strata variance from the 2001 survey, we adjusted sampling effort so that those strata with high variance in 2001 received more sampling effort in the 2002 survey.

We had funding for 6 nights of vessel hire per year sampling approximately 13 one nautical mile transects with a five metre beam trawl. Each survey was based on 70-80 trawls.

When making estimates about population size using swept area survey methods, it is necessary to make assumptions about the proportion of prawns in front of the net were actually caught and retained in the net. For our study we assumed that half of the prawns in front of the net were caught, consistent with other studies (see Joll and Penn 1990, Sparre and Venema 1992).

Survey design and population estimates (con't)

Final population size estimates…
1. Derive stratified mean catch rate (number km^{-2})
2. Estimate total area for 5 grids (km^2)
3. Population = (2) divided by (1), above
4. Adjust estimate upwards due to sampling 0.78 of population (i.e., multiply 1/0.78)
5. Adjust for capture efficiency (0.5) (i.e., multiply by 2).
6. Use stratified mean s.e. to estimate confidence intervals.

Slide 10: Survey design and population estimates – 3.

The general process was to derive estimates of the stratified mean catch rate of tiger prawns, estimate the total areas in the five grids and extrapolate upwards to get the estimate of population size within each grid. We then adjusted that estimate upwards because we were only sampling about 0.78 of the population. This figure was then adjusted upwards to take account of the selectivity and retention of the prawns in the net. Confidence intervals around the final estimates were derived from the confidence intervals around the stratified means.

Slide 11: Survey vessel and gear.

The 2001 survey used the DPI vessel (RV Warrego) and the gear pictured here. This vessel was subsequently decommissioned and for the 2002 survey we chartered a commercial vessel but we used exactly the same net, sampling gear and methods.
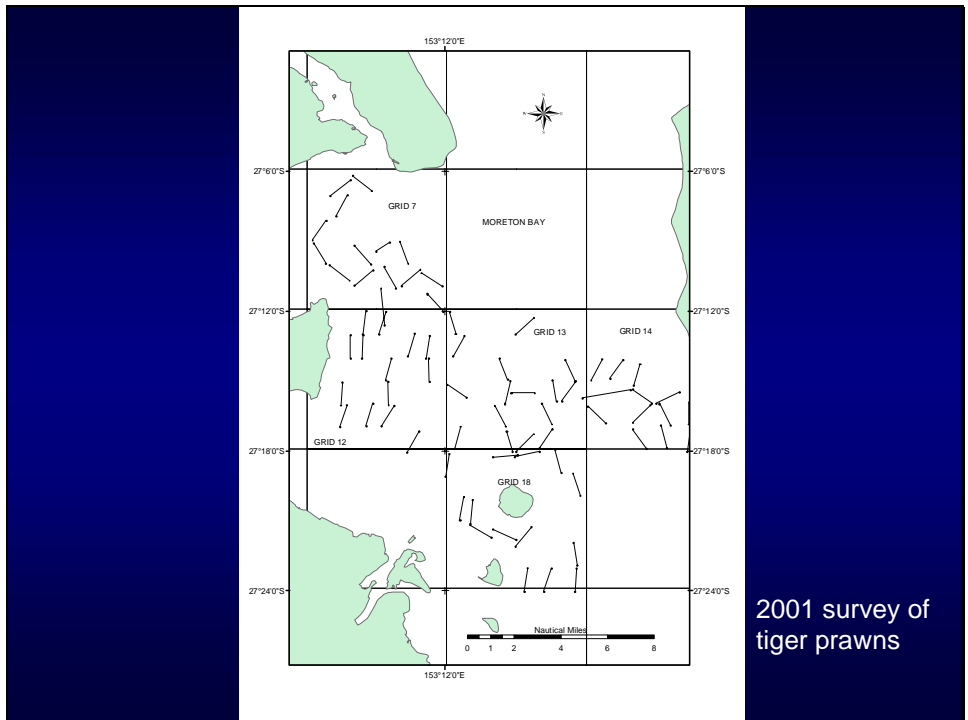
Slide 12: Trawl tracks for the 2001 survey of tiger prawns.

A chart of the 72 one nautical mile trawls conducted during the 2001 survey. The transects were located in five 6' x 6' [6 arcminutes x 6 arcminutes] logbook grids. Individual trawl tracks are represented as a black line.

Survey results for 2001. The variance estimates are contributions from each grid to the variance of the overall stratified mean. This approach was used to apportion the sampling effort across grids for the 2002 survey. (from Haddon 1997)						
Stratum (or grid)	Grid 7	Grid 12	Grid 13	Grid 14	Grid 18	Totals
Mean catch rate of prawns	2.53	2.21	7.75	3.69	3.79	
No. stations sampled	15	14	16	13	14	72stations sampled
Strata area sq km	122.4	113.1	123.5	103.8	116.5	579.3sq. km total area
SD of mean	2.21	2.21	2.21	2.21	3.96	
Strata weight	0.21	0.20	0.21	0.18	0.20	1
Weighted mean	0.53	0.43	1.65	0.66	0.76	4.04 (stratified mean)
Variance	0.0145	0.0133	0.0139	0.0121	**0.0455**	0.0992127
						Stratified SE 0.31
						Coefficient of variation 7.79
Number of stations to be sampled in 2002	11	10	10	9	**33**	

Slide 13: Survey results for 2001.

The variance from within each grid from the 2001 trawling data was used to fine-tune the 2002 survey design. Approximately 50% of the total survey variance is coming from grid 18. In the 2002 survey, grid 18 was sampled more intensively to counteract its high variance.
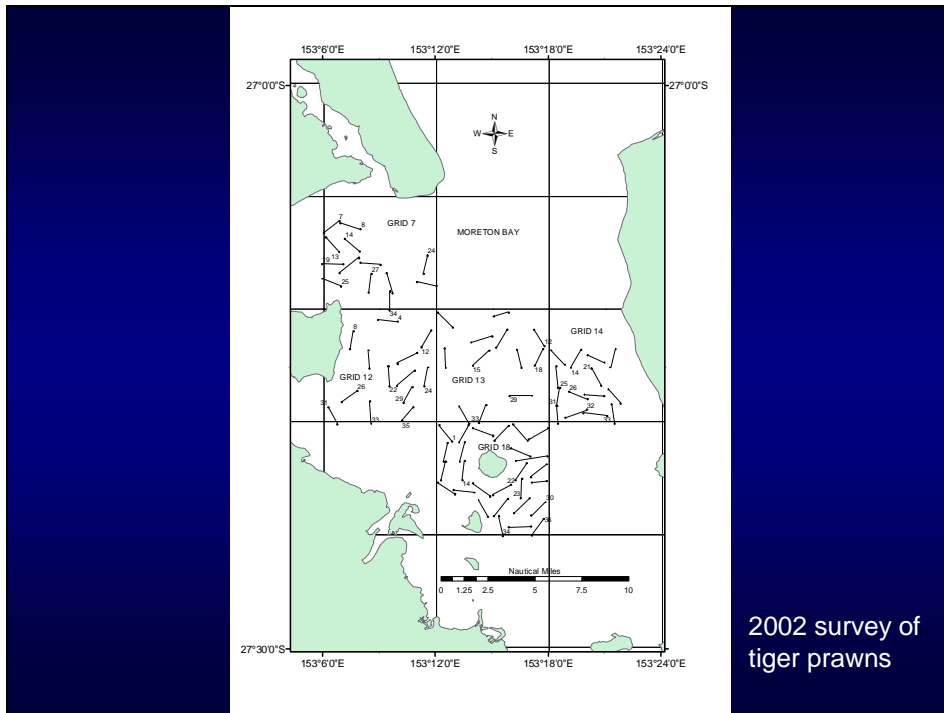
Slide 14: Trawl tracks for the 2002 survey of tiger prawns.

The 2002 survey comprised of 79 trawl samples with a larger proportion of trawls in grid 18 than in the previous year.

Mean catch rates (number 1nm^{-1}) of adult tiger prawns from the 2001 and 2002 surveys

Grid	2001 Mean catch rate (s.e.)	N	2002 Mean catch rate (s.e.)	N
7	2.53(0.40)	15	2.43(0.67)	13
12	2.21(0.47)	14	1.42(0.39)	13
13	7.75(0.66)	16	2.54(0.73)	12
14	3.69(0.74)	13	2.23(0.58)	14
18	3.79(1.06)	14	3.86(0.74)	27
		72		79

Slide 15: Mean catch rates of adult tiger prawns from the 2001 and 2002 surveys.

Presented here are the mean catch rates of the adult spawning stock (prawns with carapace length greater than or equal to 30 mm caught between October and November) from both surveys. Catch rates of the spawning stock are similar between years and generally consistent with observations from other sampling programs that have been conducted previously in the Bay at this time of year. The largest variation between years occurred in grid 13 where catch rates dropped from 8 to 3 prawns per nautical mile.

Population size estimates (Na)

- Overall stratified means:
  - 2001: 4.04 (per 1 nautical mile trawl) (95% CLs 3.43 - 4.66)
  - 2002: 2.89 (per 1 nautical mile trawl) (95% CLs 2.33 - 3.46)

- Adult tiger prawn population size
  - 2001: **649,000** (95% CLs 550,000 - 748,000)
  - 2002: **465,000** (95% CLs 373,000 - 556,000)

Slide 16: Population size estimates.

The stratified means were then used to estimate the spawning stock population size for 2001 and 2002. In 2001 an average of 4.04 prawns per nautical mile were caught which equates to a spawning stock population size of about 649,000 prawns. In 2002 slightly fewer prawns were caught per nautical mile (2.89), which equates to a smaller spawning stock population size of about 465,000 prawns.

EMPIRICAL RESULTS FROM MORETON BAY Cont.

GENETIC DATA TO ESTIMATE N_e

Jenny Ovenden

Southern Fisheries Centre, Queensland Department Primary Industries

OUTLINE

Methodology
– Sampling
– Microsatellites

Results

Gene diversity

Tests for
– Selective neutrality
– Linkage disequilibrium
– Hardy Weinberg

Estimates of N_e

N_e /Na ratio

Slide 17: Genetic data to estimate N_e.

Thanks to Tony Courtney for presenting the estimates of the actual number of spawning tiger prawns in Moreton Bay. Reports from the literature indicate that effective population size is several orders of magnitude below census size. We tested this hypothesis by going out into Moreton Bay and measuring the effective population size of tiger prawns.
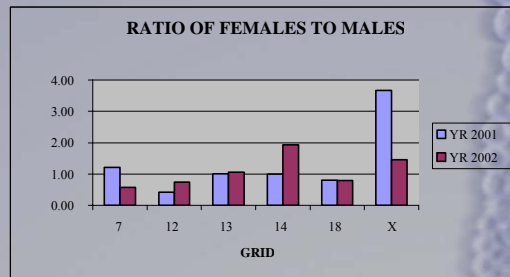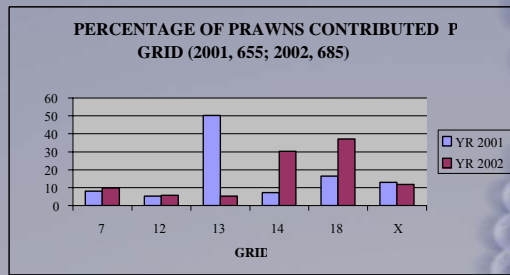
Slide 18: Sampling.

Firstly, the prawns that we genotyped were the same prawns caught in the 2001 and 2002 survey to estimate spawning population size. The survey design changed slightly from year to year to equalize variance. The top graph shows the different number of prawns caught in each grid over the sampling period. Grid X represents the few commercial samples of tiger prawns obtained from the bay to make up sample sizes. We have genotyped 665 prawns from the 2001 survey and 685 from the 2002 survey. In 2001 most of the prawns we genotyped came from grid 13 while in 2002 most of the prawns came from grid 18. A major assumption behind the estimation of effective population size is that the samples are taken randomly from the population and that the ratios of males and females are equal. The commercial sample, grid X, is biased towards females and we think this is due to fishermen selecting out the larger, more valuable prawns from their catch.

Microsatellite Loci

Locus	Repeat motif	Alleles	Het_{obs}
015	$(GAT)_{16}$	24	0.82
047	$(CTAT)_{11}$	17	0.60
120	$(TGA)_7(TGG)_8(TGA)_7$	7	0.44
176	$(GAG)_9(AAG)_{13}(GAG)_4$	15	0.87
189	$(TTC)_{10}$	7	0.73
268	$(ATG)_9(GTG)(ATG)_5$	17	0.84
1.1	$(CA)_{35}$	35	0.72
PMCD	$(TAA)_9 TAG (TAA)_5$	11	0.83

- **Eight loci (six tri-nucleotide repeats, one di-nucleotide and one tetra-nucleotide)**
- **DNA extracted & PCR's performed in Molecular Fisheries Lab**
- **Gel separation and allele scoring at AGRF Melbourne**
- **Total number of alleles = 133**

Slide 19: Microsatellite loci.

Eight microsatellite loci were used to genotype tiger prawns from Moreton Bay. These loci were under development by CSIRO Livestock Industries and they kindly allowed us to continue on with their development. They are mostly tri-nucleotide repeats although there is one tetra and one di-nucleotide repeat. These loci were highly polymorphic with 7-35 alleles each, and observed heterozygosities ranged from 0.44-0.87. We extracted DNA and performed the PCR reactions in the Molecular Fisheries Laboratory, Southern Fisheries Centre. Gel separation and genotyping was outsourced to the Australian Genome Research Facility (AGRF) in Melbourne. We performed all the data analyses after subjecting the data to quality control measures.

Test for linkage disequilibrium

- No significant associations between any pairs of loci.
- 44 haploid chromosomes
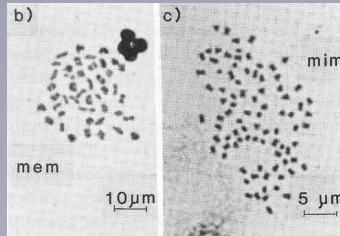
P. esculentus **chromosomes at spermatogenesis** (Xiang, Courtney and Zhou, 1996)

Slide 20: Test for linkage disequilibrium.

We tested the data for linkage disequilibrium, which tests for association between loci. If the loci are not independent then the genotypes at one locus will reflect the genotypes at the other locus. This most commonly occurs when loci are next to each other on the same chromosome and thus have a higher probability of being inherited together from parent to offspring. Linked loci need to be discarded for most kinds of population genetic analysis. We did not detect any significant associations between the 8 loci. Tiger prawns have 44 chromosomes.
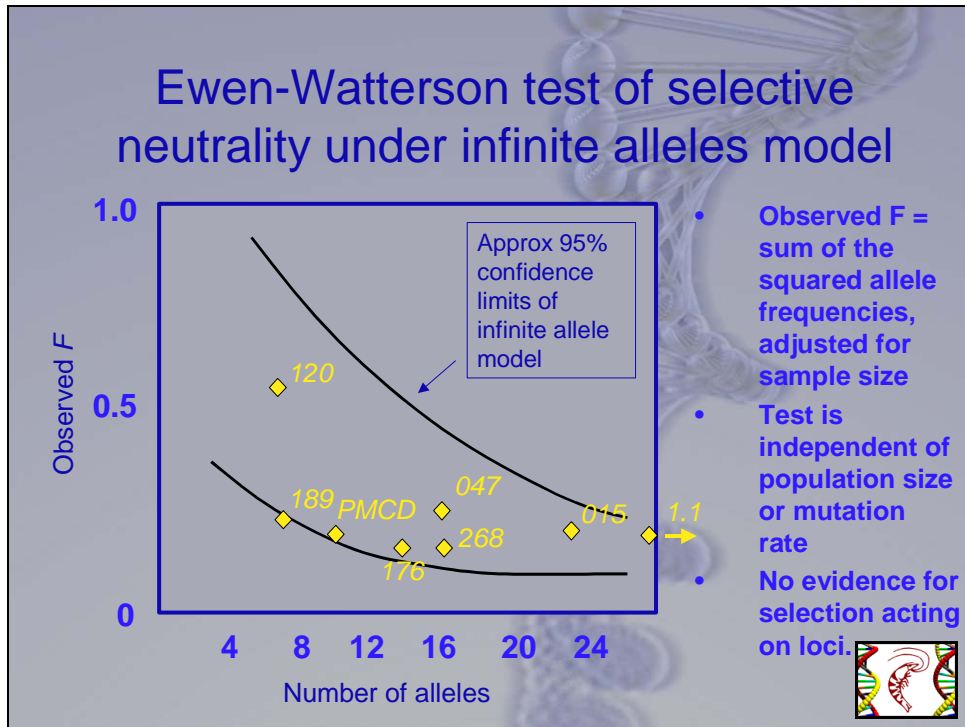
Slide 21: Ewens-Watterson test of selective neutrality under infinite alleles model.

This data set was also tested for selective neutrality; the null hypothesis was that natural selection was not operating on these loci. Most microsatellite loci are neutral but it is necessary to test this assumption. Drift cannot be measured on loci that are under selection because observed changes in allele frequency will reflect the effects of selection and not the random processes of genetic drift. The Ewens-Watterson test plots the amount of homozygosity adjusted for sample size against allele number and evaluates the observed distribution against the theoretical expectation with no selection. Each of the yellow dots relate to one of our loci and most are within the 95% confidence intervals of the predicted values under a model of no selection. These data indicate that there is no evidence for any of these loci being under selection.

Test for Hardy-Weinberg equilibrium

| Locus	Het$_{obs}$	Het$_{exp}$	Het deficit?
015	0.82	0.82	
047	0.60	0.79	yes
120	0.44	0.44	
176	0.87	0.86	
189	0.73	0.78	
268	0.84	0.83	
1.1	0.72	0.76	yes
PMCD	0.83	0.83	

- Two loci (047, 1.1) out of HW
- Selection, cryptic subdivision, null alleles?
- Observed in MB and GOC, but not two WA populations
- Kept under assumption of temporal stability of null frequency

Slide 22: Test for Hardy-Weinberg equilibrium.

Another standard procedure is to test whether the loci are in Hardy-Weinberg equilibrium (HWE). This test compares the number of observed heterozygotes (HET$_{obs}$) to the number of heterozygotes (HET$_{exp}$) expected from the allele frequency distribution. In most populations, values of HET$_{obs}$ and HET$_{exp}$ are very similar. In this case 6 loci are in HWE and 2 loci (047 & 1.1) are out of HWE. Loci out of HWE are typically discarded from data sets. However, for our purposes we have kept these loci because we understand the reason why they are out of HWE. Our previous tests for selection show that the departure from HWE is not due to selection. Cryptic subdivision of the tiger prawn population in Moreton Bay could also cause departures from HWE. We have not explicitly tested for this subdivision yet but we think it highly unlikely. We think these loci are out of equilibrium because of the presence of null alleles.   Null alleles are due to amplification failure of an allele making heterozygotes appear as though they are homozygotes. We have not discarded these loci because we assume that the effect of these null alleles is the same in both years and therefore does not affect our ability to measure genetic drift. We have genotyped other tiger prawn populations using these 8 loci and found that Western Australian populations (Shark Bay and Exmouth) were in HWE while the Gulf of Carpentaria population was out of HWE. If these loci were under selection in tiger prawns then we would expect to see departures from HWE across all populations, as this is not the case our data from other tiger prawn populations adds support to our null allele hypothesis.
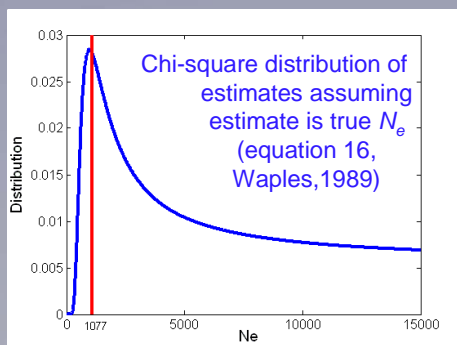
Slide 23: Moments based temporal method.

We used the moments based temporal N_e, which is Waples' (1989) original formulation, to calculate the effective population size. Our estimate of 1077 effective prawns is based on observations of genetic drift over a single generation. While the upper 95% confidence limit is infinite, the 80% upper confidence limit is non-infinite at 5225. Our first reaction to estimates with infinite upper bounds is to throw them out because of a lack of statistical power. However the distribution is informative and allows us to make statements like 90% of the estimates in Moreton Bay are below 5000 and we can make these sorts of statements with a high degree of confidence.

We repeated the analysis discarding loci that were not in HWE and generated a similar distribution of N_e. As yet, we have not been able to implement more powerful maximum likelihood and Bayesian methods because they cannot handle data with missing loci. The linkage disequilibrium method also gave us an infinite result.

N_e / N ratio

- N_e = 1077; N_a = 650,000
- Thus, N_e / N ? 0.002
- Tiger prawn N_e is about three orders of magnitude below N_a in Moreton Bay
- Context for magnitude of N_e / N_a
  - Theoretical minimum (Nunney, 1995), 0.25
  - Average after review of literature (Frankham, 1995), 0.1
  - Oysters (Hedgecock, 1993), 10^{-6}
  - Red drum (fish; Turner et al, 2002), 10^{-3}
  - New Zealand snapper (Hauser, Adcock et al, 2002), 10^{-5}
- **Why?**

Slide 24: N_e/N ratio.

Combining the estimates of spawning stock size (N_a = 650000) that Tony Courtney presented with the genetic results (N_e = 1077) we find that N_e is three orders of magnitude lower than N_a. How does this compare to other species? Nunney (1995) originally thought that the theoretical ratio would never fall below 0.25 but that study was quickly followed by a review by Frankham (1995) across a number of taxa the average ratio is about 0.1. Other estimates for type three species exist and the ratio is between 3 to 6 orders of magnitude below zero.
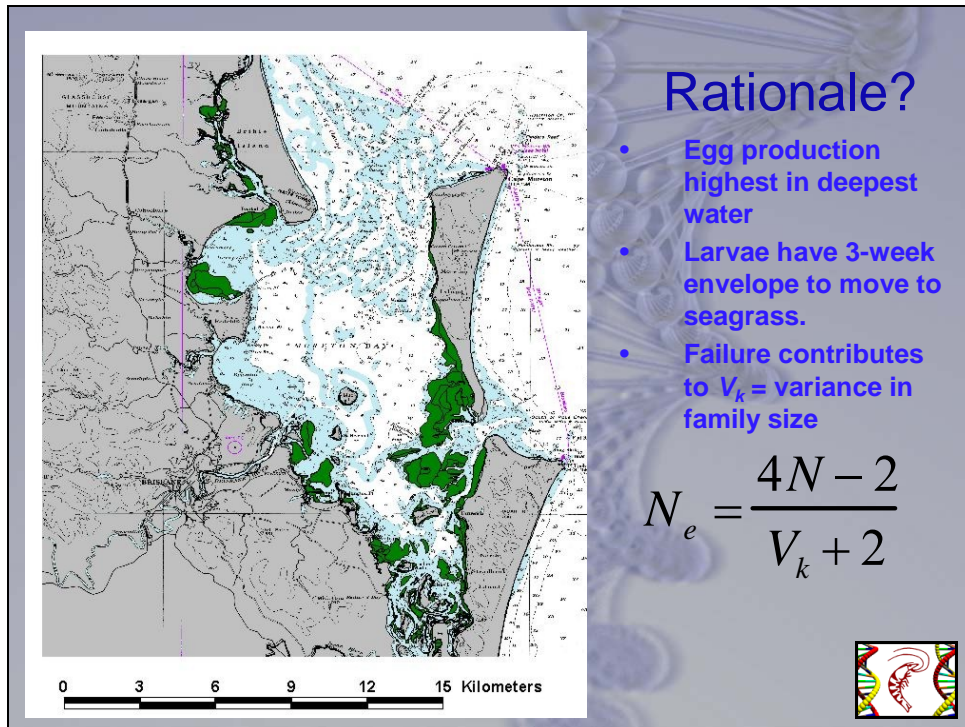
Slide 25: Rationale?

In the context of prawns, seagrass is important to their reproductive success. In Moreton Bay most of the spawning occurs in deep water but the larvae have only a short time (two to three weeks) to disperse to the shallow seagrass beds (green areas on the slide). As they have only a limited capacity to swim it is thought that most fail to find the seagrass beds within the three-week period and subsequently perish.   The formula for N_e presented here relates variance in family size, progeny number or reproductive success (V_k) with census size (N). As the variance V_k increases, the ratio between the numerator and dominator decreases resulting in a smaller estimate of effective population size.

With thanks to …..

Raewyn Street

Sally O'Connor

Sigrid Lehnert

Jennifer Meadows

Kelly Ewen-White

Melinda Ziino

Damien Broderick

Bob Ward

David Peel

Samantha Peel

Heather Podlich

Kaye Basford

Per Erik Jorde

Dave Vance

Funding agencies
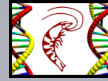
Slide 26: With thanks to.

These people have helped with the collection, development, genotyping, analysis and interpretation of the data presented here today. I would also like to acknowledge the funding bodies without whose support the work would not have been able to go ahead.

