Molecular Analysis of Genetic Variation and Relationships within the Population of Abalone (*Haliotis midae*) at the Sea Plant Products Abalone Hatchery, Hermanus, R.S.A.

Ruhan Slabbert

Thesis presented in partial fulfillment of the requirements for the degree of Master of Science at the University of Stellenbosch.

Supervisors: Dr ME D'Amato and Dr D Brink

December 2004

Stellenbosch University http://scholar.sun.ac.za

DECLARATION

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

Stellenbosch University http://scholar.sun.ac.za

ABSTRACT

The species *Haliotis midae* is the only commercially exploitable abalone species of the six found in the South African coastal waters. This species is under substantial pressure from both legal and illegal harvesters, to such an extent that it could be commercially extinct within four years. Efforts to alleviate the pressures on the natural populations of both illegal and legal harvesting are being made. The genetic management systems for abalone farming and ranching activities should be carefully evaluated. The loss of genetic diversity and the risks of contaminating the gene pools of natural populations in the vicinity of a farm should be minimized. Genetic evaluation studies will be of great importance to acquire the necessary data needed for genetic diversity and differentiation analysis.

The aim of this study was to develop species-specific microsatellite DNA markers to assess the genetic diversity and differentiation within and between the brood stock and commercial stock of the Sea Plant Products abalone farm (Hermanus, Republic of South Africa) and natural populations related to the brood stock. The species-specific DNA markers were also used for parentage assignments within the farm population (first for abalone) and preliminary QTL (quantitative trait loci)-discovery analysis studying growth rate segregation.

Samples were taken of the farm's brood stock and commercial stock (Rows 2, 3, 4) as well as from two natural populations (Saldanha Bay and Black Rock) related to the brood stock. Various statistical parameters and software packages were used to assess genetic diversity and differentiation, to infer parentage and to look for QTL's.

Eight species-specific microsatellite DNA markers were designed and used for data analysis. Data analysis showed a loss of genetic diversity from the brood stock to the commercial stock caused by the subdivision of the original brood stock into rows and the differential contributions of parents to the offspring. No genetic differentiation (F_{st}) was detected between the farm and

Stellenbosch University http://scholar.sun.ac.za

natural populations, except for the offspring of Row2. Levels of inbreeding (F_{is}) were high for all loci within the populations. Thirty-eight percent of all studied offspring were confidently assigned to a couple. The preliminary QTL-discovery suggested the segregation of a number of alleles and genotypes with growth rate.

The study concluded that the commercial abalone population of the Sea Plant Products abalone farm holds no threat to the disruption of the genetic diversity of the natural populations. It is proposed that the farm implement a rotational breeding program to increase the genetic diversity of the commercial population. Any newly acquired brood stock must be profiled before their introduction into the breeding program to assess the influence of the animals on the current levels of genetic diversity within the farm. The accuracy and reliability of parentage assignments and QTL-discovery need to be optimised by adding more loci and sampling more animals or even by trying and developing new methods.

Stellenbosch University http://scholar.sun.ac.za

OPSOMMING

Van die ses perlemoen spesies wat langs die Suid-Afrikaanse kus gevind word, is die spesie *Haliotis midae* die enigste een wat van kommersiële belang is. Wettige, sowel as onwettige versameling, plaas hierdie spesie onder sulke geweldige druk dat dit dalk binne vier jaar verlore kan wees vir die kommersiële bedryf. Verskeie strategieë word tans geïmplimenteer om hierdie druk te verlig. Die genetiese bestuurstrategieë binne perlemoen plase moet deeglik ondersoek word. Die verlies aan genetiese diversiteit en die moontlikheid vir die kontaminering van die natuurlike populasies in die omgewing van die plaas se geenpoel, moet uitgeskakel word. Genetiese evaluasies sal van groot belang wees om die nodige data vir genetiese diversiteit- en differensiasie-analises te verkry.

Die doel van die studie was om spesies-spesifieke mikrosatelliet DNA merkers te ontwikkel wat gebruik sou word om die genetiese diversiteit en differensiasie binne en tussen die broei diere en die kommersiële diere van *Sea Plant Products* se perlemoen plaas (Hermanus, Republiek van Suid-Afrika) en die natuurlike populasies wat verwant is aan die broei diere, te bepaal. Die spesies-spesifieke DNA merkers is ook vir ouerskap-bepalings binne die plaas se populasie gebruik, asook vir voorlopige *QTL* (*quantitative trait loci*) – ontdekking met betrekking tot groeitempo segregasie.

Monsters van die plaas se broei diere en kommersiële diere (Ry 2, 3, 4) asook van twee natuurlike populasies (Saldanha Baai en Black Rock) wat verwant is aan die broei diere, is geneem. 'n Verskeidenheid van statistiese parameters en sagteware pakette is vir die genetiese diversiteit- en differensiasie-analises, vir ouerskap-bepalings en vir die opspoor van *QTL's* gebruik.

Agt spesies-spesifieke mikrosatelliet DNA merkers is ontwerp en toe gebruik vir die data analises. 'n Verlies aan genetiese diversiteit vanaf die broei diere na die kommersiële diere is deur die data analises uitgewys. Dit is

Stellenbosch University http://scholar.sun.ac.za

veroorsaak deur die verdeling van die oorspronklike broei diere in rye en die differensiële bydraes deur die ouers na die nageslag. Geen genetiese differensiasie (F_{st}) is tussen die plaas se populasie en die natuurlike populasies gevind nie, maar die nageslag van Ry 2 het wel differensiasie getoon. Die vlakke van inteling (F_{is}) was hoog oor alle lokusse binne alle populasies. Agt-en-dertig persent van die nageslag wat bestudeer is, was suksesvol met 'n ouerpaartjie geassosieer. Die voorlopige *QTL*-ontdekking studie toon die potesiële segregasie van 'n aantal allele en genotipes saam met groeitempo.

Die bevindinge van die studie is dat die kommersiële populasie op die *Sea Plant Products* perlemoen plaas, geen bedreiging vir die versteuring van die natuurlike populasies se genetiese diversiteit inhou nie. Daar is voorgestel dat die plaas 'n rotasie-basis broeiprogram moet implimenteer om sodoende die genetiese diversiteit van die kommersiële populasie te verhoog. Enige nuwe broei diere se genetiese profiel moet ook voor die tyd ondersoek word om te sien wat se effek hierdie diere op die huidige genetiese diversiteit van die kommersiële populasie sal hê. Die akkuraatheid en betroubaarheid van die ouerskap-bepalings en *QTL*-ontdekking moet optimiseer word deur of meer lokusse te bestudeer, of meer diere in analises te gebruik of selfs om nuwe tegnieke te probeer of te ontwikkel.

Stellenbosch University http://scholar.sun.ac.za

TABLE OF CONTENTS

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

ACKNOWLEDGEMENTS

The author wishes thank Dr Eugenia D'Amato for her help, encouragement, contributions and comments during the study and the preparation of this manuscript. Also Dr Danie Brink for his assistance throughout the project. I also wish to thank Dr Ruvay Roodt-Wilding for her input during the preparation of this manuscript. Special thanks go to Mr Stephan Ashlin from Sea Plant Products for his enthusiasm and help, and Mr Jacques du Preez (a colleague) for his input during the experimental phase (PCRs) of the study. Finally a big thank you goes to Sea Plant Products for their monetary support.

Stellenbosch University http://scholar.sun.ac.za

ABBREVIATIONS

A	adenine
AA	acrylamide
a_e	effective number of alleles
ABI	*Applied Biosystems*
Acc. Nr.	GenBank accession number
AFLP	amplified fragment length polymorphism
$Ag(NO_3)_2$	silver nitrate
APS	ammonium persulfate
ATP	adenosine triphosphate
bp	basepairs
BLAST	basic local alignment search tool
blastn	nucleotide-nucleotide BLAST
BSA	bovine serum albumin
C	cytosine
χ^2	chi-square
CH_3COOH	acetic acid
CTAB	N-cetyl-N, N, N-trimethyl ammonium bromide
ddH_2O	double distilled water
℃	degrees Celsius
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	di-thio-threitol
EDTA	Ethylene Diamine Tetra-Acetate
EST	expressed sequence tag
Eq	equation
et al	and company
F1	first cross

Stellenbosch University http://scholar.sun.ac.za

F2	second cross
FAM	blue (R100); 5-carboxyfluorescein (ABI-fluorescent label)
FIASCO	Fast Isolation by AFLP of Sequences Containing Repeats
g	gram
G	guanine
H_o	observed heterozygosity
H_e	expected heterozygosity
H_o	null hypothesis
HCl	hydrochloric acid
IAM	infinite allele model
ID	identification
KCl	potassium chloride
λ	lambda
LB	Luria Bertani medium
LOD	natural logarithm of the likelihood ratio
m	meters
M	molar
mM	millimolar
mg	milligram
$MgCl_2$	magnesium chloride
min	minute
ml	millilitre
*Mse*I	restriction enzyme isolated from *Thermus ruber* RFL1
mtDNA	mitochondrial deoxyribonucleic acid
N	indicate one of A, T, C or G
N/A	not available
NaCl	sodium chloride

Stellenbosch University http://scholar.sun.ac.za

NaOH	sodium hydroxide
NCBI	National Center for Biotechnology Information
NED	yellow (Tamra) (ABI-fluorescent label)
ng	nanogram
NS	non-significant
p	allele frequency
P	probability
PAA	poly-acrylamide
PCR	polymerase chain reaction
pers.	personal
pers. comm.	personal communication
PET	red (ABI-fluorescent label)
pH	concentration of hydrogen ions in a solution is expressed conventionally as its pH
pmol	picomol
QTL	quantitative trait loci
RAPD	random amplified polymorphic DNA
rcf	relative centrifugal force
RFLP	restriction fragment length polymorphism
s	second
SDS	sodium dodecyl sulfate
SMM	stepwise mutation model
SNP	single nucleotide polymorphism
SPP	Sea Plant Products
SSC	standard saline citrate
SSCP	single-strand conformational polymorphism
STR	simple tandem repeat
T	thymine

Stellenbosch University http://scholar.sun.ac.za

Taq	*Thermus aquaticus* DNA polymerase
TBE	Tris, boric acid, EDTA
TD-PCR	touch-down protocol
TE	Tris, EDTA
Temed	N, N, N', N'-tetramethyl-ethylenediamine
TEN	Tris, EDTA, NaCl
Tm	annealing temperature
Tris	2-amino-2-(hydroxymethyl)-1, 3-proanediol
U	unit
U	Mann-Whitney test
*u*g	microgram
*µ*l	microlitres
*µ*M	micromolar
UV	ultra-violet
v/v	volume to volume
VIC	green (ABI-fluorescent label)

Stellenbosch University http://scholar.sun.ac.za

LIST OF EQUATIONS

Stellenbosch University http://scholar.sun.ac.za

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 4

CHAPTER 5

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 6

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

LIST OF TABLES

CHAPTER 1

CHAPTER 2

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 6

Stellenbosch University http://scholar.sun.ac.za

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 1 – INTRODUCTION

1.1) ABALONE IN GENERAL

1.1.1) TAXONOMY

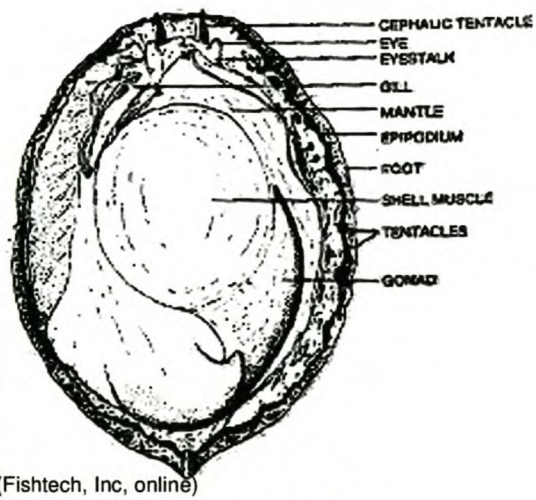
Table 1.1: Taxonomy of the abalone.

