Molecular Analysis of Genetic Variation and Relationships amongst Abalone (*Haliotis midae*) at the I&J Abalone Hatchery at Danger Point, Gansbaai, R.S.A.

Daneël Lambrechts

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science (AH) at the University of Stellenbosch.

Supervisors: Mr D Brink and Dr ME D'Amato

December 2002

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

Signature: Date:

ABSTRACT

The species Haliotis midae is the only abalone species of commercial importance to the South African fisheries industry. The species is currently under substantial pressure due to commercial harvest and illegal poaching, to the extent that genetic and biodiversity may be under threat. The species is also being cultured in commercial systems for supply to international markets. The South African production for 2002 amounted to 350 tons. The commercial production of abalone may also have implications with regard to genetic diversity of natural and commercial populations.

The aim of this project was to assess the genetic diversity of the natural and commercial populations of abalone at the I&J Abalone Farm, Danger Point, Gansbaai, in South Africa, in order to make recommendation with regard to possible impact on natural diversity as well as internal management systems. Distortion of natural genetic diversity or a loss of internal diversity will have detrimental consequences.

Representative samples have been obtained from the surrounding natural population, as well as the commercial populations, including the broodstock and various progeny groups. Both mtDNA and AFLP molecular techniques were used to assess genetic diversity.

Data analysis showed that the genetic profile of the commercial population display sufficient genetic variation. The genetic structure of the commercial population also displays no significant deviation from that of the surrounding natural population, i.e. the population of origin. The results through give indication of a small, though unsignificant loss of genetic variation from the broodstock to the subsequent progeny groups.

The investigation conclude that the commercial populations of abalone at the I&J Abalone Farm, Danger Point, Gansbaai, in its current format holds no threat to the disruption of the genetic diversity of the surrounding natural population. Further, the commercial population possesses sufficient genetic variation in view of future genetic development. There is signs of a slight loss of genetic variation in the change over from the broodstock to the progeny groups. This would necessitate careful and controlled replacement of the original broodstock with new animals from the commercial progeny groups as part of an ongoing process of domestication.

OPSOMMING

Die spesie *Haliotis midae* is die enigste perlemoen spesie van kommersiële belang vir Suid Afrikaanse visserye. Die spesie verkeer tans onder toenemende druk as gevolg van kommersiële en onwettige vangste, tot so 'n mate dat genetiese en biodiversiteit daardeur geaffekteer kan word. Die spesie word in toenemende mate kunsmatig gekweek vir voorsiening aan internasionale markte. Suid Afrikaanse produksie vir 2002 beloop reeds sowat 350 ton per jaar. Die kunsmatige produksie van perlemoen het ook moontlike implikasies op die genetiese diversiteit van natuurlike en kommersiële populasies van die perlemoen, *Haliotis midae*.

Die doel van die projek was om die aard van genetiese diversiteit van die natuurlike en kommersiële populasies van perlemoen by die I&J Perlemoen Plaas, Danger Point, Gansbaai in Suid Afrika te ondersoek, ten einde aanbevelings te maak ten opsigte van moontlike impak op natuurlike diversiteit asook ten opsigte van interne genetiese bestuurstelsels. Versteuring van natuurlike diversiteit of verlies van interne diversiteit hou nadelige gevolge in.

Verteenwoordigende monsters is versamel vanuit die omliggende natuurlike populasies, sowel as van die kommersiële populasie op die plaas, insluitend die teelmateriaal en verskillende nageslaggroepe. Beide mtDNA en AFLP-molekulêre tegnieke is gebruik tydens die ontleding van genetiese diversiteit.

Dataontleding het aangetoon dat die genetiese profiel van die kommersiële populasie voldoende genetiese variasie demonstreer. Die genetiese struktuur van die kommersiële populasie toon verder geen betekenisvolle verskil met die omliggende natuurlike populasie nie, dit wil sê met die populasie van oorsprong nie. Die resultate toon verder 'n geringe, dog nie betekenisvolle verlies van genetiese variasie wat waargeneem word met oorgang van die teelmateriaal na die onderskeie nageslaggroepe.

Die bevindinge van die ondersoek is dus dat die kommersiële populasie op die I&J Perlemoen Plaas, Danger Point, Gansbaai, in die huidige formaat, geen bedreiging inhou vir die versteuring van die omliggende natuurlike genetiese diversiteit nie. Verdere bevindinge is dat die kommersiële populasie oor voldoende genetiese diversiteit beskik met die oog op toekomstige genetiese ontwikkeling. Daar is tekens van 'n geringe afname in genetiese diversiteit met die oorgang vanaf die teelmateriaal na die nageslaggroepe. Dit beklemtoon die belang van 'n gekontroleerde vervanging van die oorspronklike teelmateriaal met nuwe teeldiere vanuit die kommersiële populasie as deel van die domestikasieproses.

TABLE OF CONTENT

ACKNOWLEDGEMENTS

The author wishes to express sincere appreciation to Mr. Brink for support and assistance throughout the project. Also Dr. D'Amato for her supervision and contributions, not only during the study, but also in the preparation of this manuscript. In addition, special thanks to Irvin & Johnson for the funding of this project, as well as Dr. P Cook and the N.R.F. for monetary support.

LIST OF ABBREVIATIONS AND SYMBOLS

AA	Acrylamide
AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
AP-PCR	Arbitrarily Primed PCR
APS	Ammonium Persulphate
ATP6/CO3	Adenosine Triphosphate 6/ Cytochrome oxidase 3 primer
BAC	Bacterial Artificial Chromosome
BLAST	Basic Local Alignment Search Tool
bps	base pairs
C	Concentration
C16sar	C16S ribosomal RNA primer
CAPS's	Cleaved Amplified Polymorphisms
CO1	Cytochrome oxidase 1 primer
CTAB	Hexadecyktrimethyl – Ammonium Bromide
CytB/C	Cytochrome B/C primer
°C	Degrees Centigrade
DAF	DNA Amplification Fingerprinting
ddH_2O	double - distilled water
DEAT	Departments of Environmental Affairs and Tourism
DNA	Deoxiribonucleic acid
dNTP	dinucleotide triphosphate
EC	East Coast
EDTA	Ethylenediaminetetraacetic acid
Et al.	And company
EtBr	Ethidium Bromide
EtOH	Ethanol
H_0	null hypothesis
HMcon	*Haliotis midae* control region primer
HMND1d	*Haliotis midae* NADH dehydrogenase 1 primer
HMND3	*Haliotis midae* NADH dehydrogenase 3 primer

hrs	hours
I&J	Irvin and Johnson
IPTG	Isopropyl-β-D-thiogalactopyranoside
kbs	kilo base pairs
λ HindIII	Lambda HindIII
λ	Lambda
LB	"Luria Bertani"
M	Molar
mg	milligram
MgCl	Magnesium Chloride
µg	microgram
µl	microlitre
min	minutes
mM	milliMolar
mt	Mitochondrion
mts	Metric Tons
mtDNA	mitochondrial DNA
NaCl	Sodium Chloride
NADH2	Nicotinamide Adenine Dinucleotide
NaOH	Sodium Hydroxide
ND1	NADH dehydrogenase 1 primer
ND5/6	NADH dehydrogenase 5/6 primer
nDNA	Nuclear DNA
ng	nanogram
OD	Optical Density
PAA	Polyacrylamide
PCR	Polymerase Chain Reaction
pH	The concentration of hydrogen ions in a solution is expressed conventionally as its pH
pmol	picomol
QTL	Quantitive Trait Loci
RAPD	Randomly Amplified Polymorphic DNA
RE	Restriction enzyme

RNase	Ribonuclease
rpm	revolutions per minute
RT	Room temperature
SA	South African
SANDF	South African National Defense Force
SAPS	South African Police Service
SCAR's	Sequence Characterized Amplified Regions
SDS	Sodium dodecyl sulfate
sec	seconds
SSCP's	Single Strand Conformation Polymorphisms
SSR	Simple Sequence Repeats
Taq	Thermis Aquaticus DNA polymerase
TBE- buffer	TRIS Borate EDTA buffer
Tm	Melting temperature
TRIS	2-amino-2- (hydroxymethyl)-1,3-propandiol
U	Units
V	Volt
Vol	Volume
W	Watts
WC	West Coast
YAC	Yeast Artificial Chromosome

INTRODUCTION

1. Background

The combined effects of environmental and genetic factors, the relative contribution of which varies from one individual to another, regulate biological productivity of animals. Efforts to improve productivity should, therefore, address both sets of criteria in order to achieve optimal efficiency. Environmental aspects are often the area of initial concern as well as the focus of short-term strategies, aimed at improvement of productivity. The contribution of genetic factors should be included into medium- and longer-term strategies for enhancement of productivity.

Knowledge on genetic factors that have been uncovered by molecular and quantitative genetic research has become essential in modern management and improvement of species and populations in animal husbandry systems. This not only refers to management and production of abalone in controlled artificial production systems, but also to alternative production methods such as farming on the basis of hatchery supplied seed stocks. Genetic information also forms the basis of the methodology for monitoring the effects that commercial hatcheries and farming systems have on the genetic composition of natural populations.

A variety of molecular techniques can be utilized to contribute towards the more effective application of conventional procedures of genetic improvement, such as optimal management of broodstock, selection, hybridization, etc. These include techniques for the determination of genetic relationships between individuals or populations, as well as for the levels of genetic variation that exist between individuals, groups and populations.

2. Objectives

The aim of this study was to compare the levels of genetic variation as well as genetic relationships that exist between the natural South African West Coast abalone (*Haliotis midae*) population, the Danger Point abalone Hatchery (on the South African West Coast) broodstock, commercial F1 (Progeny) stocks and East Coast abalone stocks. The outcome of this study will provide information that will form the basis of future

decision making with regard to the structuring, maintenance and management of broodstock at the Danger Point Abalone Hatchery in view of optimizing the short-, medium- and long-term objectives for genetic improvement. These objectives are:

1. Genetic variation needs to be maintained at optimal levels in order to ensure long-term genetic improvement. Knowledge of the genetic variation contained in the broodstock is required in order to prevent the deleterious effects of inbreeding depression within the commercial population (Progeny)
2. Knowledge of genetic relationships between different broodstocks is required to ensure optimal realization of hybrid vigor through crossbreeding.

3. Abalone

Eukaryota; Matazoa; Mollusca; Gastropoda; Archaeogastropoda; Haliotidae; *Haliotis*

3.1 General

Abalone are gastropod mollusks, or rather, marine snails. There are almost one hundred species of *Haliotis*. Abalone species occur in all the major oceans of the world, but appear to be more abundant in the temperate zones. Because of the fact that abalone prefer shallow, turbulent waters with high levels of dissolved oxygen and hard surfaces, abalone are usually most commonly found off exposed rocky headlands in cool seas. (Fallu, 1991)

The shell of abalone covers most of its body (Figure 1). Seen from above, it generally has an oval shape elongated along the anterior-posterior axis, though some species are more elongated than others. (Fallu, 1991) The abalone shell is in the form of a spiral, as is common in snails, although this may not be immediately apparent. It is shallow, only about a fifth as deep as it is long, with the spiral most clearly seen abaxially. This hard protective shell resembling the shape of a human ear is lined with mother-of-pearl, and along one side there are small holes through which the animal respires and excretes wastes. This row of holes is found just in from the left-hand side of the shell, extending anteriorly to just over the head. The anterior holes are the largest with those towards the back, usually blocked. The abalone shell grows through addition of new material on the anterior right-hand side. It grows as a spiral and, as the head must always point forward, the shell rotates on the body as it grows. (Fallu, 1991)

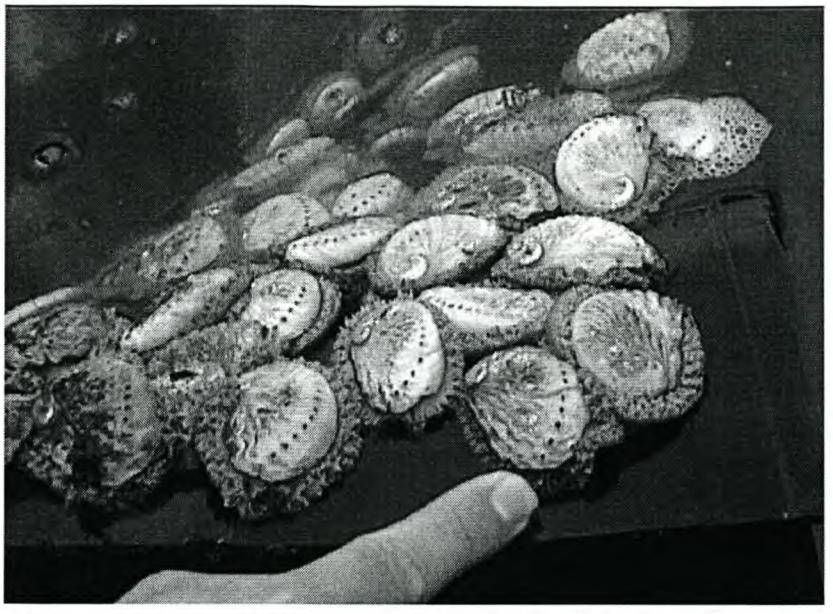


Figure 1: Photo of *Haliotis midae* at about 18 months of age.

Abalone holds on to the seabed with its foot in typical snail-like manner. Although the foot is largely hidden in most abalone species, it is clearly visible when turned upside down. The section of foot that is not completely covered by shell looks similar to a pair of lips and is often referred to as such. Lips are generally covered in very tough skin, which protects the tender inner parts of the foot against predators. In addition, lips are often coloured in a species-specific manner. *Haliotis virginia,* a New Zealand species has a yellow foot, while in Australia, colour is used to differentiate between the "brownlip" and "greenlip" abalone species. Extending from the upper part of the lips, around the outside of the foot, is a series of tentacles, which detect predators and food by touch and taste.

It is this foot part of the abalone that is so sought after as food, and comprises of the muscle in the foot as well as the short stalk joining the foot to the shell. The head of abalone is situated at the front of the foot and is also typically snail-like, with tentacles similar to those found on the lips. However, the head tentacles are larger and similar to the eyestalks of land snails. The mouth is at the base of the head underneath the lips. It is a circular orifice in the middle of a piece of flesh called the oral disc.

Abalone respire through paired gills that are located in a chamber called the mantle cavity, under the series of holes in the shell mentioned earlier. Seawater is drawn into

the mantle cavity anteriorly and passed over the gills, so that oxygen is taken up and waste gases are discarded. The used water is then passed out through the holes in the shell. The mantle cavity has other functions beside that of respiration and disposal of wastes. The reproductive glands also empty into the cavity, as do the eggs and sperm during spawning. (Fallu, 1991)

3.2 Endangered abalone

There has been growing concern over the illegal harvesting of abalone from South African waters for some time. As populations of the resource decline, increased pressure is placed on South Africa's abalone fishery to meet continuing national and international demand. This strong demand, which is not fully met through legitimate trade, creates incentives for people to supply the black market with stolen or "poached" abalone. While abalone poached from South African waters may find its way onto the domestic market the majority is destined for overseas export, especially to East Asia. (Tailby *et. al.* 2002; Hauck and Sweijd, 1999; Marshall, 2002; Hauck *et. al.*, 2002)

South Africa's law-enforcement authorities are engaged in an uphill battle to save the country's abalone stocks from extinction at the hands of poachers and international smuggling syndicates. Not only is the giant sea snail prized by gourmets as a delicacy reputed to have aphrodisiac properties, the animal's inner iridescent layer of shell is itself sought after as material for beads and other ornaments. Tons of abalone have been confiscated, scores of people arrested, and large amounts of property seized in a joint anti-poaching initiative called Operation NeptuneII. The initiative was established in August 2000 between South Africa's Departments of Environmental Affairs and Tourism (DEAT), the South African National Defense Force (SANDF) and the South African Police Service (SAPS). (Marshall 2002; Marine and Coastal management; Hauck *et. al.* 2002)

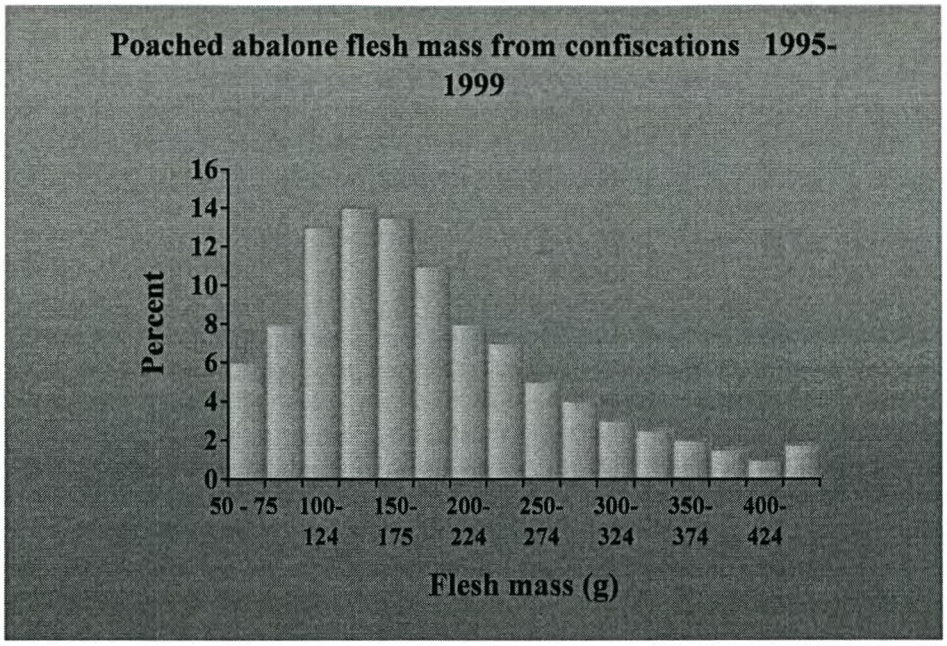


Figure 2: Poached abalone statistics, indicating poached abalone flesh mass from confiscations between 1995 and 1999. n = 82 000; 55% sublegal (Tarr, 2000)

Poaching is most intense along the rugged coastline spanning the convergence of the Indian and the Atlantic Oceans at the southern tip of the continent and along a 160-kilometer stretch of the Western Cape coast. (Marshall 2002)

Poaching, however, is not the only hazard to the country's abalone stocks. South Africa's abalone is facing an additional threat from an unexplained influx of west coast rock lobster (*Jasus lalandi*). Juvenile abalone typically shelter under spiny sea urchins until their shells are large and hard enough for protection from predators. The rock lobster prey on sea urchins and have reduced sea urchin populations to the point where juvenile abalone are now quite vulnerable to predators. (De Waal and Cook 2001; Marshall 2002)

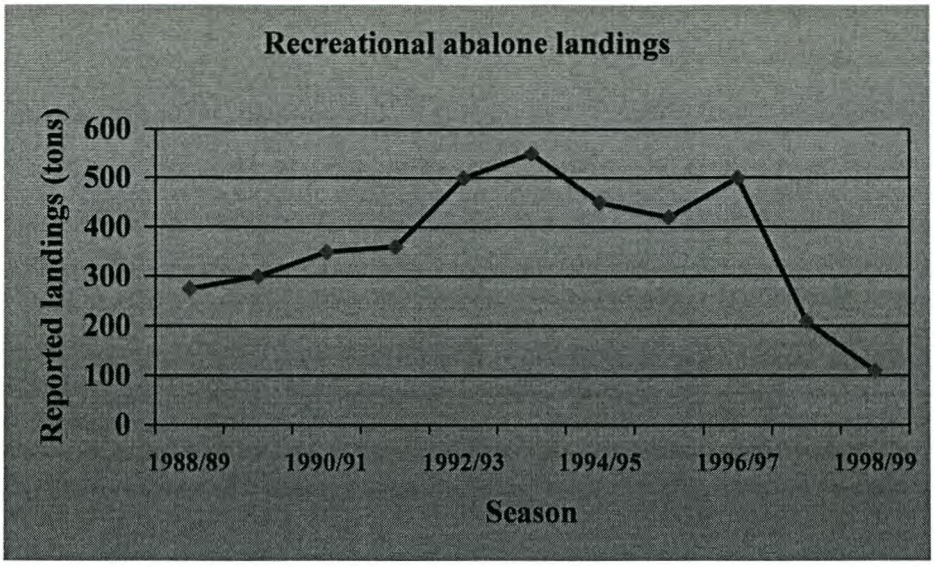


Figure 3: Graph indicating the recreational abalone landings in South African between 1988 and 1999. (Tarr, 2000)

To fulfill the high demand (Figure 3), and because of the long time that abalone take to reach marketable size in the wild, culturing of abalone has become a viable alternative to fishing wild stocks. It is seen as one way of diminishing the crisis, as it could provide a regular supply for the weighty national and international demand for this prized sea delicacy. However, abalone farms in South Africa were only recently started and it takes eight to nine years for abalone to reach the minimum legal size where they may be harvested. (Tarr, 1992) Until then, the crisis concerning depletion of abalone stocks on the South African coast is still looming.

4. Abalone Farming

4.1 World Abalone Farming

Abalone is one of the most prized sea delicacies worldwide. There are approximately a hundred species of abalone. Fifteen of which are harvested commercially. Countries with a large abalone fishery include Japan, Australia, New Zealand, South Africa, Korea, Taiwan and China. As a result of depletion of the ocean's abalone recourses commercial catches in these countries were reduced by 50 to 95 percent over the past twenty-five years. The commercial catch of abalone worldwide has declined from 18000mts to just over 10000mts over the same period. Total world abalone production was approximately 13000 tons in 1999, of which over 6500 tons could be attributed to farm production. (FISHTECH Abalone consultants)

As it is difficult to reintroduce abalone into the wild, and it seems that the most promising solution is culturing of the species. This can be done using a variety of methods. However, since there may be different policy ramifications when considering each method it is useful to divide culture methods into two categories: Land based - and Marine based hatcheries.