Questions

Ye:	Your estimate of N_e of 1077 is for 2001. Did you estimate it for 2002?
Ovenden:	The way the temporal method works and the way we arranged our sampling is that the estimate we made applies to the 2001 generation not to the 2002 generation. The other methods of estimating effective population size, the point methods like the linkage disequilibrium method and heterozygote excess, can be related to both 2001 and 2002. At least for 2001 the linkage disequilibrium method has returned an infinite result. David's [Peel] talk about the simulation experiments will show you contrasts between the way the two methods behave as sample size increases.
Ye:	For stock assessment normally we need to estimate N_e for a certain period of time. I would like to know how consistent your estimate is over time of the relationship between N_e and N?
Ovenden:	Yes. We were given a two year window by the funding agencies to perform this work and because this species had a one year generation time we were able to make just one estimate in that time. To look at the stability of the ratio through time we would need to be able to repeat this experiment over several years.
Adcock:	Have you tried to estimate N_e from demographic data?
Ovenden:	No. What Greg [Adcock] is referring to here is the incorporation of demographic parameters to convert spawning numbers to effective number of spawners. In other words if you know age specific death rates and the age specific birth rates and all kinds of other things like that you can convert one to the other. We just don't have enough information about this species' multiphase life history to be able to do that.
Adcock:	Not even a stab at it?
Ovenden:	I wouldn't like to do it here and now, no.
Moore:	It may have been a typo Jenny [Ovenden] but I thought there might be three loci out of Hardy-Weinberg equilibrium.
Ovenden:	Because differences between observed and expected heterozygosity after Bonferroni correction for probabilities, I regarded that one as actually in equilibrium. That's why it wasn't highlighted.

Effective population size - a simulation experiment

David Peel

Southern Fisheries Centre, Department of Primary Industries, Queensland

Abstract

The use of genetic methods to estimate effective population size (N_e) is well documented in the literature. However in general the populations considered have been from a conservation context and hence mainly have small effective population sizes. In this project we are concerned with the use of these methods in much larger populations.

Estimating large N_e introduces some problems, as the quantities being measured to estimate N_e (i.e. genetic drift) are much less for large effective population sizes and can be indistinguishable from the sampling error. Hence infinite estimates of upper confidence intervals are obtained.

In this presentation we examine the feasibility of the available methods to estimate large effective population size. We approached this question in terms of the current resources we have available in this project as well as the resources that we may have available in the future.

To assess the methods in various scenarios, an individual-based modelling approach was used to generate many artificial data sets. From this work we built up a much stronger understanding of the methods' behaviour for large effective population sizes. The individual-based model is also easily modified to examine further issues, such as migration and non-random breeding.

Presentation
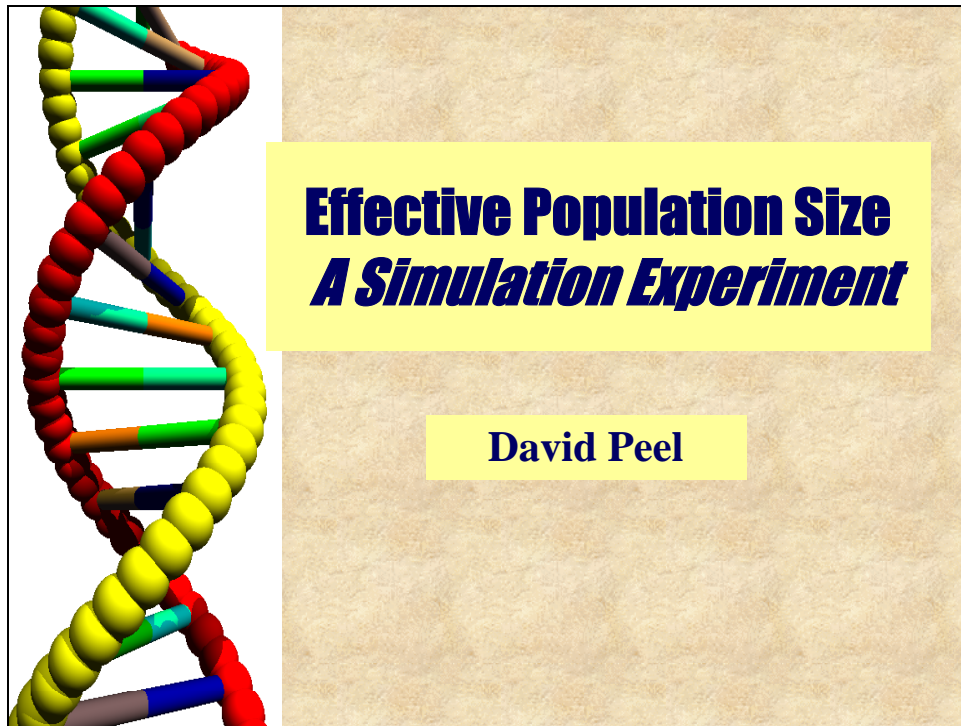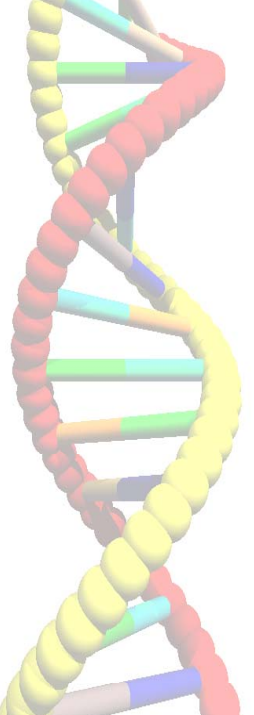
Slide 1: Effective Population Size: A Simulation Experiment.

This presentation will attempt to answer some of the more difficult questions like what power and sample sizes are required to generate sensible estimates of effective population size. A major component of this project was to examine these factors experimentally using simulation.

Slide 2: Affiliates and Sponsors.

I would like to thank all of the people who gave their time and money. Without their support this project would not have been possible.
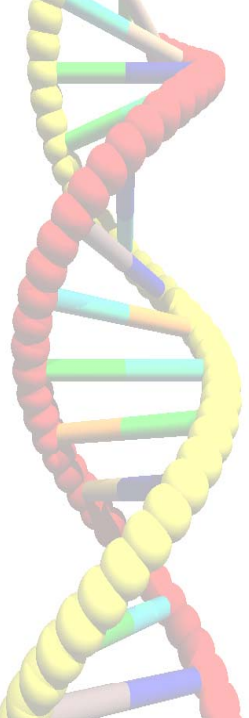
Slide 3: Introduction – 1.

Today I will talk about how we went about answering two key questions. Does the methodology work? And, what would happen if we were not constrained by resources and could increase sample sizes and number of loci genotyped?

Slide 4: Introduction – 2.
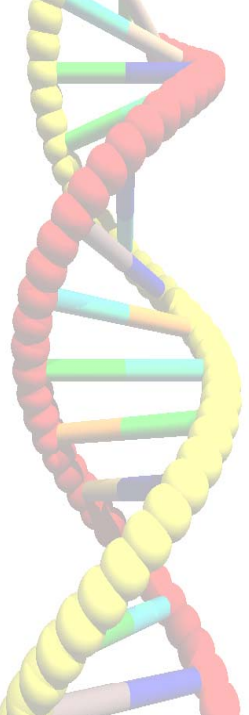
The simulations were designed to i) get a feel for how the methods perform under various conditions, ii) investigate the feasibility of the current study to estimate effective population sizes of prawns in Moreton Bay and iii) examine the resources needed to generate a certain level of power to estimate N_e.
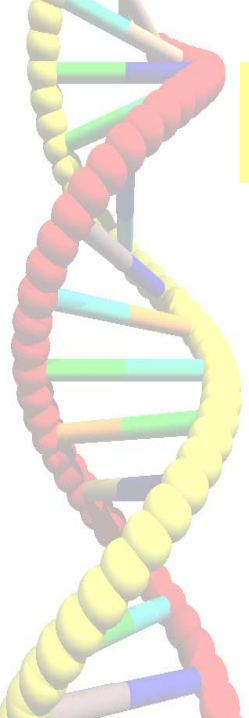
Terminology

- **Power**

  In the context of this work it is the probability of the method returning an non-infinite estimate.

- **Bias**
  Expected difference between our estimate and the true value.

Slide 5: Terminology.

In the context of this study, power refers to the probability that the method returns a finite answer and bias refers to the expected difference between our estimate and the true value.
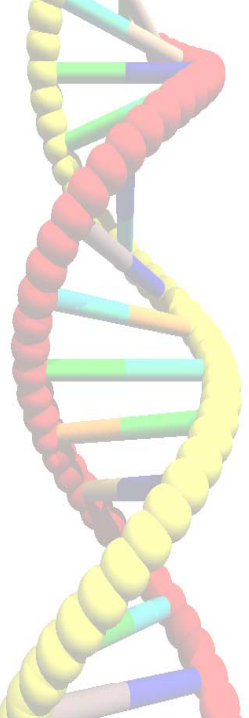
Infinite Results

- Measures variation in allele frequency between generations.
  - Genetic Drift
  - Sampling error (1/S)
- When measured drift < 1/S we get infinity

Slide 6: Infinite Results.

Jenny Ovenden has covered some of this material but one of the reasons why we get infinite answers is that sampling error is too large to measure genetic drift effectively. Measuring genetic drift has a component of sampling error. Sampling error is inversely proportional to the number of samples used to measure genetic drift. That is, our ability to accurately measure genetic drift is hampered by large sampling errors associated with small sample sizes. Infinite estimates of effective population size occur when the amount of drift measured is less than 1/S, where S is the sample size.

Methodolgy

- We simulate many (>2000) genetic data sets with a **KNOWN** true *Ne*.
- Then we apply the available methods to these synthetic data sets
- From this we builds up a profile of the method's behaviour

Slide 7: Methodology.

We simulated several data sets (>2000) across a range of known effective population sizes (2000-10,000). We then tried to estimate the effective population size from these data sets using the various methods available. The profile of results allowed us to directly compare the performance of each of the methods available.
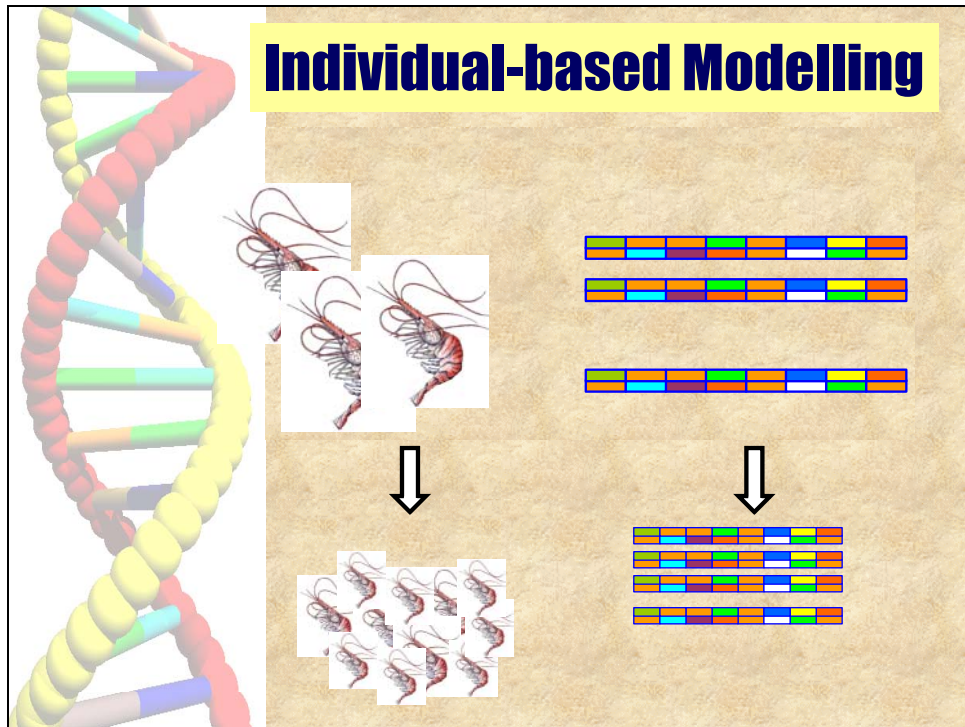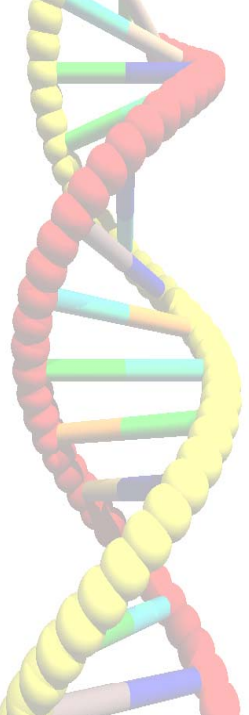
Slide 8: Individual-based Modelling.

We attempted to replicate the natural life history of prawns in Moreton Bay in our models that simulate data. The computer program represents individual prawns by their genotype, i.e. the combination of alleles that they have at each eight loci. Each generation the breeding individuals are selected at random from the entire population and mated together randomly to produce offspring. The offspring are depicted here as genotypes that have been inherited from their parents in a Mendelian fashion.
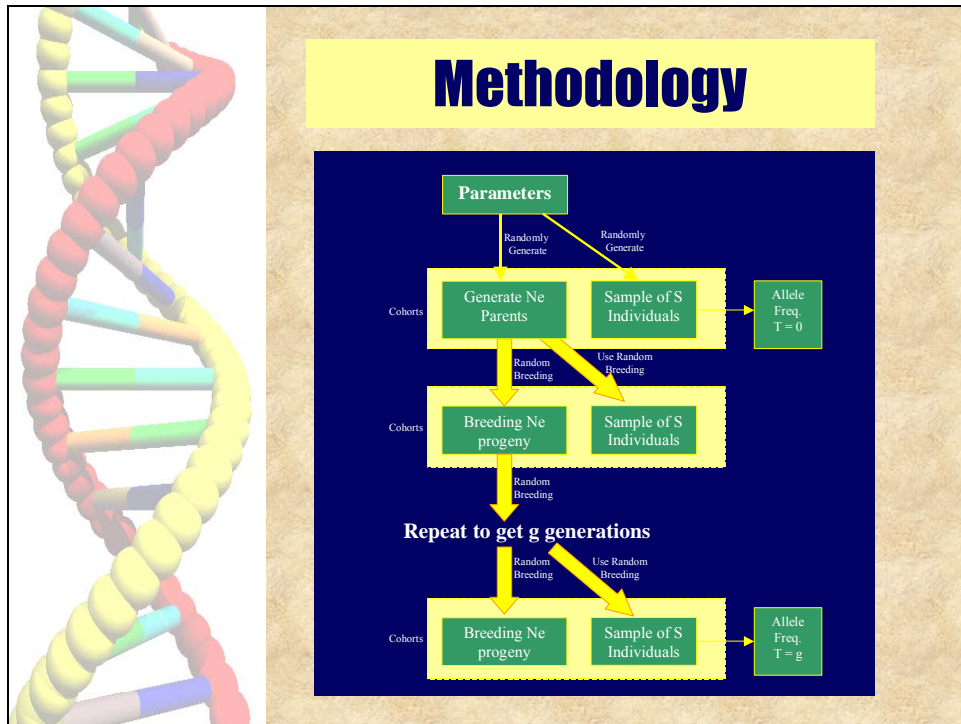
Slide 9: Methodology.

From Tony Courtney's work, estimates of the number of spawning prawns in Moreton Bay are around 650,000. Rather than simulate every single prawn, the model only considered those breeding prawns that were successful each year. Therefore, the model was not particularly computationally demanding because we considering only the successful breeders in each generation.

Slide 10: Design.
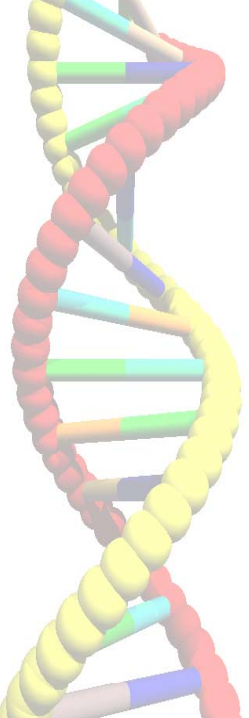
The performance of methods to estimate effective population size has been investigated in the literature, however these studies are typically restricted to populations with small effective population sizes (100-200). We are expanding upon these previous studies by investigating the performance of the methods when N_e is very large.

Several factors affect the power or performance of these methods. From the literature, we know that the larger the true N_e the worse the performance. The number of samples, loci and alleles are all positively correlated with power to estimate effective population size. The greater the number of generations elapsed between measurements of allele frequencies, the greater will be the observed amounts of drift giving rise to more statistically powerful estimates of N_e. However, it is important to remember that estimates of N_e across several generations are the harmonic mean of intervening intergenerational estimates. Other considerations that will not be discussed today include measuring the effect of rare alleles versus common alleles. We also hope to investigate the effects of migration and variation of N_e over time in future models.
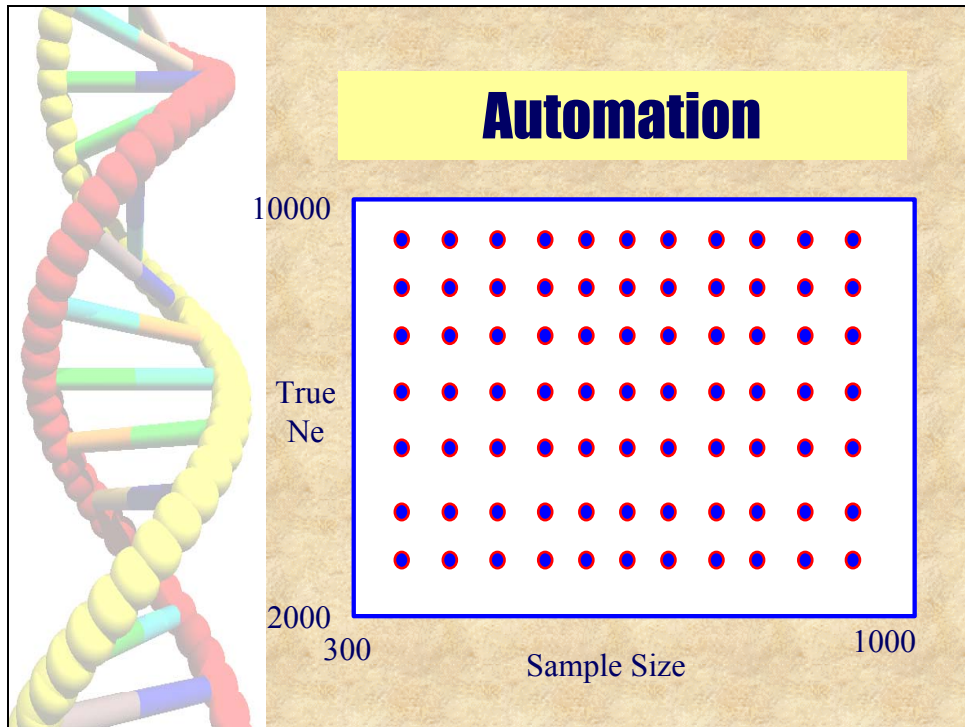
Slide 11: Automation.

To investigate the effect of these parameters on power, each parameter was investigated independently by fixing all other parameters. In this example sample size ranged from 300-1000 and the true N_e ranged from 2000-10000. Effective population size was then estimated for each point in the grid. Estimates from each grid or scenario are then replicated 2000 times to measure variance around the estimate.
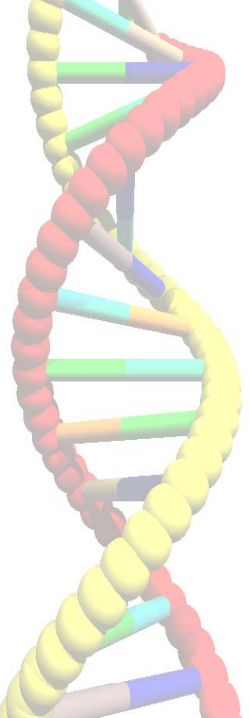
Computational Requirements

- Each set of bootstraps takes between 20min to 2 hours and typically we require 500 bootstrap sets.
- So to produce 1 grid we need 500 hours (or ~10 days (larger sample sizes would take up to 100 days).

Slide 12: Computational Requirements.

The computational requirements to explore these scenarios were daunting. Each of the bootstraps took between 20 min and two hours to run. With 500 replicates at each point on the grid, simulations typically took around ten days to run. Several months of computational time would be required to complete some of the grids. Clearly a solution needed to be found to increase the speed of the simulation process.
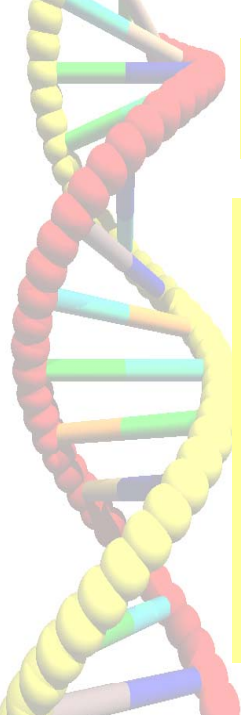
Slide 13: Solution.

A program called Condor allowed us to use a pool of about 20 computers on our local network to complete the simulations in a more feasible time frame.
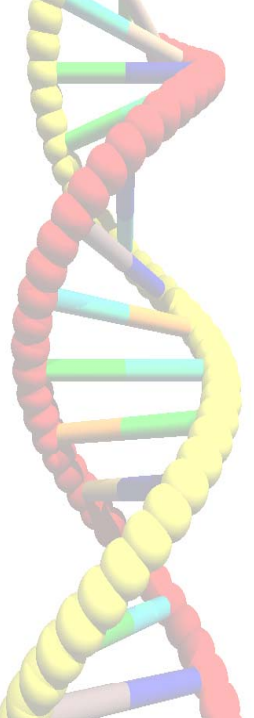
Design - Fixed parameters

- Number of Loci - 8
- No. of alleles - 14 20 7 14 6 13 22 12
- Number of Generations - 1
- Sample Size - 600 each generation
- Allele frequencies set as the same as found in Moreton Bay

Slide 14: Design - Fixed parameters.

This study is primarily concerned with estimating effective population size for tiger prawns in Moreton Bay. To make the simulations representative, the sample number of loci used and the allele frequencies were fixed to those we found empirically in a pilot study of the Moreton Bay population.

Slide 15: Results.

The first thing we varied was the number of samples taken to measure genetic drift. For these simulations N_e ranged from 10-40,000 and sample sizes ranged from 400-2000. This scenario took about 135 days of raw computing time to complete. The heterozygote excess method performed poorly, as expected, because it is more applicable to very small populations and will not be discussed further.
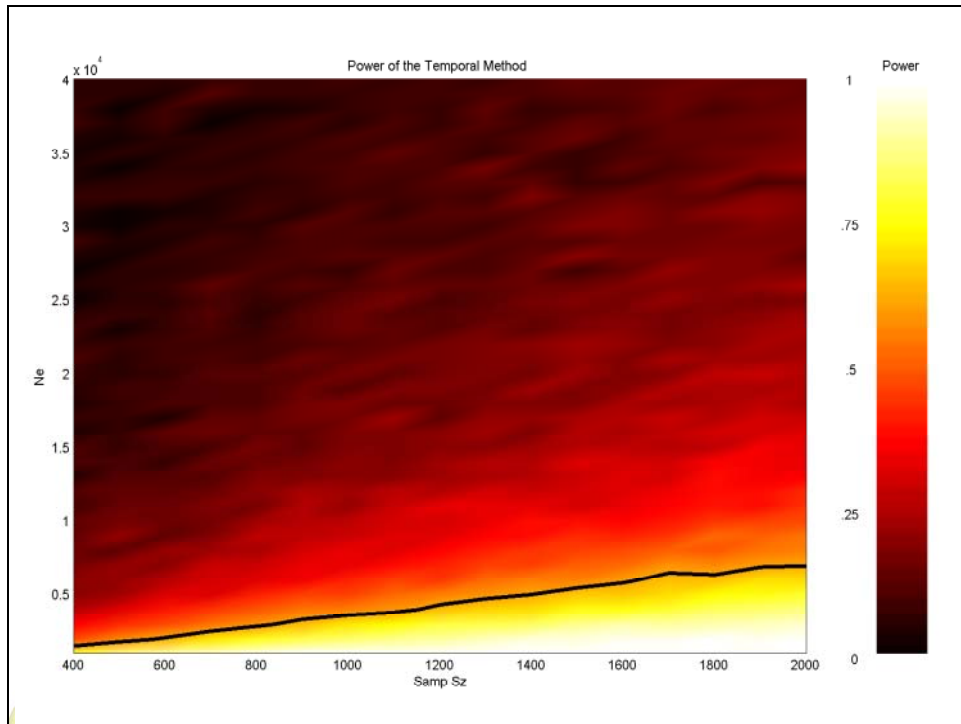
Slide 16: Power of the Temporal Method – Sample Size 1.