Phylum	**Mollusca**
Class	*Gastropoda*
Order	*Vetigastropoda*
Superfamily	*Pleurotomarioidea*
Family	*Haliotidae*
Genus	*Haliotis*

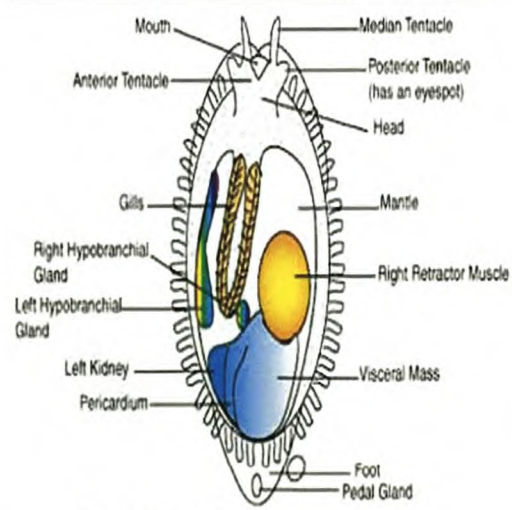
(Hardy(a), online)

1.1.2) ANATOMY

Sketches of abalone anatomy are shown in Figure 1.1. The visible parts of the abalone are the shell, the muscular foot, the mantle and the epipodium (sensory structure with tentacles). The bottom front part of the animal hosts the eyes, enlarged tentacles and the mouth that holds a tongue (radula). Respiratory pores are visible on the side of the shell (topview – on the left hand side). Internal organs consist of a heart (pumps blood through arteries, veins and sinuses), gills (under respiratory pores) and a gonad. The gonad is cream coloured in males and green or grey coloured in females (Fishtech, Inc, online).



(Fishtech, Inc, online)



(OceanLink, online)

Figure 1.1: Anatomy of the abalone.

Stellenbosch University http://scholar.sun.ac.za

1.1.3) LIFE CYCLE

The life cycle of the abalone is shown in Figure 1.2. Wild abalone occur in large groups consisting of male and female animals. Adult animals reach sexual maturity approximately at the age of 7 years. Abalone reproduce by mass spawning. Fertilization takes place externally when females release egg cells and males release sperm cells into the water. The fertilized eggs develop into trochophore larvae that remain in the plankton phase. The larvae remain in plankton phase for up to a week and are carried to different locations by the currents. The larvae then settle (veliger larvae) to the bottom and the juveniles develop into adult abalone (Coastal and Marine Life, online; Fishtech, Inc, online).

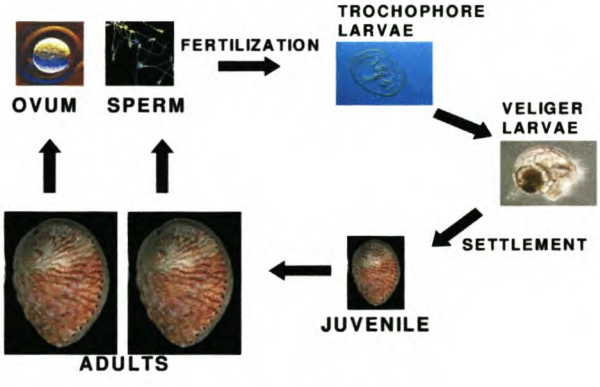


Figure 1.2: The life cycle of the abalone.

1.1.4) NATURAL ENEMIES

Filter feeders such as the mussel, *Mytilus edulis*, predate on abalone eggs and larvae. Juvenile abalone are the prey of crabs, lobsters, octopuses, starfish, fish and predatory snails. Adult abalone fall prey to eagle rays, bat rays, cabezon and sea otters. Abalone are also the prey of the most effective predator around: humans (Coastal and Marine Life, online; Fishtech, Inc, online).

Stellenbosch University http://scholar.sun.ac.za

1.1.5) FOOD

Abalone are herbivores. While juveniles feed on biofilms or algae encrusted on rocks, adults feed on drifting algae pieces such as kelp that are trapped by its foot (Coastal and Marine Life, online; Fishtech, Inc, online).

1.1.6) HABITAT

Settled larvae and juvenile abalone shelter in the spiny sea urchin fields or crevices in rocks. Adult abalone live in large groups in kelp forests. The depth and temperatures at which the animals live differ between species. Temperatures range from 2°C to 30°C and water depth from 0m to 140m. (Coastal and Marine Life, online; Fishtech, Inc, online; Seafood Watch, 2003).

1.1.7) COMMERCIAL VALUE

Abalone is an important commercial species all over the world. The different species of abalone being commercially exploited are shown in Table 1.2. During 1999 Australia for example, yielded a total of 5500 metric tons (Fishtech, Inc, online) of abalone from the wild and China yielded 3500 metric tons (Gordon & Cook, 2001) of abalone from its farming industry.

Table 1.2: Abalone species being commercially exploited internationally.

SPECIES	LOCATION
Haliotis rufescens	North America, Chile
Haliotis cracherodii	North America
Haliotis fulgens	North America
Haliotis corrugata	North America
Haliotis kamtschatkana	North America
Haliotis midae	South Africa
Haliotis laevigata	South Australia
Haliotis rubra	South Australia
Haliotis conicopora *	South Australia
Haliotis asinine #	Australia
Haliotis roei	Australia
Haliotis iris	New Zealand
Haliotis diversicolor supertexta	Taiwan
Haliotis discus hannai	Japan, China
Haliotis tuberculata +	Ireland

(Seafood Watch, 2003)
* (Government of Western Australia, online)
# (CSIROnline, 1998)
+ (Irish Sea Fisheries Board, online)

Stellenbosch University http://scholar.sun.ac.za

1.1.8) THREATS

The last 20 years shows a worldwide decrease in abalone catches of 50%. This can be attributed to the following (Fishtech, Inc, online):

1) Predators of abalone expanding their territory, for example – sea otters in California.
2) Increasing mortality rates of juvenile abalone, for example – the destruction of spiny sea urchins at Hermanus (South Africa) is exposing juvenile abalone to predators.
3) Over exploitation by commercial fisheries.
4) Competition for food from other kelp-eating organisms.
5) Poaching of abalone.
6) Pollution and coastal development programs are destroying the natural habitat of abalone.

1.2) SOUTH AFRICAN ABALONE

1.2.1) GENERAL

The South African coastal area and kelp forests are the habitat of six different members of the genus *Haliotis*. The geographical distribution of the six different species is shown in Figure 1.3 and their taxonomy is listed in Table 1.3.

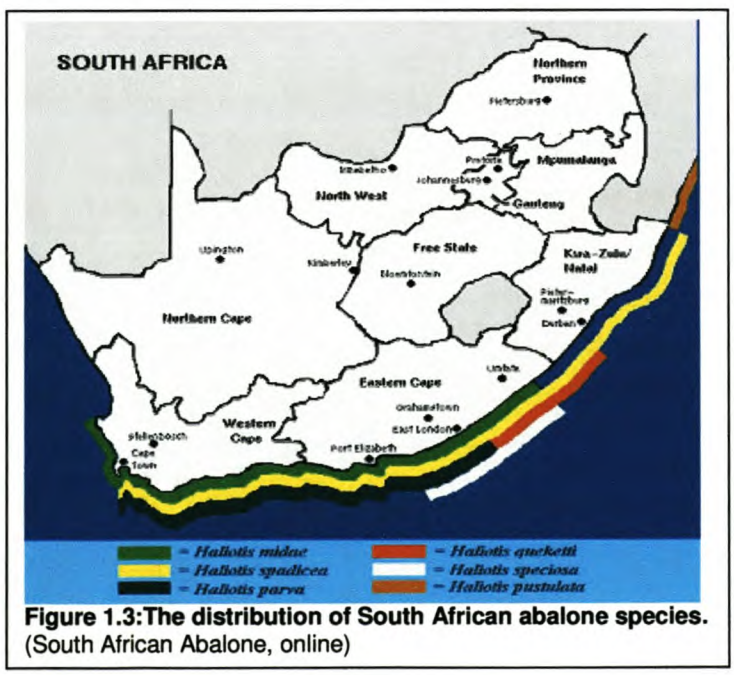


Figure 1.3:The distribution of South African abalone species. (South African Abalone, online)

Stellenbosch University http://scholar.sun.ac.za

Table 1.3: South African abalone species.

South African Abalone Species	
Haliotis midae	
Haliotis parva	
Haliotis pustulata	
Haliotis queketti	
Haliotis spadicea	
Haliotis speciosa	

(Hardy(b), online)

Stellenbosch University http://scholar.sun.ac.za

Haliotis midae is the most abundant South African species and it is also the largest growing of the six species. These factors make it suitable for commercial exploitation (South African Abalone, online). In 2002 South Africa had the fourth largest abalone industry in the world (Redmayne, 2002).

Unfortunately, it is predicted that this valuable resource will become commercially extinct within the next three to four years (Earth Crash Earth Spirit, 2003). The greatest reason for the decline is the resurgence of illegal harvesting since 1994. The abalone is a slow growing animal and only reaches sexual maturity at seven years. The poachers remove both sexually mature and immature animals. Whole communities and future generations are thus lost. In addition, an extra threat is posed by a recent influx of the west coast lobster *Jasus lalandi* into the south coast. The prey of the lobster is the spiny sea urchin, *Parechinus angulosus*, which serves as protection for settled abalone larvae. Without the much-needed cover of the sea urchins, the abalone larvae are more vulnerable to predators and mortality among juveniles will rise (Coastal and Marine Life, online; Fishtech, Inc, online).

Two strategies can be utilised to save this commercially important animal:

1) Abalone ranching
2) Abalone farming

1.2.2) ABALONE RANCHING

Abalone ranching or reseeding is the practice of introducing farmed juvenile abalone into the wild in order to increase the depleted biomass. So far, abalone have been reseeded at two locations in South Africa, McDougall's Bay near Port Nolloth (abalone does not occur naturally here) on the west coast and Gouriqua between Stilbaai and Gouritsmond on the south coast. Recovery rates vary between the locations, with McDougall's Bay showing the highest recovery rate (Sweijd *et al*, 1998; De Waal *et al*, 2003). Successful reseeding has also been reported for red abalone, *Haliotis rufescens*, in California (Gaffney *et al*, 1996), while experimental reseeding of other

Stellenbosch University http://scholar.sun.ac.za

species, such as the Gastropoda, *Trochus niloticus* (Crowe *et al*, 2002) is taking place in the Pacific Ocean.