4.2 SA Abalone Farming

The I&J abalone farm at Danger Point is a land based hatchery, with the broodstock taken randomly from the nearby coastline. Land based systems offer a higher level of control over the operation as compared to marine based systems. It does however require a high capital investment in tanks and water intake. Recent developments in tank design and management have lowered both the costs of the tanks and the water requirements.

The South African abalone fishery is one of the oldest commercial abalone fisheries in the world, with commercial catch data being recorded since 1953 (Plaganyi *et. al.*, 2001). Although six species of abalone occur on the South African coast, most are either too small or too difficult to harvest, or occur in insufficient numbers to be harvested commercially (Cook, 1998). *Haliotis midae* is the species on which the fishery industry relies, and is restricted to shallow habitats in beds of kelp, *Ecklonia maxima* (Tailby and Gant, 2002; Plaganyi *et. al.*, 2001).

Historical catch statistics show that in the mid-1960's up to 2 800 tons of fresh weight per year were landed, but since 1971 a quota system has limited catches to just over 600 tons per year (Cook, 1998; Hauck and Sweijd, 1999). Over the last few years, however, increased misuse by both recreational divers and poachers has dramatically increased the demand on the resource. These are not the only negative effects on the abalone resource. Ecological changes, including altered numbers of rock lobsters and sea urchins also seem to have an effect in some areas. (Cook, 1998; De Waal and Cook, 2001) The combined effects of these increased pressures on the abalone resource have resulted in the authorities recommending a reduced Total Allowable Catch for 1997 (Cook, 1998). Since fishing companies have little chance of increasing quotas in the near future, several of them have turned to abalone aquaculture as a possible alternative.

Present-day farming operations include recreational, subsistence and commercial activities. Since the late 1980's land-based farming of abalone has been under development. This followed after Genade et al. (1988) demonstrated that it was possible to spawn the South African abalone species in captivity. Land based farming has recently attained commercial scale production (Pitcher *et. al.*, 2001 and Cook, 1998). Even though South Africa has a large coastline, running from its border with Namibia to its East Coast border with Mozambique, commercially exploitable densities of abalone, in association with huge kelp beds, occur only along the southwestern coast. Although some farms have been established as far north as Port Nolloth on the West Coast, and as far east as Port Elizabeth on the Southeast Coast, it is not surprising that the majority of farms have been developed along the southwestern part of the coastline.

In abalone culture, wild-caught broodstock are used extensively. (Litaay *et. al.* 2001) The suitability and selection of broodstock from the wild has been based on the phenotype of mature animals. The main criteria used in such selection are size, shape of the gonad and colour. However, animals selected based on these criteria are known to vary widely in production performance. Research has therefore focused on developing easily useable indices, based on physical characteristics, for assessing the potential reproductive performance of female abalone (Litaay *et. al.* 2001).

4.3 Seeding

Enhancement of abalone stocks with cultured juveniles has been practiced for some time, especially in Japan. There does not appear to have been a sustained follow up apart from Japan where ranching is claimed to make a significant contribution to their management protocol and commercial catches. But the high cost of rearing juveniles and the high mortality of seeded stocks have created sporadic interest in seeding with larvae. Abalone larvae are relatively simple and cheap to produce in very large numbers in a basic hatchery. There is a common rule that states that large animals that are released into the wild survive better than smaller animals. The time and method of release of ranched abalone may also determine the mortality and percentage finally recovered. Bearing this in mind, it is not surprising that there have been so few instances of abalone larvae being released directly into the wild. (Fallu, 1991)

A study done by the Fisheries Research Development Corporation in Australia (Clarke, 1994) indicated that, with the best estimate of seeded abalone survival combined with current seeding costs and fishing returns, larval abalone seeding is at best of marginal economic viability. Mortality rate of seeded abalone is the main factor determining profitability, while cost of larvae is the main expense. Both of these factors may be quite variable. However, it was stated that the possibility that larval seeding of sites, in which abalone populations have become depleted may be profitable, might not be excluded.

In South Africa, seeding experiments using juvenile abalone carried out on the Northwest coast, indicated that a possibility exists for the successful introduction on this species along this stretch of coastline (Sweijd Dissertation, 1998; De Waal and Cook, 2001). The growth and survival rate of animals ranging in size between 8 and 14mm were monitored over a period of 6 months. Results showed an average minimum survival rate of 30%. A similar experiment carried out in McDougall's Bay, Port Nolloth, by De Waal and Cook (2001) also showed that seeding of juvenile abalone was a viable albeit challenging option. However, no positive controls were included and conclusions drawn on the survival of juveniles were not convincing. Natural mortality of abalone depends on different contributing factors such as predation, fishing pressure and diseases (Jamieson, 1999; Muse, 1998) and may be in the order of 20% per year. It therefore seems difficult to reintroduce abalone into the wild.

5. Molecular Markers applied to studies of genetic diversity and natural population studies

It has widely been recognized that the loss of genetic diversity is a major threat for the maintenance and adaptive potential of species. To manage genetic diversity effectively the ability to identify genetic variation is indispensable. Morphological traits have long been used in the characterization of diversity. However, morphological variation is often found to be restricted and genotype expression may be affected by environmental conditions thereby constraining the analysis of genetic variation.

Since the confirmation of DNA as the hereditary material, and indeed since scientists started investigating techniques for utilizing and manipulating DNA on a molecular basis, the repertoire of genetic markers available for population genetic studies in a number of species has increased significantly. It must be remembered however, that most estimators of genetic diversity are aimed to be neutral. Therefore, molecular genetic techniques are currently being applied as a complementary strategy to traditional approaches in the management of genetic resources. (Van Treuren, 2000)

Apart from the use of minisatellites for DNA fingerprinting in the early 1980's, the invention of PCR (Mullis et. al., 1986) laid the foundation for the development of many new genotyping techniques. Due to tremendous advancements in the field of molecular genetics, a variety of different techniques to analyze genetic variation have emerged during the last few decades, most of which use PCR for detection of fragments. No single technique is universally ideal; each available technique exhibits both strengths and weaknesses. The choice of technique, or most appropriate genetic marker, is often a compromise that depends on the genetic resolution needed and the research question pursued, as well as on financial and time constraints and the technical expertise and facilities available. (Mueller and Wolfenbarger, 1999; Vos *et. al.*, 1995; Van Treuren, 2000; Caetano-Anolles and Gresshoff, 1998) Ideally, a fingerprinting technique should require no prior investments in terms of sequence analysis, primer synthesis or characterization of DNA probes. (Vos *et. al.*, 1995) A number of fingerprinting techniques that meet these requirements have been developed over a number of years, including Randomly Amplified Polymorphic DNA (RAPD) (Williams *et. al.*, 1990) and DNA Amplification Fingerprinting (DAF) (Caetano-Anolles *et. al.*, 1991). These methods are all based on the random amplification of genomic DNA using arbitrarily selected primers. DNA fragment patterns are generated from any DNA without prior sequence knowledge. These PCR based fingerprinting methods have the major disadvantage that they are very sensitive to the quality of DNA, reaction conditions and temperature profiles, which limits their application. (Vos *et. al.*, 1995) A new technique that approaches an ideal is Amplified Fragment Length Polymorphism – Polymerase Chain Reaction (AFLP-PCR) or shortly AFLP. (Zabeau and Vos, 1993) It is a robust and consistent method that combines the reliability of the RFLP technique with the power of the PCR technique. It will be discussed in length at a later stage.

5.1 Characteristics of genetic markers

An overview of genetic markers is given in Table1, and it is according to these markers that the following discussion on characteristics is applied.

5.1.1 Level of polymorphism

In a population, a polymorphic gene is one for which the most common allele has a frequency of less than 0.95 (or even the more rigid cutoff at 0.99). This cutoff at 95% or 99% in the definition of polymorphism is arbitrary, but it serves to center attention on those genes in which allelic variation is common. The characterization of polymorphism is an attempt to focus on genes that have alleles with frequencies too high to be explained solely by recurrent mutation to harmful alleles.

The resolving power of genetic markers is determined by the level of polymorphism detected, which is primarily affected by the mutation rate. Mutations at minisatellite and microsatellite loci, mainly due to changes in the number of repeat units of the core sequence, have been estimated to occur at the relatively high frequency of 10^{-3}-10^{-2} and 10^{-5}-10^{-4} per meiosis, respectively, and thus, are highly polymorphic. (The difficulty of utilizing microsatellite markers is that knowledge of the flanking sequences either side of the microsatellite repeat is required for their use.) In contrast, variation at allozyme loci is caused by point mutations, which occur at low frequency ($<10^{-6}$ per meiosis). (Van Treuren, 2000) Moreover, only mutations altering the net electrical charge of the proteins can be detected, thereby reducing the resolving power of allozymes. In addition, sampling is destructive for allozyme analysis, whereas PCR-based techniques do not require killing of specimens. The unavoidable limitation of protein electrophoresis is its lack of ability to detect variation in a nucleotide sequence that does not alter the amino acid sequence. A polymorphism is silent if it is present in the coding region but does not alter the amino acid sequence, and many nucleotide differences in third-codon position fall in this category.

DNA polymorphisms at sequence level can be studied by either a direct approach (DNA sequencing) or an indirect approach (RE, SSCP) and both haploid and diploid data is produced. At a finer level, methods of DNA sequencing allow variation to be observed

base pair by base pair, whilst studying variation in the sites recognized by restriction enzymes provides a coarser view of base pair variation. Polymorphisms in diploid data include that produced by allozymes, microsatellites, and RFLP's of specific loci, introns, etc., whilst RAPD's and AFLP's produce polymorphisms in anonymous DNA. The traditional DNA fingerprinting technique uses polymorphisms at minisatellite loci, and fingerprints are also generated from anonymous DNA. Polymorphism of alleles that determine allozymes is particularly widespread among higher organisms. (Hartl and Clark, 1997

5.1.2 Abundance

The frequencies at which sites of interest occur in the genome largely determine the number of markers that can be generated. AFLP's are abundant markers due to the large number of restriction endonucleases available and the frequent occurrence of their recognition sites within the genomes. Microsatellites have also been found to occur frequently in the eukaryotic genome. A study done on the distribution of microsatellites in the Emu (*Dromaius novaehollandiae*), estimated that the genome contained a microsatellite repeat every 48kb. (Roots and Baker, 2002) RAPD markers are even more abundant because numerous short random sequences can be used as primers to generate anonymous patterns. In contrast, the number of allozyme markers is restricted due to the limited number (about 30) of enzyme systems available for analysis. Sequence data of the sites of interest are required for primer construction when investigating specific genomic regions by PCR-sequencing, Single Strand Conformation Polymorphisms (SSCP's), Sequence Characterized Amplified Regions (SCAR's) or Cleaved Amplified Polymorphisms (CAPS's). (Van Treuren, 2000)

5.1.3 Reproducibility

The extraction of purified, high quality DNA is a prerequisite for the majority of marker techniques. Degraded and/or unpurified DNA may, for example, affect the amplification or endonuclease restriction of fragments resulting in aspecific polymorphisms. This failure is caused by the fact that even small deviations in experimental conditions may alter band profiles.

5.1.4 Co-dominance vs. Dominance

Co-dominance means that both alleles present at a diploid (or polyploid) locus are always expressed. That is, in heterozygotes both alleles will be scorable. For population genetic questions co-dominant markers are clearly superior to dominant markers because 1) they allow estimation of allele frequencies; and 2) for a given level of analytical power, co-dominant markers require smaller sample sizes than dominant marker (Mueller and Wolfenbarger, 1999). Dominant markers are assumed to be biallelic systems, where alleles are shown as a band present or a band absent due to the presence or absence of a priming site. The non-priming allele will not be detected as a band. In the case of the dominant RAPD markers, heterozygotes cannot be distinguished from homozygotes in terms of the presence or absence of a band. In AFLP's each band is assumed to represent a biallelic locus, which may not necessarily be true.

5.1.5 Costs

Marker development may be very time-consuming and expensive when necessary probes or sequence data for primer construction are unavailable. In addition, sufficient technical skills and facilities need to be at hand. Technical equipment, laboratory facilities, wages and chemicals all contribute to the operational costs of the technologies. Relatively expensive chemicals such as Taq-polymerase, restriction enzymes and radioactive labeling all add to the expenses. Polyacrylamide gels are more expensive to run than agarose gels and require visualization of polymorphisms by autoradiography or silverstaining, which are more costly, compared to ethidium-bromide staining. Laborious and technically demanding markers, such as minisatellites, PCR-sequencing and RFLP's are therefore quite expensive. Compared to these techniques the operational costs of the other markers presented in Table1, vary from low to medium, depending on the methodological procedures followed.

5.1.6 Quantity of DNA required

Because only small amounts of template DNA (5-100ng per reaction) are required, techniques, which are based on PCR, are currently preferred. Minisatellites require the largest amount of DNA (5-10μg per reaction) but Southern blots may be reprobed several times during the course of an experiment. Intermediate quantities of DNA are needed for AFLP analysis (0.3-1μg per reaction) because endonuclease restriction of the

DNA template precedes the PCR-reaction. Application of PCR-based markers may be especially relevant when only small amounts of DNA can be extracted, e.g. when working with tiny organisms, or degraded material.

5.1.7 Labour-intensity and technical demands

DNA Fingerprinting with RFLP's and minisatellites are quite laborious and technically demanding methods, since their analysis includes the time-consuming steps of Southern-blotting, radioactive labeling of probes and hybridization. They also require amounts of DNA in the order of μg's, for which destructive sampling is needed. Therefore PCR-based techniques are presently more in favor. Unless the process is automated, throughput of PCR-sequencing sample is still low because four separate sequence reactions are needed per sample and data processing is time consuming. The labor-intensity of the other PCR-based techniques presented varies from low to medium, depending on the methodological procedures necessary in addition to PCR. RFLP's minisatellites and PCR-sequencing - including AFLP's - require technical skills and facilities to carry out radioactive labeling. In addition, Southern blot hybridization is part of the RFLP and minisatellite analysis. These techniques are therefore among the more technically demanding markers. Allozymes and PCR-based markers analyzed on agarose gels (e.g. RAPD's and SCAR's) are the least technically demanding of the markers under discussion.

5.1.8 Amenability to automation

Currently, techniques that can be automated are preferred because they enable increased sample throughput. Although considerable financial investment is still required, automation may be cost-effective when techniques are applied on a routine basis. Virtually all techniques that are based on the Polymerase Chain Reaction (PCR) are amendable to automation.

6. Mitochondrial DNA

Mitochondrial DNA (mtDNA) of higher animals is a circular molecule of some 17,000 bases. (Jiang *et. al.*, 1995) Along with chloroplast DNA, it is collectively known as cytoplasmic DNA. The mitochondrial genome is economical in information storage, as it lacks introns, using both strands to encode for two ribosomal RNA's, 22 transfer

RNA's and 13 proteins that are responsible for electron transport and ATP synthesis. (Jiang *et. al.*, 1995) The mtDNA has no repetitive DNA, spacers, or introns, and is usually the only type of DNA to survive in ancient bone specimens because of its abundance; 500-1000 copies per cell instead of only two copies of most nuclear DNA. This is notable because mitochondrial genes have been estimated to evolve at a rate approximately 5 to 10 times higher than those of the nucleus. (Brown et al., 1979) Unlike nuclear DNA that undergoes reconstruction during meiosis, mtDNA does not recombine. Therefore alteration to mtDNA is essentially due to mutation or copying errors. Mutations accumulate in a fashion dependent upon time and population growth, -contraction, -splitting, and -merging. mtDNA is therefore a potentially useful molecular marker, providing high-resolution information on the evolutionary relations between taxonomically bound families. In this way, the study of genetic variation in abalone can elucidate the relatedness of populations.

DNA from the mitochondrion is targeted in population studies for various reasons. First, mtDNA is valuable, in that it allows for the reconstruction of maternal lineages that often are correlated with geographical groups. Secondly, because of its reduced effective population size (N_e) ($^1/_4$ that of nuclear DNA (nDNA)) it is more sensitive to genetic drift, and detects diminished diversity easier than nDNA. And finally, as mentioned earlier, it has a higher mutation rate than nDNA (Smithsonian Institute Natural Museum of Natural History, 2000). The assumption that mtDNA is strictly maternally inherited however, has recently been contested. The polymerase chain reaction has been used to determine that sperm do contribute mitochondria to the fertilized egg. (Kaneda *et. al.*, 1992; Ankel-Simmons and Cummins, 1996) However, the amount of paternally inherited mtDNA in an individual is, not surprisingly, very small, given that a sperm contains only about 50 mitochondria while the egg has over 100 000.

A wide variety of animal mtDNA's, especially the metazoan, have been sequenced or analyzed, and data shows that mollusk mtDNA's have abnormal length variation (17.0-41.0 kb) (Jiang *et. al.*, 1995). However, information on mtDNA of gastropod mollusks, such as Archaeogatropoda of which abalone is a member, however, has little information available. Even though the complete mtDNA of various mollusks, including

Cepaea nemoralis, Albinaria coerulea, Pupa strigosa, Crassostrea gigas, Venerupis philippinarum, Loligo bleekeri and *Katharina tunicata* have been sequenced, the mtDNA of assorted abalone species have only been sequenced to a minimal extent. Only three genes of *Haliotis midae* have been sequenced, them being: Abalone sperm lysin mRNA complete sequence (nDNA), a partial sequence of the cytochrome c oxidase subunit I (COI) gene (mtDNA), and a complete sequence of Haliotis midae internal transcribed spacer 1, 5.8S ribosomal RNA gene (nDNA).

7. AFLP

AFLP™ is a trademark of Keygene (Wageningen, The Netherlands)

7.1 General

A new technique that approaches an ideal is Amplified Fragment Length Polymorphism – Polymerase Chain Reaction (AFLP-PCR) or shortly AFLP. It is also known under the synonyms, Selective Fragment Length Amplification (SFLA) and Selective Restriction Fragment Amplification (SRFA). It is an easy, fast, relatively cheap and reliable method to generate hundreds of highly informative genetic markers. (Mueller and Wolfenbarger, 1999) A typical AFLP fingerprint contains between 50 and 100 amplified fragments, of which up to 80% may serve as genetic markers. (Vos *et. al.*, 1995; Keygene N.V., 2001; Gillet, 1999) AFLP's are polymerase chain reaction (PCR) - based markers for the rapid screening of genetic diversity. AFLP methods rapidly generate hundreds of highly reproducible markers from the DNA of any origin or complexity, thus they allow high-resolution genotyping of fingerprinting quality. Fingerprints are produced without prior sequence knowledge using a limited set of generic primers. The number of fragments generated in a single reaction can be perfected by selection of specific primer sets (Vos *et. al.*, 1995).

7.2 Methodology

AFLP detects DNA restriction fragments by means of PCR amplification. The AFLP technology (Figure 4) usually comprises of the following steps (Mueller and Wolfenbarger, 1999; Vos *et. al.*, 1995; Zabeau and Vos, 1993; Keygene N.V, 2001):

1. Restriction of small amounts of DNA (0.3-1μg) with two restriction enzymes, preferably a hexa-cutter and a tetra-cutter;

2. Ligation of double-stranded (ds) adapters to the ends of the restriction fragments.

3. Amplification of a subset of restriction fragments using two primers complementary to the adapter and restriction site sequences and extended at their 3' ends by "selective" nucleotides;

4. Gel electrophoresis of amplified restriction fragments on denaturing polyacrylamide gels ("sequencing gels")

5. Visualization of DNA fingerprints by means of autoradiography, phospho-imaging or other methods.

Adaptor ligations are performed in the presence of restriction enzymes such that any fragment-to-fragment ligations are immediately recleaved by the restriction enzyme. The adaptor is designed so that ligation of a fragment to an adaptor does not reconstitute the restriction site. The end sequences of each adapted fragment now consist of the adaptor sequence and the remaining part of the restriction sequence.

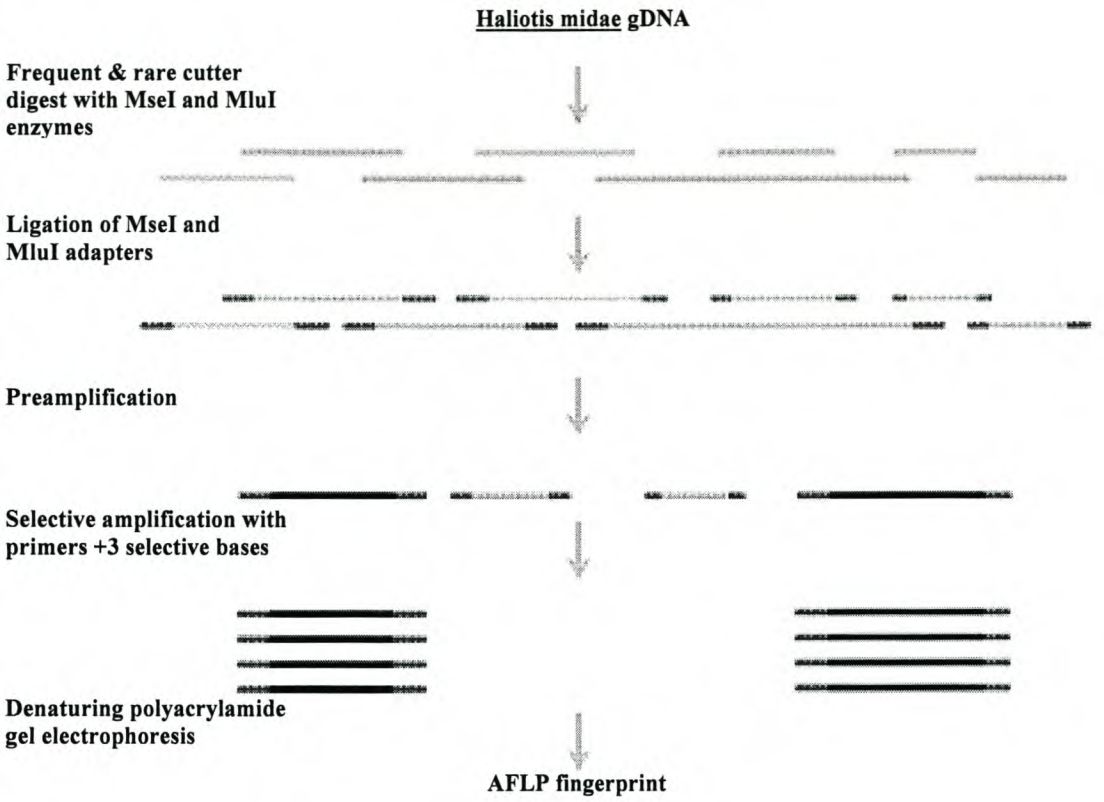


Figure 4: Schematic representation of AFLP procedure.