The different colours here represent different levels of power. Dark red represents conditions where power was low and as a result, estimates of effective population size are often infinite. The white and yellow colour spectrums represent conditions where power was very high. The black line depicts a rule of thumb where power is equal to 0.8. This graph clearly shows that power increases with sample size for the linkage disequilibrium method. For our sample size of 600 tiger prawns in Moreton Bay the power of this method to generate accurate estimates is poor for anything but very small effective population sizes. Sampling more prawns over more loci and more generations will probably increase power to acceptable levels.
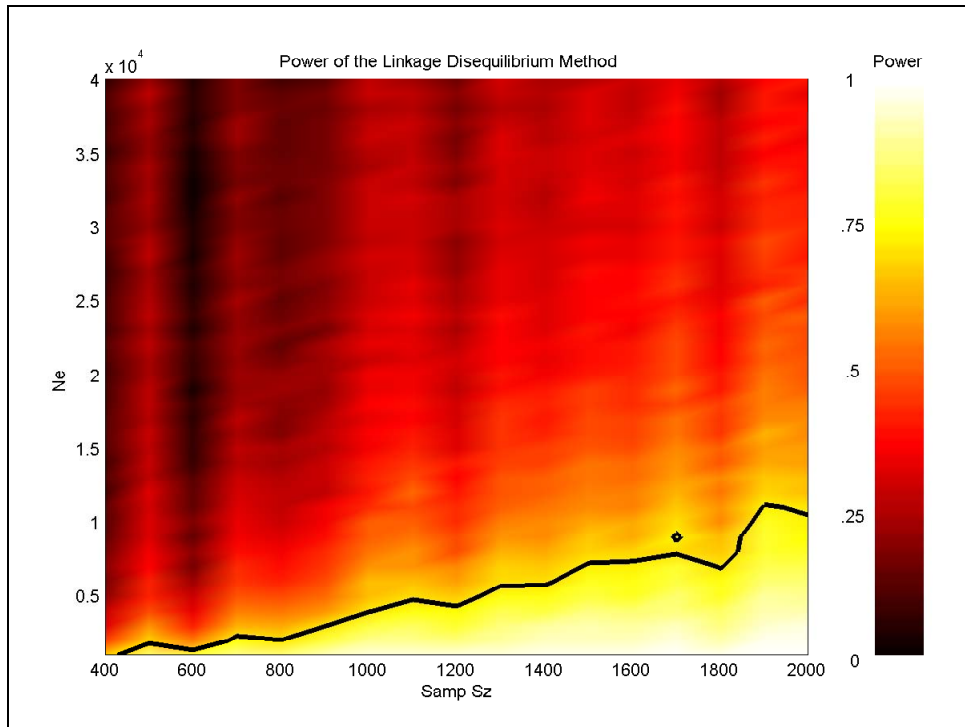
Slide 17: Power of the Linkage Disequilibrium Method - Sample Size 1.

As with the temporal method, power increases with sample size. The linkage disequilibrium method outperformed the temporal method over most values of N_e and sample size. Using this method it was possible to generate acceptable estimates of effective population size up to $N_e = 10000$ when measuring drift from 2000 individuals. However, for our sample size of 600 tiger prawns the power is still very low.
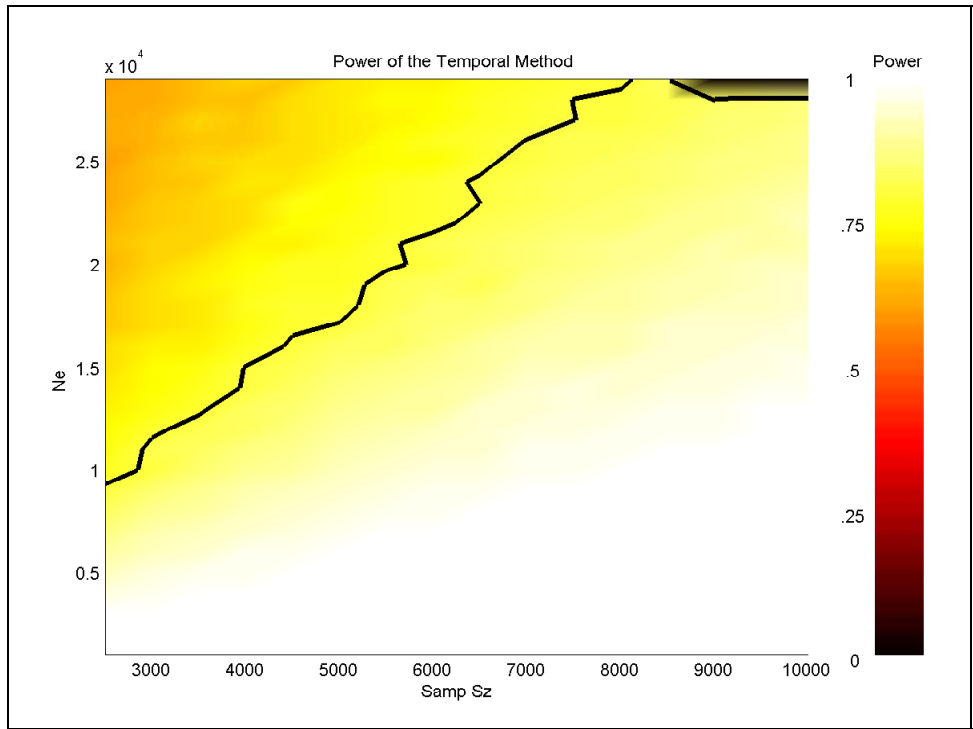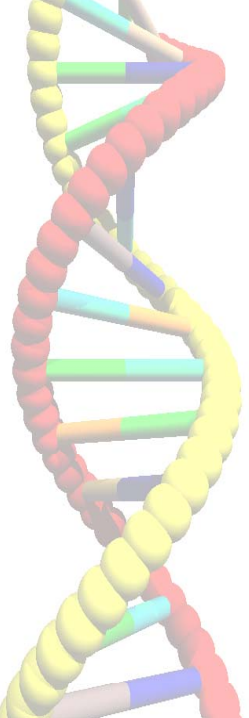
Slide 18: Power of the Temporal Method - Sample Size 2.

This graph assumes that time and money are no object and considers estimates generated from 3000-10000 samples. The power is very good under most of these scenarios. A sample size of 8000 can accurately estimate effective population sizes of 30000.
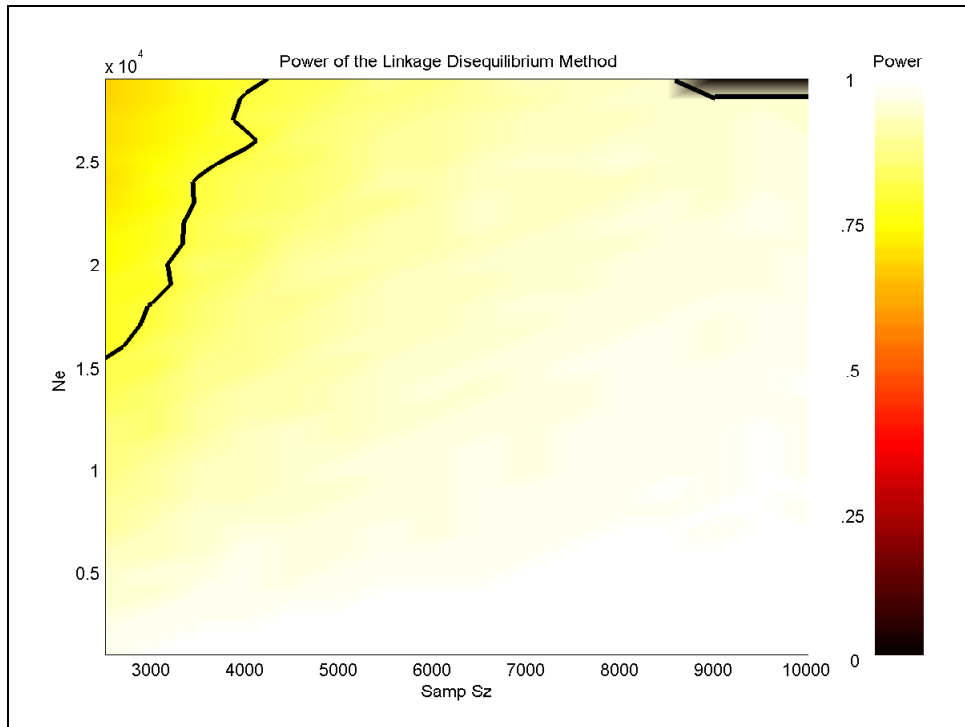
Slide 19: Power of the Linkage Disequilibrium Method - Sample Size 2.

Again the linkage disequilibrium method has more power than the temporal method. Even a sample size of 4000 can accurately estimate effective population sizes of around 30000.

Number of Generation

- Next looked at varying the number of generations between our samples.
- Only meaningful for the temporal method.
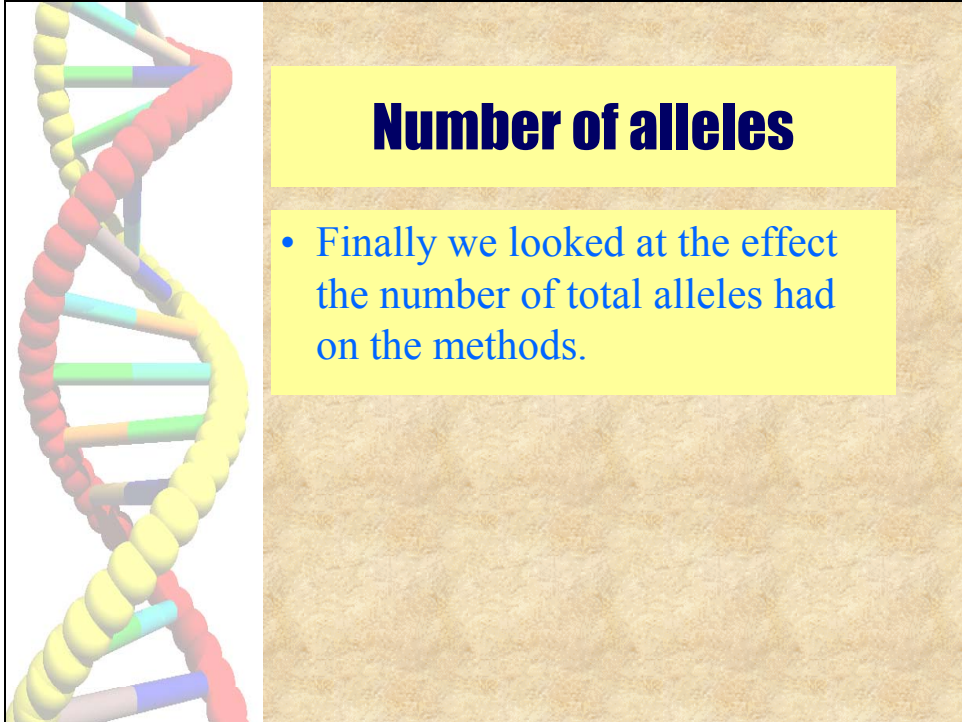
Slide 20: Number of Generations.

Varying the number of generations over which samples are taken is only applicable to the temporal method. The other methods are fixed point estimators of N_e.
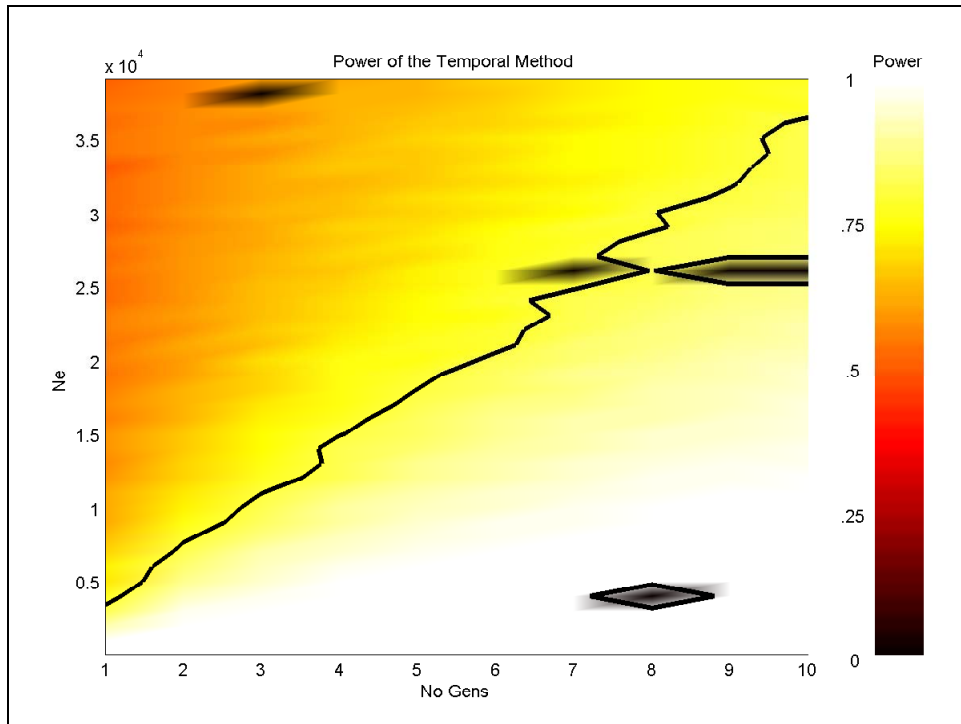
Slide 21: Power of the Temporal Method - Number of Generations.

The number of generations sampled to generate estimates increases power substantially. With a gap of ten years between measurements we have sufficient power to accurately estimate effective population size of 35000 from 600 individuals. It is important to realise that estimates of effective population size over multiple generations is calculated from the harmonic mean of genetic drift over all of the generations. A property of the harmonic mean is that it is biased towards the generation with the smallest effective population size. This is different to the arithmetic mean of several instantaneous (from one generation to the next) measures of drift over a number of generations. This method is most powerful when a sliding window approach is taken based on several sets of temporal estimates.

Number of alleles

- Finally we looked at the effect the number of total alleles had on the methods.

Slide 22: Number of alleles.

The number of alleles detected in a study depends on the levels of variability in the natural population and researchers have limited control other than to choose more variable loci.
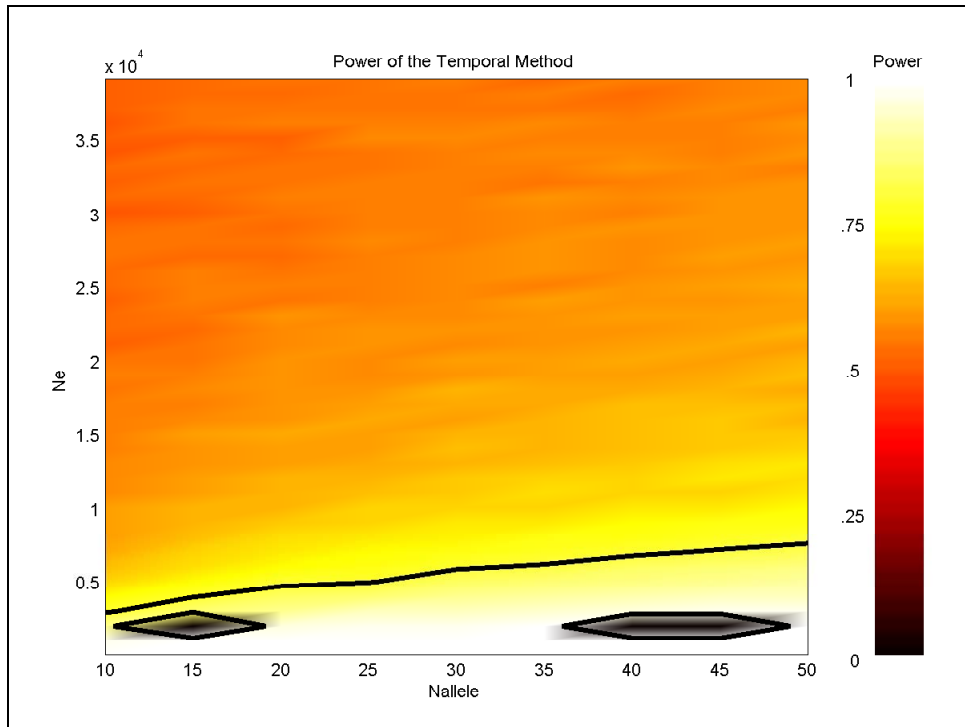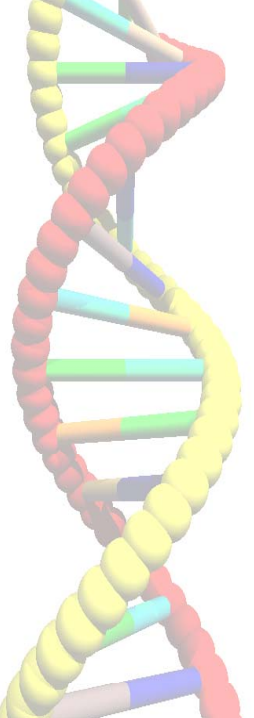
Slide 23: Power of the Temporal Method - Number of Alleles.

The effect of the number of alleles per locus on power was not as marked as with other parameters tested but it was still positively correlated with power.

Power of the Linkage Disequilibrium Method
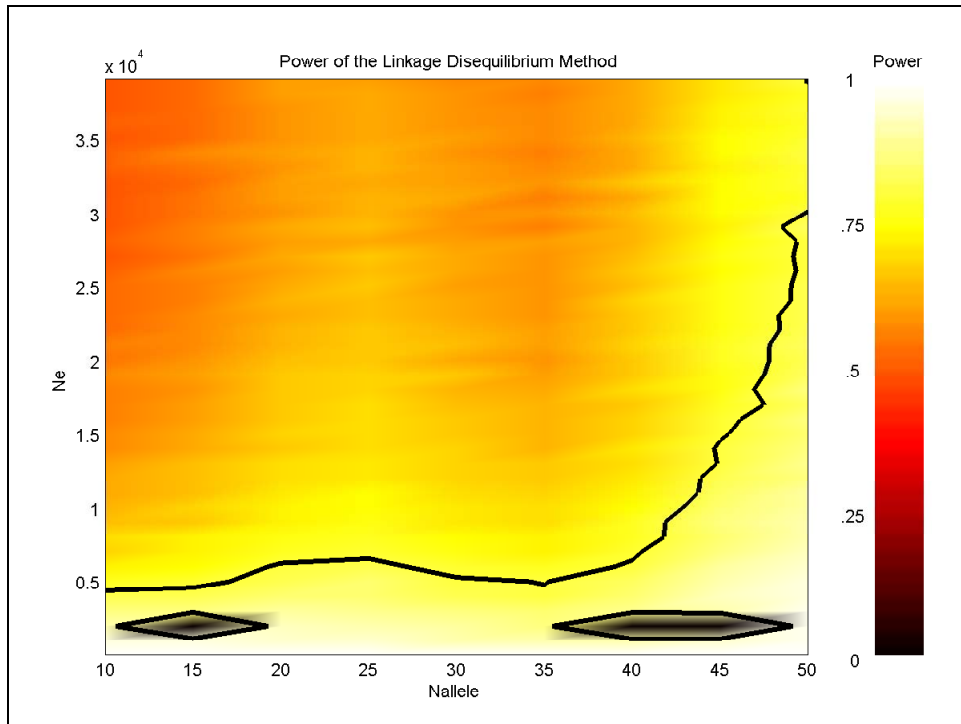
Slide 24: Power of the Linkage Disequilibrium Method - Number of Alleles.

Increasing the number of alleles per locus is more strongly correlated with power using the linkage disequilibrium method compared to the temporal method. We expected to get more power with more alleles using the linkage disequilibrium because it compares all pairwise combinations of alleles, however the sharp increase in power after 40 alleles per loci is somewhat unexpected. Further investigation is required to determine whether a glitch in the modelling procedure is responsible for this sharp increase in power or it is an inherent property of the linkage disequilibrium method.

Bias

- We also examined the actual distribution of the estimates as well as the bias.

Slide 25: Bias.

We have investigated the behaviour of several other parameters but, unfortunately, I only have time to look at the distribution of estimates in detail.
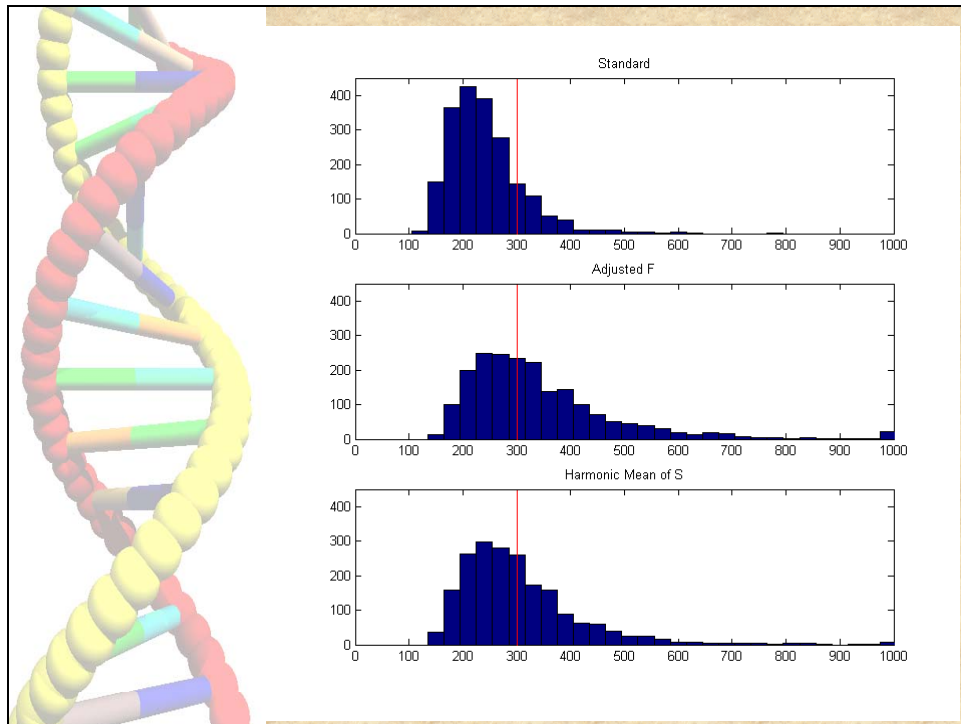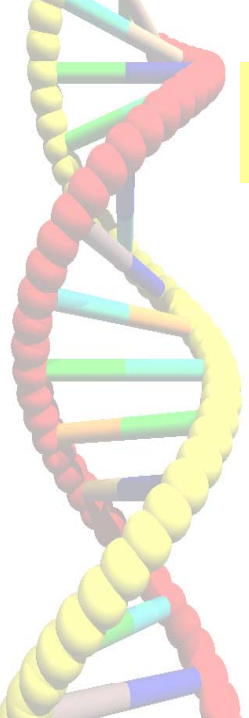
Slide 26: Example of Bias.

The available literature on estimating effective population size considers only complete data sets or at best gives only vague guidelines on how to deal with missing data. However because gaps in data sets are the norm rather than the exception we were interested in how they affected our ability to estimate effective population size. The top graph clearly shows a biased distribution of N_e estimates when uncorrected for missing data. The middle graph is corrected for missing data and is much less biased but as a consequence of the correction has a long upper tail with a few infinite estimates. The harmonic mean of sample sizes also corrects for the bias but to a lesser extent than the adjusted F method. Further simulations are planned to look more closely at appropriate correction factors for bias.

Slide 27: The Future – 1.

I have presented the simulation results to date and would now like to spend some time speculating on future directions of this research.
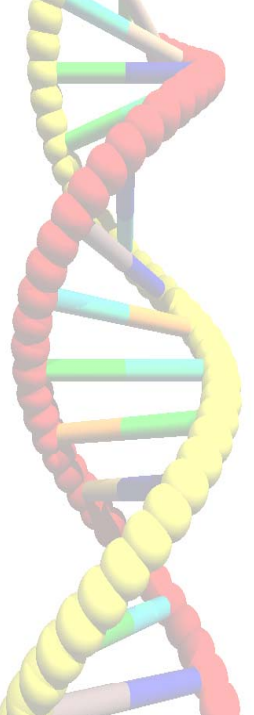
The Future

- Genetics is a fast moving field producing new innovative technologies all the time.
- This may provide much easier, faster and cheaper methods to collect genotype data.

Slide 28: The Future – 2.

The results from the pilot study showed that we do not have enough power to generate finite estimates of effective population size for tiger prawns in Moreton Bay; at least across a temporal interval of one generation with current sample sizes and loci numbers. The pilot study considers only eight microsatellite loci but if it considered 30 loci then it would have enough power. Genetics is a fast moving field with people constantly finding more innovative methods to genotype specimens more easily, quickly and cheaply. I didn't appreciate quite how quickly the field was moving and it appears as though emergent technologies may also provide suitable power to accurately estimate effective population sizes in large populations like prawns.

Single Nucleotide Polymorphisms
Snips

- Newer technology
- Provides a much larger number of loci (50-100)
- Each locus has 2 alleles
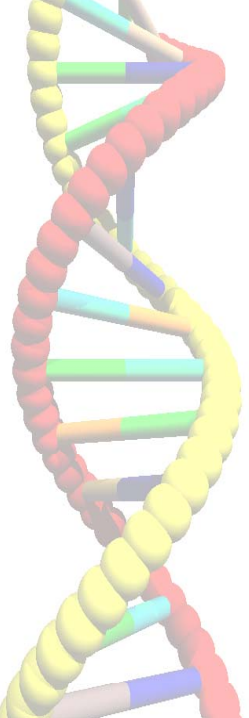
Slide 29: Single Nucleotide Polymorphisms (SNPs).