Many environmental, biological and genetic factors influence the success of reseeding programs. Environmental factors include wave and surf activity (De Waal *et al*, 2003) and competition with other species (Lachance & Magnan, 1990). Biological factors such as the age at which the juvenile abalone are reseeded (Sweijd *et al*, 1998) and the population density (reported for the Gastropoda *Trochus niloticus*, Crowe *et al*, 2002) may also influence the outcome of the reseeding. The reseeded populations may also out perform and drive out other natural populations of the same or different genus already present in the area. This situation was encountered for the coastal cutthroat trout, *Oncorhynchus clarki clarki* (Bates & McKeown, 2003). One of the genetic factors is the danger of genetic deterioration in captivity, specifically the adaptation of the organism to captivity. This will decrease the organism's ability to survive in the wild (Frankham & Loebel, 1992). Another genetic factor is the danger posed by bottlenecks, caused by the reseeding of a population with low genetic variation (Ayllon *et al*, 2004). The lack of genetic variation will lower the survival rate of the reseeded individuals, because they will be less adaptable to environmental changes (Allendorf & Phelps, 1980). Abalone ranching will only be successful if all these factors are taken into consideration to develop effective strategies.

1.2.3) ABALONE FARMING

Abalone farms may provide an answer to relieve the pressure from the declining wild populations. Large fishing companies own most of the abalone farms, because a farm requires a large capital base (Cook, online). There are currently fourteen farms registered in South Africa (pers. comm.: Stephan Ashlin, Sea Plant Products).

Most farms have both a hatchery and a grow-out facility. The hatchery is for spawning and settling of the larvae and the grow-out facility is where the

Stellenbosch University http://scholar.sun.ac.za

animals are kept until they are exported. A farm needs a brood stock to start out. Mature male and female animals are randomly collected from the wild and taken to the farm. The brood stock can be kept together or they can be divided into different new groups. A commercial stock is created by spawning the brood stock, settling the larvae (in the hatchery) and feeding them until they reach the correct export size (in the grow-out facility), usually around age three (pers. comm.: Stephan Ashlin, Sea Plant Products). The commercial stock is fed with kelp collected at sea or by artificial feed (www.fishtech.com/facts.html). More recently, South African scientists have introduced probiotics in abalone feed to increase the uptake of nutrients, thus giving an advantage to the local industry (www.scienceinafrica.co.za). The commercial stock is kept in large concrete structures called raceways. Raceways are compartmentalized with each compartment holding a certain number of animals. The compartments are also useful for keeping track of cohorts and feeding schedules. Seawater is pumped into the raceways by large pumping systems that relay water to the sea and back (pers. observation during visit to farm).

There are a few factors influencing the production success of farms such as the type of feed that is fed to the commercial abalone stock (www.fishtech.com/facts.html), the control of diseases and the maintenance of equipment, but an important factor is to have the knowledge of genetic diversity within the natural and farm populations. It all starts when the brood stock is selected from the wild. A brood stock that is representative of the genetic composition of the natural population should be obtained. This means that the brood stock must carry all or most of the alleles present in the natural population from which they were taken. This task is nearly impossible without prior genetic typing and screening of the natural population. The best way to ensure a genetic diverse brood stock is to catch animals from different locations within the zone of the natural population in order to avoid sampling animals belonging to the same family. Catch animals that are from different generations or even import animals from other geographical areas. However, the local brood stock is adapted for local conditions and if the local brood stock is hybridised with an imported brood stock, the adaptations may be

Stellenbosch University http://scholar.sun.ac.za

disrupted. This will have negative effects on the animals' survivability during future reseeding operations. These methods may reduce the chance of having too many siblings within the brood stock that will increase the likelihood of inbreeding (Ballou & Lacy, 1995; pers observation).

It is important to remember that when a brood stock is assembled a new population is created. By sub-dividing the original brood stock more new sub-populations are created. Animals with reduced genetic variation and related animals should not be used for the sub-brood stock to avoid the loss of genetic variation in the F1-population.

It becomes important to protect the genetic variation of this new population to ensure the ability of the animals to respond to environmental (Allendorf & Phelps, 1980; if they are used for later reseeding programs), production or selection pressures (Hill, 2000). The loss of genetic variation will negatively affect important commercial traits such as growth rate or meat quality (Koehn *et al*, 1988). Too little genetic variation may lead to the eventual extinction of the reseeded population (Frankham, 1995). In instances where reseeding takes place at a location where a natural population is already present, genetic variation, in conjunction with genetic similarity, plays an important role in conserving the genetic integrity of the local population. By reseeding animals that are genetically similar to the local natural population, it is ensured that no foreign alleles are introduced that could disturb the current evolutionary equilibrium existing between the animal and its environment. By reseeding animals with high genetic variation the bottleneck effect will be minimized (Ayllon *et al*, 2004) and ensure that some alleles won't dominate and eventually oust the rare alleles. A prior knowledge of the genetic composition of the reseeding stock or farm brood stock and natural population is recommended in order to assess the influence of something such as a massed escape of farm stocks into the wild for whatever reason.

The loss of genetic variation within a commercial stock can be ascribed to a few factors. A primary reason could be that the animals that were collected for the brood stock is not genetically diverse enough and therefore their

Stellenbosch University http://scholar.sun.ac.za

offspring, the commercial stock, will not be as genetically diverse. This is called a bottleneck effect. The bottleneck effect is the cause of the loss of genetic variation in farmed Pacific abalone, *Haliotis discus hannai* (Li *et al*, 2004), and Japanese flounder, *Paralichthys olivaceus* (Sekino *et al*, 2002). A second factor is ascribed to genetic drift due to small effective population sizes. This means that only a few of the sexually mature animals contribute genetic material to the offspring (Boudry *et al*, 2002). The influence of this factor is recorded for the Blacklip abalone, *Haliotis rubra*, Perlemoen, *Haliotis midae* (Evans *et al*, 2003) and the Pacific oyster, *Crassostrea gigas* (Hedgecock & Sly, 1990). A third factor is the sub-division of an original brood stock into smaller groups. This sub-division may lead to a loss of genetic variation as is found in a study on the mountain brushtail possum, *Trichosurus caninus* (Lacy & Lindenmayer, 1995).

It is therefore very important from a genetic management point of view, to be armed with the knowledge of the molecular structure of the farm populations. Many molecular markers and data analysis techniques are available with which to construct an accurate picture for farm management.

1.3) MOLECULAR MARKER SYSTEMS

1.3.1) AFLP (AMPLIFIED FRAGMENT LENGTH POLYMORPHISM)

AFLPs are dominant, polymorphic, multi-locus markers that are not limited to one species. AFLP markers were first used by Vos *et al* (1995). The technique comprises the digestion of whole genomic DNA by restriction enzymes. Adaptors of known sequence are attached to the various fragments of unknown sequence. The adaptors serve as priming sites during PCR. The restriction fragments are selected for by degenerate primers in order to lessen their number. AFLPs are analysed by labelling primers either with fluorescent labels or radioactive labels and visualising and scoring them with automated gene sequencers or more traditional electrophoretic methods (Liu & Cordes, 2004).

Stellenbosch University http://scholar.sun.ac.za

1.3.2) ALLOZYMES

Allozymes are polymorphic, codominant, single locus markers that are allelic variants of proteins. The earliest use of these markers for aquatic species was by May *et al* (1980). Kijimi *et al* (1992) used allozymes to study the Pacific abalone *Haliotis discus hannai*. The technique comprises the search for variations in polypeptide chains that will indicate allele variations at DNA level. Allozymes are separated and scored with starch gels (Liu & Cordes, 2004).

1.3.3) ESTs (EXPRESSED SEQUENCE TAGS)

ESTs are derived from sequencing of cDNA clones during transcriptomic research. ESTs are used to detect genes and study their expression under different physiological or temporal conditions. EST's are used to generate molecular markers like SNP's (see section 1.3.8). These markers are mostly used for physical and linkage mapping in animal genomics (Liu & Cordes, 2004). Tong *et al* (2002) studied ESTs in the shrimp *Penaeus monodon*. SSCP analysis was used to detect intron-length polymorphisms.

1.3.4) MINISATELLITES

Minisatellites are repeat units of >10bp. Minisatellites are polymorphic markers that could be either multi-locus or multi-allelic (single-locus probes). They are found in the coding and non-coding regions of the genome (Oura *et al*, 2003). Variation is detected by differences in genetic profiles or allele lengths. Minisatellites are analysed by using PCR and electrophoretic methods (Armour *et al*, 1999). Huang *et al* (2000) analysed Blacklip abalone (*Haliotis rubra*) using minisatellites. The more powerful microsatellite (see section 1.3.9; p13) marker is replacing minisatellites.

Stellenbosch University http://scholar.sun.ac.za

1.3.5) MITOCHONDRIAL DNA MARKERS (mtDNA)

mtDNA are single-locus (mostly only maternally inherited) and multi-allelic and polymorphisms are detected through RFLP's, SNP's (see sections 1.3.7 & 1.3.8) or sequencing. Polymorphisms occurr either within coding or non-coding regions (Liu & Cordes, 2004). The polymorphisms are visualised by electrophoretic methods or automated sequencers.

1.3.6) RAPDs (RANDOM AMPLIFIED POLYMORPHIC DNA)

RAPDs are multi-locus, bi-allelic and dominant markers. RAPDs were developed in 1990 (Welsh & McClelland, 1990; Williams *et al*, 1990). Huang *et al* (2000) used Rapd's to study Blacklip abalone (*Haliotis rubra*). The technique uses an identical primer pair (8 – 10 bp, Tm = 36 – 40°C) to amplify multiple fragments from genomic DNA. The polymorphism is measured by the presence or absence of an amplified product. RAPDs are visualised with electrophoretic methods (Liu & Cordes, 2004).

1.3.7) RFLPs (RESTRICTION FRAGMENT LENGTH POLYMORPHISM)

RFLPs are co-dominant, bi-allelic or multi-allelic and polymorphic markers. The technique uses restriction enzymes that recognise specific 4, 5, 6 or 8 bp nucleotide sequences, to digest genomic DNA into fragments. Variation is observed in the number and size of the fragments. Klinbunga *et al* (2003) analysed both nuclear and mtDNA using RFLP analysis in the Tropical Abalone *Haliotis asinina*. New analysis strategies include PCR-based techniques to detect deletions and insertions at restriction sites, as well as PCR-based techniques incorporating restriction enzymes in order to detect base substitutions. RFLPs are visualised with electrophoretic methods or Southern blotting (Liu & Cordes, 2004).

Stellenbosch University http://scholar.sun.ac.za

1.3.8) SNPs (SINGLE NUCLEOTIDE POLYMORPHISM)

SNPs are bi-allelic or tetra-allelic, co-dominant markers for which the genotyping process can be automated. SNP's are the most abundant polymorphism found in living organisms. The variation detected by SNPs lies in single nucleotide differences between individuals and many techniques for genotyping these differences are being used. Some of the techniques usually involve expensive and specialised equipment, while others using SSCP are not expensive at all. For more on the techniques employed, please see the review of Liu & Cordes (2004).

1.3.9) MICROSATELLITES

Microsatellites will be covered more extensively, because it is the marker of choice for this study (see section 1.6; p23).