7.3 Scoring and reliability of AFLP markers

A series of techniques can be used to score AFLP's, ranging from simple agarose gel electrophoresis to automated genotyping. Polyacrylamide gel electrophoresis, whether manually done or with an automated sequencer provides maximum resolution of AFLP banding patterns to the level of single-nucleotide length differences, whereas fragment length differences of less than ten nucleotides are difficult to score on agarose gels. Even so, agarose gels are user friendly, inexpensive and require minimal equipment.

Because AFLP's are dominant, multilocus markers that are scored as the presence or absence of a band, artificial amplification, or amplification failure, will reduce AFLP reliability. This erroneous presence or absence of a band originates at the restriction-ligation step. That is, it is vital to ensure complete digestion in order to thwart later amplification of uncut fragments. Use of high-quality DNA and an excess of restriction enzymes should be tried to achieve complete digestion. Enzymes that are sensitive to DNA methylation can also cause incomplete digestion. PCR-generated artifacts are

minimized through the use of long AFLP primers and the application of extremely stringent conditions such as high annealing temperatures and magnesium concentrations. High stringency ensures that the primer binds only to flawlessly matched template sequences and eliminates mispairing. Analysis of AFLP scores from duplicate test samples reveal average errors of 0-2% (Jones et al., 1997; Mueller and Wolfenbarger, 1999)

7.4 Applications

AFLP technology has broad applications in a variety of areas. Such applications include: (Mueller and Wolfenbarger, 1999; Vos *et. al.*, 1995; Keygene N.V., 2001; Global Initiative on Late Blight, 2002)

1. Use of AFLP markers in genetic studies such as
   1.1. analysis of germplasm collections,
   1.2. genotyping of individuals and genetic distance analysis,
   1.3. identification of closely-linked DNA markers,
   1.4. Biodiversity studies
2. Construction of DNA marker maps.
3. Construction of physical maps using genomic clones such as YAC's and BAC's.
4. Precision mapping of genes and subsequent isolation of these genes.
5. Generation of "transcript profiles" for gene expression analysis.
6. Monitoring the inheritance of agronomic traits.
7. Diagnosis of genetically inherited diseases.
8. Forensic typing.
9. Quantitative trait loci (QTL) mapping.
10. Pedigree analysis, parentage analysis etc.

7.5 Polymorphism

AFLP fingerprints are a rich source of restriction fragment polymorphisms. The frequency with which AFLP markers are detected depends on the level of sequence polymorphism between the tested DNA samples. Sequence polymorphisms at the nucleotide level will usually be the molecular basis of AFLP polymorphisms. Single nucleotide changes will be detected by AFLP when restriction sites are affected, or when nucleotides adjacent to restriction sites are affected. The latter would cause AFLP primers to mispair at the 3' end, thus preventing amplification. Insertions, deletions and

rearrangements affecting the presence or size of restriction fragments will result in polymorphisms detected by AFLP's. (Keygene N.V., 2001)

7.6 Dominance of AFLP markers

AFLP markers suffer from their general dominant nature compared to the popular co-dominant microsatellite (SSR) markers. However, detailed pedigree information may allow identification of co-dominant AFLP markers, which permit exact estimation of allele frequencies and more powerful population genetic analysis. If a sufficient number of co-dominant markers can be identified, standard AFLP's might approach the power of microsatellites while circumventing the lengthy setup costs in developing SSR primers.

7.7 Advantages

The many advantages to using AFLP as a marker system include the fact that AFLP analysis requires minimal amounts of DNA, and furthermore, partially degraded samples can also be used. This means that exceptionally small samples and very small organisms can be examined with AFLP (Rosendahl *et. al.*, 1997; Mueller and Wolfenbarger, 1999). Another advantage that AFLP markers offer is that it can be generated for any organism with DNA and no prior knowledge about the genomic composition is needed. Therefore, AFLP's have broad taxonomic applicability and have been used effectively in a variety of taxa, including bacteria, fungi, animals and plants. (Mueller and Wolfenbarger, 1999; Bensch *et. al.*, 2002; Liu *et. al.*, 1998) The time efficiency of AFLP markers is also remarkable. AFLP markers can be generated at vast speed, as illustrated by the high ratio of polymorphisms generated per PCR experiment. It also offers great flexibility in the number of loci that can be co-amplified in one PCR reaction. An additional advantage is that relatively few primer pairs are needed to visualize the large number of loci. Moreover AFLP amplifications are performed under conditions of high stringency, thus eliminating the artificial variation that is seen routinely in RAPD-PCR. Repeated AFLP amplifications show near perfect replicability (Jones *et. al.*, 1997; Mueller and Wolfenbarger, 1999). The capacity to examine an entire genome for polymorphism and its reproducibility are probably the most important advantages, but ease of use, its robustness and reliability, its amenability to automation, as well as the fact that it can be applied to any DNA sample, gives it the potential to become a universal DNA fingerprinting system. (Mueller and Wolfenbarger, 1999; Vos

et. al., 1995; Van Treuren, 2000; Keygene N.V., 2001; Gillet, 1999; Global Initiative on Late Blight, 2002)

7.8 Disadvantages

The main disadvantage of AFLP's, is that the method primarily generates dominant rather than co-dominant markers. (Mueller and Wolfenbarger, 1999) The difficulty in identifying homologous markers (alleles) renders this method less useful for studies that require precise assignment of allelic states, such as heterozygosity analysis. (Mueller and Wolfenbarger, 1999) Another disadvantage includes the fact that we assume each band to be a locus.

7.9 Conclusions on AFLP technology

Most AFLP fragments correspond to unique positions on the genome, and, hence, can be exploited as landmarks in genetic and physical maps, each fragment being characterized by its size and its primers required for amplification. In addition, the AFLP technique allows detection of restriction fragments in any background or complexity. (Vos *et. al.*, 1995) Therefore the AFLP technique is not simply a fingerprinting technique, it is an enabling technology in genome research, because it can bridge the gap between genetic and physical maps.

Because of their unparalleled sensitivity to slight genetic differences, PCR-based markers such as AFLP's are expected to remain key molecular tools for some time to come. The high reliability of AFLP markers could lead to the displacement of RAPD markers, and their ease of use might cause a partial replacement of other high-resolution markers such as RFLP's and microsatellites, at least for some research programs such as QTL mapping. However, because of their largely dominant character, AFLP's are unlikely to out-compete co-dominant markers, such as microsatellites or allozymes, which undoubtedly allow more powerful population-genetic analyses. Thus, AFLP and microsatellite markers, coupled with sequencing information for systematic analysis, could surface as the main tools for the analysis of genetic variation. Although AFLP-PCR in not a solution for all molecular problems in ecology and evolution, it offers many advantages and therefore will doubtless replace several standard techniques. (Mueller and Wolfenbarger, 1999; Liu *et. al.*, 1998; Jones *et. al.*, 1997; Bensch *et. al.*,

2002) Researchers interested in genetic studies, should look at the relative strengths as well as the restrictions of AFLP markers, in the context of the particular research question pursued.

8. Population Genetics

Population genetics is the algebraic description of the genetic makeup of a population and the changes in allelic frequencies in populations over time. (Hartl and Clark, 1997)

8.1 Populations and Gene Pools

A group of individuals from the same species within a geographic area, where mating is actually or potentially taking place, is known as a population. The set of genetic information carried by all interbreeding members of a population is called the gene pool. Populations are dynamic; they grow and expand or diminish and contract through changes in birth and death rates, by migration, or by integrating with other populations. This has important consequences and can lead to changes in the genetic structure of the population. (Klug & Cummings, 1997)

8.2 Hardy-Weinberg Equilibrium (HWE)

It was in 1908 that a British mathematician, G.H. Hardy, and a German physician, W. Weinberg independently discovered a rule that relates alleles and genotypic frequencies in a population. However, a few assumptions concerning the population are necessary for the equilibrium to hold. (Hedrick, 2002; Caetano-Anolles and Gresshoff, 1998) These assumptions are:

1) The population is infinitely large, which in practical terms means that the population is large enough that sampling errors and random effects are insignificant.
2) Mating within the population occurs at random.
3) There is no selective advantage for any genotype; that is, all genotypes produced by random mating are equally feasible and fertile, and no individual will have a reproductive gain or advantage over another individual.
4) There is an absence of other factors that might influence the equilibrium, including mutation, migration, and random genetic drift.
5) Organisms are diploid.

6) The allele frequencies are identical in males and females.
7) Reproduction is sexual.

The mathematical relation between allele frequencies and genotype frequencies is as follows,

A;A	A;a	a;a
p^2	2pq	q^2

and is called the Hardy Weinberg equilibrium.
AA is the homozygous dominant genotype, Aa the heterozygous genotype and aa the homozygous recessive genotype. Overall the three progeny zygotes (AA, Aa, aa) are formed in the afore mentioned proportions: p^2, 2pq, q^2.

The Hardy-Weinberg principle in the case of two alleles per locus, can also be expressed mathematically:

$$p^2 + 2pq + q^2 = 1$$

Let p = the frequency of the dominant allele and
q = the frequency of the recessive allele.

Construct a Punnett square crossing two heterozygous (pq) individuals.

	p	q
p	pp (p^2)	pq
q	pq	qq (q^2)

$p^2 + 2pq + q^2$ must equal 1 because the proportion of all alleles in a population must add up to 1 (100%) . (Hedrick, 2002; Caetano-Anolles and Gresshoff, 1998)

The crucial point is that given any set of initial genotypic frequencies, after one generation of random mating, the genotypic frequencies are in the proportions; p^2, 2pq and q^2. The equilibrium states that allele and genotype frequencies remain constant

generation after generation, and is the direct consequence of the segregation on alleles at meiosis in heterozygotes. (Griffiths *et. al.*, 1999)

The Hardy-Weinberg is a model that describes a hypothetical situation probably never to occur in nature. One or more of mutation, genetic drift, migration, non-random mating or natural selection are almost certainly acting upon natural populations at all times, which means that evolution is occurring in that population, and that it is not in HWE. A means of indicating whether a population is in HWE is the Probability value.

8.3 Probability values

A hypothesis is never proved or disproved entirely. Instead, a relative standard must be set to serve as the basis for either rejecting or failing to reject the null hypothesis. A probability (p) value of 0.05 is often chosen as this standard. When applied to any statistical test, a p value less that 0.05 means the probability is smaller than 5% that the observed deviation in the set of results could be obtained by chance alone. Such a p value indicates that the difference between the observed and predicted results is significant and thus serves as the basis for rejecting the null hypothesis (H_0). In contrast, p values of 0.05 or greater signify that the probability of the given data set occurring by chance is 5% or more. The conclusion is not to reject the null hypothesis.

The evaluation of statistical information must carefully be made on a case-by-case basis. When a null hypothesis is rejected, all underlying assumptions must be examined. If there is no concern about the validity, then other alternative hypotheses must be considered to explain the results. (Klug & Cummings, 1997)

8.4 Population parameters

8.4.1 Genetic drift

In contrast to a Hardy-Weinberg population of infinite size, restricted population size can have wide-ranging genetic consequences. A small population size causes chance alterations in allelic frequencies. The random change in allelic frequencies that results from the sampling of gametes from generation to generation is called genetic drift, and has the same anticipated effect on all loci in the genome. In a large population, on the average, only a small change in each generation in the allelic frequency will occur as a

result of genetic drift. On the other hand, if the population size is small, then the allelic frequency can undergo large fluctuations in a few generations in a seemingly random pattern. (Hedrick, 2002; Klug & Cummings, 1997).

8.4.2 Bottleneck effect

Closely linked to genetic drift, is the concept of the bottleneck effect. Natural selection or a catastrophe can cause a population bottleneck, a severe reduction in population size, which reduces the diversity of a population. The survivors have very low genetic diversity in comparison to the original population and little chance to adapt if the environment changes. They usually do not represent the full genetic variation that was present in the original population. This means that a single period of small population size can result in a serious loss in heterozygosity. Combined with inbreeding, the bottleneck effect is an especially serious problem for many endangered species, for the reason that great reductions in their numbers have reduced their genetic variability. This makes them especially vulnerable to changes in their environments and/or diseases. [One of the issues researched during this thesis, was whether or not sampling of Broodstock on the I&J hatchery might have caused a bottleneck effect.]

From time to time a population bottleneck or migration event can cause a founder effect, which occurs when a few individuals unrepresentative of the gene pool start a new population. (Hartl & Clark, 1997)

Both Genetic Drift and the Bottleneck Effect, are directly affected by effective population size (N_e)

8.5 Effective population size (N_e)

Conservation programs are often based on captive breeding, which, if based on a small number of founders and breeders, can cause an increase in the inbreeding coefficient and a consequent decrease in fitness known as inbreeding depression. In the case of aquatic species and especially marine bivalves, advancements in hatchery technology in the past decade have allowed aquaculture broodstocks to be kept as closed populations, without input of individuals from other wild or captive populations. Moreover, aquatic animals can have very high fecundity but high variance in reproductive success, and

there is evidence that hatchery seed may come from only a few successfully spawning individuals (Gaffney et. al., 1992)

For the management of genetic resources, it is critical to monitor the parameter of effective population size (N_e), as it determines the rate of increase in inbreeding ($\triangle F$), hence the rate of loss of genetic variability in a population ($\triangle F = 1/(2N_e)$) (Crow & Kimura, 1970). A reduction in the effective population size following a population bottleneck is associated with a decrease in heterozygosity and a loss of rare alleles. Launey *et. al.* (2001) concluded that small effective size and high relatedness have numerous consequences for a selection program. Inbreeding management is vital as offspring fitness is expected to vary depending on the degree of kinship of the parents.

8.6 Neutral theory

The first methodical application of protein electrophoretic methods to population genetics showed extensive genetic variation within most natural populations. (Hartl and Clark, 1997) Polymorphic alleles arose with frequencies considered to be too high to result from equilibrium between unfavourable selection and mutation. Motoo Kimura suggested in 1968 that most polymorphisms observed at the molecular level are selectively neutral, so that their frequency dynamics in a population are determined by random genetic drift. This means that at equilibrium, there is a balance between mutation and random genetic drift, so that on average, each new allele gained by mutation is balanced against an existing allele that is lost (or fixed).

However, the hypothesis was put forward more than 30 years ago, in a time when most of the genome was supposed to have a protein-coding function. Introns and other non-coding sequences were also unknown. Today we know that only about 4% of the mammalian genome codes for proteins, and that the low coding density affords sufficient scope for mutations that have little or no effect on fitness. The recent trend has been that more and more cases of departure from neutrality are being found, in part because of the increase in available data and in part because of the ever-increasing subtlety of tests that are applied (Hartl & Clark, 1997).

9. Methods of analysis

9.1 Whright's F-statistics

To quantify the inbreeding effect of population substructure, Wright defined what has come to be known as the fixation index, fixation being increased homozygosity resulting from inbreeding. The index is a valuable indicator of genetic differentiation as it allows an objective comparison of the overall effect of population substructure among different organisms without getting into details of allele frequencies, observed levels of heterozygosity, and so forth. Wright developed three fixation indices to evaluate population subdivision: F_{IS}, F_{ST} and F_{IT}, where F_{ST} correlates with the PhiPT value used in this study. (Hartl & Clark, 1997)

F_{ST} measures the effect of population subdivision, which is the reduction in heterozygosity of a subpopulation due to genetic drift. It is defined as the correlation of gametes within populations relative to gametes drawn at random from the entire population (**Sub**/**T**otal). The F_{ST} value is never negative; 0 indicates no subdivision, random mating and no genetic divergence within the population, 1 indicates extreme subdivision, and a F_{ST} value of up to 0.05 indicates insignificant genetic differentiation.

9.2 Principle Coordinate Analysis (PCA)

Principle Coordinates Analysis (PCA) is one of the simplest of the multivariate methods, of which its object is to take m variables $X_1, X_2, \ldots, X_m$ and find combinations of these to produce indices $Z_1, Z_2, \ldots, Z_m$ that are uncorrelated. The indices are prearranged so that Z_1 displays the largest amount of variation, Z_2 the second largest and so forth, where var(Z_i) indicates the variance of Z_i in the data set being considered. The Z_i are called the principle components.

When doing a PCA there is always the anticipation that the variances of most of the indices will be so low as to be negligible. In that case the variation in the data set can be adequately described by the few Z variables with variances that are not negligible i.e. the first few. However, a PCA does not always work in the sense that a large number of original variables are reduced to a smaller number of transformed variables. Indeed, if the original variables aren't correlated then the analysis is not significant.

MATERIALS AND METHODS

1. Sampling

Tissue was collected from the Irvin and Johnson (I&J) Abalone Hatchery at Danger Point, South Africa during the course of 2001. Approximately 5x5mm tissue was cut from the foot muscle of each abalone and stored in 95% Ethanol (EtOH) for further use. Four different populations were sampled, consisting of 1) the Wildstock or West Coast samples ($Wild_{(WC)}$), taken from the coastline surrounding the abalone hatchery, 2) Broodstock used on the farm, which was also collected from the nearby coastline, 3) samples from the Progeny generation on the farm, as well as 4) samples from Cape Reciefe on the East Coast ($Wild_{(EC)}$). Fifty individuals of each of the populations on the farm and from the $Wild_{(WC)}$ were sampled, whilst 13 of the $Wild_{(EC)}$ samples were acquired.

2. DNA extraction using salting-out method

(Adapted from Brufford et al., 1992, Molecular Genetic Analysis of populations, pp227-229, modified by S.B. Piertney)

After removal of the ethanol, DNA was extracted from the foot muscle of the abalone. Approximately 2x2mm tissue was dissected and coarsely diced with a scalpel blade, after which it was digested by adding 410µl extraction buffer. 4µg/ml RNase was added to each sample, and left at 37°C for 1h. 100µg/ml Proteinase K was added to the extraction buffer and inverted a few times. The samples were left in a 37°C waterbath overnight.

180µl of 5M NaCl was added after overnight lysis of the samples, and vigorously mixed by inversion x50. 60µl of 1% CTAB was added and left at 65°C for an hour. An equal volume of Chloroform:isoamyl alcohol (24:1) was added, and mixed by inversion. The samples were centrifuged at 13 000rpm for 6 min, and the supernatant transferred to a new tube. 420µl of ice-cold isopropanol was added to the supernatant and repeatedly inverted at a measured pace. The samples were left at -20°C for at least 60 min and were then centrifuged at 13 000rpm for 5 min, after which the supernatant was removed.

250µl of 70% EtOH was added to wash the DNA pellets. The samples were centrifuged at 13000rpm for 5 min and the supernatant removed. This washing step was repeated. The DNA pellets were left to dry at 55°C for 5-10 min and resuspended in 100µl dH_2O. Samples were stored at -20°C.

3. Standardization of Samples

For each of the extractions, a 0.8% Agarose gel was prepared, using 0.4g Molecular Grade Agarose (Whitehead Scientific) and 50ml 0,5 x TRIS Borate EDTA (TBE) buffer. Ethidium Bromide (EtBr) was added for visualization of the gel. Each sample was prepared for loading by adding 2µl of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol in H_2O) to 8µl of the DNA sample. 0.1-, 0.3- and 0.5µg/µl Lambda (λ) standards (Promega) were also prepared and loaded. The gel was then run at 100V for 60min on a Pharmacia LKB GNA 200, after which it was visualized under a UV light on the Geldoc™ 1000 system (BioRad). The samples were accordingly diluted to 20ng/µl using ddH_2O for later use.

4. mtDNA analysis using universal primers

The Table A2 in Appendix A, contains all information concerning primer sequences, references, DNA regions etc. that were used during this study.

4.1 Polymerase Chain Reaction (PCR) amplification

A wide range of conditions were tested for the amplification of DNA with universal primers ND5/6, CO1, CytB/C, ND1, NADH2, ATP6/COIII and CytB (shortened) (Kocher et al., 1989). These conditions included the use of gradient programs with temperatures ranging from 38°C to 58°C, Magnesium Chloride ($MgCl_2$) concentrations from 1 mM to 4 mM, as well as varying amounts of DNA, dNTP's and the primers itself. Different PCR programs were also tried, mostly using a Denaturing step of 92°C and an Extension step of 72°C. The Annealing temperatures ranged between 38°C and 58°C as mentioned previously, with the amount of cycles ranging from 35 to 37. The PCR program most often used consisted of the following steps: 92°C for 3min, followed by 37 cycles of: 92°C for 30sec, 38°C-56°C for 50sec, and 72°C for 1min. This was followed by 72°C for 3min.

4.2 Gel preparation

For each PCR reaction done, a 1% Agarose gel was prepared in 1 x TBE buffer. EtBr was added for visualization of the gel. Each sample was prepared by adding 2µl of 6 x loading buffer to 2µl of dH2O and 4µl of DNA. A Promega 100bp as well as a λHindIII ladder was used as a size marker. The gel was then run at 100V for 60min, after which the DNA was visualized under a Geldoc™ system.

5. mtDNA analysis using *Haliotis* primers

The primers used, were partially based on *Haliotis* and other mollusk sequences, and will henceforth be labeled as *Haliotis* primers. Refer to Appendix A for table indicating primers used during the research. (Table A2) These primers were obtained from Dr N Sweijd.