Single nucleotide polymorphisms (SNPs) is an emergent technology suited to automated, high throughput accurate genotyping. SNP loci can be genotyped at a fraction of the cost of microsatellite loci and this opens up the possibility of analysing more samples, providing even more power for the same cost. Each SNP loci has only two alleles but data sets typically consider 50-100 loci giving it more power than typical microsatellite data sets. Jenny Ovenden and Damien Broderick are developing this technology for various applications in Spanish mackerel and prawns.
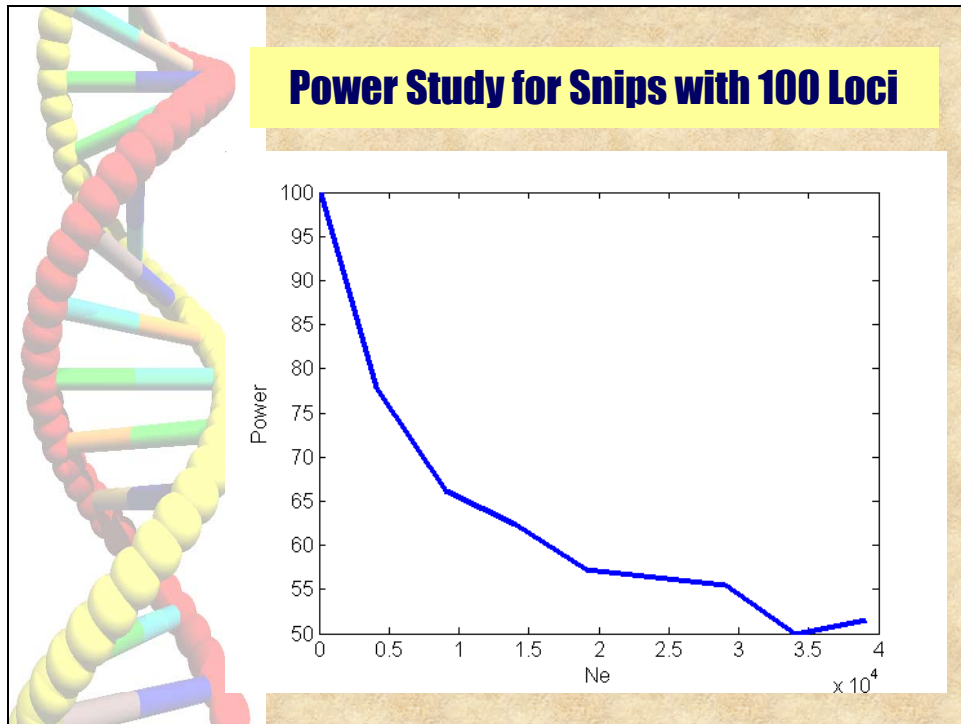
Slide 30: Power Study for SNPs with 100 Loci.

A quick power simulation shows that a typical SNP data set provides more power than the eight microsatellite loci used in the current study. SNPs can estimate effective population sizes of 5000 with 80% power compared to 2-3000 for microsatellite loci. The real advantage of SNP data sets is that you can increase power by genotyping more specimens for the same cost as a microsatellite study. The modelling that we have done establishes the framework to test the performance of existing, new and emerging methods as they become available.
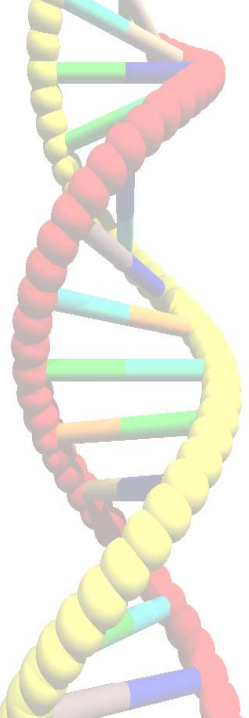
Conclusion

- How well the method works and how many resources are required depends on how large the Ne we are estimating is.
- We can assume from the census size estimates we have for our pilot study that Ne in this case will be reasonably large.

Slide 31: Conclusion - 1.

The performance of the methods discussed today depends to a large extent on how large the true effective population size is. Adjusting parameters like number of loci used has some effect, but it is inconsequential when N_e is very large.
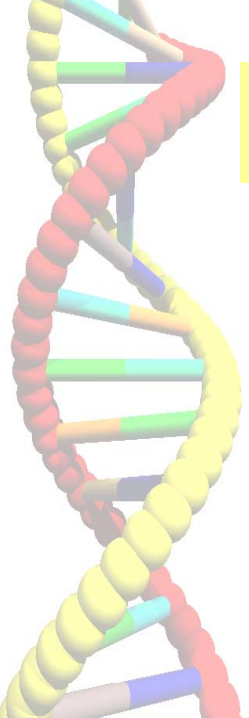
Cont...

- So the current pilot study is stretching the envelope in terms
- However we have shown that with enough resources the power becomes quite reasonable even for very large Ne.

Slide 32: Conclusion - 2.

The size of the spawning population would suggest that N_e for tiger prawns in Moreton Bay is very large, perhaps around 40,000. From the simulation data it is clear that the pilot study is pushing the envelope for effective population sizes that are currently feasible to estimate with the current resources. The simulation also shows that if more resources were available (more loci, samples etc) then it is possible to accurately estimate effective population sizes of around 40,000.

Final Word

- This work establishes a framework to examine for a given scenario if the method will be feasible.
- Can easily be modified to test other issues such as migration, non-random breeding etc.

Slide 33: Final word.

The simulation approach has enabled us to demonstrate under what circumstances it is feasible to estimate effective population size. This platform has enabled us to look at the effects of missing data and other parameters that are not easily accessible from theoretical distributions. The individual based simulation model presented here is not only a very appropriate and intuitive way to test the performance of various models, it is also a platform that is easily adapted to study new populations under different conditions. Future studies, for example, can use this individual based model to simulate the effects of migration and overlapping generations on effective population size estimates.

Questions

Donnellan:	Not so much a question but a strategy really. At the moment this is a relatively expensive tool to be using for monitoring, taking into account the cost of collecting especially. It would seem to me that time in the life of the industry on a particular species when the proportion of money spent on this analysis is lowest compared to the total value of the fishery is when you first start fishing it. A lot of our fisheries have been fished for a hell of a long time. The reason why we are monitoring them is because they are in some state of lower abundance or lower productivity so the amount of catch to do this is going to be a higher proportion of the total value of the industry. What sort of strategy were you guys thinking of in terms of trying to get this implemented on a broader scale, given the non-parlous state of some of our wild fisheries?
D. Peel:	Probably Jenny [Ovenden] is the person to talk to about that. I have actually moved off this project onto another project.
Ovenden:	You are quite right in that the funding available for developing methods like this is proportional to the value of the fishery. It was fortuitous that we could target this methodology to prawns because the northern prawn fishery (NPF) is worth a great deal of money, millions and millions of dollars annually. Implementing this methodology on less valuable fisheries would be challenging because of the limited funding available to do that.
Donnellan:	Sort of an unfairness in it too, in that presumably it would have the greatest impact on the sustainability of fisheries by implementing it earliest in the life of a fishery.
Ovenden:	Yes, although the opportunity to work on a fishery that is just under development is rare. Generally the fishermen go out there and they start making lots of money and people don't realise until several years later that research needs to be done to collect data to manage this fishery. Crispian [Ashby] do you want to comment on that?
Ashby:	No. That sounds fine to me.
Moore:	What about overlapping generations?
D. Peel:	Jenny [Ovenden] has actually done some extra work on simulation collaborating with someone from Norway, [Per Erik] Jorde.
Ovenden:	Per Erik Jorde is a population geneticist. We independently collaborated to look at the effect of overlapping generations because we know that for the NPF there is a degree of overlap. The simulation that David [Peel] has done assumes non-overlapping generations. For species that have generations for maybe over five years, if you took temporal samples over ten years, then you can apply this methodology to that.
Hoyle:	I just wondered if you wanted to talk about combining the linkage disequilibrium method with the temporal method to get more power. As those two are independent you may be able to get more precise estimates from the data.
D. Peel:	Yes exactly. One of the papers in the literature commented that a suitable strategy would be to combine these estimates. Individually they don't have the power but together they could provide you with enough power. It is something we could look at. I actually looked at adding them but have had a few problems with calculating the theoretical confidence intervals and things like that. But it would be quite trivial to stick it onto the end of the code. The other thing that I have just remembered is that all of those simulations have been done with just the moments based temporal method and there are three other more complex temporal methods. The maximum likelihood and Bayesian methods should give a better estimate. I didn't actually look at any of those because whereas the temporal method gets done in about a second, these methods take about eight hours. All of these simulations took me 130 days. To repeat them all with the other methods would probably take me a thousand years so I didn't look at them. My strategy was to look at the temporal method as being the worst case as the other methods are going to be slight improvements on it. What I have shown here is the situation for this worst case and the other methods will be the icing on the cake, as they will give slightly better power.

Ovenden:	Did you use simulation to evaluate the single nucleotide polymorphisms (SNPs) with the linkage disequilibrium method?
D. Peel:	The linkage disequilibrium method is slightly worse so I only showed that one [the temporal method]. I'm not quite sure why. The linkage diseqilibrium is concerned with the number of alleles and if we had 100 loci [SNPs] you still only have 200 alleles, whereas, with the temporal method each locus is an independent measure.
Basford:	If I understand the simulation results, which direction do you see could be the most advantageous way to go?
D. Peel:	In terms of which thing to change to get you more power? Probably a combination of all of them would give the most improvement and be the easiest to do as well. However, I would say it's three things. Obviously sample size is very easy to increase, you just go and collect more samples; it's only a monetary thing. The number of loci is also very good but from what I have seen from the people in the genetics lab trying to find those loci, it's not like I can say go out and find me twenty please. It takes a little bit of effort. That would also give a lot of improvement but it would take a lot more effort. The other one, which is easy, is the number of generations. Collect samples [over multiple generations]. So to summarise, sample size and number of generations are probably the easiest.
Basford:	Jenny [Ovenden] do you want to comment on the ability to use wax samples.
Ovenden:	Oh yes. The original project was designed to estimate drift over a period of 10-15 years for the Moreton Bay tiger prawns. We thought we were able to do this because we had tissue that had been preserved in wax as a part of Tony's [Courtney] work. We were going to extract DNA from that and compare the alleles that we are currently measuring. Preliminary experiments showed us that the DNA was good quality. Certainly people like Greg [Adcock] and others have extracted good quality DNA from scales and many other sources. However, of the eight loci we were working with we were only able to reliably amplify one of the loci for the wax specimens. I think it has something to do with the way the tissue was preserved before it was embedded in the wax that had affected the quality of the DNA. What David [Peel] says is right. If we can implement this methodology over more than one generation, (perhaps for tiger prawns that is fairly straight forward as they have a short generation time) the power of the method is significantly improved.
Donnellan:	Were they soft tissues put in wax?
Ovenden:	Yes, ovary tissues.
Donnellan:	I was just wondering if particular material, scales and things like that, would give a keratinised protection of DNA. It's really hydrophobic and acts as a better preservative under those circumstances than tissue.
Ovenden:	Accept that we thought that wax would be very good at excluding moisture and the atmosphere during the storage time.
Donnellan:	Presumably they were through formalin first.
Ovenden:	For a brief time.
Arthur:	In some of those simulations you showed, there seemed to be a non random pattern in the change of power. Some had a wave pattern. Does this suggest a worry for the simulation?
D. Peel:	Probably something I should have looked at.
Moore:	Do you think it's just an underlying numeric artefact?
D. Peel:	Probably something to do with the different machines things were run on, there might be some variation there. That's all I can think of off the top of my head. It could be some bias in the random number generator or something like that but I'm not too concerned in terms of the amount of bias, if there is some sort of correlation in there, it doesn't really effect the end result in terms of what I was looking to show. I have done it independently on different machines I have basically got these answers.
Moore:	The only one that might be a worry is the one that rocketed up.
D. Peel:	That one that rockets up [Linkage Disequilibrium] I am definitely going to investigate further. Some of these [simulations] are quite recent, some of them I have been working on this week. I'm note quite sure why it rockets up, I would have thought there would be a more linear increase. There are a

	couple of possible reasons for that and I will definitely look at those again. The way it is set up at the moment, if someone says well I've got this many samples and this situation, show me what my power would be.
Hoyle:	About increasing the number of generations and getting more power. We have had some discussions that migration tends to stabilise drift. I was wondering if anybody was able to comment on that. If that happens then you may be able to estimate effective population size well on a short term but having more generations might not help a lot because it is stabilised by immigration from neighbouring populations.
Gopurenko:	I imagine any kind of migration into your so-called isolated populations would blow out your N_e estimates over several generations after introduction.
Hoyle:	Except that on a short time scale you are still going to get those sampling effects from the variation of drift itself.
Gopurenko:	Out of Hardy-Weinberg equilibrium?
Hoyle:	That would be the case if immigration were from a population that had a different allele frequency distribution. If you have two populations that have migration between each other and as a result they have very similar allele frequencies. Within an individual population you'll get a little bit of drift and sampling effect in each small breeding population but over a long period they'll tend to drift towards each other and would tend to maintain similar allele frequencies. So you wouldn't get that long term drift but you would get short term variability.
Ovenden:	Can I add a few things? When you expect migration to be occurring that the problem has to be addressed up front and simultaneously with estimation of effective population size. The two things need to go hand in hand.
D. Peel:	There are some methods out there to estimate N_e in the presence of migration. From talking to Simon [Hoyle] we can think of cases where it would really stuff it up and cases where it wouldn't. It really depends on how the migration is occurring, whether it is a constant flow, or a spurt, a migration, a population migration.
Dichmont:	Just a comment on an observation. If you look at where stock assessment started 30 years ago, time does break down those barriers. Some stock assessment models are pretty bad. The way out of it I have always thought is to combine the resources in the two techniques.
D. Peel:	Some of the Bayesian programs for which I haven't run the simulations yet, you can actually specify the upper limit. It basically uses a uniform prior up to whatever you like and you can improve some things by incorporating that information.
Dichmont:	Even then you need independent data to set the prior.
Salini:	Is there any possibility to do a further experiment of this where you actually have a captive population.
D. Peel:	There was one paper where they had a hatchery population and varied other sorts of stock. It worked reasonably well because you had small numbers.
Ovenden:	It depends very much what your aim would be for setting up the experimental captive population. I think David's [Peel] simulations have shown very well that the methodology *actually* works. You can set up prawns that have a small effective population size and you can estimate that quite well from looking at changes in allele frequency. The method works. There is no doubt about it. Beyond that I don't know what you want to do [in setting up a captive population].
Salini:	It's like otoliths. If you don't ground truth the otoliths and show that the rings really are accurate growth rings. If you can actually do a controlled experiment, you could really nail some of those uncertainties.
Ovenden:	Moreton Bay is just about as close as we can get to an experimental population of tiger prawns.
Salini:	I was thinking about the captive breeding populations. There are some controlled lines of prawns at Cleveland, they have got about 4 generations. Really tightly controlled populations, but I don't think they would like to sacrifice any of them.
Arthur:	One of the problems with looking at this is making sure that everything is in mutation-drift-migration equilibrium. A lot of these things are being

	disturbed all the time so how sure are you that the underlying assumption of equilibrium is valid.
Broderick:	I think in the context of this you can rule out mutation and really the only big worry is migration. Those things affect long term estimates of N_e. From generation to generation you can pretty much negate that concern.
Gopurenko:	I'm not certain that you can negate mutation. I've seen an increasing number of papers on microsatellites where they show rapid rates of mutation on microsatellites. Homoplasy is a big problem with microsatellites. Mutation is probably not low as we once thought it was.
Broderick:	But mutation rates are still down to 10^{-4} to 10^{-5}, something like that.
Gopurenko:	That might not necessarily be the case. I've actually seen an alternative analysis as well showing a rise of new alleles with tight variance, so it does happen. So that is why I think it would be better to move onto SNPs.
Hale:	Is there a relationship between the number of alleles and mutation rate in a locus? The more alleles you have in a locus the higher the likelihood that you're going to get mutation. If you are dealing with loci with relatively few, say up to a repeat length of 30, then the probability of the mutation rate is around about 10^{-5}. But if your length is 50 or 60 then your mutation rate is increased dramatically.
Gopurenko:	It's also a stepwise mutation model for microsatellites too, so you do get this greater amount for a certain allele size. Homoplasy is a big problem with microsatellites.
Broderick:	But again you are measuring the variance around the frequency of those alleles in your population. So mutation, I would think, would actually have to be quite rife before it's going to affect the frequency of those alleles in the population. Sure it's there, sure it accounts for some of the changes, but as a component of the variation from one generation to the next I think it's going to be reasonably trivial.
Gopurenko:	Maybe in combination with quite large N_e's, you might be able to maintain allele frequencies over time. I wonder if that might be a potential systematic error?
Ovenden:	Yes, I think we can address that through modelling simulations.

Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*)

Greg J. Adcock[1], Lorenz Hauser[2], Peter J. Smith[3] and Gary R. Carvalho[4]

1 Department of Zoology, University of Melbourne, Melbourne 3010, Australia
2 University of Washington, Box 355020, Seattle, WA 98195-5020, USA,
3 National Institute of Water and Atmospheric Research (NIWA), PO Box 14901, Wellington, New Zealand
4 Department of Biological Sciences, University of Hull, Hull HU6 7RX, UK
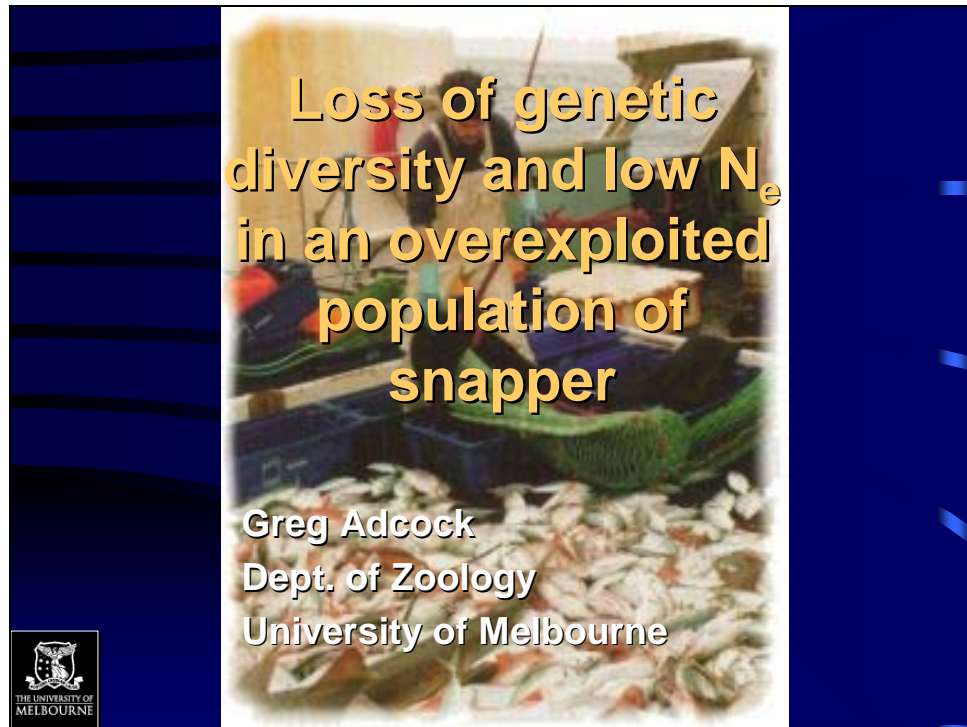
Abstract

Threats to the genetic diversity of exploited marine fish populations have been rarely considered probably because "collapsed" stocks still consist of several million individuals, and population genetics theory suggests that these should not suffer significant loss of genetic diversity. It is becoming increasingly apparent however that in marine species factors contributing to the genetically effective population size (N_e) are poorly understood and that N_e may be orders of magnitude smaller than the census population size (N). Microsatellite and mtDNA analyses of a time series of archived scales demonstrated a significant decline in genetic diversity in a New Zealand snapper population during its exploitation history. The interpretation of the results of this study was facilitated by parallel studies of population history and structure. Effective population size estimates were five orders of magnitude smaller than census population sizes from fishery data. While these results have implications for short and long-term considerations for conservation and management of fisheries resources, they also raise questions regarding the estimation of N_e and the biological meaning of those estimates.

Presentation

Slide 1: Loss of genetic diversity and low N_e in an overexploited population of snapper.

I would like to thank Jenny Ovenden and the sponsors for inviting me here today. It is a pleasure to be here and the discussion has been very stimulating so far. This presentation will depart a little from the previous presentations, as N_e will only form a small part of my presentation. Looking at N_e in New Zealand has been very instructive and the results have raised a number of interesting questions.

Slide 2: Co-investigators and affiliates.

Much of the work that I will present today was done at the University of Hull and was supported through a Leverhulme Trust grant. Peter Smith is a co-investigator and brings expertise in snapper biology to the project.

Temporal genetic studies

- **Real time test of assumptions about genetic changes and correlations with other temporal variables.**

- **Allow direct estimates of particular population parameters**

- **Problem of finding suitable comparisons**

Slide 3: Temporal genetic studies.

My work has focused on temporal genetic studies exploring population dynamics, history and evolution through time. Traditional approaches typically address these issues by taking samples at a fixed point in time. Comparing populations temporally is a relatively new area and allows us to estimate and compare particular population parameters giving fresh insights into population genetic dynamics. One of the major difficulties with temporal genetic methods is finding appropriate populations for comparisons. The extent of fishing pressure for example, may not be known or adequately documented making it challenging to identify fished versus unfished populations for comparison.

Slide 4: Snapper (Sparidae: *Pagrus auratus***).**

We were very interested to take up Peter Smith's offer of studying New Zealand snapper because this fishery is very well documented and has been the subject of a substantial amount of research effort.

New Zealand snapper is a long lived coastal demersal species, highly fecund with a relatively short pelagic phase. Fecundity is markedly higher in old fish compared to young fish making them a type three survival species. Snapper are extensively fished supporting an important export market, which also facilitated the procurement of samples.

Snapper genetics

- **Large scale collection**

- **Microsatellites:**
  - **7 loci**
  - **15 alleles per locus**
  - **mean heterozygosity = 0.75**

- **MtDNA**
  - **16S from many species**
  - **Sensitive screen for d-loop haplotypes**

Slide 5: Snapper genetics.

Historical scale samples provide an important source of DNA for a temporal comparison with contemporary samples. We developed some highly polymorphic microsatellite loci and used mtDNA 16S and D-loop sequence to make temporal and spatial genetic comparisons among populations.

Slide 6: New Zealand snapper genetic study – 1.

We took a top-down approach beginning with taxonomy and phylogeography and then onto stock structure and temporal changes of genetic composition in individual populations to study New Zealand snapper. There are levels below population structure, which we did not look at, that are particularly important. Microgeographic metapopulation factors, like where fish spawn, how fast they grow, where they feed and where they grow up, are critically important when interpreting the observed temporal genetic changes.

Slide 7: Global distribution of snapper.

The two main populations of snapper are in Australia and New Zealand. Snapper in China and Japan are considered a separate subspecies.
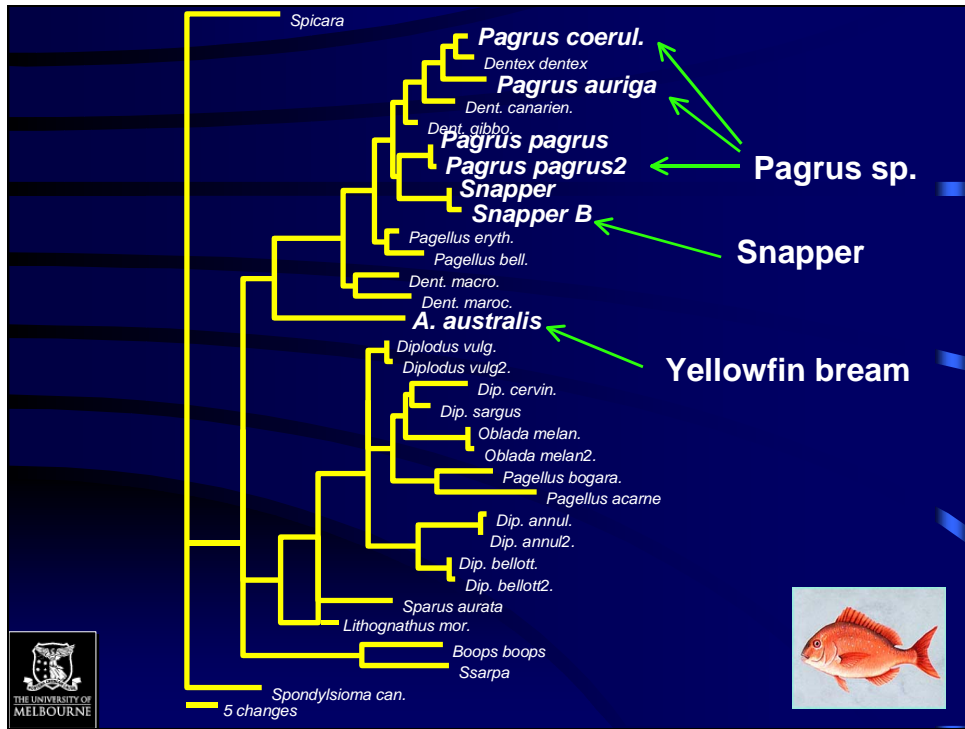
Slide 8: Snapper phylogeny.

To understand patterns of contemporary genetic diversity (like change in N_e) it is necessary to understand the patterns of variation at a number of different levels. This is necessary to ensure that observed pattern of genetic diversity is the result of contemporary processes and not due to historical processes. For example measurements of N_e would be confounded if they were taken across cryptic subspecies. MtDNA sequence data can help identify cryptic population and species structure.