1.3.9.1) Introduction

Microsatellites, also known as single sequence repats (SSRs; Edwards *et al*, 1991), are tandemly repeated motifs, usually 1 – 6 basepairs in length. They are present mostly in non-coding regions, but can also be found in coding regions, of the genome, but non-coding in themselves. Microsatellites are present in all organisms studied so far. They can be isolated from genomic libraries using oligonucleotide probes and markers for microsatellite amplification can then be designed (Hancock, 1999; Zane *et al*, 2002). They show high levels of length polymorphism. The origins of the polymorphism are still under discussion, but it could be due to DNA polymerase slippage during replication or the product of recombination events. Microsatellite markers are codominant, multi-allelic and adaptable to automation (Hancock, 1999; Huang *et al*, 2000; Liu & Cordes, 2004). They have the greatest mutation rate of any known molecular marker, which makes it ideal for population genetics and molecular ecology studies at intraspecific level. Conservation beyond species taxonomic level depends on the speciation age

Stellenbosch University http://scholar.sun.ac.za

of the taxonomic group. Since the first descriptions of microsatellites (Litt & Luty, 1989; Tautz, 1989; Weber & May, 1989), their use and usefulness increased in many fields.

1.3.9.2) Uses

The uses of microsatellite markers include the assessment of genetic variation, studies on population structure, parentage assignments, phylogeograhic studies and the search for QTL's for marker assisted selection (Estoup *et al*, 1998; Hansen *et al*, 1999; Norris *et al*, 2000; Launey *et al*, 2001; Sekino *el al*, 2002; Robinson *et al*, 2003; Evans *et al*, 2004).

1.3.9.3) Comparison to other marker systems

Microsatellite markers detect higher levels of genetic variation than allozyme markers (Li *et al*, 2004), but are more expensive. Compared to allozymes, microsatellite markers have better statistical power in statistical tests differentiating between populations (Triantafyllidis *et al*, 2002). They are simpler to score than anonymous dominant markers like RAPDs or AFLPs. The microsatellite markers of one organism can sometimes be used for other closely related species and the analysis steps can be automated (Huang *et al*, 2000). In comparison with RAPD markers, microsatellites are more adept at discerning fine-scale genetic differences between populations isolated by short distances (Lougheed *et al*, 2000).

1.3.9.4) Species-specific microsatellite markers versus non-species specific microsatellite markers

Even though microsatellite markers are conserved across species (Huang *et al*, 2000), it is more advantageous to use species-specific microsatellite markers. Table 1.4 show a few abalone species for which species-specific microsatellite markers have been designed.

Stellenbosch University http://scholar.sun.ac.za

The species-specific markers will increase the accuracy and reliability of any population related studies, because more polymorphisms will be detected in comparison to non-specific markers.

Table 1.4: Species for which species-specific microsatellite markers have been designed.

SPECIES		REFERENCE
Common Name	**Scientific Name**	
Blacklip abalone	*Haliotis rubra*	Evans *et al*, 2000
Ezzo abalone	*Haliotis discus hannai*	Li *et al*, 2002
Perlemoen	*Haliotis midae*	Bester *et al*, in press

1.3.9.4.1) Accuracy and polymorphism

The number of polymorphisms present in microsatellite loci is dependant on allele sizes. The mutational process influencing allele sizes is bias and favours expansion. Mutation rate increases with allele size. Loci with small alleles have less polymorphism than their homologous loci with larger alleles. It is possible to argue that two related lineages (one locus with small and one locus with large alleles) can arise, one characterised by low repeat number and low variability and one by high repeat number and high variability. Monomorphic and polymorphic homologues can be created between two closely related species. Crawford *et al* (1998) demonstrated this in sheep and bovine, where monomorphic loci in sheep had polymorphic homologues in bovine (Amos, 1999).

The difference in allele sizes between related species may also be influenced by rapid numerical expansion from time to time by one of the species, while the numbers of the other species stay constant. By taking the expansion-prone mutation events into account, the assumption is made that expanded populations contain larger alleles than their homologues in smaller populations (Rubinsztein *et al* 1995). More mutations will take place within the expanded population, because the mutation rate is influenced by allele size. After every expansion event more heterozygous loci is present, because

Stellenbosch University http://scholar.sun.ac.za

of all the new mutations. During the following expansion events heterozygote instability is responsible for the rapid microsatellite mutation rates. Rubinsztein *et al* (1995) saw this when they compared homologous microsatellite loci between the rapidly expanding human population and the stable chimpanzee (*Pan troglodytes*) population. The human homologues are much larger than the chimpanzee homologues (Amos, 1999).

1.3.9.4.2) Reliability

Species-specific markers are also guaranteed to amplify the desired segments it was designed to amplify within the species it was designed for. This can be seen in a study of Kondo *et al* (1997) where mouse microsatellite markers were used in genome mapping of the rat. Of 815 mouse microsatellite markers only 15% of the primer pairs amplified rat DNA, while only 4.8% of the markers showed any levels of polymorphism. Non-species specific microsatellite markers on the other hand, are useful for synteny or relatedness studies (Kondo *et al*, 1997).

1.3.9.4.3) Conclusion

The size and level of polymorphism in homologues loci differ between related species (see section 1.3.9.4.1; p15). A marker designed in the one species will detect less polymorphism in the other related species, because a microsatellite marker is usually designed in the species with the biggest population size containing larger alleles and more polymorphism per locus. To detect loci showing desirable allele size and level of polymorphism in the smaller species, microsatellite markers should be designed within that species (Amos, 1999). The same applies regarding the reliability (see section 1.3.9.4.2; p15) of the microsatellite markers. To ensure that the markers amplify within a specific species, the markers should be designed within that species.

Stellenbosch University http://scholar.sun.ac.za

1.3.9.5) Genotyping

For the present work, the genotyping process was based on scoring alleles for each locus by visually selecting and assigning alleles from electropherograms generated by labelled primers using automated gene sequencers (author's def). This is the most important step in microsatellite analysis, because without the correct genotypes of individuals, all other statistical and visual analysis will be incorrect (pers. observation). This step should therefore be optimised to achieve the best quality and most accurate electropherograms. The quality and accuracy of the electropherograms are influenced by a few factors that should be given attention during optimisation:

1) The quality and amount of material used for DNA extractions (Goossens *et al*, 1998).
2) Optimum PCR conditions are necessary for the amplification of the correct alleles of the analysed locus (Rahman *et al*, 2000; Fernando *et al*, 2001).
3) The concentration of the PCR product can also affect the analysis of the ABI-series sequencers and may lead to false identification of heterozygotes or incorrect allele assignment (Fernando *et al*, 2001).
4) Allele dropout occurs when the amplification of one allele is weaker than the other allele. This will influence the intensity of the peaks on the electropherogram and may cause the false identification of homozygotes (Gagneux *et al*, 1997).
5) Amplification artefacts caused by slippage events during PCR amplification may cause trouble with allele identification (Schlotterer & Tautz, 1992).
6) Mutations within the priming sites of any one primer of a microsatellite primer pair can cause the non-amplification of an allele or even both alleles. This phenomenon is known as null alleles. These null alleles will cause the identification of true heterozygotes as homozygotes. Null alleles could also affect parentage analysis in a positive or negative way. The missing data caused by null alleles could also have

Stellenbosch University http://scholar.sun.ac.za

quantitative effects on QTL-marker discovery, as well as population studies (Callen *et al*, 1993; Pemberton *et al*, 1995).

1.4) MOLECULAR STUDIES

1.4.1) COMPARATIVE STUDIES OF FARMED AND WILD POPULATIONS

A few molecular studies have been done on farmed aquatic species to assess their levels of genetic variation and genetic similarities towards natural populations. A few examples using some of the markers discussed under section 1.3; p10 – 17, will be described here. The markers are used to detect genetic variation and differentiation within and between various populations.

1.4.1.1) Allozyme markers

Allozyme markers detected no genetic differentiation or loss of genetic variation between the wild and farmed populations of the Green ormer, *Haliotis tuberculata* (Mgaya *et al*, 1995) and the Pacific oyster, *Crassostrea gigas* (English *et al*, 2000).

1.4.1.2) Microsatellite DNA markers

Microsatellite markers indicated a loss of the observed number of alleles between the wild and farmed populations of Atlantic salmon, *Salmo salar* (Norris *et al*, 2000), Brown trout, *Salmo trutta* (Was & Wenne, 2002), Japanese flounder, *Paralichthys olivaceus* (Sekino *et al*, 2002), Blacklip abalone, *Haliotis rubra* (Evans *et al*, 2004), Perlemoen, *Haliotis midae* (Evans *et al*, 2004) and Ezzo abalone, *Haliotis discus hannai* (Li *et al*, 2004). Results for average expected heterozygosity differed between studies (see section 4.4; p100 – 103). Atlantic salmon, *Salmo salar* (Norris *et al*, 2000) and Ezzo abalone, *Haliotis discus hannai* (Li *et al*, 2004) had differences between wild and farmed populations while the opposite was true for Brown trout, *Salmo trutta* (Was & Wenne, 2002), Japanese flounder, *Paralichthys olivaceus*

Stellenbosch University http://scholar.sun.ac.za

(Sekino *et al*, 2002), Blacklip abalone, *Haliotis rubra* (Evans *et al*, 2003) and Perlemoen, *Haliotis midae* (Evans *et al*, 2004).

1.4.1.3) Mitochondrial DNA markers

Studies using mtDNA markers showed a loss of genetic variation between the wild and farmed populations of the Japanese freshwater fish Ayu, *Plecoglossus altivelis* (Iguchi *et al*, 1999) and another study detected genetic differentiation between the wild and farmed populations of Brown trout, *Salmo trutta* (Morán *et al*, 1996).

1.4.2) STATISTICAL PARAMETERS FOR MOLECULAR STUDIES

1.4.2.1) Allele frequencies

The most important numerical data to be obtained from genotype data is the frequency of the occurrence of every allele in the data set. All statistical formulae used to calculate genetic variation, genetic differentiation, parentage analysis and phylogenetic analysis, uses allele frequency data.

1.4.2.2) Genetic variability

The most common parameters for measuring the genetic variation of diploid organisms are the number of observed alleles due to the loss of low-frequency alleles. This parameter is also known as allelic diversity and is dependent on the size of a data set (Sekino *et al*, 2002). Another parameter called allelic richness is independent from the size of the data set. Allelic richness measurements are influenced by rare alleles (El Mousadik & Petit, 1996). Genetic variation is also measured by the differences of average expected heterozygosities due to the frequency shifts of common alleles (Smith & Conroy, 1992). Heterozygosity values can also be used to detect the presence of null alleles within data sets (Brookfield, 1996).

Stellenbosch University http://scholar.sun.ac.za

A reduction in either allelic diversity or heterozygosity will indicate a loss of genetic variation.

1.4.2.3) Genetic differentiation and similarity

Genetic differentiation or population genetic structure is the differential distribution of alleles across the subpopulations of species. The reasons for limited gene flow among groups range from the presence of physical barriers (ex: The Great Wall of China; Su *et al*, 2003), local inbreeding (for example: Plateau pikas; Dobson *et al*, 2000), and geographical isolation (for example: *Hediste diversicolor*; Virgilio & Abbiati, in press).