5.1 Polymerase Chain Reaction (PCR) amplification

An Applied Biosystems GeneAmp® PCR System 2700 was used during the amplification of the *Haliotis* primers, HMND1d, c16sar, HMcon and HMND3 (Sweijd, 1999). As with the universal primers, wide ranges of conditions were used during the amplification. Gradient programs were used, with temperatures starting at 38.5°C and increasing to 57.6°C. $MgCl_2$ concentrations were tried in a range between 1mM and 6mM, but 2.5mM – 3mM seemed to give nominal results. 0.1µM Primer was used, although the one forward primer, HMND1d was degenerate and the concentration was doubled (N.Sweijd, personal communication). 2.5mM dNTP's were used together with a range of DNA concentrations. The PCR cycles varied between 35 and 37. Different PCR programs were also used during the research, but the following conditions wore most often applied. A Denaturing step of 92°C for 3min was followed by 37 cycles of: 92°C for 30sec, a gradient step; 38°C-56°C for 50sec and an Annealing step of 72°C for 1min. The program ended with a 72°C step for 3min and 4°C indefinitely.

5.2 Gel preparation

For each PCR reaction done, a 1% Agarose gel was prepared by adding 1.2g Molecular Grade Agarose (Whitehead Scientific) to 120ml 1 x TBE buffer. EtBr was added for visualization of the gel. Each sample was prepared by adding 1.5µl of 6 x loading buffer to 2µl of dH_2O and 3.5µl of DNA. A Promega 100bp as well as a λHindIII ladder was

used as a size marker. The gel was run at 100V for 60min on a Pharmacia LKB GNA 200 Gel apparatus, after which it was visualized under aUV light on the Geldoc™ system.

5.3 Restriction enzyme (RE) digest

The PCR products with primers HMND1d & c16sar (1600bp) and HMcon & HMND3 (1700bp) (combinations 1 and 2 respectively) were cut with restriction enzymes HhaI and HaeIII (Promega), both 4-base cutters, for confirmation. (Sweijd, 1999) 8µl of PCR product was added to 2µl of RE Buffer, 2µl of the respective enzymes and 8µl of ddH_2O. The cocktail was left in a 37°C waterbath overnight, after which it was visualized on a 2% Agarose gel prepared with 1 x TBE and EtBr. A 100bp Promega ladder was used for fragment size analysis.

5.4 BLAST and CLUSTAL analysis of primer sequences

Alignment of the primer sequences to those of the clone sequences (Sweijd, 1999) (Appendix A, Figure A1) was done using BioEdit Sequence Alignment Editor. (Clone sequences were developed (Sweijd, 1999), by aligning mollusk sequences from GenBank (www.ncbi.nlm.nih.gov), for the development of *Haliotis* primers.)

This was followed by a CLUSTAL alignment (BioEdit Sequence Alignment Editor), of the four primer sequences to the mitochondrial genomes of seven mollusks that were downloaded from GenBank. They were: *Cepaea nemoralis* (NC_001816), *Albinaria coerulea* (NC_001761), *Pupa strigosa* (NC_002176), *Crassostrea gigas* (NC_001276), *Venerupis philippinarum* (NC_003354), *Loligo bleekeri* (NC_002507) and *Katharina tunicata* (NC_001636). Because no significant similarity was found in any of the CLUSTAL results a Nucleotide BLAST was done for short nearly exact sequences, using the mitochondrial database. It must be noted however that these alignments were done with a general mitochondrial database, and not specifically for that of molluscs, because of the fact that the CLUSTAL analysis had already been done.

After the alignment of primer sequences to mollusk sequences, the HMND1d and HMND3 primers were aligned - using BioEdit Sequence Alignment Editor, against *ND1* and *ND3* gene sequences. The gene sequences were downloaded from four

mollusk sequences: *Cepaea nemoralis* (NC_001816), *Albinaria coerulea* (NC_001761), *Pupa strigosa* (NC_002176) and *Venerupis philippinarum* (NC_003354), respectively.

6. AFLP analysis using Mse/Mlu primer combination

6.1 Digestion

A cocktail, containing One-phor-all buffer, 4U/µl MseI, 10U/µl MluI (New England Biolabs) and dH_2O was prepared for the total number of reaction plus one extra. The cocktail was then added to 100 - 200ng DNA to a final volume of 25µl. The samples were incubated at 37°C for 1h before adding ligation solution.

6.2 Ligation

A cocktail, containing One-phor-all buffer, 10mM ATP, 50pmol MseI adapter, 5pmol MluI adapter, T4 ligase (USB Corporation) and dH_2O to a final volume of 5µl was prepared for the total number of reactions plus one extra. The cocktail was added to the 25µl digestion solutions, to a volume of 30µl and incubated in a 37°C waterbath overnight.

6.3 Collection of biotinylated Mlu-fragments with Dynabeads

A volume of beads, equivalent to the number of samples (10µl per sample) was pipetted into an Eppendorf tube. Wash solution, comprising of 10mM Tris-HCl and 0.1mM EDTA, and 100mM NaCl in a volume equivalent to that of the beads, was added. The beads were then collected with a magnet and the wash and collection steps were repeated 3 more times. The beads were resuspended in the same volume of wash solution. 10µl of beads were added to each 30µl digested/ligated DNA sample, and lightly mixed. The mixture was then incubated on ice for 30 min with gentle agitation every 5-10 min. 60µl of wash solution (TE, 100mM NaCl) was added. The beads were once again collected using a magnet, and washed using100µl of wash solution. The washing and collection of beads were repeated 3 more times. The remaining Mlu-Mse and Mlu-Mlu fragments were suspended in 50µl of TE buffer and stored at -20°C.

6.4 Labeling

A cocktail, containing 10x Kinase Buffer, 300ng Mse Primer, Polynucleotide Kinase:Dilution Buffer 1:3 (USB Corporation), ^{33}P (10µCi/µl, 25.4 days half life) and

dH_2O to a final volume of 21µl, was prepared. The cocktail was incubated at 37°C for 1-2h and then at 65°C for 10min. The labeled primers were then stored at -20°C for later use.

6.5 PCR amplification

The reactions were performed on a MJ Research DNA Engine Peltier Thermal Cycler PTC-200. Two cocktails were prepared, both for the total number of reactions plus one extra. The first containing 10x Taq buffer (Bioline), 2.5mM $MgCl_2$, 30ng unlabeled primer and dH_2O to a final volume of 315µl. 7.5µl cocktail was added to the 0.55µl DNA samples, as well as a negative control. The second cocktail, containing 50µM dNTP's, 30ng labeled primer, 0.5U Taq polymerase (Bioline) and dH_2O to a final volume of 84µl. 2µl of the second cocktail was added to the DNA and negative control samples, and placed in the PCR machine. Thus: 7,5µl Cocktail A was added to 0.55µl of DNA after which 2µl Cocktail B was added to the reaction mixture.

The PCR program consisted of the following steps: 72°C for 1min and 94°C for 2min, was followed by a 12 cycle step of 94°C for 20sec, 65°C for 30sec, 72°C for 2min. This was followed by a 25 cycle step of 94°C for 20sec, 56°C for 30sec and 72°C for 2min. Reaction then went to a 72°C step for 30min. In the end, 5µl of formamide loading buffer (98% formamide, 10 mM EDTA, 1 mg/ml of each bromophenol blue and xylene cyanol) was added to each PCR sample and stored at -20°C.

6.6 Gel preparation

6.6.1 Glass plates

The steps that were taken with the first use of the plates consisted of placing them in a large rectangular plastic container with 2M NaOH overnight to strip them. The plates were washed with Alconox (B&M Scientific), thoroughly rinsed with dH_2O and left to dry. Both the inner sides of the plates were wiped with Repelcote (Saarchem) and left for 10min. The plates were once again rinsed and left to dry, ready to be assembled with 4mm spacers on the side, and casting boat around the plates. Clean plates were essential, and were wiped with 100% EtOH prior to assemblage.

6.6.2 4% denaturing polyacrylamide (PAA) gel

25g Urea (Riedel-de-Haën), 6ml 10 x TBE, 6ml 40% PAA Gel mix (Promega) and dH_2O was added to a final volume of 60ml. The mixture was dissolved using a magnetic stirrer, after which 320µl 10% ammonium persulphate (APS) and 40µl TEMED was added to polymerize the gel mixture. The mixture was poured between the glass plates, avoiding the formation of bubbles, and the combs inserted with the flat side in the gel. The opening of the gel was clamped using two large paper clamps and left to polymerize for 45min. If the gel was left overnight, care was taken to cover the wells with tissue paper soaked in 1 x TBE and sealed with wrap, to ensure that the gel did not dry out.

6.6.3 Running gel

Once the gel had polymerized and the clamps and casting boat removed, the PAA residue was rinsed off the outside of the plates using running tap water. The plates were then dried and fitted onto the gel apparatus. 1 x TBE was added to the bottom and top chamber ensuring that the top chamber was tightly shut. The combs were removed and the wells flushed of any urea and PAA residue. The gel was pre-run at 60W for 30min. The PCR samples were denatured at 90°C for 5min and directly chilled on ice. The gel run was stopped and the wells were once again flushed. The combs were then inserted with teeth ±2mm below the gel surface. 4µl of each PCR sample was loaded, and the gel run at 60W for approximately 2h.

6.6.4 Gel exposure

The plates were separated using a plate separator and 3MM Whatman chromatographic paper was placed on top of the gel. The entire surface of the filter paper was wiped to ensure that the gel adhered to it. The excess paper was cut of and wrap was placed over the entire surface of the gel. The gel was then placed onto a gel drier, between two sheets of Whatman paper, and dried at 80°C for 1½h. Once the gel was dry, a Geiger counter was used to determine the radioactivity and was then placed into an X-ray cassette. In a dark room, an X-ray film was placed on top of the gel, making sure that the film was in correct orientation. The gel was then exposed in relation to the amount of radioactivity shown. Refer to Table A3 in Appendix A for day on which P^{33} was used, and corresponding decay factor.

6.6.5 Gel development

Film (Sigma BioMax MR-1 30×40cm) was placed in the developer according to manufacturers instructions, followed by a brief wash in the H_2O. The film was then gently shaken in fixative and once again washed in the H_2O. The film was hanged to dry, ready for analysis.

7. Statistical analysis of AFLP data

7.1 Binary data in Microsoft Excel

The AFLP gels were analyzed for polymorphic bands between the two or three populations compared. Bands that were present were indicated with a 1 and those absent with a 0. A binary data sheet was prepared for each of the primer combinations tested between the different populations in Microsoft Office Excel 2000, and the bands present for each marker, per population, was calculated. This data was then used to construct contingency tables for Probability calculations in RxC (Miller, 1997)

.

7.2 Contingency tables and exact tests

Probability values were calculated using RxC (Miller, 1997), a Windows program that performs Fisher's Exact test on any sized contingency table through the use of the Metropolis algorithm. The calculations were done in 20 batches of 2500 replicates per batch performed, with 50000 total replicates.

7.3 Analysis of Molecular Variane (AMOVA)

The Analysis of Molecular Variance (AMOVA) was performed, using the software program GenAlEx (Peakall and Smouse, 2001). The estimation of pairwise binary genetic distances for dominant data followed the method of Huff et al. (1993) in which any comparison with the same allele has a value of 0, both 0 vs. 0 comparisons and 1 vs. 1 comparisons are counted, while any comparison with a different allele 0 vs. 1, has a value of 1. When calculated across multiple loci for a given pair of samples, this is equal to the tally of band differences between the two DNA profiles:

$$D = n\left[1 - (2n_{xy}/2n)\right]$$

where $2n_{xy}$ equals the number of shared bands and n equals the total number of banding positions. This is a Euclidean distance metric, unlike related methods such as Nei's 1 – F, and was therefore appropriate for a subsequent Analysis of Molecular Variance, which required a Euclidean metric (Excoffier et al. 1992, Huff et al. 1993).

The Analysis of Molecular Variance (AMOVA) procedure followed the methods of Excoffier *et al.* (1992), Huff *et al.* (1993), Peakall *et al.* (1995), and Michalakis and Excoffier (1996). In the study by Excoffier *et. al.* (1992), the analysis of molecular variance (AMOVA) produced estimates of variance components and F-statistic analogs, designated as Phi-statistics. The method was flexible enough to accommodate a number of alternative input matrices, corresponding to different types of molecular data, as well as different types of evolutionary assumptions, without modifying the basic structure of the analysis.

Binary Distance Matrixes were used for the calculation of PhiPT values. The significance of the variance components and Phi-statistics was tested using a permutational approach, eliminating the normality assumption that is conventional for analysis of variance but unsuitable for molecular data. Among- and Within Population variance was also calculated.

7.4 Principal Coordinates Analysis (PCA)

The relationship between distance matrix elements based on their first two principal coordinates were plotted. Five different groups were plotted, consisting of the East Coast/West Coast, West Coast/Broodstock, Broodstock/Progeny, West Coast/Progeny and West Coast/Broodstock/Progeny populations. GenAlEx allows PCA analysis of pairwise PhiPT matrices, and a scatter plot of the first two coordinates is generated. The procedure is based on an algorithm published by Orloci (1978).

7.5 Similarity index analysis

NTSYSpc 2.0 (Rohlf, 1997) was used to perform Similarity Index analyses between the different populations: West/East Coast, West Coast/Broodstock, Broodstock/Progeny and West Coast/Progeny. The binary data of the different populations were converted into triangular matrixes using the Kulczynski (1927) formula,

$$\frac{1}{2}\left(\frac{a}{a+b}+\frac{a}{a+c}\right)$$

Where *a*, *b* and *c* are defined as follows for a two way frequency table for all pairs of two object i and *j* :

		j	
		+	-
i	+	**a**	**b**
	-	**c**	**d**

The triangular matrix was then used in Microsoft Excel, to calculate an average Similarity per analysis.

RESULTS

1. Extractions

Although original extraction results seemed good, with high quantities of DNA present when visualized on Agarose gels (0.01µg/µl - 0.1µg/µl), difficulties were experienced during the amplification of the universal- (Kocher *et. al.*, 1989), as well as the Sweijd, 1999 *Haliotis* primers. It was thought that the quality of DNA might be one of the problems, and the samples were extracted for a second time, using a CTAB extraction protocol. After repeated failures during the amplification procedures, and consultation with Drs. Sweijd and D'Amato, DNA extracted from gill tissue was considered a better source, and would possibly present improved results. Once again the extraction procedure was redone for a few samples, this time using gill tissue. High quality and quantity of DNA was extracted.

2. mtDNA amplification using universal primers

Universal primers (Kocher et.al. 1989); ND5/6, CO1, CytB/C, ND1, NADH2 and ATP6/COIII as well as a shortened CytB primer (Kocher et.al. 1989), were used in the amplification of mtDNA for population analysis. Minimal results were obtained during the use of these primers, and the CO1 and CytB primers were the only two where optimal conditions were established. The CO1 product was expected to be around 600-800bp in length, and amplified at a MgCl concentration of 2mM and 56°C. The CytB product, which was 360bp in length, amplified at a MgCl concentration of 2.5mM and 56°C. The CO1 primer was invariable and was not considered suitable for population analysis (D'Amato, personal communication). Whilst attempting to optimise the conditions, most of the primers amplified unspecific product, but never product of the correct size.

3. mtDNA amplification using *Haliotis* primers

Haliotis primers, HMND1d, c16sar, HMcon and HMND3, obtained from Dr. N. Sweijd, were used in the amplification of mtDNA of abalone. HMND1d and c16sar (primer combination1), was expected to give a product with a size of 1550bp and the HMcon and HMND3 combination (primer combination2), a 1650bp product. Initial experiments, as done by Sweijd (1999), did not show any results, with no amplification

of DNA at all. It was at this point that extractions were redone for a second time, and the DNA used in further experiments. Occasional results were mostly unspecific product, with the correct fragment sizes being amplified only four times out of about 400 reactions, thus only four individuals amplified out of 400 tested. Amplification with the primers was not consistent and not all the individuals amplified, regardless of the quality of DNA.

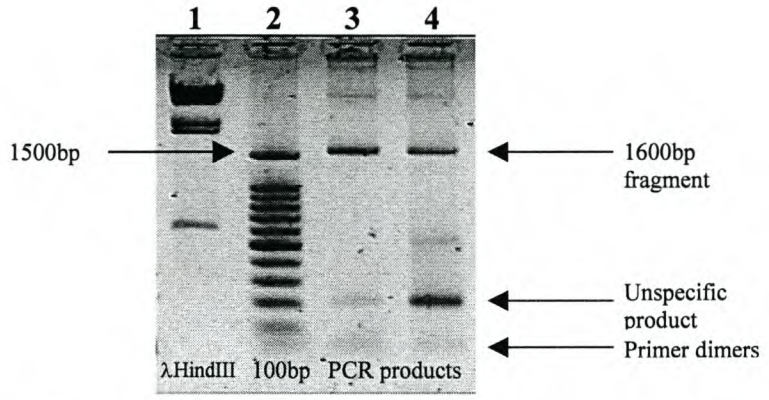


Figure 5: Example of PCR products obtained with primer-combination 1. Lanes 3 and 4 are the PCR products with the supposed 1550bp fragments.

Verification, that failure to amplify was not because of DNA template quality, was needed, and extracted DNA from gill tissue was used in a small amount of PCR reactions. However, no product was seen. A few random samples showed a strong PCR product using the *Haliotis* CO1 primers from Metz *et al.* 1998 (D'Amato, personal communication), which confirmed that the template quality was indeed good. Restriction digestion, with HhaI and HaeIII, of one of the 1550bp fragments (Figure 5), was done. When run on a higher resolution gel (Figure 6) (2% Agarose) the three resulting fragments seemed to run at 100bp, 300bp, and 1250bp with HhaI, and at 100bp, 300bp and 1150bp with HaeIII. The HaeIII digest results was contrary to that obtained from Sweijd (1999) as fragments of 955pb, 522pb, and 74bp were expected.

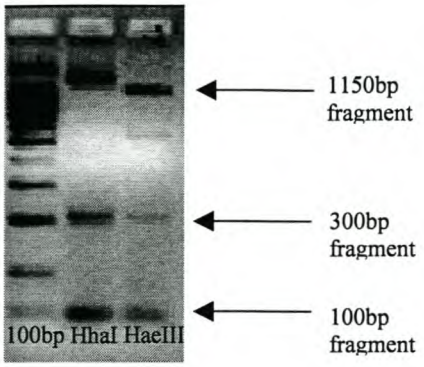


Figure 6: Restriction Digest of *Haliotis midae* 1550bp fragments with HhaI and HaeIII.

Primer design and amplification of *Haliotis* primers, as discussed by Sweijd (1989), was used as a reference during this study, and although all steps were taken to ensure amplification, failure to do so led to verification of sequence similarity and position through BLAST and CLUSTAL analyses

With the alignment of the primer sequences to those of the clone sequences obtained from Dr. Sweijd (Appendix A, Figure1) the HMND1d, c16sar, as well as the HMcon primers aligned with clone 1 when in complement reverse. The HMND3 primer was the only one to align with clone 2.

Table 1: *Haliotis* primer sequence alignments with that of clone sequences used during primer development (Sweijd, 1999). Clone alignment position indicated next to clone sequence.

Sequence	Name	Alignment of sequences
Clone 1	**Hmmt-Cl1**	428 3'**GTGTTTTAGAGGTGAAAATACCGATTGGT**'5 458
Haliotis F	HMND1d	**---TTGCWGAAG-GAGAATCYGAGYTAGT**
Haliotis R	c16sar	**ATGTTTTTGCTA--AACAGGCGAGG----**
Clone 1	**Hmmt-Cl1**	151 3'**GGGGGGGTTTTGGTGT-AGTA--------**'5 179
Haliotis F	HMcon	**GGGGGGGTTTTGGTGTTAGTA--------**
Clone 2	**Hmmt-Cl2**	61 3'**GAGCCTTCATTTCATTCGTGGA-------**'5 88
Haliotis R	HMND3	**GAGCCTTCATTTCATTCGTGGA-------**

CLUSTAL multiple allignments, on BioEdit Sequence Alignment Editor, was used to align the four primer sequences to the mitochondrial genomes of seven molluscs. They were: *Cepaea nemoralis* (NC_001816), *Albinaria coerulea* (NC_001761), *Pupa strigosa* (NC_002176), *Crassostrea gigas* (NC_001276), *Venerupis philippinarum* (NC_003354), *Loligo bleekeri* (NC_002507) and *Katharina tunicata* (NC_001636). No

significant similarity was found in any of the CLUSTAL results. Following the CLUSTAL analysis, a Nucleotide BLAST was done for short nearly exact sequences, using the mitochondrial database. All of the primers, except HMcon, gave the expected results, i.e. aligned with the expected gene, when BLASTed to the different mitochondrial genomes. However, these results varied in the amount of bases that aligned. It must be noted however that these alignments were done with a general mitochondrial database, and not specifically for that of molluscs.

CLUSTAL multiple alignments were done once again. The HMND1d and HMND3 primers were aligned against ND1 and ND3 gene sequences of four molluscs: *Cepaea nemoralis, Albinaria coerulea, Pupa strigosa* and *Venerupis philippinarum,* respectively. The alignments showed poor sequence similarity, with primer bases matching with sequence bases ranging from 28 – 54% (Table 3 and Table 4). These alignments can be viewed in Table 2.