This phylogeny is based on 16S mtDNA sequence data. We included a large number of Atlantic sparid sequences for comparison. *Pagrus auratus* is shown on this phylogeny as snapper and it groups with other closely related Pagrids that appear to have recently speciated. However an Australian Sparid, yellow fin bream, is clearly outside this group. My interpretation of this is that the radiation of Sparids in the Indo-Pacific predates the arrival of snapper from the north Atlantic region. That is, a recent ancestor of snapper evolved in the North Atlantic Ocean well after the diversification of Sparids in the Indo-Pacfic.

Slide 9: New Zealand snapper genetic study – 2.

Phylogeographic structure is perhaps more relevant in the context of the talk than taxonomic structure.

Slide 10: Historical and current processes: phylogeography and population structure.

Australian sequences were obtained from Steve Donnellan at the Evolutionary Biology Unit, Adelaide Museum and compared with our New Zealand collection. The samples used in Steve Donnellan's study were predominantly from eastern South Australian and Victorian coastal sites. The main question we wanted to address using mtDNA markers was the degree of past and present connectivity between Australian and New Zealand snapper populations. We were also interested in comparing the process of colonisation of snapper in New Zealand and Australia.
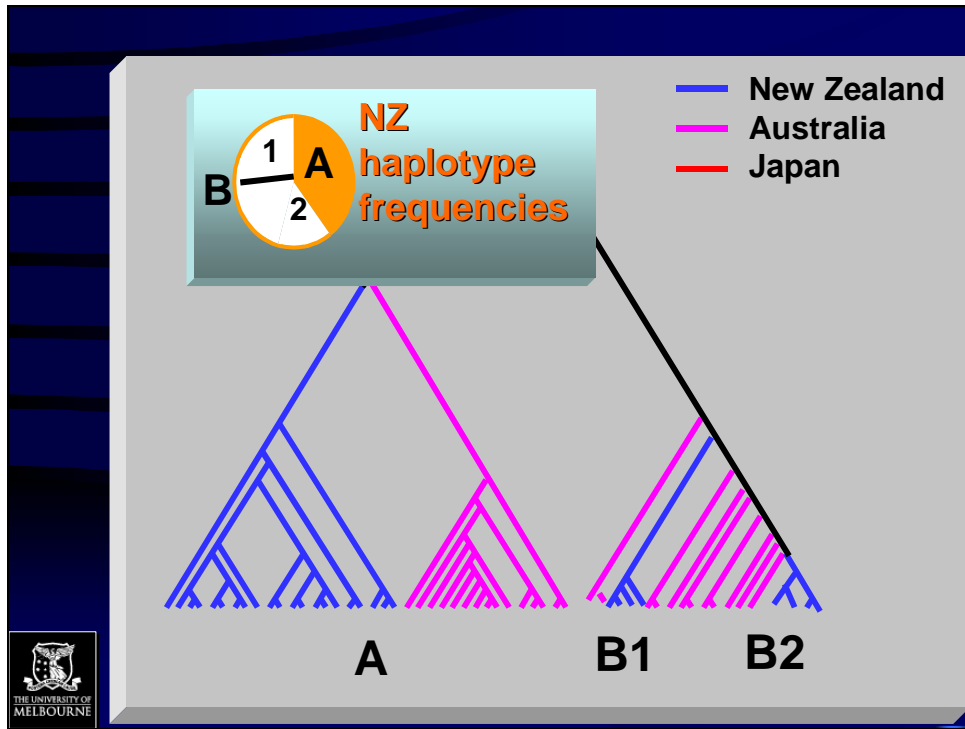
Slide 11: Snapper phylogenetic cladogram.

Australian sequences are in pink and New Zealand sequences are in blue. The sequences cluster into three distinct clades. The first clade contains only sequences from New Zealand, the second clade contains only sequences from Australia and the third clade is a mixture of both sequences from both New Zealand and Australia. About 60% of snapper in New Zealand have sequences that belong to the less diverse B group while only 40% belong to the more diverse A group. New Zealand snapper in the B group appear to have derived relatively recently from Australian snapper populations. New Zealand snapper in the A group appear to be more anciently derived from Australian snapper in the same group and may be representative of the initial colonisation of Australia and New Zealand.
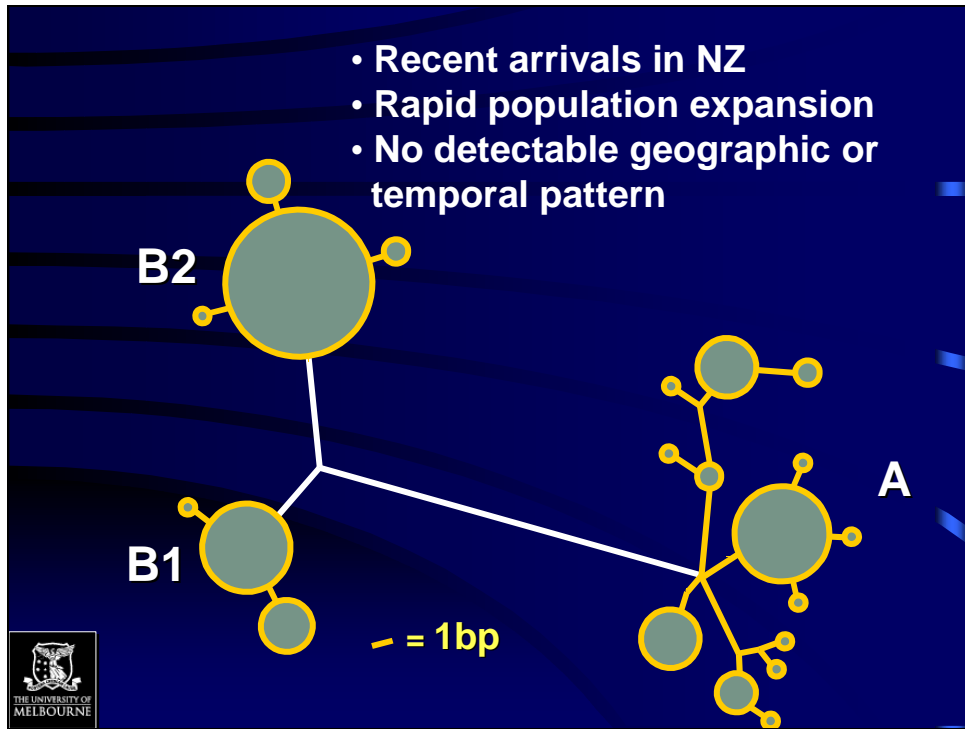
Slide 12: Haplotype network tree of snapper in New Zealand.

This haplotypic network represents the relationships among the different sequences and their frequencies. Haplotypes are represented by circles. The diameter of the circle is proportional to the frequency of the haplotype. The length of the lines connecting haplotypes is proportional to their genetic divergence. It is possible to infer aspects of population history from the structure of the haplotype network. Sequences in group A have a signature of an ancient colonisation (~ 4 mya) and subsequent expansion of snapper in New Zealand. Sequences in group B appear to have arrived in New Zealand more recently, say in the late Pleistocene – early Holocene, because they have fewer derived sequences from the common internal and presumably ancestral haplotypes. Snapper in group B presumably had a selective advantage over the original snapper in group A because they now represent 60% of snapper in New Zealand. It appears as though selection is no longer acting as the frequency of group A and B and has been stable since the 1950's. Despite clear evidence for historical connectivity between Australia and New Zealand Snapper, the mtDNA data show that there is no evidence of contemporary gene flow.

Slide 13: New Zealand snapper genetic study – 3.

While snapper in New Zealand and Australia are strongly structured, we could not detect any phylogeographic structuring of snapper within New Zealand.
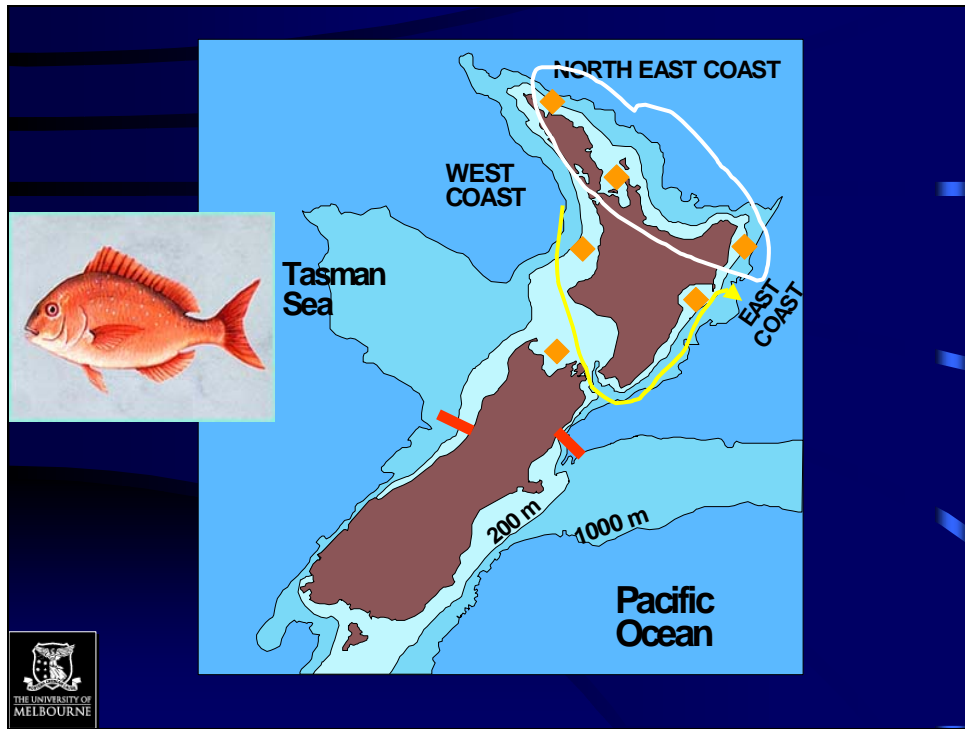
Slide 14: Study site.

We investigated genetic structure at microsatellite loci from six New Zealand sampling locations. The red lines represent the southern limit of snapper distribution in New Zealand. We failed to find any genetic differences among the three populations sampled in northeast New Zealand and we consider them as belonging to the same demographic population. This observation is consistent with a 1978 allozyme study by Peter Smith. The remaining three populations differed from each other in that the populations on the east and west coasts were more similar to each other than either was to the Tasman Bay population. This finding is also consistent with Peter Smith's 1978 allozyme study. The east and west coast populations appear to be connected by a current that runs through Cook Strait and another that runs up the west coast. The population in Tasman Bay is the most distinct and we presume that this is due to eddying currents that limit gene flow. More samples from intermediate locations are required to adequately test this hypothesis.

Slide 15: Temporal study of two populations.

Scale samples collected 50 years ago from snapper in Tasman Bay and the Hauraki Gulf were used to compare historical versus contemporary patterns of genetic diversity.
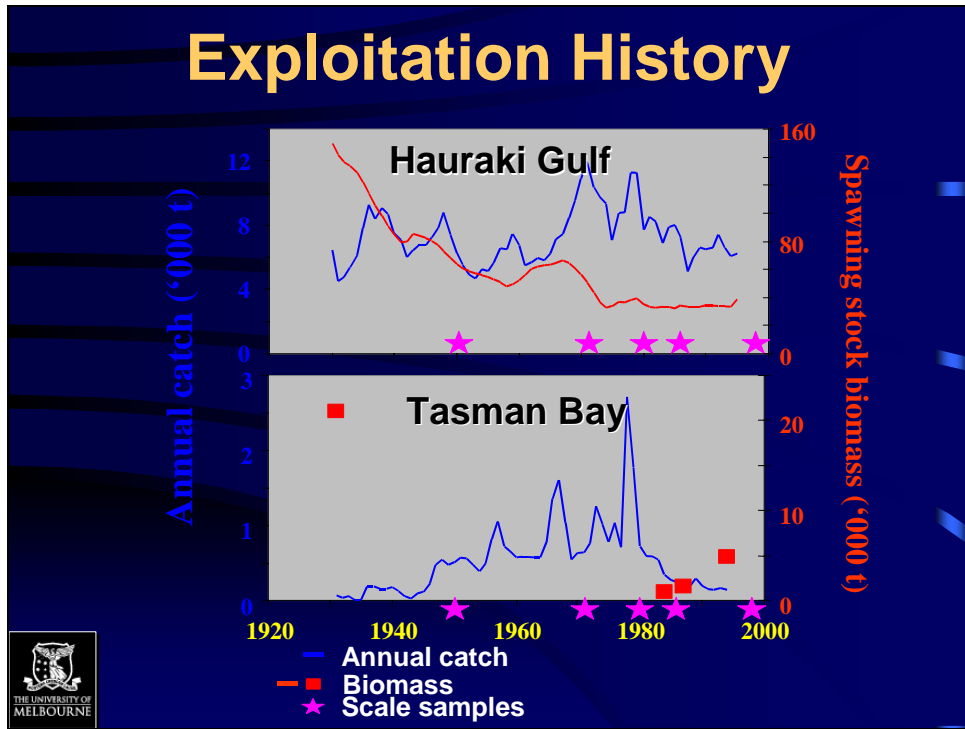
Slide 16: Exploitation history.

The exploitation history of Hauraki Gulf and Tasman Bay are different allowing us to explore the effects of fishing pressure on patterns of genetic variation. The blue line represents annual catch and the red line represents spawning stock biomass. The fishing intensity at Hauraki Gulf has been relatively constant from the 1850's and has only recently intensified since the 1970's. In contrast, the Tasman Bay fishery commenced in the 1950's. In the early 1980's strong management was implemented to arrest a decline in stock size.

- **Hauraki Gulf**
  - fished since pre-1900
  - B_0 =254,000 t
  - B_{80s}=33,000 t (13% of B_0)

- **Tasman Bay**
  - "significant" exploitation only since late 1940s
  - B_0 =21,000 t
  - B_{80s}=1,500 t (7% of B_0)
  - Est. No. of spawning stock = 500,000 fish

Slide 17: Exploitation history (cont.).

Virgin biomass in the Hauraki Gulf was around 254000 tonnes and was reduced to about 13% of that in the 1980's. Virgin biomass in Tasman Bay is much lower at 21000 tonnes and that was reduced to about 7% (or 500000 fish) of that in the 1980's.

Slide 18: MtDNA D-loop results.

We compared the mtDNA data from the 1950's through to 1998 by observing the changes in two diversity measures, the number of haplotypes and haplotype diversity or heterozygosity. There are no clear patterns emerging from these data. Perhaps the large variances are hindering our ability to detect changes in genetic diversity over time.
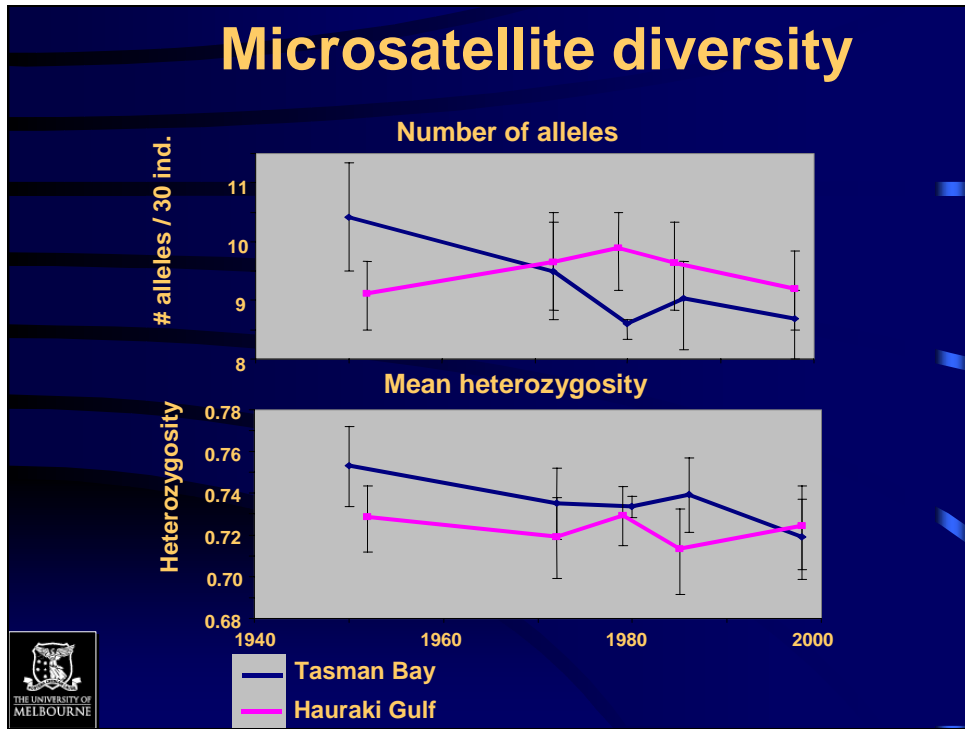
Slide 19: Microsatellite diversity.

Changes in the total number of alleles and mean heterozygosities over time for Hauraki Gulf and Tasman Bay populations are presented here. In contrast to the mtDNA data, with the microsatellite data we observed significant declines in genetic diversity over time for both Hauraki Gulf and Tasman Bay populations.

Comparison between 1950 and 1998

- **Tasman Bay**
  - Significant trend for number of alleles: $P=0.028$ (Wilcoxon signed rank test)

- **Hauraki Gulf**
  - No clear trend

	Tasman Bay		Hauraki Gulf	
	# alleles	He	# alleles	He
GT2	-	-	-	+
GT4	-	-	+	-
GA2A	-	-	+	+
GA2B	-	-	-	-
Pma1	+	+	+	-
Pma2	-	-	-	-
Pma5	-	-	+	-
MtDNA	-	-	+	+

Slide 20: Comparison between 1950 and 1998.

This table compares the microsatellite data for Hauraki Gulf and Tasman Bay locus by locus and shows the direction of change in genetic diversity measures from 1950 to 1998. A significant decline (Wilcoxon signed rank tests) in genetic diversity was only detected in the Tasman Bay population.
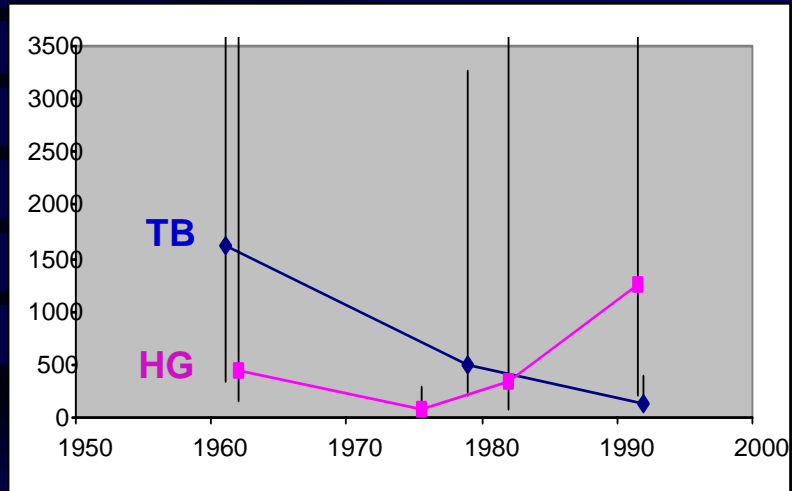
Slide 21: Effective population size.

This chart shows the change in estimates of effective population size through time for both Hauraki Gulf (HG) and Tasman Bay (TB). We expect that there is a very low chance of losing genetic diversity over time in very large populations. Thus we expect to be able to observe changes in genetic diversity over time only in small populations, say less than 1000 breeding individuals. The upper confidence limits for effective population size are all infinite in the Hauraki Gulf population but are all finite in the Tasman Bay population. The downward trend in effective population size (from Waples' moments estimator) for Tasman Bay is matched with a reduction in genetic diversity and stock biomass.

Slide 22: Conclusions.

Genetic diversity has been lost in Tasman Bay due to small effective population sizes where the ratio of N_e/N is around 10^{-5}. Combinations of demography (highly fecund long-lived fish) and fishing mortality have contributed to these low values of N_e. Perhaps more important than the actual value of N_e, we have clear evidence that genetic diversity has been lost in New Zealand snapper populations that are currently supporting fisheries.

Does it matter ?

- **Loss of genetic diversity in quantitative traits?**

- **Alleles which may be important for adaptation to environmental change may be lost**

- **Assumptions about "safe" catch sizes may be incorrect**

- **What have we learned about and from N_e?**

Slide 23: Does it matter?

We have not tested for the loss of quantitative genetic traits (those under selection) in New Zealand snapper but we have demonstrated a loss of genetic diversity in neutral markers. Does this matter? The intensity of the decline in genetic diversity suggests that some quantitative genetic traits under weak selection may have also been lost.

The fact that N_e was measured as being 10^{-5} lower than N, whether or not this led to a loss of diversity, is not expected from our present knowledge of the biology and dynamics of this species. This knowledge is what is used to calculate various parameters on which the fishery is assessed. Clearly if there are aspects of this species that cause such a low N_e compared with N, these will impact on fisheries models. While it could be argued that snapper is not a model for other fisheries, this finding still points to a general concern. Snapper is very well studied in comparison with most fished species. If we could not foresee the loss of diversity attributed to the low N_e in snapper, how can we be confident that other species are not suffering a similar hit to their genetic diversity?

While the loss in genetic diversity is true, we are still not sure as to what factors are driving the low N_e in New Zealand snapper. If we knew what factors influence N_e then we could adjust fisheries management practices so that it maximises effective population size.

Questions

Gopurenko:	Just remind me in terms of Tasman Bay samples, you had a loss of diversity through time with microsatellites but what happened to mtDNA?
Adcock:	It reduced. I included it in the plus-minus graph. It was one of the minuses among them. There was a trend downwards but it was not significant.
Gopurenko:	That's unusual that [the loss of diversity] for mtDNA would not be more apparent than microsatellites.
Adcock:	Yes. I think it's to do with the power. First of all it's one locus. The microsatellite's trend downwards was a combination of eight highly variable loci. The variability in the mtDNA test was not as great. If we plotted any one of those microsatellites on a similar graph it would not have given a significant trend. So it was the combination of those that gave a significant trend. So while mtDNA is supposed to have an effective population size of 1/4 of a nuclear gene, but I think it's all to do with the power of that locus.
Gopurenko:	I was just wondering if it was a rare thing?
Adcock:	Maybe the population history might be having some sort of impact that we haven't detected as well.
Donnellan:	The effective population size of mtDNA depends on the sex ratio to some degree.
Adcock:	Yes. I think they are fairly equivalent. This is where Peter Smith's knowledge would be useful. I recall that there is no disparity in the sex ratios.
Donnellan:	Could you be getting some weird sampling effect?
Adcock:	Like if we had sampled all from one sex and not the other.
Donnellan:	In a really small population?
Adcock:	I wouldn't discount that. We certainly wanted to look at a finer scale and we could possibly sex the scales too, which I haven't thought about.
Page:	From catch data it seems like the Hauraki Gulf is a much larger fishery. So, is it a larger population?
Adcock:	Yes.
Page:	Why does it look like N_e was a lot higher in the Tasman Bay than in the Hauraki Gulf?
Adcock:	You've given me the opportunity to mention something I've forgotten. What you expect with a population with a much larger population size biomass is the retention of much higher diversity. So you expect this population [Hauraki Gulf] to be more diverse than this one [Tasman Bay]. The Tasman Bay population actually has a lot more diversity than the Hauraki Gulf [slide 19]. You expect from the population size that the pink line should be above the blue line because it should maintain more diversity. That suggests to us that possibly the Hauraki Gulf has already lost diversity.
Page:	Did you find A and B mtDNA haplotypes at both?
Adcock:	Yes it's all the same. That would have been an exciting thing had they been different. In answer to your question why is it lower? I don't know. What we think is that the Hauraki Gulf estimates are severely compromised by migration. It's a very big population. We are sampling one spot in a population that's continuous. There is going to be lots of inter-annual effects, I think.
Page:	Could there also be some confusion (fishing not as much before compared to now)?
Adcock:	It could be movement into the population that hasn't reached some sort of equilibrium. Anything that is non-equilibrium is going to affect those. The Tasman Bay is the one that we have got most confidence in because of its history and it population status.
Page:	Still fairly long confidence intervals.
Adcock:	Yes they are. Perhaps we can get some Bayesian stuff going to reduce those error rates.
Ovenden:	I just wanted to comment that perhaps the Tasman Bay population is receiving a small number of migrants from the North Island. If that was occurring that would give you the impression of more measurable genetic drift, which would actually drive your N_e estimates down to this level. Did you try and do maybe some assignment testing to see if individual snapper where sourced from other populations?