Measures of genetic differentiation for microsatellites are based on models of mutation. Two available models are the infinite allele model and the stepwise mutation model. Interpretable numerical values are given for each model in order to analyse genetic differentiation or similarity.

1.4.2.3.1) Infinite allele model

In this model new mutations, are generated at random and alleles of all sizes are equally likely to originate in one mutation step. The size of the new mutant allele is independent of its progenitor (Goodman, 1997). An example of a histogram of allele distributions is shown in Figure 1.4. The alleles shown in red are the new mutant alleles and the green allele is the progenitor.

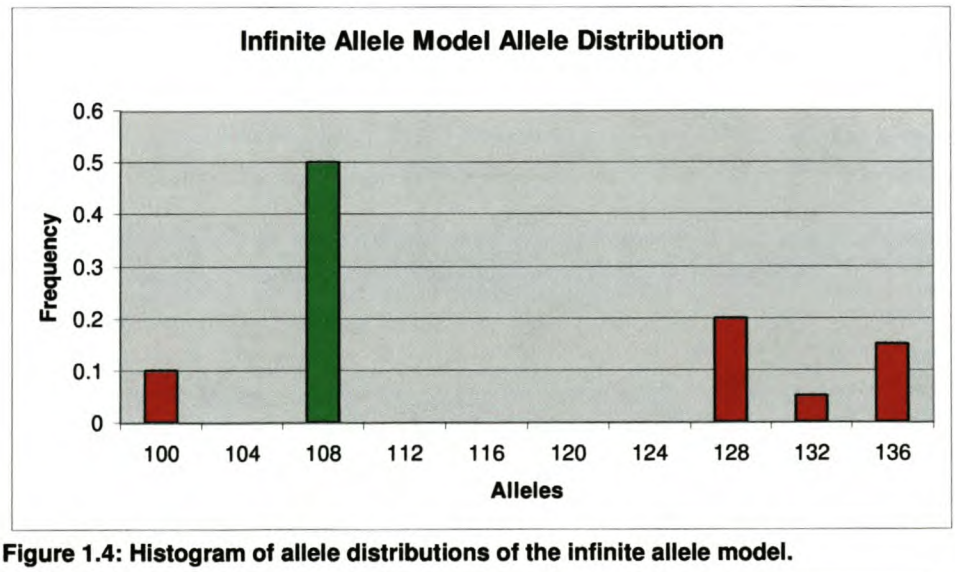


Figure 1.4: Histogram of allele distributions of the infinite allele model.

Stellenbosch University http://scholar.sun.ac.za

1.4.2.3.2) Stepwise mutation model

New mutant alleles are created by small nucleotide or repeat unit deletions or additions to the original allele (Goodman, 1997). Theoretically only two new alleles can arise per generation. An example of a histogram of allele distributions is shown in Figure 1.5. The green allele is the progenitor. The red alleles are the potential mutant alleles that could be created in the next generation. The yellow alleles are the potential mutant alleles that could be created two generations on. The blue and black alleles could be created respectively three and four generations later. The histogram of the stepwise mutation model differs from the histogram of the infinite allele model in the respect that the first model covers the whole allelic size range, while the latter allows for missing allelic sizes.

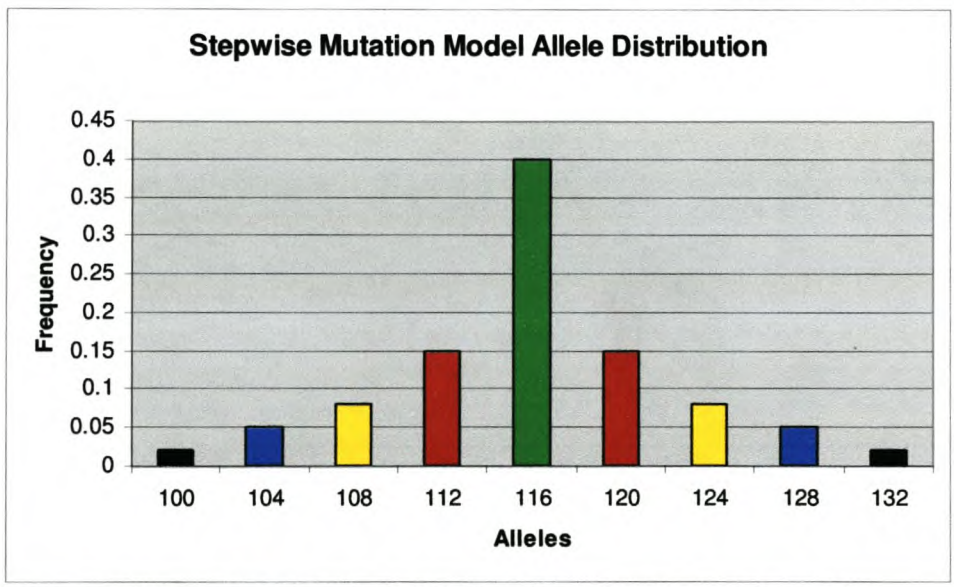


Figure 1.5: Histogram of the allele distribution of the stepwise mutation model.

1.4.2.3.3) F_{st} and R_{st}

F_{st} values (Weir & Cockerham, 1984) are given for the infinite allele model and R_{st} values (Slatkin, 1995) are given for the stepwise mutation model.

Stellenbosch University http://scholar.sun.ac.za

Small F_{st} and R_{st} values suggest less genetic differentiation and more genetic similarity between data sets.

F_{st} and R_{st} values are numerical parameters for population structure. Both F_{st} and R_{st} differentiate between pairs of populations. F_{st} is calculated by taking different allele frequencies into account, while R_{st} is calculated by taking different allele sizes into account. Both parameters are affected by sampling methods and sample size. A recent study by O'Reilly *et al* (2004) suggested that the higher the level of polymorphism within a locus, the less the power of differentiation will be when F_{st} is used. R_{st} is a useful parameter for studies using microsatellite loci, because microsatellite evolution is hypothesised to follow a stepwise mutation model (see section 1.4.2.3.2; p21) and is size dependant (see section 1.3.9.4.1; p15). F_{st} could be used for microsatellite analysis, but it depends on the allele distribution of the locus being studied (see section 1.4.2.3.1; p 15; Weir & Cockerham, 1984; Slatkin, 1995).

1.5) QUANTITATIVE TRAIT LOCI

Traits that cause the differences observed in phenotypes that can't be explained by a simple model of inheritance are called quantitative traits. Many genes control quantitative traits and each gene will have a small effect on the trait. Quantitative traits may also be influenced by environmental factors.

The identification of the genes that control the quantitative traits is important for animal breeders, because many economically important traits, such as growth rate in abalone (Robinson *et al*, 2003), are influenced by many genes. The loci that control the traits are commonly called quantitative trait loci (QTL's). Strategies for discovering QTL's involve the construction of genetic maps and searching for a relationship between the desired trait and a locus by using polymorphic markers. Significant associations between a locus and a specific trait may be evidence of a QTL near that trait (Liu, 1997). This QTL can be used by the animal breeder to pre-select breeding animals for stock-

Stellenbosch University http://scholar.sun.ac.za

improvement breeding programs. For example: if a QTL for fast growth rate in abalone was available, fast growing animals could be selected for breeding and potentially produce fast growing offspring. The fast growing offspring will reach market size at an earlier age, thus saving the breeder time and money (Robinson *et al*, 2003).

QTL's are available for some fish species like rainbow trout, *Oncorhynchus mykiss* (Sakamoto *et al*, 1999), but none are available for abalone. Robinson *et al* (2003) have done preliminary studies using *Haliotis rubra* and *Haliotis laevigata* to search for QTL's segregating with the growth rate trait.

It is obvious that potential use of QTL's for the genetic improvement of cultured abalone is enormous. The only limiting resource necessary to develop this useful molecular tool is money!

1.6) AIMS OF THIS STUDY

The aims of this study are the analysis of the genetic structure of the Sea Plant Products (SPP) farm population with species-specific microsatellite DNA markers to:

1. Assess the levels of diversity and differentiation within the natural populations, the brood stock and their F1.
2. Parentage assignment in the F1 produced by mass-spawning practices.
3. Conduct a preliminary study to assess their potential for QTL discovery.

Stellenbosch University http://scholar.sun.ac.za

1.7) REFERENCES

Allendorf, F.W., Phelps, S.R., 1980. Loss of genetic variation in a hatchery stock of cutthroat trout. *Transactions of the American Fisheries Society* **109**, 537 – 543.

Amos, W., 1999. A comparative approach to the study of microsatellite evolution. In Goldstein, D.B., Schlötterer, C. *Microsatellites, Evolution and Applications*. Oxford: Oxford University Press, 66 – 79.

Armour, J.A.L., Alegre, S.A., Miles, S., Williams, L.J., Badge, R.M., 1999. Minisatellite and mutation processes in tandemly repetitive DNA. In Goldstein, D.B., Schlötterer, C. *Microsatellites, Evolution and Applications*. Oxford: Oxford University Press, 24 – 33.

Ayllon, F., Patrick Davaineb, P., Beallb, E., Martineza, J.L., Garcia-Vazqueza, E., 2004. Bottlenecks and genetic changes in Atlantic salmon (*Salmo salar* L.) stocks introduced in the Subantarctic Kerguelen Islands. *Aquaculture* **237**, 103 – 116.

Ballou, J.D., Lacy, R.C., 1995. Identifying genetically important individuals for management of genetic diversity in pedigreed populations. In Ballou, J.D., Gilpin, M., Foose, T. *Analytical Methods and Strategies in Small Population Conservation*. New York: Columbia University Press, 76 – 111.

Bates, D.J., McKeown, B.A., 2003. Growth in stream-stocked juvenile hatchery-reared coastal cutthroat trout (*Oncorhynchus clarki clarki*) and the implications for wild populations. *Aquaculture* **222**, 215 – 228.

Bester, A.E., Slabbert, R., D'Amato, M.E. Isolation and characterisation of microsatellite markers in South African abalone (*Haliotis midae*). In press. *Molecular Ecology Notes*. (Appendix A)

Stellenbosch University http://scholar.sun.ac.za

Boudry, P., Collet, B., Cornette, F., Hervouet, V., Bonhomme, F., 2002. High variance in reproductive success of the Pacific oyster *Crassostrea gigas*, Thunberg revealed by microsatellite-based parentage analysis of multifactorial crosses. *Aquaculture* **204**, 283 – 296.

Brookfield, J.F.Y., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**, 453 – 455.

Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C., Sutherland, G.R., 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics* **52**, 922 – 927.

Coastal and marine life [Online]. Available: sacoast.uwc.ac.za/publications/3A-Classification.pdf. [2004, August]

Cook, P. Abalone farms in South Africa. [Online]. Available: http://web.uct.ac.za/depts/zoology/abnet/newslet1/saupdate.html. [2004, August]

Crawford, A., Knappes, S.M., Peterson, K.A., De Gotari, M.J., Dodds, K.G., Freking, R.T., Stone, R.T., Beattie, C.W. 1998. Microsatellite evolution: testing the ascertainment bias hypothesis. *Journal of Molecular Evolution* **46**, 256 – 260.