Table 2: *Haliotis* HMND1d and HMND3 primer alignments with various mollusc gene sequences

HMND1d	**----ACTARCTCRGATTCTCCTTCWGCAA---**
Albinaria coerulea	ATCA**TCTAACTCTAAATATTCTTTTTTAG**GGG
Cepaea nemoralis	GGCG**TCAAACTCAGCCTACTCTTTCTTAG**GGG
Pupa strigosa	GGCC**TCTAACTCTAAATACGCGTTTTTGG**GAG
Venerupis philippinarum	ATCT**TGTGATTCACGATATGGCTTTTTAG**GGG
HMND3	**----G---AGCCTTCATTTCATTCGTGGA---**
Albinaria coerulea	ATCT**TAACAGTATTATTTTTAATTTTTGA**CGT
Cepaea nemoralis	TACA**TTAGTTATCTATTCCACGTGTTTGA**CGT
Pupa strigosa	CTTT**TAGTAATCTTGTTCTTGGTATTTGA**CGT
Venerupis philippinarum	TTAC**TAGCCCTCCTTTTTTTAGTTTTTGA**CTT

Table 3: Summary of the amount of base alignments of mollusc sequences with primer sequence, HMND1d.

Sequence	Amount of bases showing similarity to primer sequence	Total bases	% Similarity
HMND1d		25	
Albinaria coerulea	13		52
Cepaea nemoralis	11		44
Pupa strigosa	10		40
Venerupis philippinarum	7		28

Table 4: Summary of amount of base alignments of mollusc sequences with primer sequence, HMND3.

Sequence	Amount of bases showing similarity to primer sequence	Total bases	Percentage Similarity
HMND3		22	
Albinaria coerulea	12		54.5
Cepaea nemoralis	6		27
Pupa strigosa	9		41
Venerupis philippinarum	9		41

This confirmed the suspicion of the degenerative nature of both the primers.

4. AFLP analysis using Mse/Mlu primer combination

4.1 Scoring of AFLP gels

In the AFLP analysis, a total of 19 primer combinations were studied: 3 for each of the West Coast and Broodstock analysis and the Broodstock and F1 analysis. 7 primer combinations were used for the West Coast and F1 analysis, and 6 combinations for the East- and West Coast analysis. These primer combinations can be viewed in Table 5, all consisting of *Mlu* and *Mse* primers with 3 selective bases each. The number of amplified products varied between 30 and 100, with an average of 64.3 products per primer combination of which 64.18% were polymorphic. (Table A4 – Appendix A) The

bands scored (demonstrating variation amongst individuals) per analysis ranged from 19 to 44, although the analysis that presented only 19 polymorphic bands, *Mlu*p8/*Mse*p5, was redone with another primer combination *Mlu*p5/*Mse*p3.

Table 5: Primer combinations used during AFLP analysis.

East/West Coast	West Coast/Broodstock	Broodstock/F1	West Coast/F1
Mlup2/Msep8	Mlup2/Msep8	Mlup2/Msep8	Mlup2/Msep8
Mlup3/Msep7	Mlup3/Msep7	Mlup3/Msep7	Mlup3/Msep7
Mlup5/Msep5	Mlup5/Msep5	Mlup5/Msep5	Mlup5/Msep5
Mlup4/Msep3			Mlup8/Msep2
Mlup4/Msep8			Mlup8/Msep5
Mlup2/Msep5			Mlup3/Msep2
			Mlup5/Msep3

A more complete version of Table 6 is available in the Appendix as Table A4.

West/East Coast			West/Broodstock			Broodstock/F1			West Coast/F1				Population
5/5	3/7	2/8	5/5	3/7	2/8	5/5	3/7	2/8	3/2	8/2	5/3	8/5	Primer Comb. Mlu/Mse
22	21	40	23	25	27	26	39	44	35	31	26	19	Polymorphic bands
30	40	57	66	66	63	40	84	62	100	62	71	67	Total number of bands
73.3	52.5	70.2	34.8	37.9	42.9	65	46.4	71	35	50	36.6	28.4	% Polymorphic bands
		65.3			38.5			60.8				37.5	Average % per pop.
												47.8	Average total %

Table 6: Analysis of AFLP assays

Twenty individuals from each of the West Coast, Broodstock and F1 populations were included in the AFLP analysis. However, only thirteen samples of the East Coast population were available for comparison with the West Coast. For the West Coast/Broodstock/F1 as well as the East/West Coast analysis, where the data of three primer combinations were pooled, only 16 polymorphic bands were comparable over the gels.

4.2 Exact test calculations

From all these analyses the Probability value (p) was calculated on RxC. A significance level of 0.05 was used, and according to results obtained, there was no difference between pairwise comparisons. All the p values (except one) were considerably higher

than 0.05. Table 7 summarizes these results. See Appendix for complete calculations of *p* values.

Table 7: Probability values of different primer combinations.

Populations compared	Primer combination	Probability value
West/East Coast	Mlu2/Mse8	0.7236
	Mlu3/Mse7	0.1850
	Mlu5/Mse5	0.8461
	Mlu2/Mse5	0.9585
	Mlu4/Mse3	0.9619
	Mlu4/Mse8	1.0000
West Coast/Broodstock	Mlu2/Mse8	0.7618
	Mlu3/Mse7	0.7399
	Mlu5/Mse5	**0.0676**
Broodstock/F1	Mlu2/Mse8	0.1247
	Mlu3/Mse7	0.9441
	Mlu5/Mse5	0.8716
West Coast/F1	Mlu3/Mse2	0.6147
	Mlu5/Mse3	0.9794
	Mlu8/Mse2	0.9724
	Mlu8/Mse5	0.9399

The *p* value calculated for West Coast/Broodstock Mlu5/Mse5 combination was a great deal lower than that of the other combinations, at ~0.068.

4.3 AMOVA and PCA analyses

Analysis of Molecular Variance (AMOVA) genetic distance calculations with GenAlEx, gave PhiPT values that varied between 0.000 and 0.149 and indicated very low (0% - 15%) Among Population variance, with the most variation found Within Populations. The West Coast/Broodstock/F1 population analysis gave a PhiPT value of 0.081 indicating an 8% Among Population variance. The values are summarized in Table 8a and 8b. See Appendix for complete AMOVA and PCA analyses.

Table 8a: AMOVA distance values of pairwise population comparisons, with individual primer combinations.

	West/East Coast			**West/Broodstock**			**Broodstock/F1**		
	5/5	3/7	2/8	5/5	3/7	2/8	5/5	3/7	2/8
PhiPT value	0.000	0.117	0.081	0.149	0.131	0.004	0.035	0.012	0.042
PhiPT probability.	0.547	0.007	0.183	0.001	0.001	0.345	0.037	0.206	0.003
Among pop. variance	0%	12%	2%	15%	13%	0%	3%	1%	4%
Within pop. variance	100%	88%	98%	85%	87%	100%	97%	99%	96%

Table 8b: AMOVA distance values of population comparisons, with pooled results of three primer combinations: Mlu/Mse - 5/5, 3/7 and 2/8

	West/East Coast	West/Broodstock	Broodstock/F1	West/F1	West/Broodstock/F1
PhiPT value	0.000	0.071	0.081	0.092	0.081
PhiPT probability	0.544	0.001	0.001	0.001	0.001
Among pop. var.	0%	7%	8%	9%	8%
Within pop. var.	100%	93%	92%	91%	92%

Table 9: Principal Coordinates Analysis data of the West Coast/Broodstock/F1 study.

Axis	Percentage	Cumulative Percentage	Eigen value Total	Total of first 15 Eigen values	Percentage
1	12.18	12.8			
2	11.31	23.49			
3	6.65	**30.14**			
			43.765	35.328	
					80.722

Table 9 summarizes the PCA results of the West Coast/Broodstock/F1 study, where the data of all three primer combinations: Mlu/Mse - 5/5, 3/7, 2/8 were used. The cumulative percentage was very low, only 30.14%, and the first 15 components accounted for 80.772% of the total variance.

Principle Coordinate Analyses (PCA) and Analysis of Molecular Variance (AMOVA) pie chart of all the populations analysed done, can be viewed in the Appendix with that of the West/Broodstock/F1 analysis (Figure 7 and Figure 8) shown below.

Figure 7: PCA of West/Broodstock/F1 population analysis. Data series 1 - 20 represents data from the West Coast, 21 - 40 that of the Broodstock and 41 - 60 the F1 population.

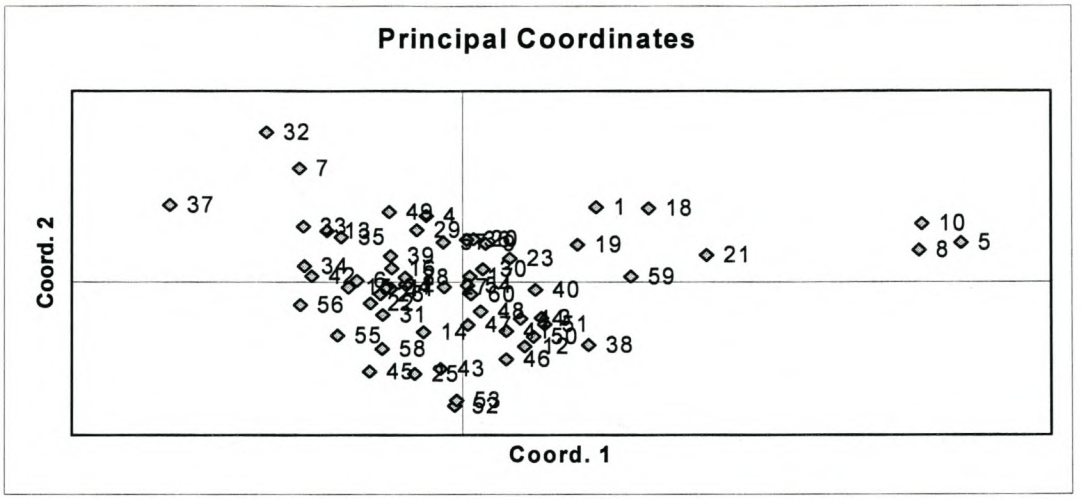
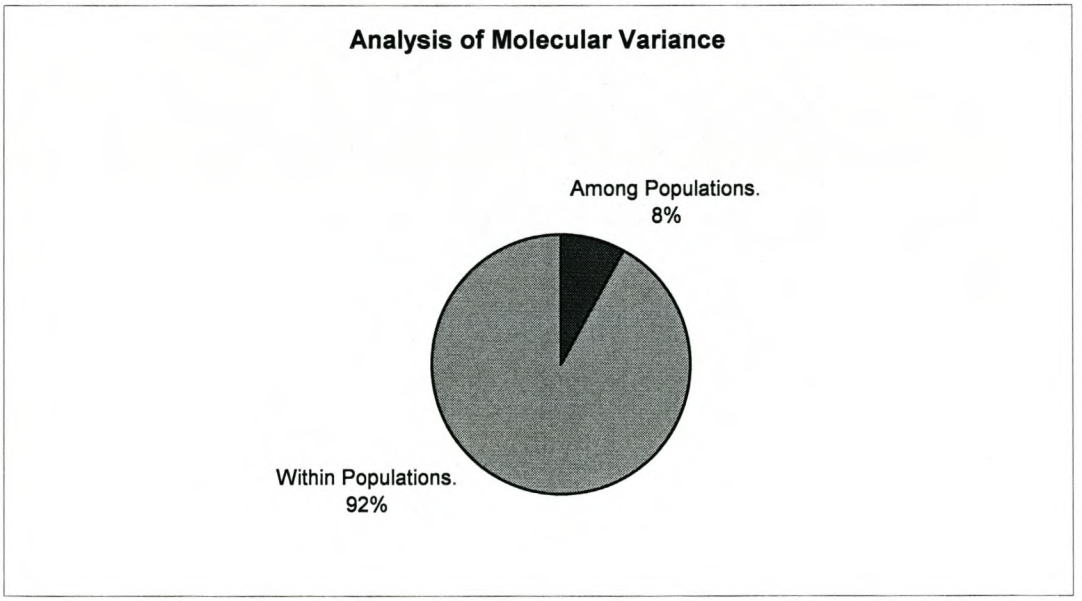


Figure 8: AMOVA of West/Broodstock/F1 population analysis.



4.4 Similarity index analysis

The individual Similarity Index results of the populations analysed: West/East Coast, West Coast/Broodstock, Broodstock/F1 and West Coast/F1 all lay between 0.7293 and 0.7978 with the individual values as follows:

Table 10: Similarity values of pairwise population comparisons

West/East Coast	0.729252269
West Coast/Broodstock	0.797847526
Broodstock/F1	0.774147154
West Coast/F1	0.774552047

DISCUSSION AND CONCLUSIONS

1. Amplification failure during mtDNA analysis

1.1 Universal primers

Because the primers (Kocher *et. al.* 1989) were developed as universal primers, and not primers specific for abalone, *Haliotis midae*, the optimisation was difficult, and seldom gave clear bands of the correct size. A large number of undesirable, unspecific products were amplified, even though a wide range of conditions was tried. Amplification failure could have been caused by presence of mutations in the priming sites, mitochondrial rearrangements or pseudogenes.

1.2 *Haliotis* primers

1.2.1 Extraction – template quality

Failure to amplify was originally thought to be because of DNA template quality, as extractions were done using a salting-out method that was different to that used by Sweijd (1989) during primer development and amplification. The protocol was adapted to include 1% CTAB (sigma) in saturated NaCl instead of SDS, as CTAB removes mucopolysaccharides that are abundant in molluscs, and interfere with DNA preparation protocols. (Rumpho et al. 1994.) Amplification was still not achieved, and gill tissue was used for a repeated extraction as a tissue free of mucopolysaccharides. However, even this did not produce acceptable results. As an alternative control to test the quality of DNA template, a CO1 primer (Metz et. al. 1998), specific to *Haliotis midae*, was amplified using original extractions (D'Amato, unpublished results). The results were positive, which indicated that the DNA template quality was not the contributing factor in the amplification failure.

1.2.2 CLUSTAL and BLAST alignments

Even though all steps were taken to ensure amplification, failure to do so led to verification of sequence similarity and position through BLAST and CLUSTAL analyses. The clone sequences that were used during the development of the *Haliotis* primers (Sweijd, 1999) were aligned against the 4 *Haliotis* primer sequences, and it was found that 3 of the sequences aligned with Clone 1, and 1 with Clone 2. As Clone 1 contained sequences of the 16sRNA region, and Clone 2 sequences from the Co1 and

ND3 regions, the HM16sar primer would align with Clone 1 and the HMcon and the HMND3 with Clone 2, and this was not the case. Further CLUSTAL analysis showed no significant similarity, which pointed towards the fact that the primers would not amplify under the desired stringent conditions. This was confirmed when the primer sequences were aligned with their corresponding gene sequences, and only 28-54% similarity was found.

2. Population structure analysis

2.1 Hypotheses tested

Two hypotheses (H_o) were tested; the first H_o being concerned with the population genetic structure of the $Wild_{(WC)}$ samples vs. the $Wild_{(EC)}$ samples, stating that there is no genetic difference between the two wild populations. The second H_o was focused on the potential loss of genetic diversity from the Wildstock (West Coast) samples, through the Broodstock to the Progeny populations on the farm. It was suspected that small population size, selective management practices and the founder effect might influence these genetic structures. These hypotheses were tested with the AFLP technique.

2.2 AFLP profiles

19 primer combinations were used during the analysis, as verification of original results was needed, and all possible pairwise comparisons were made. As discussed earlier, the number of overall different amplified products varied between 30 and 100 per primer combination, with an average of 64.3 products per primer combination, of which 64.18% were polymorphic. Even though individual results varied between 28% – 73% polymorphism (Table A4), only small differences were observed between different individuals within samples. The number of variable bands scored per pairwise comparison ranged from 19 to 44. This can be compared with a study done by Wilding *et. al.* (2001) where 5 AFLP primer combinations were tested on 50 samples each of 5 groups of *Littorina saxatilis,* an intertidal snail. The number of amplified products varied between 43 and 83 per primer combination, and levels of polymorphism were particularly high with 94,8% of the bands being polymorphic.

2.3 AFLP's polymorphisms

Non-neutral amplification was a problem considered during this study, as AFLP's are aimed to be neutral, and even though they allegedly are, the genome is scanned at

random and this may not be the case. The AFLP's were considered sufficient for preliminary analysis of data. However, due to the aspect of non-neutral amplification it is recommended that either a combination of systems giving complementary information or another system needs to be utilized for additional information or further analyses.

2.4 Exact test calculations of genetic differentiation

Fisher's Exact test was used to assess genetic similarity between populations by calculating Probability-values of pairwise comparisons between populations. The *p* value is an indicator of distribution of genetic variation, with a value less than 0.05 as an indicator of significant genetic difference. All the results (Table 7) ranged between 0.1247 – 1.0000, except one at 0.0676, and was much higher than the significance level. It was therefore concluded that the range of populations showed no significant difference in genetic variation and diversity when compared to one another.

One *p* value at 0.0676 for the $Wild_{(WC)}$/Broodstock Mlup5/Msep5 analysis was an exception. It approached the significance level of 0.05, but in view of the fact that all other results consistently illustrated that there was no loss of diversity, it was not taken into regard. Even though the AFLP technique is assumed to be anonymous, it is not known what the nature of the amplified product is, and the value could have been affected by selection. It could also be ascribed to random sampling effects. It is recommended that for further assessments more markers should be used.

Another important observation that was highlighted by this analysis (and later the AMOVA and PCA results) was that there was no difference observed in genetic variation between the two geographically distinct wild populations from the East Coast and the West Coast. A similar conclusion was drawn by Sweijd (1999) on the basis of molecular marker assessment. He proposed that the eastern populations of *H. midae* are the result of an isolated introduction from established populations in the west, and a subsequent easterly range expansion. Evidence from allozyme, mitochondrial DNA and microsatellite data by Evans *et. al.* (unpublished) on these populations supports this isolated introduction event to the east of Cape Agulhas, and subsequent range expansion in an easterly direction. This range expansion is expected to be because of wind driven

currents along the East Coast steering the abalone in a more northerly direction. This phenomenon indicates a Founder effect, with fairly recent and random distribution.

The H_o's are therefore confirmed, where the similarity between the two wild populations; West Coast and East Coast, can be ascribed to the founder affect. Similarity between the $Wild_{(WC)}$ and the Broodstock could be due to random sampling, sufficient numbers sampled whilst the Broodstock was acquired and the fact that the two populations are from the same geographical area. The Broodstock/Progeny similarity can be due random mating systems, unequal sex ratio and the contribution of mass spawning seems unfounded. Repetitive spawning could also play a role, as it ensures that all the males contribute equally to the offspring. The $Wild_{(WC)}$/Progeny data should by implication then also indicate no loss of diversity.

2.5 Assesment of genetic variation through AMOVA calculations and F-statistics

The Analysis of Molecular Variance (AMOVA) further assessed the genetic differentiation between populations. The calculations gave PhiPT values (Table 8b) that varied between 0.000 and 0.149. These results also consistently, but independently, indicated very low (0% - 15%) Among Population variance, with the most variation found Within Populations. This means that most of the observed variation is found between individuals of a population rather than between different populations. With a value of 0 indicating no difference, this observation further confirmed the lack of significant difference between the populations as assessed. In addition, the AMOVA results supported the conclusion drawn i.e. that of similar genetic profile, concerning the two wild populations, East- and West Coast as discussed earlier.

2.6 Grouping of populations through Principal Coordinates Analysis (PCA)

The results presented in the PCA figures in Appendix A repeatedly disclosed no observed significant population structure amongst the populations. In the case of differentiation between populations, data would have clustered in distinctive zones with little overlay. This is confirmed by PCA data (e.g. Table 9) where the cumulative percentage onto the 3rd axis explain in the range of 30% of the total variance as apposed to 75% as an indicator of significant population structure, i.e. genetic differentiation.

2.7 Similarity Index analysis

The Similarity Index is a further indicator of genetic variation within and between populations. It is generally defined as the fraction of shared bands expressed as a value that varies between 0 and 1. Analysis of the pairwise comparisons between the $Wild_{(WC)}$/$Wild_{(EC)}$ (0.729), $Wild_{(WC)}$/Broodstock (0.797), Broodstock/Progeny (0.774) and $Wild_{(WC)}$/Progeny (0.774), all indicated a very high similarity. The high similarity as observed is an indication of genetic similarity between populations. This once again supported previously drawn conclusions.

2.8 Microsatellite data on farm populations

In a parallel study, done by Van Zyl (unpublished), the genetic differentiation between *Haliotis midae* populations (Wild $_{(WC)}$, Broodstock and Progeny) was studied by the use of 5 microsatellite loci. The primers were non-specific however, as they were originally designed for the focal species, *Haliotis rubra*. These populations were analysed with the same samples that were used during the AFLP study, calculating the departure from Hardy-Weinberg Equilibrium (HWE), F_{IS}-, F_{ST}- and Heterozygosity values. The HWE Probability values were significant, indicating that allele & genotype frequencies were not in equilibrium. Significant F_{IS} values also indicated a departure from HW equilibrium, and can be a measure of either inbreeding, non-random mating, null-alleles or the Wahlund effect. Only one locus showed HWE, signifying that the overall system was in disequilibrium, and only two of the overall Heterozygosity values indicated deficiency. Even though the test of overall F_{ST} values supports the finding of the AFLP analysis - illustrating no loss of variance, it must be noted that the microsatellite data might not be accurate, because of suspected null-alleles, and that further studies need to be done with new isolated microsatellite markers specific for *Haliotis midae*.

2.9 Conclusions drawn on commercial management and genetic improvement strategies.

1) There is no subsequent loss of genetic variation observed from the $Wild_{(WC)}$ through Broodstock to the commercial stock. This reflects that current management and operational procedures is sufficient in terms of the maintenance of genetic variation over the short term.

2) Levels if genetic variation is high and compares favourable to the Wild $_{(WC)}$. The current Broodstock provides ample potential as base material for future genetic improvement programs.
3) Crossbreeding between geographically distinct populations i.e. the West Coast and East Coast populations is not expected to yield improvement due to lack of observed genetic variation.
4) Future genetic selection programs should focus on the exploitation of genetic variation within populations

It is concluded that no significant genetic differentiation between the two Wildstock (West- and East Coast) populations. Data also confirms no stepwise loss of genetic variation among populations. Biological explanations to follow confirm the H_o.