Adcock:	No, it's a good idea. I once plotted everything into Genetix, a French program that produces a 3D plot, and there was one genotype that was way out there. It's always a possibility that it is an Australian one. It is possible that some sort of principle component analysis will bring out something, but I'm not sure there is the power in the genotypes for that.
Anderson:	Do you think that the difference in the mtDNA results and the microsatellite results, where you have noticed a big difference in change in genetic diversity, may be a result of the sampling size because microsatellites are actually more polymorphic and you are not picking up some of the more rare alleles?
Adcock:	Yes that's true.
Anderson:	Did you actually mention the sample sizes in there?
Adcock:	No I did not mention them. They are all around 50 accept for the 1980 population which was 30. That's why, for all our estimates, we had to resample the other populations down to 30 to get a comparable estimate for that same population size for each sample.
Page:	Back to your phylogeny tree in the very beginning you talked about the New Zealand stock and the Australian stock. It looked as if there was a really deep node. I would have thought that if they came from Australia they would be nested within the Australian ones. As opposed to being quite separate [Slide 11].
Basford:	What data was actually used to produce that tree? What measurements were you using on the fish?
Adcock:	These were mtDNA sequences.
Page:	I would have thought B1 and B2 would have been nested within the Australian ones.
Adcock:	They are. This one certainly is. But also this [clade B] was not exhaustively sampled for the pinks. I would predict that you would get something as close to this one [B2] as that is [B1]. Definitely if they come from the east coast of Australia, none of these pink ones are very far up the east coast. They are mostly southern Australian, southern Victorian and their origin may not be represented here.
Donnellan:	Could selection explain the mtDNA results?
Adcock:	Yes of course.
Donnellan:	With your change in frequency or lack of change in frequencies?
Adcock:	I am not sure how that would happen.
Donnellan:	Could be some weird balancing selection?
Gopurenko:	One clade would be selected for all of its derivatives.
Donnellan:	That would be for directional selection.
Adcock:	The main selective forces you could imagine would be ecologically based and there are no ecological differences between southern and northern samples for the two types, for group A or group B, or any subdivision within those so it wouldn't seem as though that is happening, unless there is something else that I couldn't think of.
Ovenden:	Towards the end of your talk you talked about the management implications of knowing the effective population size for the snapper populations was remarkably small. There are two ways of looking at this. One is to manage the population so that it remains large enough so that in the long term it doesn't loose genetic diversity. But there is another side to this and that is that the population is currently being sustained with remarkably few spawners. The same interpretation can be put on the Moreton Bay tiger prawns. The population is stable in size from year to year but the offspring are coming from a very small number of individuals. Ideally from a management point of view you would want to provide protection for those spawners because they are the critical spawners that are providing recruitment. In most fisheries you can't do that because those critical spawners are equally dispersed amongst all of the other spawners.
Adcock:	It might be more tractable with snapper. What we don't know is whether there is some other anthropomorphic pattern. This whole pattern could be explained by the loss of spawning grounds over the same period of time, which is possibly an indirect effect of fishing. That's what I mean by not understanding the biology. Do you protect grass bed or what ever or do you just reduce fishing. What is more important?

Ovenden: I'm concerned because looking at these results through a fisherman's eyes they would say well very few individuals are needed to propagate this species so why can't we increase our catch and still maintain the same level [of both biomass and of genetic diversity].

Adcock: Many population geneticists would agree with them. The Mauritius kestrel began from two individuals twenty years ago and it is now a thriving population and it has no genetic variation at all. It's growing in numbers with no visible signs of any problems and it's not a species that had coped with low genetic variation in the past. It was a highly diverse one. The assumption that a loss of diversity is a bad thing is not completely well tested so I don't know how to answer that.

Dichmont: This is a little off the topic but it is with regards to management. The interesting thing in the NPF in the gulf is that we actually looked at effective population size in a different way. We took an ecological point of view compared to a genetic point of view. You've got a limitation that the larvae have to get to the seagrass beds under certain oceanographic conditions and that they are generally fairly passive swimmers, they migrate up and down but can't really swim much in any direction. You can actually simulate how far they have to be away from a seagrass bed to be an effective population. We worked out that a substantial proportion of our population in the NPF was outside that boundary layer. That does have management implications. If for example a genetic difference showed that the effective population size was much smaller than we thought, then that would already be quite useful information. There are other ways of linking in the genetic techniques to other techniques to identify the numbers of spawners in the population.

The prawns who say 'N$_e$': how and when to use N$_e$ in stock assessment

Simon Hoyle

Southern Fisheries Centre, Department of Primary Industries, Queensland

Abstract

Fish stock assessment is at the cutting edge of methods for quantitative risk assessment in general, but new methods for data collection have lagged behind. Stock assessment is greatly limited by data availability. Effective population size may provide a new data source with which to increase the precision and accuracy of stock assessment, and improve fishery management decisions. Advantages include fishery independence, focus on individual species, and a unique ecological perspective.

Effective population size may be included in stock assessment models by deconstructing the stock recruitment relationship into a two-stage 'genetic stock recruitment relationship': the stock to N$_e$ relationship, and the N$_e$ to recruit relationship. Equations are proposed and placed in a state-space modelling context. This context permits estimation of the requisite parameters and hypothesis testing of alternative models, and uses the power of the additional information to reduce uncertainty in management outputs. Results from a simple simulation show increased precision of parameter estimates in some scenarios but not others. Further model development and simulation are recommended.

Factors that affect the usefulness of N$_e$ estimates are complex and context-dependent, and include observation error and process error in N$_e$, the length and contrast of the N$_e$ and fishery time series, and the quantity and quality of other fishery data.

Fisheries more likely to benefit from N$_e$ estimates will have short generation time, non-overlapping generations, small stock size, low fecundity, and/or a high N$_e$ to N ratio. Fisheries with a longer generation time and overlapping generations, but reliable ageing, may also benefit given further development of theory and practice.

Management of the northern prawn fishery is likely to benefit in the long term from a time series of fishery-independent, single species, substock-based N$_e$ estimates. The sooner sample collection and storage begins, the sooner the data will become useful. Samples should be stored long-term to obtain the benefits of future technology advances, which will reduce observation error. This fishery may contain a series of sub-stocks. Lack of awareness of existing stock substructure will bias N$_e$ estimates upwards, which will reduce their utility (for some purposes) and increase observation error. For this reason, stock discrimination analyses are recommended. A method is suggested which may permit stock discrimination even where migration has equalised allele frequencies.

The following steps are suggested to further develop the use of N$_e$ in stock assessment: 1) develop SNP technology, since it reduces annual costs and observation error in N$_e$; 2) further develop the integrated stock assessment model, and simulate alternative fishery and life history scenarios; 3) begin sampling in the northern prawn fishery; and 4) begin cataloguing existing sampling in suitable candidate fisheries, and identify fisheries where additional data collection may provide greatest benefit.

Presentation

Slide 1: The prawns who say 'N$_e$'.

Today I am going to try and answer the question of when and how to use effective population size in stock assessment.

Slide 2: Introduction.

There is a long list of things to go through in this talk and the structure of the talk is outlined here.

1. Objectives

- Overall goal is more accurate stock assessment models
- Project objective: determine conditions where Ne can be used to improve stock assessments

3

Slide 3: Objectives.

The overall goal of this project is to get more accurate stock assessment models and better management outcomes. In this presentation I will discuss the various conditions where effective population size can be used to improve stock assessment models.

Slide 4: Stock assessment.

Stock assessment models describe populations as they travel through time. The model we use to describe that process contains various kinds of uncertainty. Model error is where you are unsure as to the exact nature of the processes that are going on in the population and are therefore unsure as how to best model the population. Observation error, or sampling error, is a well-known source of error and involves the things, like age structure or catch per unit effort, that we assume we can observe. Process error is the variation in the process from year to year, for example natural mortality may be different each year or there may be a different ratio between stock size and effective population size each year. Estimation error is a combination of the three sources of error above. It is the uncertainty in estimating the parameters within your model because of the combined effect of those sources of error. What you are trying to do as part of your model is determine whether the raw effective population size data will give you useful information.

Slide 5: N_e as data.

N_e as raw data has a lot of good features. It is a new and independent data source going beyond the usual kinds of data like catch and CPUE data that are typically available for stock assessment. N_e data gives a unique perspective on the fishery that has not been previously available. It is fishery independent data enabling it to avoid some of the problems with fishery dependent data like effort creep using CPUE data; Cathy Dichmont talked in detail earlier about some of these problems. It is also independent of the fishing sector so that you can still use your effective population size data even if the fishery completely transforms itself from using say trawling gear to traps. It is also single species data meaning that it is possible to focus on a particular species even in multi-species fisheries. In the long term the technology may become relatively cheap to use as a monitoring tool on a yearly basis.

4. Role of Ne in stock assessment

- Absolute abundance?
  - No: N ≠ Ne
- Index of abundance?
  - Sometimes, but probably non-linear + noise
- Index of recruitment?
  - Sometimes, but probably non-linear + noise
- More than this?
  - Information about ecology of recruitment

6

Slide 6: Role of N_e in stock assessment.

How can stock assessment use effective population size? As we have already discussed it is not the same as absolute abundance. Sometimes, effective population size may have some of the features of an index of abundance but as I will discuss later the relationship between N_e and N is probably non-linear and it is a fairly noisy relationship. There may be some relationship between N_e and recruitment but it is probably also non-linear and noisy. Beyond these two relationships, N_e provides some additional information about the ecology of recruitment.
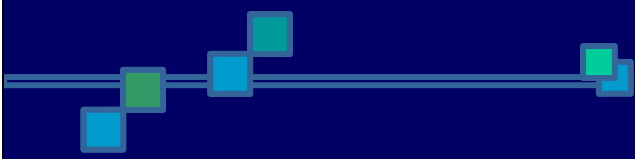
Slide 7: 'Life history' perspective.

A good way to look at effective population size is that it gives unique information about reproduction, because the main factors effecting N_e are the number of individuals reproducing and the family related mortality of the offspring. N_e is most useful in a fisheries context when it is looked at as a component of the stock recruitment relationship.

Slide 8: Mortality types.

When eggs are first spawned they are in close proximity to their siblings, resulting in strong spatial patchiness of relatedness. Over time this patchiness will reduce, but on some scales it may linger for a long time. Anything that affects one group of related individuals (eggs / larvae / juveniles) more than another group of related individuals is going to reduce the effective population size. It is difficult to conceive how the offspring of different individuals could have the same survivorship, given variable environmental conditions and the patchiness in space and time of recruitment in a natural population. So this is the first type of mortality – family-related mortality. In addition there is the general background mortality across the entire population. Sources of mortality from spawning right through to eventual recruitment back into the spawning population can be classified into general and family-related mortality.

Standard stock-recruitment relationship
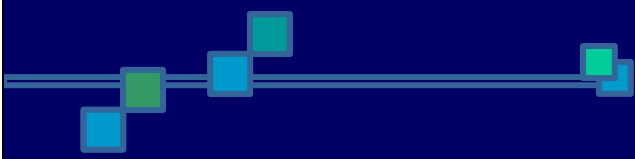
- A well-established perspective on fish reproduction
  - Recruits as function of stock / spawning biomass

- Standard Beverton-Holt model (q.v. Ricker)

$$R = \frac{S}{\alpha + \beta S} . e^{\varepsilon}$$
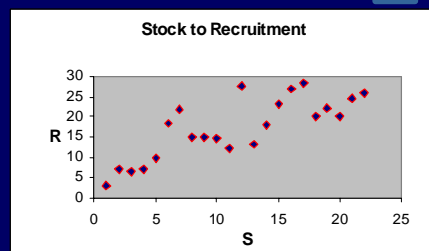
Slide 9: Standard stock-recruitment relationship.

One of the standard models used for looking at stock recruitment relationships in fisheries is the Beverton-Holt model (q.v. Ricker) where recruitment is a function of stock and spawning biomass. The model has a fairly flexible shape that describes various possible relationships between stock and recruitment. The error parameter describes the degree of variability in the relationship. A typical relationship between stock and recruitment is shown in the chart. In this example there is a good relationship between stock and recruitment at low stock sizes, but large amounts of noise quickly flatten out this relationship for larger stock sizes.

Slide 10: Proposed model – genetic SRR.

The genetic stock recruitment relationship (SRR) divides the stock recruitment process into two components. The first component is the spawning variation and family related mortality, denoted here as the function f, where N_e = f(spawning biomass). The second component is the general mortality function g, where recruitment = g(N_e). Both of these model components have the same form as the standard stock recruitment relationship where N_e is a function of stock size and recruitment is a function of N_e. The level of spawning and family related mortality will determine the effective population size, and the general background mortality will produce the recruitment that results from that effective population size.

Slide 11: Substitution.

Here I have just substituted one formula into the other, showing that it still retains the same form as the standard stock recruitment relationship. The standard Beverton-Holt stock recruitment model can now be considered as a special case of the genetic stock recruitment model.
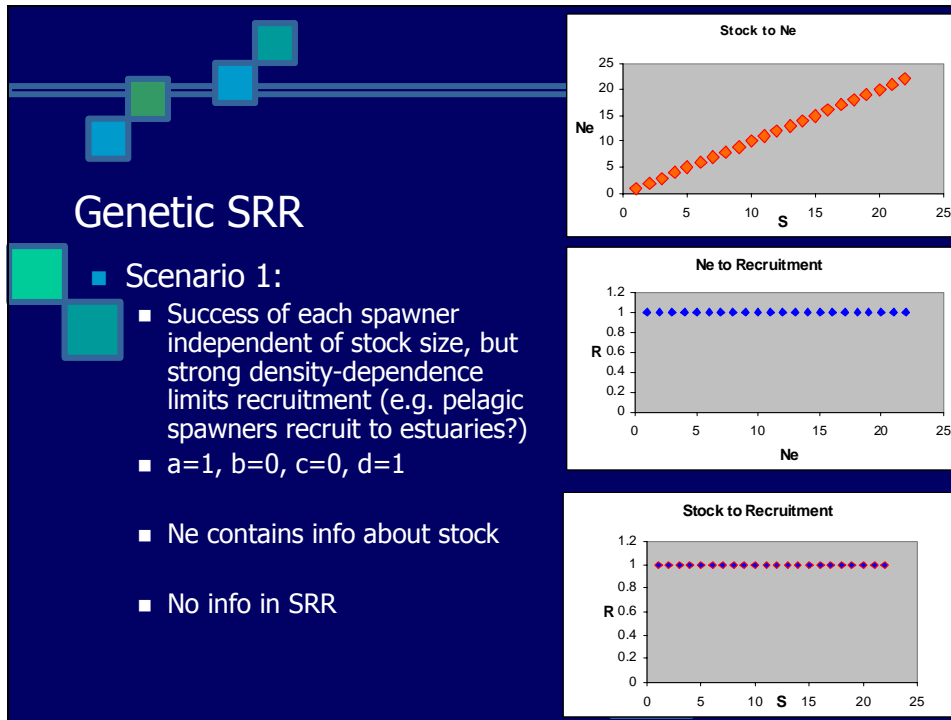
Slide 12: Genetic SRR – Scenario 1.

I'll now go through a couple of scenarios, exploring what effect different values of the parameters have on the stock recruitment relationship. The values of these parameters represent extreme cases. In this first example N_e is directly proportional to stock size, effectively making it an index of abundance, and there is no relationship between recruitment and N_e. Under this scenario the success of each spawner is independent of stock size so that the greater the spawning biomass the bigger the effective population size. Very strong density dependent factors limit recruitment so that no matter how many individuals are spawning recruitment is constant. This scenario may be applicable to pelagic spawners where once the larvae are fully mixed they recruit to estuaries, where their survivorship becomes density dependent. In this case N_e contains information about stock size but bears no information about recruitment or the stock recruitment relationship.
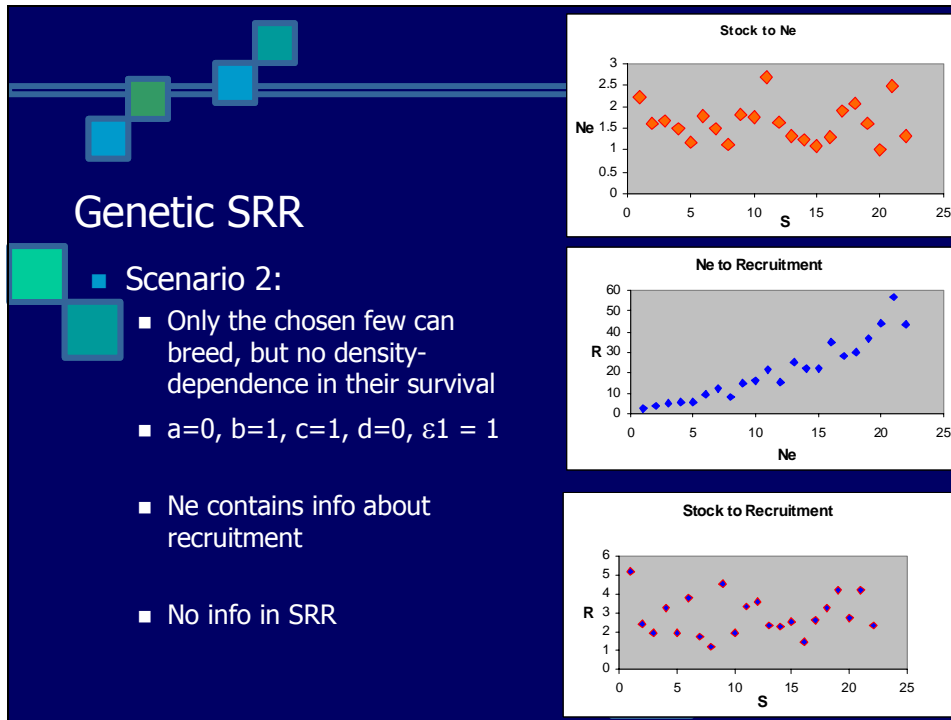
Slide 13: Genetic SRR – Scenario 2.

This scenario describes another extreme example where only the chosen few are able to breed, perhaps due to some sort of territoriality. However, once they have bred there is no density dependence acting on their survivorship. In this case there is no relationship between stock and effective population size, because only a few animals in any given situation are able to breed. Because there is no density dependence influencing survivorship, more recruits result in higher effective population sizes. Under this scenario N_e contains information about recruitment but not about stock size or the stock recruitment relationship.

Slide 14: Is there N_e N_e relationship?

Recent meta analyses suggest that there is a relationship between stock and recruitment for most fisheries. If that were the case then we would logically imagine that there should be even stronger relationships between N_e and stock and between recruitment and N_e, if the relationships are as I have described. However this may not always be the case, as there are mechanisms where stock can affect recruitment indirectly without going through the normal path of N_e. For example, if an adult stock preys on another species that competes with its juveniles then a higher stock size will result in higher recruitment, but this stock recruitment relationship is not mediated through N_e.

Slide 15: Modelling context.

The genetic stock recruitment relationship in a modelling context can be used as a confirmation of other stock recruitment models based on traditional information. However a more powerful approach is to unify all of the models into one. Among the benefits are that this permits statistical testing of various hypotheses, to see if different models are suitable for the data. Uncertainty is also substantially reduced, and can be estimated better when all of the data are combined into a single model. However, there are more parameters (a, b, c, d, $\epsilon 1$ & $\epsilon 2$) in the genetic stock recruitment relationship than in the standard SRR. There is a basic statistical principle that estimating more parameters requires more information. N_e estimates do provide new information, but additional information from the fishery such as longer time series of fishery data may also be needed. We need to simulate to examine this issue further.

Slide 16: Modelling solution.

A solution is to use an integrated state space model, in which the genetic stock recruitment relationship replaces the standard stock recruitment relationship. We have partly developed this model in a Bayesian framework. We believe that further development of this model through simulations is highly desirable, but was unfortunately outside the scope of the current project.

Slide 17: What affects its usefulness?

There are several things that are likely to affect the usefulness of N_e as part of an integrated stock assessment model. In the next few slides I'll go through some of those things. In general, data quality and the amount of data affect the amount of information about the genetic stock recruitment relationship, which obviously affects the speed and precision with which genetic parameters can be estimated. The genetic stock recruitment relationship may be more useful in certain kinds of fisheries than in others. We really need to use simulation to explore in more detail the behaviour of the genetic stock recruitment relationship parameters, and the way features of the fishery and data affect the usefulness of N_e.

Slide 18: Observation error in N_e.

Observation error in N_e is a very big issue, and David Peel's simulations have shown ways to reduce and quantify this source of error. It appears as though while single nucleotide polymorphisms (SNPs) have more power than microsatellites, even they may not have sufficient power for many fisheries that have large effective population sizes. Future developments will improve the power of these methods and reduce costs so that more samples can be analysed. It is difficult to speculate on how long it will take until observational error is reduced to acceptable levels but it is already moving rapidly in the right direction.

Slide 19: Process errors in N_e.

Process error in the estimate of effective population size is the randomness in the true N_e relationship each year. Standard stock recruitment relationships are traditionally subject to large process error. The larger the process error the longer it will take to estimate the genetic stock recruitment parameters and the longer it will take before estimates of N_e will be useful to the fishery. The example shown here is data from a tiger prawn fishery, including both estimation and process error. Reducing the fishery biomass to very low levels as in this example would allow us to get good estimates of the parameters; but we would be less certain of those relationships in a fishery maintained at higher levels of biomass. There are two components to process error in the genetic stock recruitment relationship model: ε1 & ε2. One may be higher than the other and their relative magnitudes will be unknown until you have a time series of effective population size estimates. Comparing species with similar life histories may provide priors for novel fisheries but this is a long way off and will require N_e estimates from a large number of fisheries.

Slide 20: Process errors in N_e (cont).

Process errors are generally unknown and are species and stock dependent. If the process errors are too large then it may not be possible to use N_e as an index of abundance or as an index of recruitment. However, this does not mean that N_e is not useful because information about the life history on the species may be uncovered during the process of investigating the causes of process error.
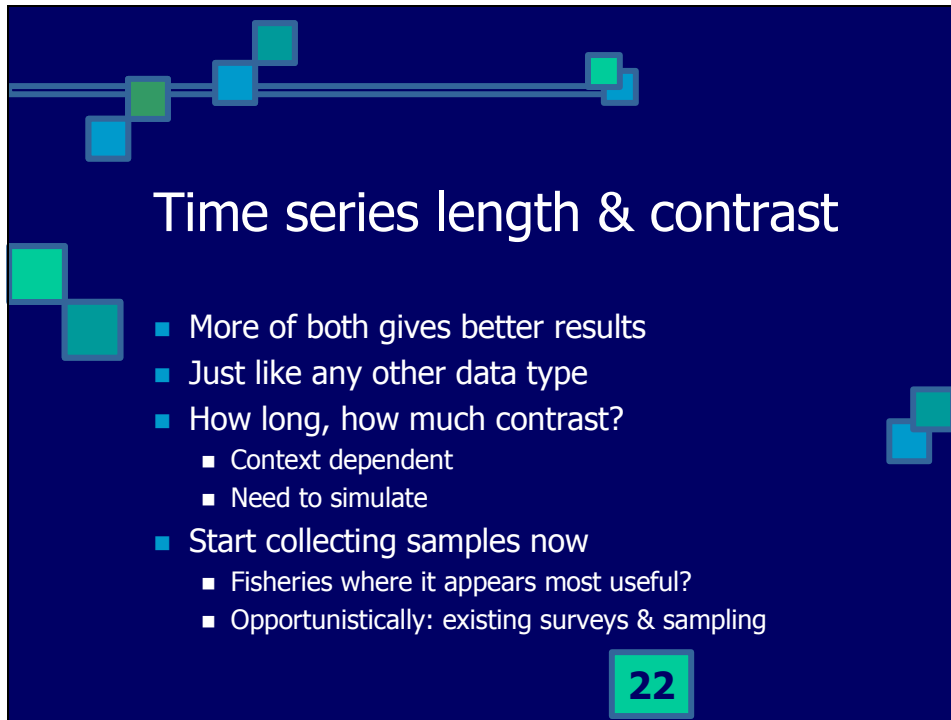
Existence of comparable data

- If abundance or recruitment indices exist
  - Faster to estimate Genetic SRR parameters
  - Ne not redundant
  - Validate, reject, or adjust other index
  - E.g. show that hyperstability affects CPUE
  - Reduce model error
- If such indices don't exist
  - Slower to estimate Genetic SRR parameters

21

Slide 21: Existence of comparable data.

The existence of comparable data improves the usefulness of N_e. The more that is known about a stock the more quickly the relationship with N_e can be deduced. If you already have alternative abundance or recruitment indices then N_e estimates may help to validate the information you have or reject some of the assumptions you are making about the fishery. Knowledge of N_e can therefore reduce your model error and help to ensure that you are using the right sort of model. If alternative indices do not exist then the genetic parameters will be much slower to estimate. Once the genetic parameters are estimated the model can then be effectively used for stock assessment and management.

Slide 22: Time series length and contrast.

As with any data type, results are much better with longer time series and more contrasting data. By contrasting data I mean that the power to uncover a signal in the data is limited unless you can compare data from stocks in both good and bad situations. Poorly managed or highly variable fisheries can provide this sort of contrast. How much time and how much contrast you need is very context dependent, given the other parameters in the model, and can only be answered through simulation of each fishery. Because long time series are so valuable it would be a good idea to start collecting tissue samples now, particularly in fisheries where we think N_e may be most useful. Opportunistic sampling during existing surveys and at the market place is perhaps the most efficient way to do this.

Slide 23: Types of fisheries.

The types of fisheries where N_e may be most useful include those with non-overlapping and short generation times. N_e is easiest to estimate in these fisheries and results can be forthcoming in reasonable timescales. Fisheries with overlapping and long generation time are more problematic and the theory is still being developed. Accurate ageing of individuals will be important in these sorts of fisheries. N_e is likely to be more useful in fisheries with smaller process error because the relationship between N_e and stock size will be easier to elucidate. Likewise, N_e will be more useful in fisheries that have high N_e/N ratios and in fisheries with lower stock sizes, since this will reduce effective population size. Fisheries with low observation error, such as those with smaller effective population sizes and those with more molecular markers giving more precision, are likely be good candidates for the genetic stock recruitment models.

8. Further simulation testing

1. Set up 'operating model' structure and parameters (including process error and observation error)
2. Generate data (Ne, CPUE, catch)
3. Use 'estimating model' to estimate the parameters of interest
4. Repeat many times

24

Slide 24: Further simulation testing.