Crowe, T.P., Lee, C.L., McGuinness, K.A., Amos, M.J., Dangeubun, J., Dwiono, S.A.P., Makatipu, P.C., Manuputty, J., N'guyen, F., Pakoa, K., Tetelepta, J., 2002. Experimental evaluation of the use of hatchery-reared juveniles to enhance stocks of the topshel *Trochus niloticus* in Australia, Indonesia and Vanuatu. *Aquaculture* **206**, 175 – 197.

CSIROnline, 1998. 'Cocktail abalone to boost exports to Asia [Online]. Available:

Stellenbosch University http://scholar.sun.ac.za

http://www.csiro.au/index.asp?type=mediaRelease&id=CocktailAbaloneToBoostExports&stylesheet=mediaRelease. [2004, August]

De Waal, S.W.D., Branch, G.M., Navarro, R., 2003. Interpreting evidence of dispersal by *Haliotis midae* juveniles seeded in the wild. *Aquaculture* **221**, 299 – 310.

Dobson, F.S., Smith, A.T., Gao, W.X., 2000. The mating system and gene dynamics of plateau pikas. *Behavioural Processes* **51**, 101 – 110.

Earth Crash Earth Spirit, 2003. Scientists warn that South Africa's perlemoen (abalone) could be commercially extinct within 3-4 years due to rampant poaching by criminal gangs to supply the demand for the delicacy in parts of Asia [Online]. Available: http://www.eces.org/articles/000263.php. [2004, August]

Edwards, A., Civitello, A., Hammond, H.A., Caskey, C.T., 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* **49**, 146 – 756.

El Mousadik, A., Petit, R.J., 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* **92**, 832 – 839.

English, L.J., McGuire, G.B., Ward, R.D., 2000. Genetic variation of wild and hatchery populations of Pacific oyster, *Crassostrea gigas* (Thunberg) in Australia. *Aquaculture* **187**, 283 – 298.

Estoup, A., Gharbi, K., SanCristobal, M., Chevalet, C., Haffray, P., Guyomard, R., 1998. Parentage assignment using microsatellites in turbot (Scophthalmus maximus) and rainbow trout (Oncorhynchus mykiss) hatchery populations. *Canadian Journal of Fisheries and Aquatic Sciences* **55**, 715 – 725.

Stellenbosch University http://scholar.sun.ac.za

Evans, B., White, R.W.G., Elliott, N.G., 2000. Characterisation of microsatellite loci in the Australian Blacklip abalone (*Haliotis rubra*, Leach). *Molecular Ecology* **9**, 1183 – 1184.

Evans, B., Bartlett, J., Sweijd, N., Cook, P., Elliott, N.G., 2004. Loss of genetic variation at microsatellite loci in hatchery produced abalone in Australia (*Haliotis rubra*) and South Africa (*Haliotis midae*). *Aquaculture* **233**, 109 – 127.

Fernando, P., Evans, B.J., Morales, J.C., Melnick, D.J., 2001. Electrophoresis artefacts – a previously unrecognised cause of error in microsatellite analysis. *Molecular Ecology* **1**, 325 – 328.

Fishtech, Inc. Facts about abalone [Online]. Available: http://www.fishtech.com/facts.html. [2004, August]

Frankham, R., 1995. Inbreeding and extinction: a threshold effect. *Conservation Biology* **9**, 792 – 799.

Frankham, R., Loebel, D.A., 1992. Modeling problems in conservation genetics using captive *Drosophila* populations. *Zoo Biology* **11**, 333 – 342.

Gafney, P.M., Rubin, V.P., Hedgecock, D., Powers, D.A., Morris, G., Hereford, L., 1996. Genetic effects of artificial propagation: signals from wild and hatchery populations of red abalone in California. *Aquaculture* **143**, 257 – 266.

Gagneux, P., Boesch, C., Woodruff, D.S., 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* **6**, 861 – 868.

Stellenbosch University http://scholar.sun.ac.za

Goodman, S.J., 1997. RstCalc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* **6**, 881 – 885.

Goossens, B., Waits, L.P., Taberlet, P., 1998. Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology* **7**, 1237 – 1241.

Gordon, H.R., Cook,P., 2001. World abalone supply, markets and pricing: historical, current and future. *Journal of Shellfish Research* **20**, 567 – 570.

Government of Western Australia, Department of Fisheries. The fishery [Online]. Available: http://www.fish.wa.gov.au/comm/broc/abalone. [2004, August]

Hancock, J.M., 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In Goldstein, D.B., Schlötterer, C. *Microsatellites, Evolution and Applications*. Oxford: Oxford University Press, 1 – 9.

Hansen, M.M., Mensberg, K.L.D., Berg, S., 1999. Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus* sp.) populations in Denmark, inferred from mitochondrial and microsatellite markers. *Molecular Ecology* **8**, 239 – 252.

Hardy, E. (a). *Haliotis midae* [Online]. Available: http://www.gastropod.com/4/Shell_964.html. [2004, August]

Hardy, E. (b). Hardy's internet guide to marine gastropods 13.03 [Online]. Available: http://www.gastropods.com. [2004, August]

Hedgecock, D., Sly, F., 1990. Genetic drift and effective population sizes of hatchery-propagated stocks of the Pacific oyster, *Crassostrea gigas*. *Aquaculture* **88**, 21 – 38.

Stellenbosch University http://scholar.sun.ac.za

Hill, W.G., 2000. Maintenance of quantitative genetic variation in animal breeding programmes. *Livestock Production Science* **63**, 99 – 109.

Huang, X., Peakall, R., Hanna, P.J., 2000. Analysis of genetic structure of blacklip abalone (*Haliotis rubra*) populations using RAPD, minisatellite and microsatellite markers. *Marine Biology* **136**, 207 – 216.

Iguchi, K., Watanabe, K., Nishida, M., 1999. Reduced mitochondrial DNA variation in hatchery populations of Ayu (*Plecoglossus altivelis*) cultured for multiple generations. *Aquaculture* **178**, 235 – 243.

Irish Sea Fisheries Board. New species – Novel shellfish [Online]. Available: http://www.bim.ie/templates/text_content.asp?node_id=256. [2004, August]

Kijima, A., Ikeda, M., Fujio, Y., 1992. Genetic characteristics of the artificial seed populations of abalone. *Fisheries, Genetic and Breeding Sciences* **18**, 53 – 63.

Klinbunga, S., Pripue, P., Khamnamtong, N., Puanglarp, N., Tassanakajon, A., Jarayabhand, P., Hirono, I., Aoki, T., Menasveta, P., 2003. Genetic diversity and molecular markers of the Tropical abalone (*Haliotis asinina*) in Thailand. *Marine Biotechnology* **5**, 505 – 517.

Koehn, R.K., Diehl, W.J., Scott, T.M., 1988. The differential contribution by individual enzymes of glycolysis and protein catabolism to the relationship between heterozygosity and growth rate in the Coot clam, *Mulinia lateralis*. *Genetics* **118**, 121 – 130.

Kondo, Y., Sato, K., Kitada, K., Serikawa, T., 1997. Application of mouse microsatellite markers to rat genome mapping. *Transplantation Proceedings* **29**, 1766 – 1767.

Stellenbosch University http://scholar.sun.ac.za

Lachance, S., Magnan, P., 1990. Performance of domestic, hybrid and wild strains of brook trout, *Salvelinus fontinalis*, after stocking: the impact of intra- and interspecific competition. *Canadian Journal of Fisheries and Aquatic Sciences* **47**, 2278 – 2284.

Lacy, R.C., Lindenmayer, D.B., 1995. A simulation of the impacts of population subdivision on the mountain brushtail possum *Trichosurus caninus* ogilby (Phalangeridae: Marsupialia), in South-Eastern Australia. II. Loss of genetic variation within and between subpopulations. *Biological Conservation* **73**, 131 – 142.

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. *Genetical Research of Cambridge* **78**, 259 – 270.

Li, Q., Park, C., Endo, T., Kijima, A., 2002. Isolation and characterisation of microsatellite loci in the Pacific abalone, *Haliotis discus hannai*. *Journal of Shellfish Research* **21**, 811 – 815.

Li, Q., Park, C., Endo, T., Kijima, A., 2004. Loss of genetic variation at microsatellite loci in hatchery strains of the Pacific abalone (*Haliotis discus hannai*). *Aquaculture* **237**, 207 – 222.

Litt, M., Luty, J.A., 1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics* **44**, 397 – 401.

Liu, B.-H., 1997. Chapter 12: QTL mapping: Introduction. In Liu, B.-H. *Statistical Genomics: Linkage, Mapping, and QTL Analysis*. Boca Raton, London, New York, Washington, D.C.: CRC Press, 175 – 177.

Liu, Z.J., Cordes, J.F., 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* **238**, 1 – 37.

Stellenbosch University http://scholar.sun.ac.za

Lougheed, S.C., Gibbs, H.L., Prior, K.A., Weatherhead, P.J., 2000. A comparison of RAPD versus microsatellite DNA markers in population studies of the Massasauga rattlesnake. *The Journal of Heredity* **91**, 458 – 463.

May, B., Stoneking, M., Wright, J.E., 1980. Joint segregation of biochemical loci in Salmonidae: II Linkage associations from a hybridized *Salvelinus* genome (*S. namaycush* x *S. fontinalis*). *Genetics* **95**, 707 – 726.

Mgaya, Y.D., Gosling, E.M., Mercer, J.P., Donlon, J., 1995. Genetic variation at three polymorphic loci in wild and hatchery stocks of the abalone, *Haliotis tuberculata* Linnaeus. *Aquaculture* **136**, 71 – 80.

Morán, P., Pendás, A.M., García-Vásquez, E., 1996. Mitochondrial DNA variation in wild and hatchery broun trout (*Salmo trutta* L.) populations from Spain. *Aquaculture* **141**, 59 – 65.

Norris, A.T., Bradley, D.G., Cunningham, E.P., 2000. Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* **182**, 73 – 83.

OceanLink. The amazing abalone [Online]. Available: http://oceanlink.island.net/oinfo/Abalone/abalone.html. [2004, August]

O'Reilly, P.T., Canino, M.F., Bailey, M., Bentzen, P., 2004. Inverse relationship between *F*st and microsatellite polymorphism in the marine fish, walleye Pollock (*Theragra chalcogramma*): implications for resolving weak population structure. *Molecular Ecology* **13**, 1799 – 1814.

Oura, C.A.L., Odongo, D.O., Lubega, G.W., Spooner, P.R., Tait, A., Bishop, R.P., 2003. A panel of microsatellite and minisatellite for the characterisation of field isolates of *Theileria parva*. *International Journal for Parasitology* **33**, 1641 – 1653.

Stellenbosch University http://scholar.sun.ac.za

Pemberton, J.M., Slate, J., Bancroft, D.R., Barrett, J.A., 1995. Nonamplification alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**, 249 – 252.

Rahman, M.H., Jaquish, B., Khasa, P.D., 2000. Optimization of PCR protocol in microsatellite analysis with silver and SYBR stains. *Plant Molecular Biology Reporter* **18**, 339 – 348.

Redmayne, P., 2002. Abalone: A devoted following keeps prices high for this marvelous mollusk [Online]. Available: http://www.seafoodbusiness.com/archives/searchframe.asp. [2004, August]

Robinson, N., Baranski, M., Rourke, M., Hayes, B., 2003. Application of advanced genetic techniques to abalone breeding.10th Abalone Aquaculture Workshop.