3. <u>Biological reasons for which there is no observed loss of diversity</u>

The fact that there is no loss of diversity in the populations might be because of various reasons, the first being that a representative sample from the wild was taken when the Broodstock was acquired. Over 150 animals were sampled for breeding purposes, and factors such as inbreeding and genetic drift did not take place, at least not at a detectable level by AFLP's.

The second reason for maintained diversity is that the hatchery is not very old (~10 years), and that the generation time of abalone is quite long (~6 years). This means that the offspring of the original Broodstock are not old enough to be used in spawning, and inbreeding through backcrosses with parents has, as yet, not taken place.

In addition, the Broodstock is essentially a sub-sample of the Natural Population (West Coast), as the individuals were randomly selected from the wild - although from the same bay - and not from another hatchery or farmed group, which might have reduced the effective population size (N_e).

Lack of differentiation between the $Wild_{(WC)}$ and $Wild_{(EC)}$ is explained to a degree by the random introduction event discussed earlier.

4. Importance of this work

During the original discussions on the subject of the project, concerns were raised with regard to the impact of escaped animals on the natural population in the bay leading off from the farm. Conclusions drawn indicated that there be no immediate need for concern in this regard, but from a conservationist point of view, care needs to be taken to prevent future escape of larvae, especially if breeding practices are to change to include selection of fast-growing animals from the Progeny population.

The concepts discussed during the course of this thesis are only of value to the Danger Point Hatchery, if they are used to manage a population. Genetic aspects of Broodstock management are of the utmost importance for the reason that the genetic composition of a population determines its potential. Hatchery managers need to be aware of the techniques that can be used to manage, conserve, or exploit this potential.

5. Conclusions

5.1 Conservation

Care needs to be taken in future Broodstock management and breeding practices. Even though no loss of diversity was observed during this study, inexpert breeding practices in future might still lead to this loss. This is especially true if fast-growing Progeny individuals are to be used in future backcrosses with the original Broodstock. The loss of diversity might have a negative impact on the natural populations (West Coast population) if ever there is an escape of larvae through circulation system.

Operational procedures to limit geneflow from the farm back to the natural population in the form of live escapees, larvae and spat is still recommended as standard farming practices. This will become more important in light of future genetic changes in farmed populations through genetic improvement strategies such as crossbreeding, selection and genetic engineering.

5.2 Commercial management practices

Spawning techniques can be altered to improve on current techniques, where only certain individuals might contribute during mass spawning. Postponement of the initiation of a modified breeding program will only lead to further delay, and since

breeding is the aspect of management that manages the biological potential of the animal, yields will never be maximized.

Quotation

"The genetic aspects of broodstock management must become an integral part of hatchery management. The alleles in a population determine that population's potential. If you damage the potential, you will make it difficult to achieve production quotas. But if you manage and exploit it properly, you will exceed production quotas. Poor quality fish perform poorly; good quality fish give good results." – *Douglas Tave, 1993

*Douglas Tave is co-editor of *Proceedings: Auburn Symposium on Fisheries and Aquaculture*. He is editor of the *Journal of Applied Aquaculture* and is the genetics columnist for *Aquaculture Magazine*. He has authored and co-authored over 70 papers, and wrote and produced the video, *Aquaculture: Its Time Has Come*.

REFERENCES

1. Ankel-Simons, F., Cummins, J.M. 1996. *Misconceptions about mitochondria and mammalian fertilization: Implications for theories on human evolution.* Proceedings of the National Academy of Sciences USA 93:13895-13863
2. Bensch, S., Helbig, A.J., Salomon, M., Seibold, I. 2002. *Amplified fragment length polymorphism analysis identifies hybrids between two subspecies of warblers*. Molecular Ecology 11: 473-481
3. Brown, W.M., George, M., Wilson, A.C. 1979 *Rapid evolution of animal mitochondrial DNA*. Genetics 76(4): 1967-1971
4. Caetano-Anolles G., Bassam B.J., Gresshoff P.M. 1991 *DNA amplification fingerprinting using very short arbitrary oligonucleotide primers*. Bio/Technology 9: 553-557.
5. Caetano-Anolles, G., Gresshoff, P.M., 1998 *DNA Markers: protocols, applications, and overviews*. New York, Wiley-Liss, Inc.
6. Clarke, S. 1994. *Feasibility of Enhancing Abalone Stocks by Larval Seeding*. Fisheries Research and Development Corporation, Australia 94/005
7. Cook, P. 1998. *The current status of abalone farming in South Africa*. Journal of Shellfish Research 17(3): 601-602
8. De Waal, S., Cook, P. 2001. *Quantifying the physical and biological attributes of successful ocean seeding sites for farm-reared juvenile abalone (Haliotis midae)*. Journal of Shellfish Research 20(2): 857-861
9. Elliot, N.G., 2000. *Genetic improvement programmes in abalone: what is the future?* Aquaculture Research 31: 51-59
10. Evans, B., Sweijd, N., White, R.W.G., Elliott N.G., Cook, P. A marked genetic discontinuity in stocks of the South African abalone, Haliotis midae separated by Cape Agulhas: Evidence from allozymes, mitochondrial DNA and microsatellites.
11. Excoffier, L., Smouse, P.E., Quattro, J.M. 1992. *Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data.* Genetics 131: 479 - 491
12. Fallu, R. 1991. *Abalone Farming*. Oxford: Fishing News Books

13. Genade, A.B., Hirst, A.L., Smit, C.J. 1988. *Observations on the spawning, development and rearing of the South African abalone Haliotis midae Linn.* South African Journal of Marine Science 6: 3-12
14. Gillet, E.M. 1999. *Which DNA marker for which purpose?* [Online]. Available: http://webdoc.gwdg.de/ebook/1999/whichmarker/m14/Chap14.htm. [2002, August]
15. *Global Initiative on Late Blight* [Online]. 2002, Available: http://www.cipotato.org/gilb/research_tools/AFLP.htm. [2002, August]
16. Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C.; Gelbart, W.M. 1999. *Introduction to Genetic Analysis.* 7th edition. New York: W.H. Freeman & Co
17. Hartl, D.L. 1999. *A primer of population genetics.* 3rd edition. Mass.: Sinauer Associates, Inc.
18. Hartl, D.L., Clarck, A.G. 1997. *Principles of Population Genetics.* 3rd edition. Mass.: Sinauer Associates Inc.
19. Hauck, M., Hector, R. 2002. *An analysis of Operation Neptune: Government's response to marine poaching.* (Occasional Paper Series). Cape Town: Institute of Criminology, University of Cape Town.
20. Hauck, M., Sweijd, N.A. 1999. *A case study of abalone poaching in South Africa and its impact on fisheries management.* ICES Journal of Marine Science 65(6): 1024-1032
21. Hedrick, P.W. 2002. *Genetics of Populations.* 2nd edition. London: Jones and Bartlett Publishers
22. Huff, D.R., Peakall, R., Smouse, P.E. (1993) *RAPD variation within and among populations of outcrossing buffalograss (Buchloë dactyloides (Nutt.) Engelm).* Theoretical. Applied. Genetics 86: 927-934
23. Innan, I., Terauchi, R., Kahl, G., Tajima, F. 1998. *A method for estimating nucleotide diversity from AFLP data.* Genetics 151: 1157-1164
24. Jamieson, G.S. 1999. *Review of Status of Northern, or Pinto, Abalone, Haliotis Kamtschatkana, in Canada.* Canadian Stock Assessment Secretariat Research Document 99/190 Fisheries and Oceans, Canada

25. Jiang, L., Wu, W.L., Huang, P.C. 1995. *The mitochondrial DNA of Taiwan abalone Haliotis diversicolor Reeve, 1846 (Gastropoda: Archaeogastropoda: Haliotidae).* Molecular Marine Biology and Biotechnology 4(4): 353-364

26. Jones C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Van de Wiel, C., Bredemeijer, G., Vosman, B., Metthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volkaert, G., Rueda, J., Linacero, R., Vazquez, A., Karp, A. 1997. *Reproducibility testing of RAPD, AFLP and SSR markers in plant by network of European laboratories.* Mol. Breed. 3, 381-390

27. Kaneda, H., Hayashi, J., Takahama, S., Taya, C., Lindahl, K.F., Yonekawa, H. 1992. *Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis*. Proceedings of the National Academy of Sciences USA 92: 4542-4546

28. *Keygene N.V.* [Online]. 2001, Available: http://keygene.com/html/aflp.htm. [2002, August]

29. Klug, W.S., Cummings, M.R. 1997 *Concepts of Genetics.* 5th edition. New Jersey: Prentice-Hall Inc.

30. Kocher, T.D., Thomas, W.K., Meyer, A., Edwards, S.V., Pääbo, S., Villablanca, F.X. and Wilson, A.C. 1989. *Dynamics of mitochodrial DNA evolution in animals: amplification and sequencing with conserved primers*. Proceedings of the National Academy of Sciences USA, 86: 6196-6200.

31. Litaay, M., de Silva, S.S. 2001. *Reproductive performance indices based on physical characteristics of female Blacklip abalone Haliotis rubra L.* Journal of Shellfish Research 20(2): 673-677

32. Launey, S., Barre, M., Gerard, A., Naciri-Graven, Y. 2001. *Population bottleneck and effective size in Bonamia ostreae-resistant populations of Ostrea edulis as inferred by microsatellite markers.* Genetic Research Cambridge, 78: 259-270.

33. Liu, Z., Nichols, A., Li, P., Dunham, R.A. 1998. *Inheritance and usefulness of AFLP markers in channel catfish (Ictalurus punctatus), blue catfish (I. furcatus), and their F1, F2 and backcross hybrids*. Molecular General Genetics 258: 260-268

34. Lynch, M. 1990. *The similarity index and DNA fingerprinting.* Molecular Biology of Evolution 7: 478-484.

35. Lynch, M., Milligan, B.G. 1994. *Analysis of population genetic structure with RAPD markers*. Molecular Ecology 3: 91-99

36. Marshall, L. 2002. *Poaching, smuggling threaten abalone colonies in South Africa*. National Geographic News April 3, 2002

37. Melcher, U. *Molecular Genetics* [Online]. 2001, Available: http://opbs.okstate.edu/~melcher/MG/MGW1/MG11128.html#Info. [2002, August]

38. Metz, E.C., Robles-Sikisaka, R., Vacquier, V.D. 1998 *Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA* Proceedings of the National Academy of Sciences USA 95(18): 10676-10681.

39. Michalakis, Y., and Excoffier, L. 1996 *A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci*. Genetics 142: 1061 - 1064.

40. Miller, M.P. 1997. RxC Analytical Software for Genetic Data. Nothern Arizona University, Arizona, USA.

41. Mitchell-Olds, T. 1995. *The molecular basis of quantitative genetic variation in natural populations*. TREE 10(8): 324-327

42. Mueller, U.G., LaReesa Wolfenbarger, L. 1999 *AFLP genotyping and fingerprinting*. TREE 14(10): 389-394

43. Mullis, K.B., Faloona, F.A., Scharf, S., Saiki, R.K., Horn, G., Erlich, H.A. 1986. *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harbor Symposia on Quantitative Biology

44. Muse, B. 1998. *Management of the British Columbia Abalone fishery*. Alaska Commercial Fisheries Commission, CFEC 98-1N

45. Orloci, L. 1978. *Multivariate analysis in vegetation research.* Dr W. Junk B. V. The Hague – Boston.

46. Patzak, J. 2001. *Comparison of RAPD, STS, ISSR and AFLP molecular methods used for assessment of genetic diversity in hop (Humulus lupulus L.).* Euphytica 121: 9-18

47. Peakall, R., Smouse, P.E. 2001 GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia.

48. Peakall, R., Smouse, P.E., and Huff, D.R. 1995 *Evolutionary implications of allozyme and RAPD variation in diploid populations of Buffalograss. (Buchloë dactyloides (Nutt. Engelm.).* Molecular Ecology 4: 135 - 147.

49. Pitcher, G.C., Franco, J.M., Doucette, G.J., Powell, C.L., Mouton, A. 2001. *Paralytic shellfish poisoning in the abalone Haliotis midae on the West Coast of South Africa.* Journal of Shellfish Research 20(2): 895-904

50. Plagányi, E.E., Butterworth, D.S., Brandão, A. 2001. *Toward assessing the South African Abalone Haliotis midae stock using an age-structured production model.* Journal of Shellfish Research 20(2): 813-827

51. Rohlf, F.J. 1997. NTSYSpc version 2.0. State University of New York, New York, USA.

52. Roots, E.H., Baker, R.K., 2002. *Distribution and characterization of microsatellites in the Emu (Dromaius novaehollandiae) genome.* Journal of Hereditory 93(2): 100-106

53. Rosendahl, S., Taylor, J.W. 1997. *Development of multiple genetic markers for studies of genetic variation in arbuscular mycorrhizal fungi using AFLP.* Molecular Ecology 6: 821-829

54. Rumpho, M.E., Andrews, D.L., Manhart, J.R., Mujer, C.V., Pierce, S.K. 1994. *Extraction of DNA from mucilaginous tissues of a sea slug.* BioTechniques 17: 1097-1101

55. *Smithsonian Institute Natural Museum of Natural History* [Online]. 2000, Available: http://www.mnh.is.edu/anthro/humanorigins/faq/Encarta/diversity.htm. [2002, September]

56. Sweijd, N.A. 1999. Molecular Markers and Abalone Seeding as Tools for the Conservation and Management of the South African Abalone (Perlemoen), Haliotis midae Linn. Resource. PhD. thesis, University of Cape Town, Cape Town.

57. Tailby, R., Gant, F. 2002. *The illegal market in Australian Abalone*. Trends & Issues in Crime and Criminal Justice nr. 225, Australian Institute of Criminology, Canberra.

58. Tarr, R. *Marine and Coastal Management* [Online]. 2000, Available: http://www.environment.gov.za/mcm/inshore/abstory.htm. [2002, September]

59. Tarr, R.J.Q. 1992. The abalone fishery of South Africa. Shepherd S.A., Tegner, M.J., Guzmán del Próo, S.A. (eds.). *Abalone of the World. Biology, Fisheries and Culture.* Fishing News Books. Oxford.

60. Tave, D. 1993. *Genetics for fish hatchery managers*. 2nd edition. New York: Van Nostrand Reinhold.

61. Van Treuren, R. *CGN molecular markers* [Online]. 2000, Available: http://www.cgn.wageningen-ur.nl/pgr/research/molgen/default.asp. [2002, June]

62. Van Zyl, B. 2003. Molecular Investigation of the Genetic Variation of Populations of Abalone, Haliotis Midae, at the Danger Point Abalone Hatchery. M.Sc. thesis, University of Stellenbosch, Cape Town.

63. Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeau, M. 1995. *AFLP: AFLP new technique for DNA fingerprinting*. Nucleic Acids Research 23(21): 4407-4414

64. Wilding, C.S., Butlin R.K., Grahame J. 2001. *Differential gene exchange between parapatric morphs of Littorina saxatilis detected using AFLP markers*. Journal of Evolutionary Biology 14(4): 611

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

66. Zabeau, M., Vos, P. 1993. *Selective restriction fragment amplification: general method for DNA fingerprinting*. European Patent Application EP0534858.

APPENDIX A

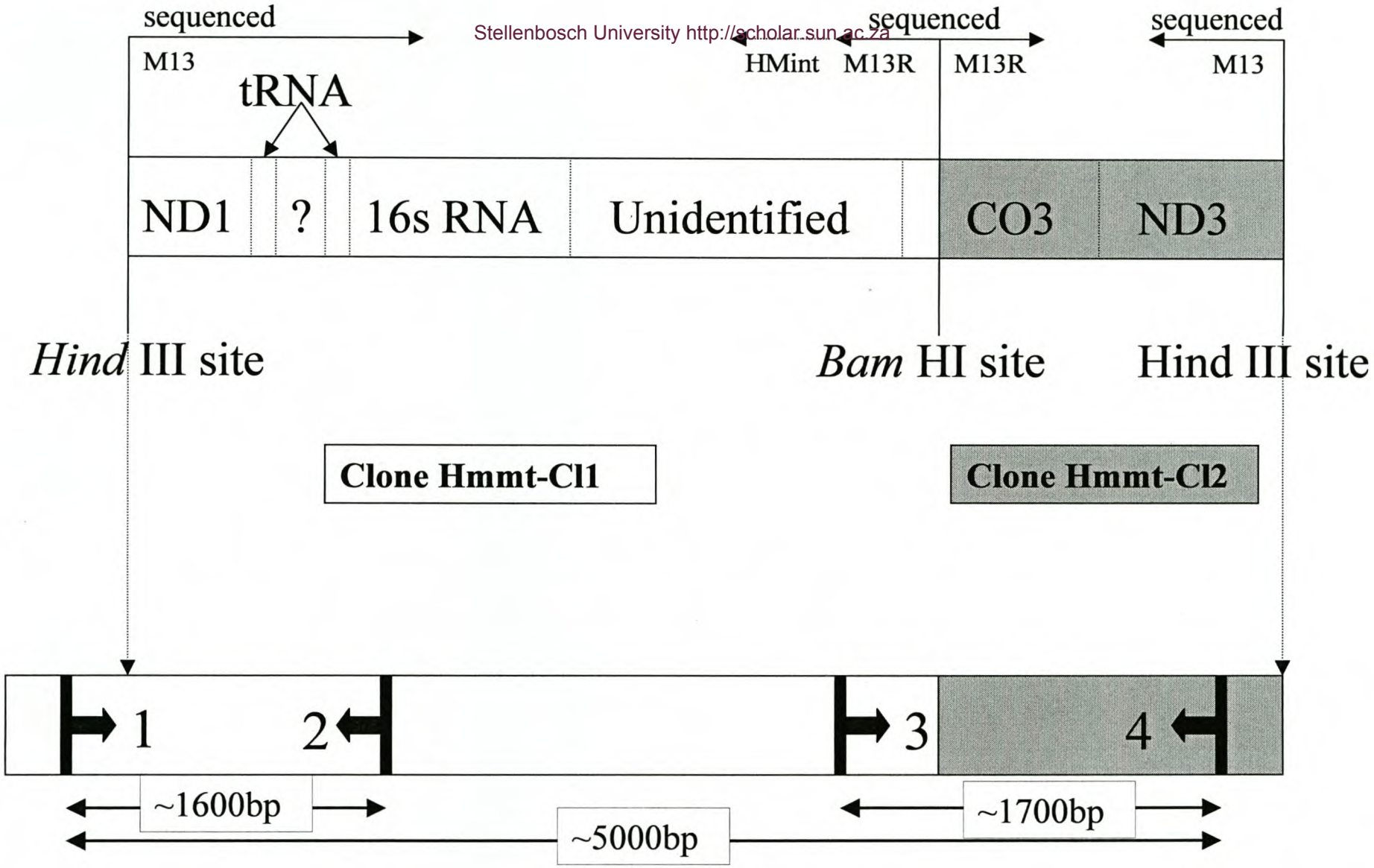


Figure A.1: Schematic representation of the gene order and location in the clone mtDNA fragments. Regions of the clones that were sequenced and sequencing primers used are shown. The relative positions of the designed PCR primers are indicated. Dotted lines indicate the position of the gene borders. Primers: 1 = HMND1d, 2 = c16sar, 3 = HMcon, 4 = HMND3. Not drawn to scale.

Appendix A

Table A.1: Table summarizing the characteristics of genetic markers

	Abundance	Level of polymorphism	Locus specificity	Codominance of alleles	Reproducibility	Labor-intensity	Technical demands	Operational costs	Development costs	Quantity of DNA required	Amenability to automation
Allozymes	low	low	yes	yes	high	low	low	low	low	-	no
RFLP	high	medium	yes	yes	high	high	high	high	medium-high	high	no
Minisatellites	medium	high	no/yes	no/yes	high	high	high	high	medium-high	high	no
PCR-sequencing	low	low	yes	yes	high	high	high	high	high	low	yes
RAPD	high	medium	no	no	low	low	low	low	low	low	yes
Microsatellites	high	high	yes	yes	high	low	low-medium	low-medium	high	low	yes
ISSR	medium-high	medium	no	no	medium-high	low	low-medium	low-medium	low	low	yes
SSCP	low	low	yes	yes	medium	low-medium	medium	low-medium	high	low	no
CAPS	low	low-medium	yes	yes	high	low-medium	low-medium	low-medium	high	low	yes
SCAR	low	medium	yes	yes/no	high	low	low	low	high	low	yes
AFLP	high	medium	no	no/yes	high	medium	medium	medium	low	medium	yes

Modified from CGN molecular markers: R. van Treuren.

Appendix A

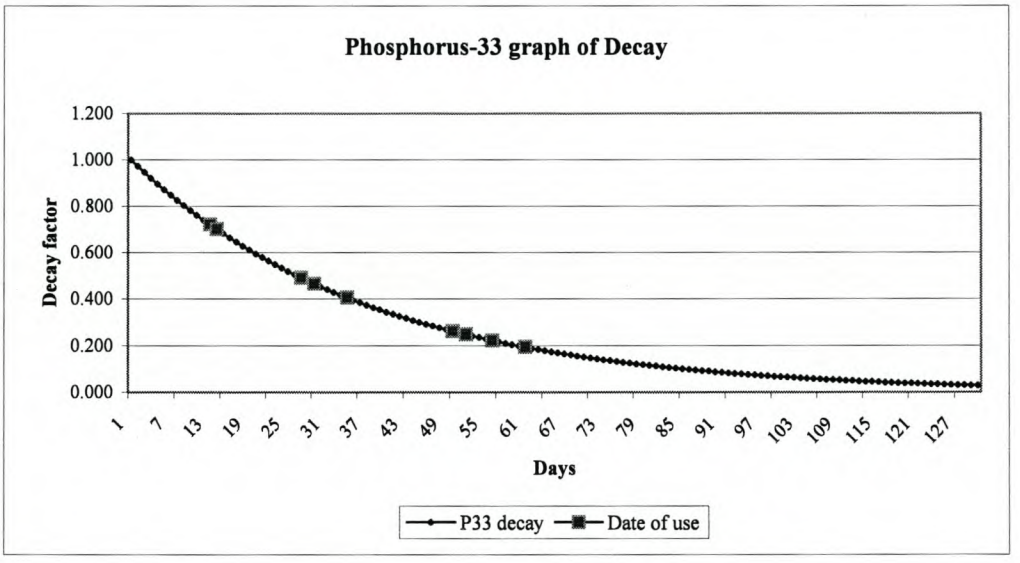
Table A.2: Table indicating primer sequences and other information of primers used during the study.