This is the proposed simulation method that we will use to validate our genetic stock recruitment models. Unfortunately there is no simulation data as yet to present.

9. Possible spin-offs: stock discrimination

- Separate but cryptic sub-stocks may exist
- Managing separately gives better outcomes
- Sampling for Ne gives power to detect small allele frequency differences
- If > 1 migrant per generation, allele frequencies will not differ: stocks may be separate on management but not on evolutionary timescales
- May use Ne estimates to infer stock subdivision

25

Slide 25: Possible spin-offs: stock discrimination.

There are possible spin-offs beyond estimating N_e from collecting genetic data. Cryptic stocks may exist in the fishery that adversely influences the efficacy of management. Management outcomes are much better if separate stocks are managed separately, and it is therefore advantageous that cryptic stocks be recognised. Sample sizes needed to estimate N_e are very large and therefore typically have the power to detect very small allele frequency differences between potential cryptic stocks. However differences in allele frequencies between stocks will not be observed if migration is greater than one migrant per generation. In these cases stocks may be separate on management time scales but not on evolutionary timescales. While beyond the scope of this presentation, we have been considering ways to recognise this kind of stock subdivision using changes in the effective population size estimates among the cryptic stocks.

Slide 26: Northern prawn fishery.

One of the aims of this project was extension into the northern prawn fishery. Sample collection began in 2001. An advantage of extending the effective population size work into the northern prawn fishery is that it provides species specific information in a mixed species fishery. It is a fishery independent data source and could be very useful at identifying cryptic stocks. Additional life history information could result from estimating N_e and this would then in turn help to reduce model error. Given the expected size of N_e in the northern prawn fishery we could do a cost benefit analysis and simulation to determine the number of samples and loci required to estimate N_e. This information could be used to make judgements about what technologies to use or whether it is even worthwhile to pursue estimating the effective population size in the northern prawn fishery. These simulations are needed before proceeding to analyse northern prawn samples.

Slide 27: Next steps.

Finally, future directions of this project include:

i.   The development of the SNP technology, as it appears to have more potential than microsatellites. Higher levels of automation and their straightforward scoring mean that SNPs markers have substantially less observation error than microsatellites.

ii. Further develop integrated fisheries models and use simulation of different fisheries scenarios to investigate under what conditions N_e may be useful.

iii. Conduct cost benefit simulations in the northern prawn fishery to establish whether it is useful to pursue estimating N_e.

iv. Collect tissue samples opportunistically in candidate fisheries to establish important time series data.

v.   Develop stock discrimination methodology.

Questions

Bartlett:	My understanding of SNPs is that you usually make them from expressed sequence tag (EST) libraries, which generally are not neutral in terms of selection. So wouldn't that then undermine most of the assumptions of your models?
Ovenden:	Yes, you would have to try and develop SNP loci that were neutral to selection. Apart from using ESTs there are many other ways of developing SNPs. You would collect data and then do a test for selective neutrality exactly the same way I did for microsatellite neutrality.
Bartlett:	With multi-age or multi-generational stocks it is easy to go in and use the actual fishermen and collect fish off them than take samples. In real terms it would probably be much easier to take small samples, like fin clips and that kind of thing as opposed to a fishermen giving you say two thousand bits of fish. This technique [N_e] could be useful but if you had the whole fish to age then it will be much more useful for single generational species in the long run.
Hoyle:	Yes that's possibly true. We are already collecting otolith samples from the commercial fisheries so you could collect fish frames and in that case you have both the otolith and the genetic sample. That is the sort of thing I meant by tagging onto existing sampling programs.
Moore:	Picking up on some of the new techniques. You can use telomeres to age fish. Something by just taking a genetic sample, extract the DNA, maybe you don't need the otolith in the future. Take it back to the lab where you can age the fish and do the analysis required.
Donnellan:	I think it would still be worth collecting otoliths from a microchemistry point of view. It is potentially very useful to get high resolution data on a spatial scale using microchemistry.
Dichmont:	Listening to all the talks today it interested me whether it isn't worth discussing going into a prawn fishery, as much as I would love you to go into the NPF, would perhaps not be the right way to go. It seems that, unless I have misunderstood the talks today, one of the things that seems to really kill this method is large population sizes. We've got estimates here that are going to be hundreds of thousands of animals within the northern prawn fishery. The Queensland and NPF initiative wanted to show some kind of benefit to both Queensland and NPF, it's very understandable. I can't help thinking that if you just said well let's target initially populations that are pretty small, that have longer generation time and also have very good data. I can't help thinking that straight away you start looking at the large pelagics, southern blue fin tuna. I bet you that they are probably the most [valuable] animals in the southern hemisphere. That is, a fishery that is very well studied, they have individual tags and they most likely have these kinds of loci genetically. I was wondering if it's worth having that kind of discussion today.
Hoyle:	If you could age them well?
Dichmont:	Yes I am sure they can. Maybe someone from FRDC who has read all of the reports can tell us. The other one I thought of was the mackerel in the Northern Territory where the Genetag work is being done which means they are already taking genetic samples as well.
Hoyle:	That's right, we have definitely been thinking about that. Another one is the Fitzroy barramundi where there is a lot of environmental variation from year to year and N_e could be very useful to look at that, especially because we already have samples with ages going back a fair way.
Dichmont:	Doubtless said that if you have a stock assessment relationship that is quite noisy it's very hard to get that.
Hoyle:	If the noise is driven by an environmental variable then that relationship between N_e and that variable may be very informative.
Adcock:	What's the need to measure N_e's of 40000?
Hoyle:	How do you mean what's the need?
Adcock:	Well it would seem that they are pretty happy populations. Many methods can detect whether it falls below some measure and therefore you don't need to identify 500 SNPs.

Hoyle:	If the method can tell you when it is falling below a certain amount and if you have the power then that is very useful. Also, if you can say this fishery is of a certain size, even if it's 40000, then that's really useful information to help you tune your fishery model. I guess that's what I am talking about here. You always want to know whereabouts you are; even if you are really doing well.
Adcock:	I imaging looking at krill or something where your N_e is 500000 in one year and 700000 in another year is that really worth pushing the method to find that out.
Hoyle:	Well it's a cost benefit thing always. It's probably never going to be worth it for krill. It's always useful to know if you've got 700000 and your management technique is one where you always leave 200000 then you can say OK that year we can take 500000 out, whereas last year we could only take 350000 out. I'm exaggerating the kind of precision you can get but any information that will improve the precision of your stock assessment model is useful because it will give you more precise management information, and when you have more precise management information you can push the fishery harder and get higher yields.
Ovenden:	What Greg [Adcock] is saying is right - if you have a large effective population size then the population is going to perpetuate itself into the future with a high level of genetic diversity to meet environmental challenges etc. But what Simon [Hoyle] is saying is also right because the stock assessment process needs this data, or it would be highly beneficial if it had this data, to validate the models, which it is producing. You are both right on that. As I said earlier this project is providing a new data source for stock assessment methods and if we are going to target that then we are going to have to reliably estimate these large effective population sizes.
Ye:	I have two questions for you. First, including N_e doubled the number of parameters in your stock recruitment relationship. I think this would make model fitting much more difficult because of parameter confounding and process errors which you have already talked about a bit. Have you ever thought about the difficulties of model fitting?
Hoyle:	I think you are right. When you have six parameters to estimate instead of three you are going to need a lot more data to fit the model but then currently we don't have very reliable estimates of stock or recruitment and we don't have very reliable stock recruitment models in a lot of fisheries. How much data do you need? I don't know. It depends on the fishery and it depends on a lot of things. I would like to be able to simulate it to look at that issue in detail but I find it hard to say exactly what is going to happen without actually doing that.
Ye:	My second question. It seems to me quite uncertain at this stage about the relationship between N_e and N_a. Maybe you can do some integrated simulation, like combining the fishery harvesting part with what David [Peel] presented in the morning.
Hoyle:	That's the way you would go about it. If you had a year with a high N_e for example you would have more uncertainty in the N_e estimate than a year with a low N_e, when you would have a more precise estimate. You could carry out a simulation where you said in this simulation we are taking 500 samples each year and estimating N_e with this amount of precision. I think that's a good approach.

General discussion

Basford: One of the things that has been raised quite clearly throughout the day is the precision in which N_e is estimated. Is it sufficient precision to make it worthwhile in subsequent models?

Nelson: I came along today mainly to learn more about what was going on and to understand more about how genetic methods to estimate effective population size. I hear is that industry tends to be a little disappointed at the costs that are associated with fisheries management. It partly might be us. No one really likes paying for that sort of stuff. It could potentially increase the costs of doing stock assessment. I'm hearing that we expect that we would still need to do fisheries independent surveys, we'd still have to do some of the other surveys and this is something that we would do and it would add additional information and could help us better understand the biology and variation in the stocks. From a purely management perspective it could be argued that while those questions might be nice to answer, do we really need to answer them? Industry may argue the same point. Could it just be an additional layer of cost on our current practices?

Ovenden: Perhaps I could say a few words and then maybe Simon [Hoyle] could say a few words. Simon [Hoyle] has been able to integrate effective population size into this genetic stock recruitment relationship and by doing so has generated another four parameters that need estimation. I assume that that's the expensive extra data that you are referring to.

Hoyle: This kind of data isn't going to be useful immediately. You do need a time series before it starts to pay benefits, to get the benefits of reducing the uncertainty. Reducing the uncertainty does have benefits for industry because you can fish harder and have good sustainability, whereas, if you have got less certainty you can't fish as hard for the same level of sustainability. So reducing uncertainty has monetary benefits for industry.

Nelson: I'm not sure though. Some where in the discussion about the N_e term and how it links back to the spawning numbers and then their recruitment into the population there is that really large jump towards either of those directions that are going to be hard to overcome. Do you think we can do something about that?

Hoyle: It's hard to know. In some fisheries it is and in other fisheries it won't be useful. I think we do need to carry out simulations to look at what kind of observation error and what kind of precision is needed before it's going to be useful. I think that's an important step to take.

Ovenden: I understand that there is existing data for a wide range of fisheries and using the modelling approach that Simon [Hoyle] outlined, that existing data could be put into that model and for each fishery we could answer the question how useful is N_e going to be for this one, or this one, or this one. Then we can narrow down our efforts to those particular fisheries.

Dichmont: One of the things we can do is take Simon's [Hoyle] point further. There has been quite a lot of simulation work done on fisheries independent surveys. In a way, thinking about adding independent data sources is just one. We've talked a lot about stock recruitment relationships but not everybody in the world tries to estimate stock recruitment relationships and they manage their fisheries quite adequately without them. Their simulations have shown for example you can use survey data sets as an index of abundance. Even precisions of 50% CV start helping over and above your CPUE data. There is an opposite of that. Survey data has been shown to complicate matters. It particularly complicates matters when it goes in the opposite directions to your CPUE data. It tends to pose more questions than it really answers. So the reality is that there isn't a simple way and it probably will be another layer on. Always, when you talk about additional data, you are talking about doing it relative to your risk and your reward. For some fisheries that is always extremely justified. The perfect example, I believe, is the northern prawn fishery that has just spent $600000 on two surveys. The industry think that they weren't very justified. That's a lot of money but this is also a 200 million dollar fishery. The third thing, which I think we have to remember, is that it is a new technique and what we are trying to do is align it with a fishery that already has a lot of

information so as to test the technique. It would be very hard to jump this technique into a place where we know nothing because we are not going to know if we are doing it right or wrong. So at the moment you will tend to see that the technique is aligned with fisheries that have good knowledge about them. The ultimate, if you are going to talk about the one error of fisheries data would be the ones where you go in and you don't have good catch statistics and you don't have CPUE. I was talking to Jenny [Ovenden] about a classic one which we have worked on together, the red snapper fishery up north, which is a shared stock between Indonesia and Australia. It would be very hard to do surveys between Indonesia and us. To try and get our boats to go there would be very hard. They do not have a log book system to speak of so, we have estimates of their red snapper catches of anywhere between 2000 and 16000 tonnes per year. This means that any stock assessment is a little pie in the sky but what we have managed to do is collect genetic samples in Indonesia. There would be once place that you wouldn't immediately want to start the technique but you would think maybe this is one place where in ten years time it would be very beneficial.

Basford: We are particularly interested in other comments from people who are not so close to the project but have been here today listening to what's going on.

Nelson: Would another year benefit surveys in Moreton Bay?

Ovenden: Yes, we have already the 2002 census data. If we were able to collect another genetic sample from Moreton Bay say in the spawning season this year in 2003 we wouldn't necessarily have to go to the expense of doing it in a fisheries independent way. We could sample the commercial catch for example. We would then have a second and independent chance of measuring the effective size versus the census size. It would be particularly interesting to see if it came up with the same orders of magnitude. Maybe we could look at the same time at environmental influences on the bay in that intervening period to try and explain the difference.

D. Peel: Also we would have two generations between the two samples. So we would have more power in the estimate even though it is an estimate over the last two years it would give us a much tighter [confidence] interval.

Ashby: Which do you think would be more beneficial, survey or simulation?

Ovenden: It is very important to continue on the simulation work. However doing another batch of genotyping would also be beneficial. I would just hate to say one over the other. The benefit for the genotyping is that you wouldn't have to pay for the expensive charter of private vessels to do the survey. We could buy the samples or indeed ask the fishermen to donate the samples. We could just go to the wharf and pull legs of prawns if we had to, we would not actually have to take the whole product.

Hoyle: There is the option of having the samples and storing them until say three years time so you have three years of extra data to analyse all at one time. So it's a more expensive way [than simulation] to do that but would probably be cheaper [than fisheries independent surveys].

Broderick: There are no data sets out there with time slices where we can validate the particular models that you are putting on top of it. Both things need to move forward in concert I believe.

Basford: I think realistically people won't believe it unless they actually see it happen to their data. No matter how much you can justify things with simulation, unless you believe simulation models, people want to see it against real data. They want to know what happens in practice.

Salini: That's what I was getting at back there earlier on about the controlled experiment. I don't know whether it's something someone wants to do a PhD on. It sounds like a fantastic project to me if you want a guaranteed result where you can reduce the unknown factors and assumptions. But by the same token that's probably how industry views it too. They want to be convinced that it will deliver. It's my own personal bias that controlled experiments are the way to go, as you can point to exactly what happened. Other than doing it another way I think that's the barrier at the moment.

Broderick: I understand how that would be a really fantastic educational tool for the method. But if most of the sources of errors are due to wild populations with different variances, then maybe the controlled experiment won't work. We know probably from simulation and genetic theory that there is a

	relationship between sample size and variance of alleles from generation to generation, we know that and we can predict that within a certain [degree] of certainty. What we don't know is how to design our fisheries experiments to collect this sort of data, how much do we need and under a whole gamut of environmental and fisheries interactions is this going to work?
Ashby:	I guess following on from that point where does the extra information stop? The number of effective spawners will vary continually due to temperature, due to currents, etc. Every time we do the survey are we also going to have to get that environmental information and plug that into the model, no?
Hoyle:	You don't have to. It may help but we won't have it in most cases.
Ovenden:	This may be resolved when more data is accumulated to allow us to try and investigate this idea of process error. Once that has been evaluated then the field can go forward having that information already. I'm not sure how transferable the measurements of process error will be from species to species. Currently we are geared up to do tiger prawns. To switch to another species involves investment in development of new genetic markers for other species. Nowadays, there are readily available off the shelf markers for southern blue fin tuna, I am sure that variable microsatellite loci already exist. We know they do for Spanish mackerel so it may not be such a big issue. In my opinion a major issue, which we have yet to address, is this issue of migration and the connectiveness between populations. Here I am extending upon what Simon [Hoyle] was alluding to in his talk that temporal measurements of allele frequencies give us a lot more power to distinguish substocks when migration is occurring at fairly high levels. Do people think that it would be important to look at this empirically? To look at the effect of migration on effective population size estimates?
Nelson:	Is that something that may be better to do in a controlled way? For example, if you have populations that you are keeping separated and then introduce individuals into them. If you take that sort of thing to the field, particularly at the sort of scale you are talking about, and without having done all the mapping of the populations' prior, it could be tricky.
Basford:	Maybe there is too much uncertainty at this stage to be able to really know what the impact of migration is going to be. Do we know enough about the parameter estimates?
Dichmont:	Wouldn't the problem be here that you are going for a limited fishery with a really large number? Judging from the simulations, you are going to get estimates of infinity and the effects of migration on infinity is infinity.
Hoyle:	Yes that's right. The method we were talking about relies on being able to compare the genetic drift in two populations and if you can't measure the genetic drift then you can't use it. So it wouldn't be useful if the populations were that large [as the NPF].
Donnellan:	You might have to do it experimentally too if you have such large effective population sizes. That's a huge number of migrants to make them effective migrants.
Hoyle:	It may have to wait until we have SNPs or very large numbers of microsatellite loci.
Broderick:	I think in response to one of David's [Peel] ideas earlier about how do we improve on these things, one is more loci and the other is more samples. Just sample a greater proportion of your population. With the current microsatellites we seem to be around 10000, if the true N_e is anything around 10000 then it is going to be acting as though it is infinite.
D. Peel:	Probably slightly less than that actually.
Broderick:	Even a bit less than that. But with a sample size of lets say instead of 600, if we go for 5000 or something like that, with the current markers we've got now we should be able to get estimates of 40000.
D. Peel:	No, I don't think it is that high. Its definitely higher but it wasn't as high as 40000. That depends on what level of power you are willing to go to. What chance do you want to accept that you won't get an infinite answer? If you want a 100% or 99% chance that you won't get an infinite answer then you'd have to go a lot larger [sample size].
Broderick:	I guess that's one of the things that excites me about SNPs. Here you have a capacity to increase your sampling by an order of magnitude for the same price. By doing so you are also increasing your power, not by an order of

	magnitude but pretty close to it. At the moment we seem to think that we are quite restricted by the kind of fishery, we have to be very careful about the fisheries we choose to use this method on. If the N_e is too high then we are not going to get there. Down the line the genetic applications will improve and that will open it up to a lot more fisheries than what is currently available to us.
Ovenden:	Implementing SNPs does involve the routine lab work to try and find the markers and get them working etcetera, but its no more work than say developing a new microsatellite library for another species. The other advantage of SNPs is that screening is incredibly quick and much cheaper and probably more reliable, less prone to user error than microsatellites. The microsatellites for the prawns are excellent loci, highly polymorphic but they have been problematic to implement for the Moreton Bay population.
Donnellan:	I think the real advantage of SNPs is the high level of automation. Getting the operator out of the system is much easier for SNPs than microsatellites.
Ovenden:	We are on the verge of that with SNPs. SNPs claim that they are able to determine the frequency of alleles in a pool of 100 or 300 individuals. Essentially that means that in a single test tube you have the DNA of 300 prawns and you can instantaneously determine the allele frequencies of say 133 alleles at once in a single reaction, which amazingly increases the speed at which you can do the analysis. This is a claim made by the manufacturer of the apparatus that has just been installed at UQ. Everybody is highly sceptical about it but apparently it's been tested and tried overseas and it definitely works. With those kinds of advances in genetic technology, you achieve orders of magnitude improvement on throughput and cost.
Nelson:	At that point maybe you could collect material from processors and just run it on that.
Donnellan:	You would still need to know how many individuals and equalise the amount of DNA going into it.
Ovenden:	Yes, you have to set the pools up in the lab. These samples can be taken, not non-invasively or non-lethally because the animals are already dead, but when I said taking individuals legs I was not joking. We can do DNA analysis from individual legs. The fishermen, the processors don't actually loose any product because of our method. What we have to invest in is some technical time to recover the legs or whatever off the prawns. There are frames that are thrown away from fish processors that are also a good source of DNA for projects like this.
Street:	The only down side of that is you don't have any other information, you don't know sex, weight and length. It is just a quick and easy method of getting a heap of samples cheaply.
Ovenden:	With this allele pooling method you end up with allele frequencies that relate to a population sample. What we have been collecting from the Moreton Bay tiger prawns is individual genotypes. That data can be useful for the linkage disequilibrium method. The allele frequency data we would be collecting from the pooling method would only be able to be analysed with the temporal methods. That would be one of the problems with it. Do you know if New Zealand fisheries geneticists are pursuing this question of effective population sizes?
Adcock:	Peter [Smith] is trying to get funding for that. It would be especially valuable to take samples from smaller spatial scales from around Tasman Bay to see what heterogeneity there is and see what some of the effects are of sampling related, seasonality, and the basic biology of snapper. Is there any natal homing going on? Are the females going back to the same grounds every year? Do the bigger ones knock off the smaller ones to secure the best locations? Those sorts of things have quite an effect on effective population size, I think. It's not just a matter of variation in fecundity it's behavioural variation as well that is quite significant. Do prawns do anything like that?
Ovenden:	Probably do, we just don't know about it.
Adcock:	While I am talking, you asked before about impressions of whether this research should go on or whether the simulations should continue. From an outsider with a broad interest in conserving marine species in general it

	would seem vital that the work continues. You can't have one without the other.
Ovenden:	You mean the simulation and the empirical work.
Adcock:	Yes absolutely, because they are both incredibly instructive and will inspire other people to do similar work as well. I know that funding agencies are not necessarily altruistic like that but then again if this work is not done or it is delayed then some species within Queensland fisheries or the same area that the prawn fisheries are in, might miss the opportunity of having this sort of study done on it that will stop collapse in some way if the monitoring improves. I think it is very important that this thing gets going if the funding is there.
Basford:	So the question would be, if you continue the work, what should the emphasis be on? Should the emphasis be on continuing with Moreton Bay? Should it be on looking at the other stock in the northern region, which has originally been suggested, or should it be changing to something that has a much smaller population size? What are peoples' views?
Adcock:	You can't stay with one species forever that's going to model all situations. The Moreton Bay stuff has momentum going so it has to be completed to a good degree so that it's some sort of starting point for other studies. I don't know how you mix your emphasis and apportion your funds for those different [goals].
Ovenden:	A compromise may be to also include ground breaking work on a species that has overlapping generations. We have tiger prawns that just have an annual crop and that is by far the entry level if you wanted to estimate effective population size. If you are going to deal with a long lived species like mackerel or tuna that presents another set of both theoretical and empirical challenges. I think it would be very good to make a start on it right now so that we can at least evaluate whether effective population size is going to be useful on those kinds of species given that only a very small proportion of our fisheries resources have non-overlapping generations.
Adcock:	It does seem to me that a lot of those fisheries that you talked about are further south. It seems like something where you would need a broader scale effort in some way. It's not just someone sending fish. It's a bit more than that.
Dichmont:	I know its good to choose fisheries independently but at the same time it would be nice to choose research that goes back and benefits research agencies. Although southern blue fin tuna is very attractive it may be just as attractive to look at Spanish mackerel.
Hoyle:	The population may be too large. It's quite a large population.
Adcock:	Population genetics might help.
Hoyle:	We've got a fair bit of population genetics.
Adcock:	Some kind of historical samples would be useful. Any stored samples that aren't formalin preserved.
Ovenden:	We've been working on Spanish mackerel since 1998 when we sampled six spatially distributed samples through northern Australia. We have samples from 1998, 1999 and 2000 in our lab. Perhaps we should replicate some of that sampling for that species for some of the populations that we think are spatially confined. Perhaps the Northern Territory population and perhaps the Torres Strait population could take advantage of those historical samples that we already have.
Dichmont:	That GeneTag project that Rik Buckworth is working on, the very valuable Spanish mackerel. There is a lot of technology already in this fishery that would value add.
Hoyle:	The samples are coming in.
Dichmont:	The samples are coming in, the method of collecting the genetics [samples] that wouldn't kill the fish and the genetics is almost resolved.
Ovenden:	Crispian [Ashby] made that point earlier on that it would be good to make multiple use of the data that was being collected for gene tagging. Not only for making estimates of harvest rates that we plan to do but perhaps for laying the groundwork for effective population size estimates also.
Hoyle:	There is a problem with the Spanish mackerel in that we are not killing the fish so we do not have otoliths so we can't age them. We will be getting fin samples and there might be some information in the fin rays. There might

	be some information in the telomeres that Andy [Moore] was talking about earlier, that's quite speculative so there's a problem.
Basford:	So what about with any N_e comparison with the prawns in the gulf area.
Ovenden:	The mackerel population that is being studied under this Genetag program is adjacent to Darwin. It's actually physically separate from the Gulf.
Basford:	I was meaning choosing to go with that species rather than to look at the overlapping time [generations] of the prawns in the Gulf of Carpentaria area which is what I thought was originally the way you were going to expand that. I guess I am really asking the group, the pros and cons of going either way. Going to say the mackerel versus going to the overlapping time frame of prawns in the Gulf of Carpentaria. You were talking about the size of the population. I think you said before that the size of that population in the Gulf is very large, I mean which one is bigger? Are they similar sorts of orders?
Hoyle:	They are bigger. Then you've got this issue that it's not the total population size that's the issue it's the N_e that's the issue. We don't know what the N_e/N_a ratio is. Based on previous examples they can range a lot between two orders of magnitude and six orders of magnitude. It's pretty uncertain.
D. Peel:	It would be more for the prawns than for the mackerel, wouldn't it?
Hoyle:	Possibly yes. Though do you [Jenny Ovenden] have that list of ratios [N_e/N], they looked good for the fish.
Ovenden:	Greg's [Adcock] work with the snapper indicates five orders of magnitude. Work with the red drum, which is essentially a snapper species in the Gulf of Mexico, is three orders of magnitude. Our work with prawns also seems to be three orders of magnitude. There doesn't seem to be any consistency there between invertebrates and vertebrates or crustaceans versus non-crustaceans. This issue of the order of magnitude between N_e and N_a I think is very important. If we had less error on that then N_e could be used more efficiently in the genetic stock recruitment model we proposed.
Hoyle:	One of the things is that we probably don't know that ratio very well. If N_e is too big we can't measure it, so those ratio estimates are always going to be biased downwards since high N_e/N_a ratios are less likely to be estimable than low N_e/N_a ratios.
Ovenden:	Well that of course presupposes that you can estimate N_e in a non-infinite way. I think Kaye Basford's way of evaluating the pros and cons. The advantages of going with a mackerel are that we already have the genetic markers developed. We have processes in place to sample the naturally occurring populations. Ultimately as a result of this harvest rate work we will have mark recapture estimates of population size as well. Instead of having census estimates of tiger prawns that we have got here in the Bay through Tony's [Courtney] work, with the mackerel work we will have mark recapture estimates to compare to the effective population sizes.
Hoyle:	But the lack of ages with the fish almost rules out using effective population size methods.
Ovenden:	Well there may be a way around that. If there was a spread in the size of the fish that were landed we could maybe choose small fish versus large fish and do a comparison and make some assumptions about the age differences.
Hoyle:	We would be working on fins so that we don't have information on the size of the particular fish.
Ovenden:	Yes but the size of the fin is proportional to the size of the fish isn't it?
Hoyle:	I don't think so. The growth curve for the Spanish mackerel tends to level off a bit.
Adcock:	Difference between males and females?
Dichmont:	The worst case scenario would be like a shark fishery. What makes a big shark?
Ovenden:	Even so, we could take a random sample of the fins in the first year of the gene tagging project and compare that to a random sample taken in the last year of the GeneTag project, approximately estimate how many generations have elapsed and make your temporal estimates on that.
S. Peel:	I just want to clarify, doesn't the method look at the variation given the genetics in the population. If you take one sample and then five years later take another sample, why is the age of each individual important if you are just looking at the difference in the genetics between the two populations?