Rubinsztein, D.C., Amos, W., Leggo, J., Goodburn, S., Jain, S. Li, S.H., Margolis, R.L., Ross, C.A., Ferguson-Smith, M.A., 1995. Microsatellite evolution—evidence for directionality and variation in rate between species. *Nature Genetics* **10**, 337 – 343.

Sakamoto, T., Danzmann, R.G., Okamoto, N., Ferguson, M.M., Ihssen, P.E., 1999. Linkage analysis of quantitative trait loci associated with spawning time in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* **173**, 33 – 43.

Schlotterer, C., Tautz, D., 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* **20**, 211 – 215.

Science in Africa, 2004. A gut feeling about abalone [Online]. *Science in Africa* **33**, http://www.scienceinafrica.co.za. [2004, March]

Stellenbosch University http://scholar.sun.ac.za

Seafood Watch., 2003. Seafood Report, Cultured abalone (*Haliotis* spp): Final Report 26 March 2003 [Online]. Available: www.mbayaq.org/cr/cr_seafoodwatch/content/media/Abalone_Farmed_SFW.pdf. [2004, August]

Sekino, M., Hara, M., Taniguchi, N., 2002. Loss of microsatellite and mitochondrial DNA variation in hatchery strains of Japanese flounder *Paralichthys olivaceus*. *Aquaculture* **213**, 101 – 122.

Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457 – 462.

Smith, P.J., Conroy, A.M., 1992. Loss of genetic variation in hatchery produced abalone, *Haliotis iris*. *New Zealand Journal of Marine and Freshwater Research* **26**, 81 – 85.

South African Abalone [Online]. Available: http://web.uct.ac.za/depts/zoology/abnet/safrica.html. [2004, August]

Su, H., Qu, L-J., He, K., Zhang, Z., Wang, J., Chen, Z., Gu, H., 2003. The Great Wall of China: a physical barrier to gene flow? *Heredity* **90**, 212 – 219.

Sweijd, N., Snethlage, Q., Harvey, D., Cook, P., 1998. Experimental abalone (*Haliotis midae*) seeding in South Africa. *Journal of Shellfish Research* **17**, 897 – 904.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**, 6463 – 6471.

Triantafyllidis, A., Abatzopoulos, T.J., Leonardos, R.G., Guyomard, R., 2002. Microsatellite analysis of the genetic population structure of native and translocated Aristotle's catfish (*Silurus aristotelis*). *Aquatic Living Resources* **15**, 351 – 359.

Stellenbosch University http://scholar.sun.ac.za

Virgillo, M., Abbiati, M. Habitat discontinuity and genetic structure in populations of the estuarine species *Hediste diversicolor* (Polychaeta: Nereididae). In press. *Estuarine, Coastal and Shelf Science*.

Vos , P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., Zabeay, M., 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407 – 4414.

Was, A., Wenne, R., 2002. Genetic differentiation in hatchery and wild sea trout (*Salmo trutta*) in the Southern Baltic at microsatellite loci. *Aquaculture* **204**, 493 – 506.

Weber, J.L., May, P.E., 1989. Abundant Class of Human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics* **44**, 388 – 396.

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

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

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.

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1 – 16.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 2 – MATERIALS AND METHODS

2.1) SAMPLE COLLECTION

Samples were taken from captivity and from natural populations. The captive animals were sampled from the Sea Plant Products abalone farm near Hermanus. The choice of which natural populations to sample was based on the composition of the brood stock of the farm. The brood stock contained individuals of both the East (Port Elizabeth) and the West (Saldanha Bay) coast. Therefore one East coast and one West coast population were sampled to investigate the compatibility of the farm population with the natural populations they originated from. The geographical locations of the farm and natural populations are shown in Figure 2.1.

The geographical sites from where the samples were taken, along with the number of individuals per sample are shown in Table 2.1.

Samples were taken in a destructive way by killing the animal to sample either muscle or gill tissue (Figure 2.2) and in a non-destructive way by cutting off 1 or 2 tentacles (Figure 2.2), leaving the animal alive. Samples from the brood animals (Tables 2.2 to 2.4; SPP) were taken in the non-destructive way, while samples from the commercial (Tables 2.5 to 2.7; SPP) and natural populations (Table 2.8; Saldanha Bay), were taken in a destructive way. Animals were measured at the longest and widest point of the shell. All samples were preserved in 99% Ethanol.

Stellenbosch University http://scholar.sun.ac.za

Figure 2.1: Map of South Africa showing the locations where the two natural populations were sampled and the location of Sea Plant Products.
(www.adventureaddicts.co.za/…/maps/map_sa.jpg)

Stellenbosch University http://scholar.sun.ac.za

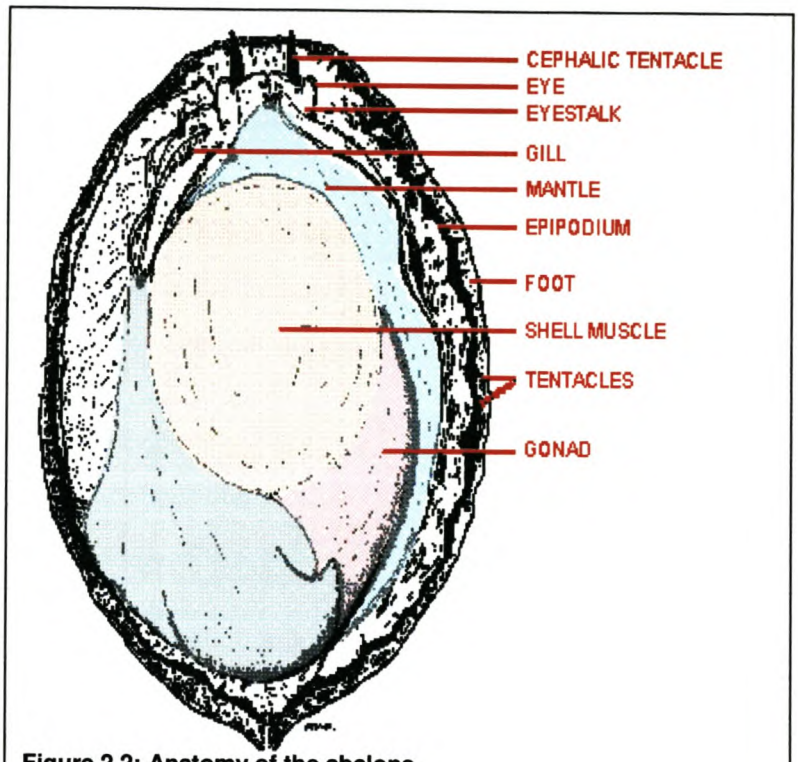


Figure 2.2: Anatomy of the abalone.
(Fishtech, Inc, online)

Table 2.1: The number of samples per population.

POPULATION	NUMBER OF INDIVIDUALS	ORIGIN	TABLE
Brood stock: Row 2	37	SPP, Hermanus	2.2
Commercial stock: Row2	31	SPP, Hermanus	2.5
Brood stock: Row 3	24	SPP, Hermanus	2.3
Commercial stock: Row 3	33	SPP, Hermanus	2.6
Brood stock: Row 4	27	SPP, Hermanus	2.4
Commercial stock: Row 4	34	SPP, Hermanus	2.7
Saldanha Bay	32	Saldanha Bay	2.8
Black Rock	24	Black Rock	-

SPP = Sea Plant Products

Stellenbosch University http://scholar.sun.ac.za

Table 2.2: Brood Stock: Row 2.

Animal ID	Gender	Sample
44	Female	Tentacle
146	Female	Tentacle
300	Male	Tentacle
301	Male	Tentacle
302	Male	Tentacle
303	Male	Tentacle
304	Male	Tentacle
305	Male	Tentacle
306	Male	Tentacle
307	Male	Tentacle
308	Male	Tentacle
309	Male	Tentacle
310	Male	Tentacle
311	Female	Tentacle
312	Female	Tentacle
313	Female	Tentacle
314	Female	Tentacle
315	Female	Tentacle
316	Female	Tentacle
317	Female	Tentacle
318	Female	Tentacle
319	Female	Tentacle
320	Female	Tentacle
321	Female	Tentacle
322	Female	Tentacle
323	Female	Tentacle
324	Male	Tentacle
353	Male	Tentacle
354	Male	Tentacle
423	Female	Tentacle
424	Female	Tentacle
425	Female	Tentacle
426	Female	Tentacle
427	Female	Tentacle
428	Female	Tentacle
429	Female	Tentacle
430	Female	Tentacle

Table 2.3: Brood stock: Row 3.

Animal ID	Gender	Sample
1	Female	Tentacle
5	Male	Tentacle
24	Male	Tentacle
29	Female	Tentacle
40	Male	Tentacle
53	Female	Tentacle
55	Female	Tentacle
71	Female	Tentacle
113	Female	Tentacle
130	Male	Tentacle
164	Male	Tentacle
170	Male	Tentacle
197	Male	Tentacle
355	Male	Tentacle
356	Male	Tentacle
357	Male	Tentacle
358	Male	Tentacle
359	Male	Tentacle
431	Female	Tentacle
432	Female	Tentacle
433	Female	Tentacle
434	Female	Tentacle
435	Female	Tentacle
436	Female	Tentacle

Table 2.4: Brood stock: Row 4.

Animal ID	Gender	Sample
3	Male	Tentacle
13	Female	Tentacle
35	Male	Tentacle
46	Male	Tentacle
57	Male	Tentacle
70	Male	Tentacle
78	Female	Tentacle
100	Female	Tentacle
105	Female	Tentacle
118	Female	Tentacle
131	Female	Tentacle
140	Female	Tentacle
142	Male	Tentacle
148	Male	Tentacle
152	Male	Tentacle
179	Female	Tentacle
182	Male	Tentacle
190	Male	Tentacle
438	Female	Tentacle
439	Female	Tentacle
440	Female	Tentacle
441	Female	Tentacle
442	Female	Tentacle
443	Female	Tentacle
444	Female	Tentacle
445	Female	Tentacle
446	Female	Tentacle

Stellenbosch University http://scholar.sun.ac.za

Table 2.5: Commercial stock: Row 2 (365/02* - 150/03#).

Animal ID	Row	Length (mm)	Width (mm)	Alcohol Sample	Extraction Sample
2	2	11	7	Whole	Muscle
4	2	14	8	Whole	Muscle
6	2	12	7	Whole	Muscle
8	2	10	6	Whole	Muscle
9	2	11	7	Whole	Muscle
10	2	11	7	Whole	Muscle
11	2	6	4	Whole	Muscle
12	2	15	10	Whole	Muscle
13	2	14	9	Whole	Muscle
15	2	13	8	Whole	Muscle
16	2	12	8	Whole	Muscle
17	2	10	6	Whole	Muscle
18	2	10	6	Whole	Muscle
19	2	10	6	Whole	Muscle
20	2	6	5	Whole	Muscle
21	2	9	5	Whole	Muscle
22	2	9	5	Whole	Muscle
23	2	9	5	Whole	Muscle
24	2	9	5	Whole	Muscle
25	2	7	5	Whole	Muscle
26	2	7	4	Whole	Muscle
27	2	7	4	Whole	Muscle
28	2	7	4	Whole	Muscle
29	2	8	5	Whole	Muscle
30	2	7	4	Whole	Muscle
31	2	7	5	Whole	Muscle
32	2	7	4	Whole	Muscle
33	2	6	4	Whole	Muscle
34	2	6	4	Whole	Muscle
35	2	6	4	Whole	Muscle
36	2	5	3	Whole	Muscle

* Date of Spawning (31 December 2002)
# Date of Sample Collection (30 May 2003)

Stellenbosch University http://scholar.sun.ac.za

Table 2.6: Commercial stock: Row 3 (262/02* - 150/03#).