Primer	F/R	Sequence	DNA region	Taxa	Literature
ND1	Forward	GCCTCGCCTGTTTACCAAAAACAT	NADH1	*O. mykiss*	Nielsen et al. 1998.
ND1	Reverse	GGTATGGGCCCGAAAGCTTA	NADH1	*O. mykiss*	Nielsen et al. 1998.
ND5/6	Forward	AATAGCTCATCCATTGGTCTTAGG	NADH5/6	*O. mykiss*	Nielsen et al. 1998.
ND5/6	Reverse	TAACAACGGTGGTTTTTCAAGTCA	NADH5/6	*O. mykiss*	Nielsen et al. 1998.
CYTB/CR	Forward	TGACTTGAAAAACCACCGTTGTTA	Cyt B	*O. mykiss*	Bernatchez and Osinov 1995.
CYTB/CR	Reverse	GTGTTATGCTTTAGTTAAGC	Cyt B	*O. mykiss*	Bernatchez and Osinov 1995.
NADH2	Forward	ACATCTTAGCCATTCTCCTC	NADH2	*O. niloticus*	http://srs.ebi.ac.uk/srs6bin/cgi-bin/w
NADH2	Reverse	AGAAGACAAAGTGTCATTGC	NADH2	*O. niloticus*	http://srs.ebi.ac.uk/srs6bin/cgi-bin/w
COI	Forward	ACAATCACAAAGACATTGG	CO1	universal	Normark et al.1991.
COI	Reverse	AGAAAATGTTGAGGGAAGAA	CO1	universal	Normark et al.1991.
ATP/COIII	Forward	CACCAAGCACACGCATACCACAT	ATP6/CO3	universal	Meyer. 1993
ATP/COIII	Reverse	GGGGTTCCTTCAGGCAATAAATG	ATP6/CO3	universal	Baldwin 1992.
HMND1d	Forward	ACTARCTCRGATTCTCCTTCWGCAA	NADH1	*Haliotis*	N. Sweid unpublished 1999
c16sar	Reverse	CCTCGCCTGTTTAGCAAAAACAT	CO3	*Haliotis*	N. Sweid unpublished 1999
HMcon	Forward	TACTAACACCAAAACCCCCCC	NADH/tRNA/Cyt 3	*Haliotis*	N. Sweid unpublished 1999
HMND3	Reverse	GAGCCTTCATTTCATTCGTGG	NADH/tRNA/Cyt 3	*Haliotis*	N. Sweid unpublished 1999

Appendix A

Table A.3: Phosphorous-33 Decay Table and Graph of Decay, indicating day of use of P33.

Days	Days 0	1	2	3	4	5	6	7	8	9
0	1.000	0.973	0.947	0.921	0.897	0.872	0.849	0.826	0.804	0.782
10	0.761	0.741	0.721	0.701	0.683	0.664	0.646	0.629	0.612	0.595
20	0.579	0.564	0.549	0.534	0.520	0.506	0.492	0.479	0.466	0.453
30	0.441	0.429	0.418	0.406	0.395	0.385	0.374	0.364	0.355	0.345
40	0.336	0.327	0.318	0.309	0.301	0.293	0.285	0.277	0.270	0.263
50	0.256	0.249	0.242	0.236	0.229	0.223	0.217	0.211	0.205	0.200
60	0.195	0.189	0.184	0.179	0.174	0.170	0.165	0.161	0.156	0.152
70	0.148	0.144	0.140	0.136	0.133	0.129	0.126	0.122	0.119	0.116
80	0.113	0.110	0.107	0.104	0.101	0.098	0.096	0.093	0.091	0.088
90	0.086	0.084	0.081	0.079	0.077	0.075	0.073	0.071	0.069	0.067
100	0.065	0.064	0.062	0.060	0.059	0.057	0.055	0.054	0.053	0.051
110	0.050	0.048	0.047	0.046	0.045	0.043	0.042	0.041	0.040	0.039
120	0.038	0.037	0.036	0.035	0.034	0.033	0.032	0.031	0.030	0.030



Appendix A

Table A.4: Table indicating polymorphism data between pairwise comparisons.

West/East Coast						West/Broodstock			Broodstock/F1			West Coast/F1				Population
4/3	4/8	2/5	5/5	3/7	2/8	5/5	3/7	2/8	5/5	3/7	2/8	3/2	8/2	5/3	8/5	Primer Combination Mlu/Mse
39	30	37	22	21	40	23	25	27	26	39	44	35	31	26	19	Polymorphic bands
93	59	52	30	40	57	66	66	63	40	84	62	100	62	71	67	Total bands
41.9	50.8	71.2	73.33	52.5	70.18	34.8	37.9	42.9	65	46.43	70.97	35	50	36.6	28.4	% Polymorphic
		54.6			65.34			38.5			60.8				37.5	Average % per population
															64.18	Average total % polymorphism

Appendix A

Exact test calculation of Broodstock/Progeny Mlu2/Mse8

```
2 44
1  3  3  2  4  7  0  3  1  7  6 13  3 18 17  6  0  0  1  0  6  0  2  3 13  1  2
2  2 16  8  0  3  2  4  2  2  0 16  1 16  0  0 10
1  2  6  0  5  4  1  0  0  5  4 15  2 18  3  2  3  2  3  2  5  3  1  0 10  2  2
5  4 17  8  2  0  7  4  3  3  1 19  0 16  1  2 12
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
40	41	42	43	44	Total				
Row									
1	1	3	3	2	4	7	0	3	1
7	6	13	3	18	17	6	0	0	1
0	6	0	2	3	13	1	2	2	2
16	8	0	3	2	4	2	2	0	16
1	16	0	0	10	206				
2	1	2	6	0	5	4	1	0	0
5	4	15	2	18	3	2	3	2	3
2	5	3	1	0	10	2	2	5	4
17	8	2	0	7	4	3	3	1	19
0	16	1	2	12	205				
Total	2	5	9	2	9	11	1	3	1
12	10	28	5	36	20	8	3	2	4
2	11	3	3	3	23	3	4	7	6
33	16	2	3	9	8	5	5	1	35
1	32	1	2	22	411				

p = 0.124700 S.E. 0.011817
25997 switches made during analysis.

Appendix A

Exact test calculation of Broodstock/Progeny Mlu3/Mse7

```
2 39
13  3  0  1  1  0 14  7 11 14  1  6 11  3 10 17 18  6  1  0  8  3  5 18  2  2  2
 1  0  2  4  4 15  3 18  8  1  0  3
14  5  1  0 10  1 16  8 18 17  0  9 14  4 12 19 19  6  2  1  9  4  4 17  3  1  2
 0  2  2  0  4 18  3 18  3  0  1  5
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	32	33	34	35	36	37	38	39
Total									
Row									
1	13	3	0	1	1	0	14	7	11
14	1	6	11	3	10	17	18	6	1
0	8	3	5	18	2	2	2	1	0
2	4	4	15	3	18	8	1	0	3
236									
2	14	5	1	0	10	1	16	8	18
17	0	9	14	4	12	19	19	6	2
1	9	4	4	17	3	1	2	0	2
2	0	4	18	3	18	3	0	1	5
272									
Total	27	8	1	1	11	1	30	15	29
31	1	15	25	7	22	36	37	12	3
1	17	7	9	35	5	3	4	1	2
4	4	8	33	6	36	11	1	1	8
508									

p = 0.944120 S.E. 0.008738
27657 switches made during analysis.

Appendix A

Exact test calculationWest Coast/F1 Mlu8/Mse2

2 31
3 3 2 3 6 6 12 2 16 13 8 4 7 19 1 6 13 5 1 14 1 19 2 4 19 2 15
16 11 2 12
2 6 3 0 5 5 17 2 18 14 5 2 3 20 2 4 14 0 3 10 2 20 4 4 20 3 16
15 13 1 13

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
30	31	Total							
Row									
1	3	3	2	3	6	6	12	2	16
13	8	4	7	19	1	6	13	5	1
14	1	19	2	4	19	2	15	16	11
2	12	247							
2	2	6	3	0	5	5	17	2	18
14	5	2	3	20	2	4	14	0	3
10	2	20	4	4	20	3	16	15	13
1	13	246							
Total	5	9	5	3	11	11	29	4	34
27	13	6	10	39	3	10	27	5	4
24	3	39	6	8	39	5	31	31	24
3	25	493							

p = 0.972360 S.E. 0.004167
32595 switches made during analysis.

Appendix A

Exact test calculation of Wild(WC)/Broodstock Mlu5/Mse5

```
2 23
14  4 19 20  3 16 15  3  7 12 18  1 10 12 19 14 18 20  0  5 19  4 20
13 14 18 18 15 12 13  4 17 17 17  0 18 15 14 14 12 17  1  0 18  3 16
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

```
    Col 1       2       3       4       5       6       7       8       9
10      11      12      13      14      15      16      17      18      19
20      21      22      23     Total
Row
 1      14      4       19      20      3       16      15      3       7
12      18      1       10      12      19      14      18      20      0
5       19      4       20      273
 2      13      14      18      18      15      12      13      4       17
17      17      0       18      15      14      14      12      17      1
0       18      3       16      286

Total   27      18      37      38      18      28      28      7       24
29      35      1       28      27      33      28      30      37      1
5       37      7       36      559
```

p = 0.067560 S.E. 0.015073
35589 switches made during analysis.

Analysis of C:\MYDOCU~1\DANE_L~1\TESIS\DATA\TEXT\WBSUM55.TXT
20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

```
    Col 1       2       3       4       5       6       7       8       9
10      11      12      13      14      15      16      17      18      19
20      21      22      23     Total
Row
 1      14      4       19      20      3       16      15      3       7
12      18      1       10      12      19      14      18      20      0
5       19      4       20      273
 2      13      14      18      18      15      12      13      4       17
17      17      0       18      15      14      14      12      17      1
0       18      3       16      286

Total   27      18      37      38      18      28      28      7       24
29      35      1       28      27      33      28      30      37      1
5       37      7       36      559
```

p = 0.061560 S.E. 0.011949
35510 switches made during analysis.

Appendix A

Exact test calculation of Wild(WC)/Progeny Mlu3/Mse2

```
2 35
1  0  2  7  8  7  3  2  3  0  5  2  7 10 15  0  8  5 10  2  5 14  3  5 11 19  1
2  5 20 19  2  6  7  1
4  4  0  7  8  4  4  0  5  2  6  1  1  9  8  3  9 10 11  1  6 18  4  7  8 13  1
1  8 19 14  0  9  7  3
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

```
    Col 1       2       3       4       5       6       7       8       9
10      11      12      13      14      15      16      17      18      19
20      21      22      23      24      25      26      27      28      29
30      31      32      33      34      35   Total
Row
 1      1       0       2       7       8       7       3       2       3
0       5       2       7       10      15      0       8       5       10
2       5       14      3       5       11      19      1       2       5
20      19      2       6       7       1       217
 2      4       4       0       7       8       4       4       0       5
2       6       1       1       9       8       3       9       10      11
1       6       18      4       7       8       13      1       1       8
19      14      0       9       7       3       215

Total   5       4       2       14      16      11      7       2       8
2       11      3       8       19      23      3       17      15      21
3       11      32      7       12      19      32      2       3       13
39      33      2       15      14      4       432
```

p = 0.614660 S.E. 0.027206
30296 switches made during analysis.

Appendix A

Exact test calculation Wild(WC)/Progeny Mlu5/Mse3

2 26
3 14 3 15 2 0 1 15 10 6 14 12 2 9 3 19 19 19 9 16 10 6 2 8 2 17
3 13 0 15 1 2 1 14 8 5 15 8 1 7 3 9 18 18 10 15 7 8 4 3 3 19

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	Total		
Row									
1	3	14	3	15	2	0	1	15	10
6	14	12	2	9	3	19	19	19	9
16	10	6	2	8	2	17	236		
2	3	13	0	15	1	2	1	14	8
5	15	8	1	7	3	9	18	18	10
15	7	8	4	3	3	19	210		
Total	6	27	3	30	3	2	2	29	18
11	29	20	3	16	6	28	37	37	19
31	17	14	6	11	5	36	446		

p = 0.979360 S.E. 0.003670
32866 switches made during analysis.

Appendix A

Exact test calculation of Wild(WC)/Broodstock Mlu3/Mse7

2 25
14 15 14 1 5 1 3 5 12 1 2 9 9 3 3 3 3 1 7 1 1 1 0 1 1
6 5 12 0 1 0 0 4 4 0 1 3 3 3 2 4 0 0 1 0 0 0 2 1 0

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	Total			
Row									
1	14	15	14	1	5	1	3	5	12
1	2	9	9	3	3	3	3	1	7
1	1	1	0	1	1	116			
2	6	5	12	0	1	0	0	4	4
0	1	3	3	3	2	4	0	0	1
0	0	0	2	1	0	52			
Total	20	20	26	1	6	1	3	9	16
1	3	12	12	6	5	7	3	1	8
1	1	1	2	2	1	168			

p = 0.739880 S.E. 0.013023
20055 switches made during analysis.

WildBrood28b

Appendix A

Exact test calculation of Wild(WC)/Broodstock Mlu2/Mse8

```
2 27
1 13  1 13  0  6  7  0  1  3 10  1  1 16  2  2  1  2 16  1  2 14  1 14 12  9  5
3 13  2 10  7  5  2  1  0  0 11  1  0 14  2  3  0  2 14  1  2 15  2 11  8 11  3
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col 1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	Total	
Row									
1	1	13	1	13	0	6	7	0	1
3	10	1	1	16	2	2	1	2	16
1	2	14	1	14	12	9	5	154	
2	3	13	2	10	7	5	2	1	0
0	11	1	0	14	2	3	0	2	14
1	2	15	2	11	8	11	3	143	
Total	4	26	3	23	7	11	9	1	1
3	21	2	1	30	4	5	1	4	30
2	4	29	3	25	20	20	8	297	

p = 0.761840 S.E. 0.020046
26798 switches made during analysis.

WildProg85p

Appendix A

Exact test calculation of Wild(WC)/Progeny Mlu8/Mse5

```
2 19
3  2  4  2 16  3  2  2  1  8  3 16 14 10  4  2 15 15 15
0  1  5  1 13  6  0  0  2 11  4 12 10  8  4  3 17 11 13
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
Row																				
1	3	2	4	2	16	3	2	2	1	8	3	16	14	10	4	2	15	15	15	137
2	0	1	5	1	13	6	0	0	2	11	4	12	10	8	4	3	17	11	13	121
Total	3	3	9	3	29	9	2	2	3	19	7	28	24	18	8	5	32	26	28	258

p = 0.939940 S.E. 0.005798
30257 switches made during analysis.

Appendix A

Exact test calculation of Broodstock/Progeny Mlu5/Mse5

```
2 26
10  2 19  1  1  2  9 14  6  1 16  3 12  1  4 12 19  2 19  5 20  1  2  3  0  0
16  1 20  2  0  0 17 17  9  1 18  1 14  0  8 12 17  1 18  1 19  0  1  1  1  1
```

20 batches of 2500 replicates per batch performed.
50000 total replicates.

Contingency Table

Col	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	Total		
Row									
1	10	2	19	1	1	2	9	14	6
1	16	3	12	1	4	12	19	2	19
5	20	1	2	3	0	0	184		
2	16	1	20	2	0	0	17	17	9
1	18	1	14	0	8	12	17	1	18
1	19	0	1	1	1	1	196		
Total	26	3	39	3	1	2	26	31	15
2	34	4	26	1	12	24	36	3	37
6	39	1	3	4	1	1	380		

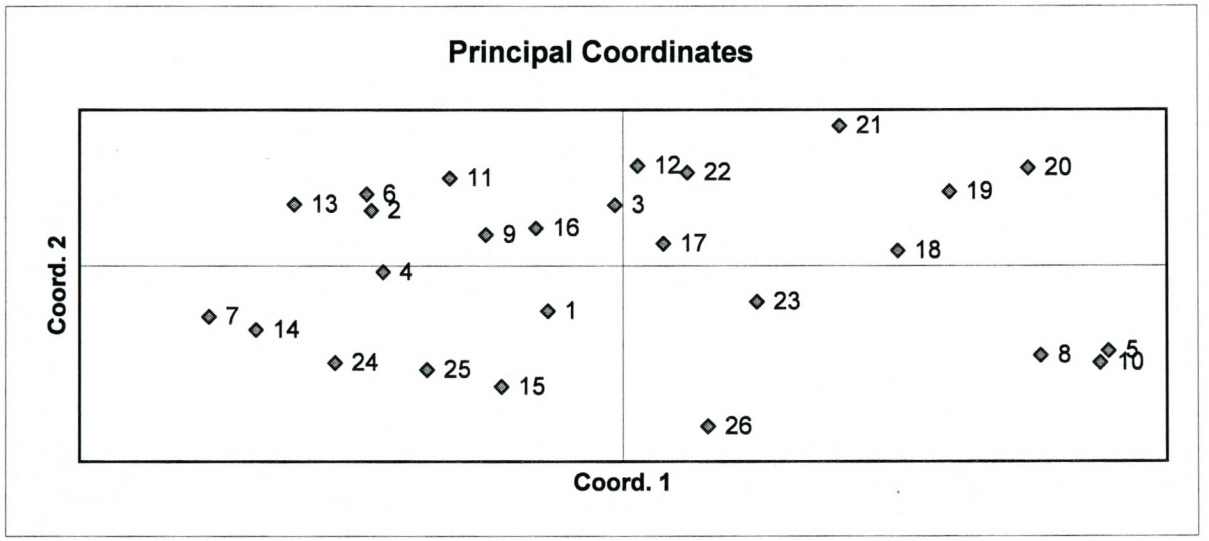
p = 0.871580 S.E. 0.011579
26791 switches made during analysis.

Appendix A

Principal Coordinates Analysis (PCA) of Wild(WC)/Wild(EC) populations
PCA via Covariance matrix with data standardization
No. Sampl 26

Percentage of variation explained by the first 3 axes

Axis	%	Cum %
1	19.08	19.08
2	13.91	32.99
3	12.87	45.86



Appendix A

Results of Analysis of Molecular Variance of Wild(WC)/Wild(EC) populations
Input as Binary Distance Matrix for Calculation of PhiPT

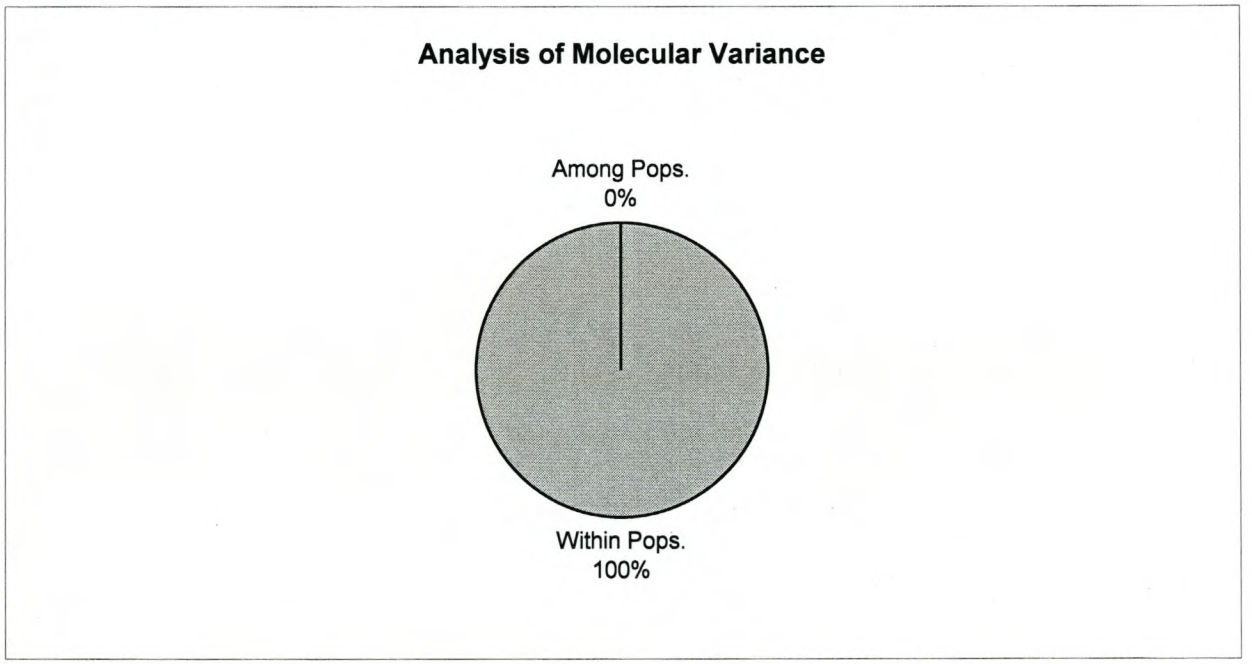
No. Samples	26
No. Pops.	2
No. Regions	1
N0	13.000
SSTOT	165.038

Pop. No.	**1**	**2**
n	13	13
SSWP	65.538	93.231

Summary

Source	**df**	**SS**	**MS**	**Est. Var.**	**Stat**	**Value**	**Prob**
Among Pops.	1	6.269	6.269	0.000			
Within Pops.	24	158.769	6.615	6.615	**PhiPT**	0.000	0.544

Negative Among Pops. Est. Var. of -0.027 converted to zero!
Negative PhiPT of -0.004 converted to zero!