D. Peel:	It's because of the overlapping generations. It's going to get muddled up because you have multiple different ages happening.
Courtney:	If you want to estimate N_e and you also want to estimate spawning stock size for the mackerels then spawning stock size for the mackerels isn't necessarily specific to the year classes. When you do stock assessment and come up with an estimate of spawning biomass and say OK the spawning biomass is 50 tonnes. You don't know really whether 30% of that is two year olds, 40% is for five year olds etcetera, etcetera. When you collect your fin clips you don't know what age they are and for traditional stock assessment models won't know what the relative age of the spawning biomass is.
Donnellan:	Could you just discard those fish that are too small for sexual maturity?
Broderick:	If your fishery has a proportion of fish that are juvenile and they can be identified by size, they are only going to be juveniles once so you could come back five years later say and sample the juveniles again only his time they will belong to a separate generation of spawners.
Adcock:	That's not exactly effective population size is it?
Hoyle:	The effective population size is all the fish of all the ages that are breeding and then from those numbers you get the variation in number of recruits.
Adcock:	Consideration over a period of time is lifetime reproductive success not what happens in a particular year. Sometimes you can have complete recruitment failure therefore N_e is zero but it doesn't mean that much towards the fishery perhaps because the previous years recruits are getting bigger and the fishery never catches one year of fish.
Hoyle:	No but what we are interested in, in this context, is the recruitment in a particular year. That's what we are looking at, the spawning in one year. The effective population size is just the number of fish that are breeding in that year and then recruitment in that year.
Dichmont:	To me it sounds as though you are suggesting that maybe mackerel won't be a good example. There must be examples of overlapping generation resources in Australia that you do have age information for that you can also effectively sample. To me it looks almost like what you should be doing is saying OK these are the criteria we like to work with, and which fishery has those and conforms mostly to them rather than compromise results from the beginning. Personally, if I wrote a stock assessment model of mackerel I certainly would want to know my age classes. I do think they are actually important because in long lived species if you do have some kind of failure of catch in one year you want to know is that because you have a single age class that hasn't come through or is it because of the general problem in that you are looking at several years of bad catches. That's fundamental to the assessment of that animal, whereas, prawns you really are only working with stock because that stock certainly isn't going to be alive in the next year, certainly not in the northern prawn fishery.
Ovenden:	I don't think I entirely agree with you that the implementation of N_e on mackerel is not possible. I do think it is possible with the samples we are taking with the GeneTag project. We have talked about it in many ways here but I am pretty sure that we can do it. The great advantage with the mackerel project is that it is underway already and that for a small amount of extra expenditure we can maximise the returns from that project by adding effective population size estimates to that project. I also understand what you say about picking and choosing amongst the other Australian fisheries resources to find species that are ideal in a whole lot of other ways. Nothing immediately springs to mind. The barramundi might be a good one, perhaps. To me it's important to have active stock assessment occurring on that species at the same time. It's no good going and doing it on turtles or dugongs which aren't subject to stock assessment processes because the outputs of effective population size work has to be integrated in the same time frame with the stock assessment work.
Dichmont:	Sure, you actually want more than that. You want stock assessment and survey data. You want stock assessment that already has an independent data source. It really looks like the funding agency and industry are saying what is the worth of these extra data sets. There is a perception that they are competing, well not competing but comparing surveys versus genetic

	techniques although focus on one or both etcetera. In that regard, and I hate to say it, that's probably going to be in the very valuable fisheries. What are those commonwealth fisheries or in WA?
Broderick:	I just remembered something about the Spanish mackerel fishery; a proportion of it is frozen fillet. So at the moment we are choosing to take fin clips just because it is easy and it suits our purpose. If we wanted to slightly modify our sampling methods we could access that fillet market and then we would have frames with otoliths and then we would have age data. The whole sampling design has been set up for a different purpose but it is possible to modify the collection strategy and get the information we need.
Hoyle:	Just to confuse things a bit more. On the east coast with the Spanish mackerel fishery there is an age structure monitoring program. We could collect samples from there. It may just be adding another complication, as I don't know if the numbers are enough. Some of the numbers are enough and certainly the population size is higher than in the Northern Territory then it's going to affect the quality on the east coast.
Courtney:	Can I just ask Greg [Adcock] to elaborate a little bit on the New Zealand snapper? The relationship between N_e and spawning biomass there, can you elaborate on your estimate of spawning biomass? Did that just come out of a stock assessment model that said OK in those two bays the spawning biomass in 1999 is 1000 tonnes and then from the sampling program came up with the estimate of N_e for animals in the same year?
Adcock:	We estimated numbers of spawning stock size for 50 years and we have re-evaluated some of those with new models. So there are some retrospective estimates as well. It's all taken from fisheries data. It would be an ideal one for testing on, I would have thought.
Courtney:	Your estimate of spawning biomass, that would have been based on the aggregative year classes that contribute to the spawning stock?
Adcock:	Over a period of time concerned it would be the harmonic mean.
Ovenden:	I think what you are getting at is the nature of the ratio that you derived is slightly different from the nature of the ratio that we derived. We did ecological census techniques to work out the number of spawners by direct measurement. Whereas, the spawner numbers you obtained would have been from fisheries stock assessment models.
Adcock:	As far as I am aware, this is where Peter Smith would have been useful, there are two things going on there. There is the monitoring of the fishery with catch numbers with CPUE and their ongoing scientific research program. I don't think that both are done every year.
Ovenden:	These kinds of ratios have been criticised in the literature because if the ratio is calculated over a large number of generations then the effective population size estimates you obtain are harmonic means of all the intervening N_e estimates. That's just the nature of what they are. Whereas, the average of the spawning numbers are actually arithmetic means and that biases the ratio that you obtain. The methodology we adopted for tiger prawns circumvents all of that. It just has to be factored in when you talk about these ratios. It's very unsatisfactory.
Adcock:	The take home message was meant to be that N_e is very low and something has happened to diversity as a result of fishing. Obviously there is a lot more to look at and the estimates of N_e have improved. Certainly our confidence is not at a level that you are looking at trying to use N_e to extrapolate into other areas. I think the potential is there to develop that in New Zealand if they provided the money. I did actually hunt around for something comparable to NZ snapper in Australia. It does seem as though the monitoring isn't so great, the fishery is not as intensively studied because it's not as an intensive fishery. Also the snapper have a much higher rate of dispersal.
Ovenden:	Than in Australia, so there is no spatial structure?
Adcock:	No good ones, there is an isolation by distance thing happening possibly.
Donnellan:	Shark Bay has got a snapper fishery.
Adcock:	Yes Shark Bay is the only one, I had forgotten about that one.
Donnellan:	It has got at least two substocks in there, maybe three.
Adcock:	That could cause complications.

Donnellan:	Unfortunately they did have material that might help to establish the stock structure from the early eighties but they threw it out about three years ago when they shifted buildings. That's an issue of archiving material.
Basford:	Are there any other issues people wanted to raise, anything we have missed?
Donnellan:	Just something that you brought up earlier Jenny [Ovenden] about knowing where the archived sources of materials are. It would be easy to get some sort of fisheries associated web site that would lead you into the collection but how you go about ferreting them out in the first place?
Adcock:	Aren't the museums collaborating on that? Though it's not fisheries specific.
Donnellan:	Not very well.
Adcock:	Museum people I speak to speak highly of it, maybe they were trying to impress me.
Donnellan:	The herbariums have done it very well but the museums have not done it very well. There is a conventional pilot fish collection project being developed but that's unrelated to genetics samples, more for conventional taxonomy.
Donnellan:	It wouldn't be very difficult to [access]. I guess the other thing is that the materials we are talking about are not already archived. They are sitting in someone's research area and they are in danger of being thrown out right now.
Adcock:	That's the thing. We need legislation.
Donnellan:	It's not legislation. We need some resources to have the material properly catalogued in the first place so that it can be stored somewhere else longer term.
Adcock:	What I mean by legislation is that the funding bodies at some point early on, say when you collect samples, you are required to do X with them at the end.
Donnellan:	And give you $5000 for 10000 samples to do it, something like that.
Ovenden:	It is costly, especially if they have to be maintained in -80°C freezers.
Adcock:	Hopefully we have gone past [the use] of allozymes.
Donnellan:	This is not so much the issue. It's getting it out of the research lab, where it is in danger of being lost, and put into a centralised archiving facility.
Basford:	Just the data themselves.
Ovenden:	The data and the samples.
Donnellan:	Some of that could be circumvented too by having the sorts of measures Greg [Adcock] is talking about putting into place where the archiving facility supply the databases right at the start of the project. So that it is all in one sort of database software or some sort of database format that can be easily acquired into a centralised database.
Ashby:	Is that done anywhere? In the Smithsonian or one of those highly organised well funded places? It would be nice to have a model.
Donnellan:	I suppose it wouldn't be very hard to put together a model
Broderick:	The [Museum of Vertebrate Zoology] at Berkeley is not so bad at that. They have got the whole collection since the 19th century all catalogued and it is accessible by the web. As I understand it that is one of its strong points, archiving of material, tissue bank and accessibility. It's one thing to have the stuff there. It's another to know where to look and then be able to find it.
Adcock:	It's a big issue. You can't predict what fishery or what region is going to be necessarily fished. Although, anything that is commercially viable is being fished at the moment. You want to put these things away for the future.
Ovenden:	Especially when you recognise the value of those samples. Once they get into the -80°C freezers in your lab, in terms of the fieldwork, the personnel the sampling costs. By the time they hit your -80°C freezer they are worth an enormous amount.
Donnellan:	The actual cost of maintaining them is trivial.
Ovenden:	That's right.
Donnellan:	Especially by the time you have half a million specimens in the facility. Is there something sensible we can start about this? You guys are actually in the fisheries sector, I'm not. So, you probably have more of an idea about where these resources are and what sort of resources you would need to maintain them at all.

Hoyle:	We would need a professional archivist to work out all the issues.
Ovenden:	I feel that it has to be handled through a museum.
Donnellan:	Well that's their role in government, to maintain archives. So it's the perfect place to have it.
Ovenden:	I could bring up the issue with the Queensland Museum. Your South Australian Museum maintains archives of mammal tissues don't you?
Donnellan:	No we have tissues from 80000 individual animals.
Ovenden:	All different taxa?
Donnellan:	Yes, from invertebrates, vertebrates, plants and bacteria. We just got an ARC linkage infrastructure grant to take on this archiving work. It's in the early phases at the moment but, if it takes off, additional funding could be found.
Ovenden:	The use of temporal methods really emphasises the importance of archiving samples. This could be another rationale to proceed with your archiving work, not that you need one.
Donnellan:	We already have quite a large number of fishery stock samples already archived in the system from local projects. It wouldn't be difficult to do this on a wider basis.
Basford:	Getting back to the project area itself are there any other points people would like to bring up in the general discussion.
Ovenden:	Can I just quickly mention Albatross Bay in the Northern Prawn Fishery which was proposed a while back as being an isolated population of tiger prawns.
Dichmont:	It would be a very good population to do but the problem is there is not a lot of *P.esculentus* there. It would be good from the point of view of being a control for the Moreton Bay stuff that's very appropriate because at the moment Albatross Bay has had some trouble for quite a few years in a row. There is a large question as to what is actually happening there. It's mostly a banana prawn issue but it's also happening with tiger prawns. The species down there is actually *P.semisulcatus.* It is highly likely, from a time point of view, an isolated stock compared to its nearest neighbour, which is down in Mornington.
Ovenden:	Albatross bay is the bay adjacent to Weipa, which is on the far north eastern side of the Gulf and is spatially very separate from the rest of the tiger prawn fishery.
Dichmont:	Spatially very separate unless animals march across the centre of the Gulf it is spatially very, very separate. Although it is very close to Torres Strait it is a very isolated stock. Also much smaller numbers than the rest of the Gulf.
Ovenden:	What proportion of the catch would be *P.esculentus* and how would we go about getting sample sizes?
Salini:	Minor, minor size. Lucky if it's 10%.
Ovenden:	Still large numbers though.
Dichmont:	But then Albatross Bay total catch of prawns in terms of biomass is a much smaller fishery all together. If you are looking at 2-3000 tonnes of prawns coming out of the Gulf, excluding bananas of course, they are not nearly as large.
Donnellan:	You've got a couple of thousand specimens.
Adcock:	Small N_e's and densities.
Dichmont:	[It's] probably about a 300 tonne fishery. The other thing is that we have a project that we're sampling 4 times this year and four times next year in Albatross Bay.
Ovenden:	Taking fisheries independent surveys?
Dichmont:	Taking fisheries independent surveys and some of the samples you have in your own freezer come from that area.
Ovenden:	You have a stock assessment model, probably a mixed species model for that area?
Dichmont:	If you ask me for a stock assessment model on *P.esculentus* in Albatross Bay I will struggle because its numbers are so small.
Ovenden:	So the model more accurately applies to *P.semisulcatus*?
Dichmont:	Yes. For *P.semisulcatus* I would be able to do it for sure.
Donnellan:	Is *P.semisulcatus* isolated too?
Dichmont:	Yes. The whole area is very isolated.
Donnellan:	Could you transfer markers?

Ovenden:	Sometimes microsatellites transfer very well from species to species. We have tested our Moreton Bay markers on the Gulf population for browns and they are fine. What we have tried not our markers on is *P.semisulcatus.*
Dichmont:	If you translated it and went to *P.semisulcatus* in Albatross Bay it would be a completely different story. That would be different if you were saying lets go to the Gulf and do *P.esculentus* then you would probably be looking at going to Mornington and the Vanderlin group and looking at overlapping stocks you are looking at a huge area, massive numbers of animals.
Ovenden:	For *P.semisulcatus* the stocks are just as diffuse and overlapping in the Gulf?
Dichmont:	Yes. The areas are literally continuous. There is some isolation but it is not discrete. It's just isolated.
Salini:	Years ago Mulley and Latter reported on the population structure of *P.semisulcatus* within the Gulf of Carpentaria using allozymes.
Dichmont:	Within the Gulf.
Ovenden:	But that was using allozymes, fairly low power.
Donnellan:	Would that reflect how isolated they are?
Ovenden:	That would be part of the study.
Donnellan:	You have the stuff in the freezer?
Dichmont:	That would be extraordinarily useful.
Ovenden:	To assess the degree of genetic separation.
Dichmont:	We at the moment have enough problems trying to get assessments at what levels we should record. It's likely that we will show that we need to do it at stock level. Of course the study that was done doesn't actually know where the stock boundaries are. It's one of those solutions where you say we should be looking at this but we don't know where they are.
Basford:	Are there any other issues before we close the meeting?

References

Anderson, E. and Williamson, E. (2000). MCLEEPS software. For more information, visit the MCLEEPS web page at http://www.stat.washington.edu/thompson/Genepi/Mcleeps.shtml.

Beaumont, M.A. (2002). TM3 software. For more information about TM3, visit Mark Beaumont's software web page at http://www.rubic.rdg.ac.uk/~mab/software.html.

Courtney, A.J. and Masel, J.M. (1997). Spawning stock dynamics of two penaeid prawns, *Metapenaeus bennettae* and *Penaeus esculentus*, in Moreton Bay, Queensland, Australia. *Mar. Ecol. Prog. Ser.* **148**, 37-47.

Frankham, R. (1995). Effective population size/adult population size ratios in wildlife: a review. *Genet. Research Cambridge* **66**, 95-107.

Hauser, L., Adcock, G.J., Smith, P.J., Ramirez, J.H.B. and Carvalho, G.R. (2002). Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper *(Pagrus auratus).* *Proceedings National Academy Science USA* **99**, 11742-11747.

Hilborn, R. and Walters, C. J. (1992). *Quantitative fisheries stock assessment: Choice, dynamics, and uncertainty*. Chapman and Hall, New York, 570 p.

Hill, W.G. (1981). Estimation of effective population size from data on linkage disequilibrium. *Genetical Research* **38**, 209-216.

Joll, L.M. and Penn, J.W. (1990). The application of high-resolution navigation systems to Leslie-DeLury depletion experiments for the measurement of trawl efficiency under open-sea conditions. *Fisheries Research* **9**, 41-55.

Krimbas, C.B. and Tsakas, S. (1971). The genetics of *Dacus oleae*. V. Changes of esterase polymorphism in a natural population following insecticide control: Selection of drift? *Evolution* **15**, 264-276.

Nei, M. and Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics* **98**, 625-640.

Nunney, L. (1995). Measuring the ratio of effective population size to adult numbers using genetic and ecological data. *Evolution* **49**, 389-392.

Pollock, E. (1983). A new method for estimating the effective population size from allele frequency changes. *Genetics* **104**, 531-548.

Pudovkin, A.I., Zaykin, D.V. and Hedgecock, D. (1996). On the potential for estimating the effective number of breeders from heterozygote excess in progeny. *Genetics* **144**, 383-387.

Sparre, P. and Venema, S.C. (1992). *Introduction to tropical fish stock assessment: Part 1 - manual*. Food and Agriculture Organization of the United Nations, Rome, 376p.

Turner, T.F., Wares, J.P. and Gold, J.R. (2002). Genetic effective size in three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* **162**, 1329-1339.

Wang, J. (2003). MLNE software. For more information, visit Jinliang Wang's web page at http://www.zoo.cam.ac.uk/ioz/people/wang.htm.

Waples, R.S. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* **121**, 379-391.

Waples, R.S. (2002). Definition and estimation of effective population size in the conservation of endangered species. In *Population Viability Analysis* (Eds Beissenger, S.R. and McCullough, D.R.). University of Chicago Press, Chicago, 147-168.

Appendix 1

Results of Questionnaire

There were 20 responses to the questionnaire regarding the Forum.

The work I do falls into the following classifications:
a)	Researcher	15%
b)	Fisheries manager	5%
c)	Geneticist	25%
d)	Fisheries stock assessment	15%
e)	Funding body	5%
f)	Researcher/fisheries stock assessment	5%
g)	Researcher/geneticist	5%
h)	Researcher/geneticist/Uni employee	10%
i)	Researcher/Gov. employee/geneticist	5%
j)	Researcher/stock assessment/biologist	5%
k)	Researcher/geneticists/biologist/Uni	5%

Were you provided with opportunities to participate in the discussion during the day?
a)	Yes	100%

Did you think your contributions were valued by the forum convenors?
a)	Yes	70%
b)	No	0%
c)	Sometimes	20%
d)	No response	10%

Did you acquire new knowledge as a result of the forum?
a)	Yes	95%
b)	No	0%
c)	A little	5%

Do you expect to be able to use this knowledge in your work?
a)	Yes	50%
b)	Possibly	50%
c)	No	0%

How do you think the research team (Moreton Bay tiger prawn N_e) has addressed the majority of concepts involved in using N_e for stock assessment?
a)	Adequately	65%
b)	Fully	30%
c)	Inadequately	0%
d)	Comments:	There's always more to do.
		Not (b) because of many unknowns. Many of these were discussed in group discussions at breaks

What are the major questions that remain to be investigated?
- Need more time to combine in an assessment – more simulations
- More on the biases of not having the "perfect" population eg migration, overlapping generations, large populations, error
- Increasing power by combining LD and temporal methods
- More about implementing N_e in stock assessment models
- Precise estimates of N_e
- A controlled study of the methods
- Work out the problems that need to be addressed but not necessarily how to do it
- How will a method that is based on classical assumptions of population genetics handle deviations from these i.e. migration, selection
- Will it work on other species
- Unsure of the benefits of N_e in stock assessment
- Use and opportunity to use SNPs

What do you think was the most important aspect of the workshop?
- Interaction with researchers and sharing of ideas on
- Discussion around the reasons for variability in N_e
- This questionnaire – prize
- Software, discussions and process
- Communication – get people to think and maybe apply
- Discussion with others about N_e estimations and their usefulness
- Getting a better understanding of the simulation models and possible benefits
- N_e estimates
- An airing of all the relevant issues. The pros and mostly the cons
- The need for this type of multi-year approval in fisheries/genetic analysis
- Demonstration of the application to Moreton Bay
- Getting feedback from colleagues with good discussions after every speaker
- N_e estimation problems – simulations
- Snapshot of current state of field
- Provide a potential way to improve stock assessment
- Explaining the methods and its weaknesses
- Simulations
- Provided a forum for discussion on N_e and genetic stock assessment
- Détente (the relaxing or easing of tension between parties)

How would you rate the standard of presentations?
a) Excellent   75%
b) Good   20%
c) Average   0%
d) Bad   0%
e) No response   5%

Based on the information from the forum and your knowledge of fisheries stock assessment, do you believe the field is developed enough to apply genetic methods to tiger prawns in the Gulf of Carpentaria?
a) Yes   20%
b) No   20%
c) Maybe   40%
d) Don't know   10%
e) No response   10%

What other fisheries do you believe could benefit from using genetic methods to estimate N_e or stock assessment?
- Fitzroy barramundi
- Southern blue fin tuna
- Endangered populations
- Small populations eg line fisheries
- Stocks that have a smaller population of effective spawners where smaller samples required
- Large pelagics
- Barramundi, snapper and mackerel
- Abalone
- All harvested and imperilled populations (freshwater populations impacted by river degradation)
- Aggregating spawners – Gemfish
- Species worked over by fishermen outside of Australian national waters
- All species with ambiguous stock perimeters
- Freshwater
- Most

Would you prefer to attend a conference talk on this research or read about it in journals?
a) Conference   65%
b) Journal   20%
c) Both   15%
d) Neither   0%

Any further comments?

- More time was needed in demonstrating the software
- Very well organised and presented – impressed with how the collaborative research partners communicated and worked together
- Well organised, good information and great discussions
- The field of calculating N_e is in its early days and the team has done an excellent job in addressing the complex and many issues involved. Future work and technological advances will improve our ability to generate better estimates of N_e.
- Where were the cheesy puffs, scotch and chocolates?

Appendix 2

Interested delegates unable to attend

John Benzie
University of New South Wales
NSW, AUSTRALIA
j.benzie@unsw.edu.au

Caleb Gardner
TAFI
TAS, AUSTRALIA

Norman Good
Southern Fisheries Centre
Department of Primary Industries, Queensland
QLD, AUSTRALIA

Dennis Hedgecock
Bodega Marine Laboratory
University of California, Davis
USA

Brigid Kerrigan
Queensland Fisheries Service
QLD, AUSTRALIA
brigid.kerrigan@dpi.qld.gov.au

Colin Limpus
Environmental Protection Agency
QLD, AUSTRALIA
col.limpus@epa.qld.gov.au

Neil Loneragan
CSIRO Marine Research
QLD, AUSTRALIA
neil.loneragan@csiro.au

Anne McMillen-Jackson
Florida Fish and Wildlife Conservation
Commission
USA
anne.jackson@fwc.state.fl.us

Catherine Nock
Centre for Animal Conservation Genetics
Southern Cross University
NSW, AUSTRALIA
cnock@scu.edu.au

Heather Podlich
The University of Queensland
QLD, AUSTRALIA

William Sherwin
School of Biological, Earth and Environmental
Sciences
University of New South Wales
NSW, AUSTRALIA

Peter Smith
NIWA
NEW ZEALAND
p.smith@niwa.co.nz

Terry Walker
Offshore Fisheries, Marine and Freshwater
Resources Institute
VIC, AUSTRALIA

Rohan Wallace
Centre for Environmental Management
University of Queensland
QLD, AUSTRALIA

Appendix 3

NeEstimator Software