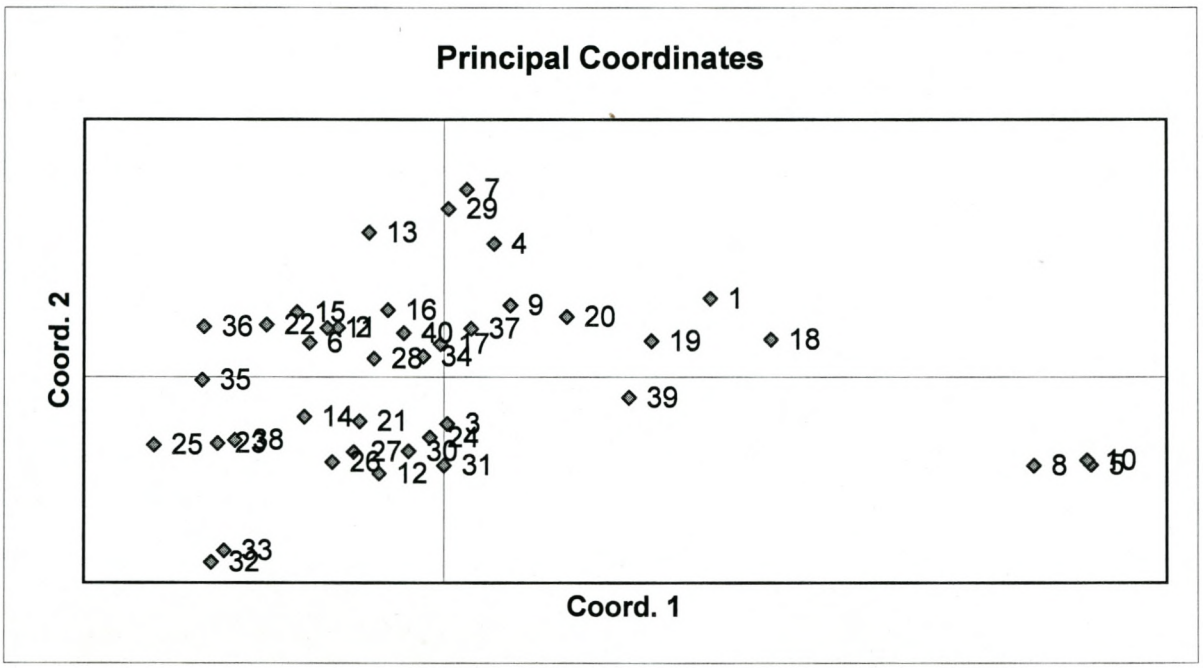
Appendix A

Principal Coordinates Analysis (PCA) of Wild(WC)/Progeny populations

PCA via Covariance matrix with data standardization
No. Sampl 40

Percentage of variation explained by the first 3 axes

Axis	%	Cum %
1	15.96	15.96
2	12.52	28.48
3	8.25	36.73



Appendix A

Results of Analysis of Molecular Variance of Wild(WC)/Progeny populations
Input as Binary Distance Matrix for Calculation of PhiPT

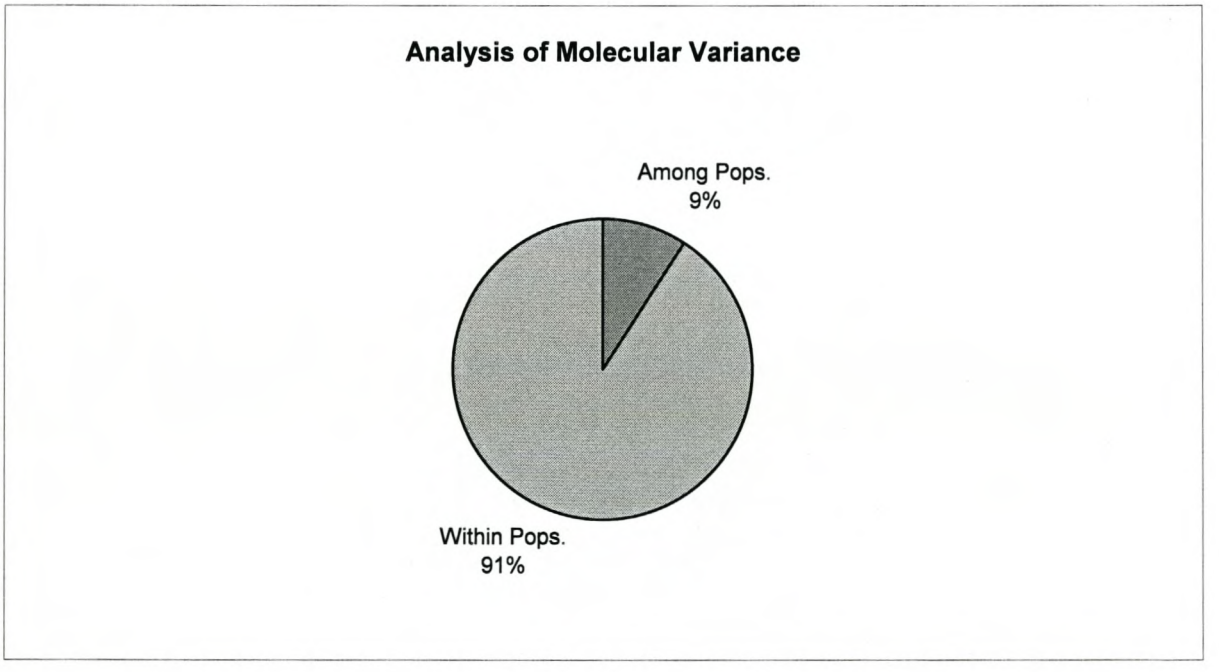
No. Sampl	40
No. Pops.	2
No. Regio	1

N0	20.000
SSTOT	218.975

Pop. No.	**1**	**2**
n	20	20
SSWP	102.900	99.950

Summary

Source	**df**	**SS**	**MS**	**Est. Var.**	**Stat**	**Value**	**Prob**
Among Po	1	16.125	16.125	0.539			
Within Po	38	202.850	5.338	5.338	**PhiPT**	0.092	0.001

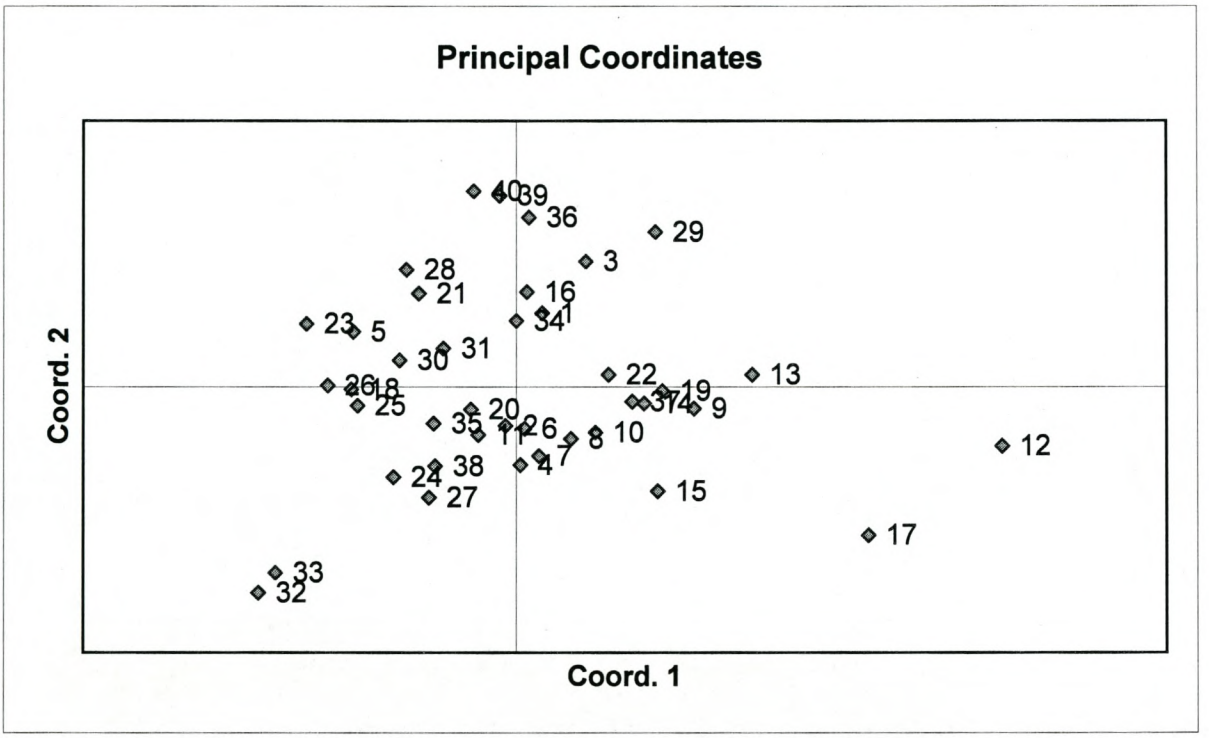


Appendix A

Principal Coordinates Analysis (PCA) of Broodstock/Progeny populations
PCA via Covariance matrix with data standardization
No. Sampl 40

Percentage of variation explained by the first 3 axes

Axis	%	Cum %
1	14.15	14.15
2	9.28	23.42
3	7.79	31.22



Appendix A

Results of Analysis of Molecular Variance of Broodtock/Progeny populations
Input as Binary Distance Matrix for Calculation of PhiPT

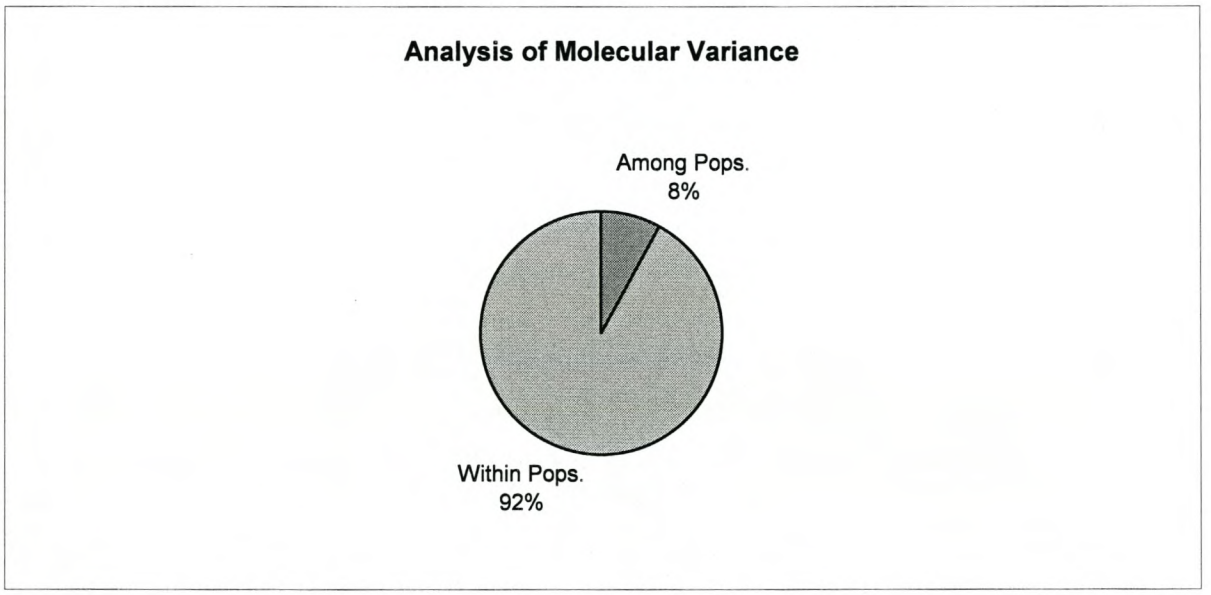
No. Sampl	40
No. Pops.	2
No. Regio	1

N0	20.000
SSTOT	218.300

Pop. No.	**1**	**2**
n	20	20
SSWP	103.550	99.950

Summary

Source	**df**	**SS**	**MS**	**Est. Var.**	**Stat**	**Value**	**Prob**
Among Po	1	14.800	14.800	0.472			
Within Po	38	203.500	5.355	5.355	**PhiPT**	0.081	0.001

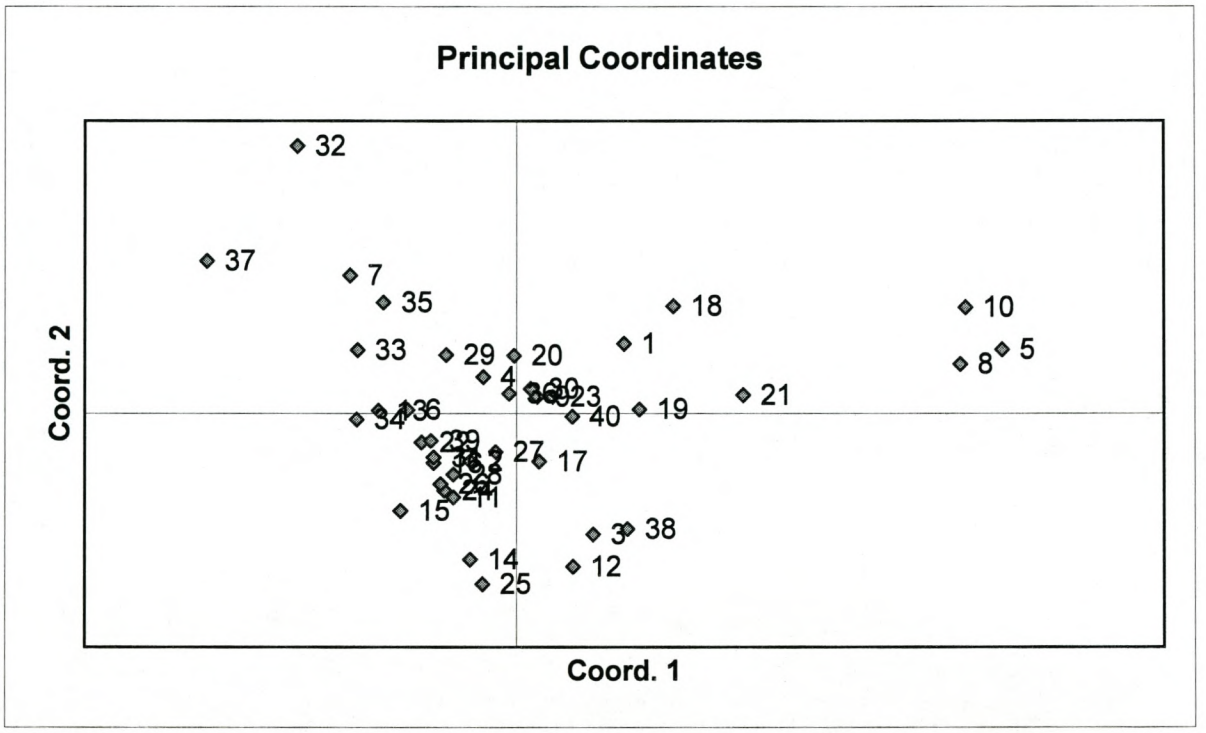


Appendix A

Principal Coordinates Analysis (PCA) of Wild(WC)/Broodstock populations
PCA via Covariance matrix with data standardization
No. Sampl 40

Percentage of variation explained by the first 3 axes

Axis	%	Cum %
1	16.90	16.90
2	10.60	27.50
3	8.25	35.75



Appendix A

Results of Analysis of Molecular Variance of Wild(WC)t/Broodstock populations
Input as Binary Distance Matrix for Calculation of PhiPT

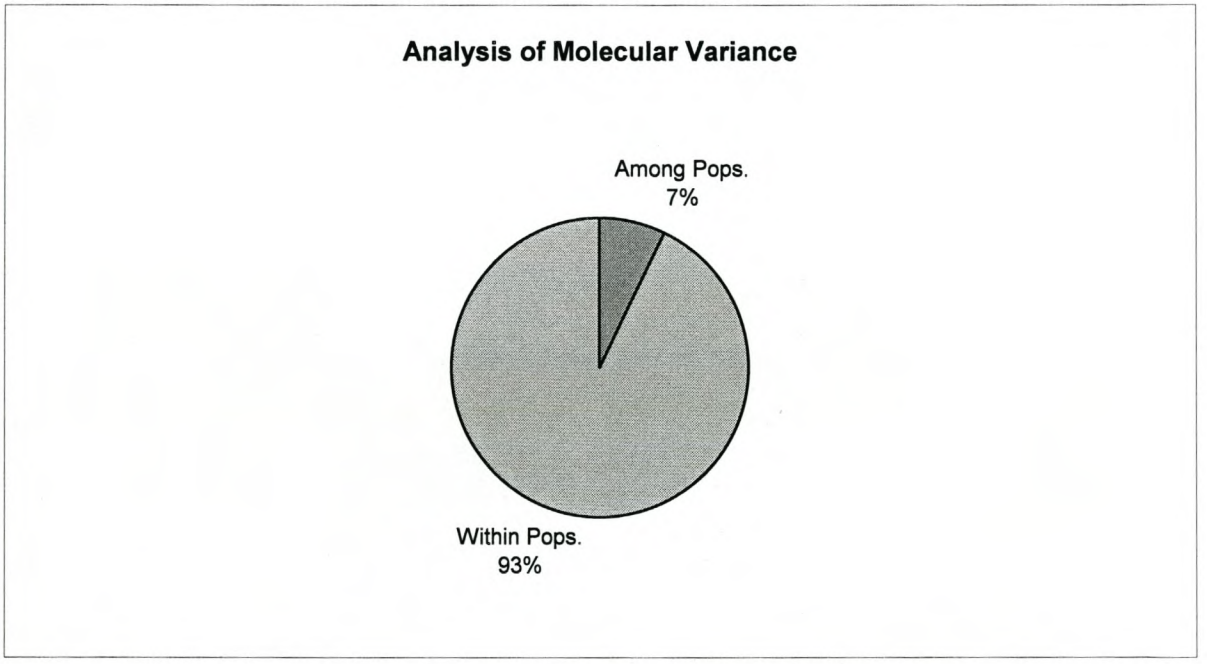
No. Sampl	40
No. Pops.	2
No. Regio	1

N0	20.000
SSTOT	220.175

Pop. No.	**1**	**2**
n	20	20
SSWP	102.900	103.550

Summary

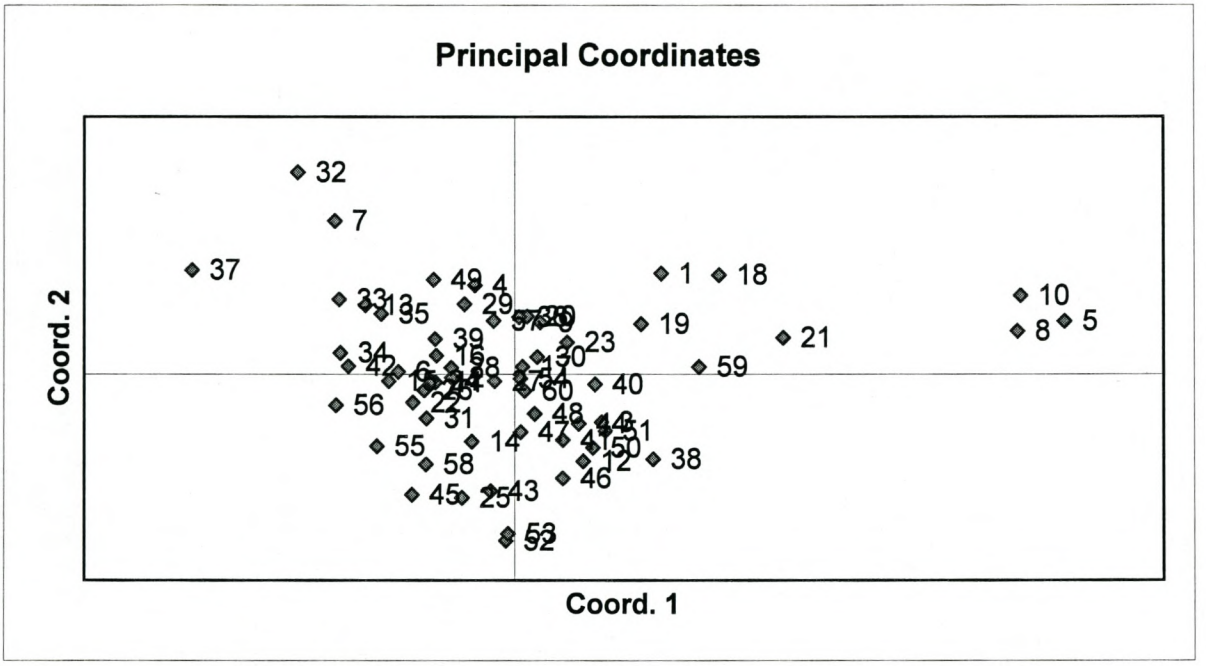
Source	**df**	**SS**	**MS**	**Est. Var.**	**Stat**	**Value**	**Prob**
Among Po	1	13.725	13.725	0.415			
Within Po	38	206.450	5.433	5.433	**PhiPT**	0.071	0.001



Overall Principal Coordinates Analysis (PCA)
PCA via Covariance matrix with data standardization
No. Sampl 60

Percentage of variation explained by the first 3 axes

Axis	%	Cum %
1	12.18	12.18
2	11.31	23.49
3	6.65	30.14



Overall Results of Analysis of Molecular Variance
Input as Binary Distance Matrix for Calculation of PhiPT

No. Sampl	60
No. Pops.	3
No. Regio	1
N0	20.000
SSTOT	336.167

Pop. No.	**1**	**2**	**3**
n	20	20	20
SSWP	102.900	103.550	99.950

Summary

Source	**df**	**SS**	**MS**	**Est. Var.**	**Stat**	**Value**	**Prob**
Among Po	2	29.767	14.883	0.475			
Within Po	57	306.400	5.375	5.375	**PhiPT**	0.081	0.001