Animal ID	Row	Length (mm)	Width (mm)	Alcohol Sample	Extraction Sample
3	3	18	11	Whole	Muscle
4	3	20	12	Whole	Muscle
5	3	15	9	Whole	Muscle
6	3	18	10	Whole	Muscle
7	3	15	10	Whole	Muscle
8	3	14	9	Whole	Muscle
9	3	14	9	Whole	Muscle
10	3	15	9	Whole	Muscle
11	3	15	8	Whole	Muscle
12	3	19	11	Whole	Muscle
13	3	15	9	Whole	Muscle
14	3	15	10	Whole	Muscle
15	3	25	15	Whole	Muscle
16	3	14	9	Whole	Muscle
17	3	15	9	Whole	Muscle
18	3	15	9	Whole	Muscle
19	3	13	9	Whole	Muscle
20	3	20	12	Whole	Muscle
21	3	18	10	Whole	Muscle
22	3	19	13	Whole	Muscle
23	3	18	11	Whole	Muscle
24	3	15	9	Whole	Muscle
25	3	14	9	Whole	Muscle
26	3	15	9	Whole	Muscle
27	3	15	9	Whole	Muscle
28	3	18	11	Whole	Muscle
29	3	20	13	Whole	Muscle
30	3	22	13	Whole	Muscle
31	3	15	10	Whole	Muscle
32	3	14	9	Whole	Muscle
33	3	15	9	Whole	Muscle
34	3	14	8	Whole	Muscle
35	3	17	12	Whole	Muscle

* Date of Spawning (19 September 2002)
# Date of Sample Collection (30 May 2003)

Stellenbosch University http://scholar.sun.ac.za

Table 2.7: Commercial stock: Row 4 (268/02* - 150/03#).

Animal ID	Row	Length (mm)	Width (mm)	Alcohol Sample	Extraction Sample
1	4	13	8	Whole	Muscle
2	4	14	9	Whole	Muscle
3	4	14	8	Whole	Muscle
4	4	13	7	Whole	Muscle
5	4	14	8	Whole	Muscle
6	4	16	11	Whole	Muscle
7	4	14	9	Whole	Muscle
8	4	13	8	Whole	Muscle
9	4	12	7	Whole	Muscle
10	4	13	8	Whole	Muscle
11	4	15	9	Whole	Muscle
12	4	22	13	Whole	Muscle
13	4	15	9	Whole	Muscle
14	4	14	7	Whole	Muscle
15	4	14	7	Whole	Muscle
16	4	15	8	Whole	Muscle
17	4	14	7	Whole	Muscle
18	4	17	10	Whole	Muscle
19	4	15	9	Whole	Muscle
20	4	16	11	Whole	Muscle
21	4	15	10	Whole	Muscle
22	4	14	8	Whole	Muscle
24	4	30	20	Gill + Muscle	Muscle
25	4	18	11	Whole	Muscle
26	4	18	11	Whole	Muscle
27	4	18	11	Whole	Muscle
28	4	17	10	Whole	Muscle
29	4	29	17	Gill + Muscle	Muscle
30	4	29	19	Gill + Muscle	Muscle
31	4	17	10	Whole	Muscle
32	4	16	11	Whole	Muscle
33	4	12	8	Whole	Muscle
34	4	11	7	Whole	Muscle
35	4	12	8	Whole	Muscle

* Date of Spawning (25 September 2002)
\# Date of Sample Collection (30 May 2003)

Stellenbosch University http://scholar.sun.ac.za

Table 2.8: Natural population: Saldanha Bay.

Animal ID	Length (mm)	Width (mm)	Alcohol Sample	Extraction Sample
1	140	115	Gill	Gonad
2	115	85	Gill	Gill
3	120	90	Gill	Gill
4	120	90	Gill	Gill
5	120	90	Gill	Gill
6	105	80	Gill	Gill
7	125	100	Gill	Gill
9	105	75	Gill	Gill
10	100	70	Gill	Gill
11	95	70	Gill	Gill
12	120	95	Gill	Gill
13	105	80	Gill	Gill
14	120	90	Gill	Gill
15	100	70	Gill	Gill
16	110	85	Gill	Gill
17	100	70	Gill	Gill
18	110	80	Gill	Gill
19	105	80	Gill	Gill
20	90	70	Gill	Gill
21	125	95	Gill	Gill
22	110	90	Gill	Gill
23	120	90	Gill	Gill
24	110	85	Gill	Gill
25	120	100	Gill	Gill
26	130	110	Gill	Gill
27	130	100	Gill+Muscle	Gill
28	140	115	Gill	Gill
29	125	100	Gill+Muscle	Gill
30	135	100	Gill+Muscle	Gill
31	135	110	Gill+Muscle	Gill
32	145	115	Gill+Muscle	Gill
33	135	105	Gill+Muscle	Gill

2.1.1) BLACK ROCK SAMPLES

Twenty-four destructive samples were collected by Brian Godfrey. Gill tissue was used for the extractions.

Stellenbosch University http://scholar.sun.ac.za

2.2) ISOLATION OF GENOMIC DNA

A total of 32 individuals from Saldanha Bay, 24 from Black Rock and 186 individuals from the farm were extracted. Details are shown in Tables 2.1 – 2.7 and section 2.1.1; p42. Aletta Bester (Department of Genetics, University of Stellenbosch, Aquaculture Division) performed the DNA extractions for the Black Rock population.

DNA was extracted from either gill or muscle tissue using a CTAB DNA extraction method (Saghai-Maroof *et al*, 1984). A small piece of tissue was placed into an eppendorf tube containing 500µl of digestion buffer [2% CTAB, 1.4M NaCl, 0.2% β-mercapto-ethanol, 20mM Ethylene Diamine Tetra-Acetate (EDTA, pH 8), 100mM Tris-HCl, pH 8] and 2µl Proteinase K (10mg/ml; *Roche*) was added. The mixture was shaken and incubated overnight at 60˚C in a waterbath.

An equal volume of chloroform:iso-amyl alcohol (24:1) was added to each eppendorf tube and shaken for 5 minutes using a Vortex-Genie 2 (*Scientific Industries*). The aqueous phase and organic phase were separated by centrifugation at high speed (16.1rcf) for 5 minutes using the Eppendorf Centrifuge 5415D. The clear aqueous phase was transferred to a new eppendorf tube and the chloroform:iso-amylalcohol step was repeated. DNA was precipitated overnight at -20°C by adding 2/3 volume of isopropanol.

The tubes were centrifuged at high speed (16.1rcf) for 20 minutes. The DNA pellet was washed with 70% ethanol, left for a few minutes and centrifuged again at high speed (16.1rcf) for 10 minutes. The alcohol was removed and the pellet was dried in an oven for 15 minutes at 55˚C. The pellet was resuspended in 100*u*l of ddH_2O.

DNA was quantified by running the samples on an agarose gel (2%; 1xTBE [1.08% Tris, 0.54% Boric Acid, 0.058% EDTA]; Ethidium Bromide) along with DNA standards (50ng, 100ng, 200ng) of Lambda (λ) to estimate the

Stellenbosch University http://scholar.sun.ac.za

concentration. The gel was visualised by Beta 4.0.2 of Scion Image while being exposed to UV-light. The DNA was then stored at -20°C until further use.

2.3) MICROSATELLITE ISOLATION

The newly isolated microsatellite DNA markers for this study are detailed in Bester *et al*, in press (Appendix A).

The Fast Isolation by AFLP of Sequences Containing Repeats (FIASCO, Zane *et al*, 2002) method was used to isolate microsatellites from *Haliotis midae* genomic DNA. DNA from two individuals from Sea Plant Products and one individual from Saldanha Bay were initially used: Cohort 16 Individual 20 (16-20), Cohort 16 Individual 21 (16-21) and Saldanha Bay Individual 15 (S15).

2.3.1) RESTRICTION AND LIGATION

*Mse*I-restriction enzyme digestion and the ligation of the *Mse*I AFLP-adaptor (5'-TAC TCA GGA CTC AT-3' / 5'-GAC GAT GAG TCC TGA G-3') were done simultaneously in a 25µl reaction [250ng Genomic DNA, 1x One-Phor-All Buffer, 5mM DTT (Di-Thio-Threitol), 50ug/ml BSA (Bovine Serum Albumin), 1µM *Mse* Adaptor, 200µM ATP, 2.5U *Mse*I, 1.0U T4-Ligase]. The reaction was incubated at 37°C in a waterbath for 3 hours. A 1:10 dilution was made of the reaction.

2.3.2) 1ST AFLP AMPLIFICATION

An AFLP-amplification was done using AFLP-adaptor specific primers (5'-GAT GAG TCC TGA GTA A**N**-3' referred to as *Mse*-N). A 20µl reaction volume (1x$MgCl_2$ free buffer, 1.5mM $MgCl_2$, 200uM of each dNTP, 120ng *Mse*I-N, 0.4U Taq DNA polymerase, 5µl ligation reaction) was used to amplify the fragments. The PCR-program was set for 94°C 30s, 53°C 1min, 72°C

Stellenbosch University http://scholar.sun.ac.za

1min, for 17, 20, 23 and 26 cycles. The results were analysed on an agarose gel (1%, 1xTBE, Ethidium Bromide) and visualised by Beta 4.0.2 of Scion Image while being exposed to UV-light.

2.3.3) HYBRIDISATION

The DNA was hybridised with a biotinylated $(AC)_{12}$ probe in a 100µl reaction containing 250 – 500ng of DNA, 50 – 80pmol $(AC)_{12}$ probe, 4.2x SSC (Standard Saline Citrate) and 0.07% SDS (Sodium Dodecyl Sulfate). The DNA was denatured at 95°C, 3min and hybridised at 25°C, 15min. The hybridisation reaction was diluted by adding 300µl TEN_{100} [10mM Tris-HCl (pH 8), 1mM EDTA (pH 8), 100mM NaCl, pH 8].

2.3.4) SELECTIVE CAPTURING OF HYBRIDISED DNA

1mg (100µl) streptavidin coated beads were washed four times with an equal volume of TEN_{100} and resuspended in 40µl of the same buffer. A volume of 10µl of an unrelated PCR-product (plant DNA, supplied by Helena Gardner, Department of Genetics, University of Stellenbosch, Vitis Lab) was added to the beads to minimize non-specific binding of genomic DNA. Forty micro litres of the beads were then added to 400µl of the hybridisation reaction and incubated at room temperature for 30 minutes while it was slowly shaken on a Vortex Genie 2 (*Scientific Industries*).

The beads-probe-DNA complex was separated using a magnetic field and the hybridisation buffer was discarded. Non-specific DNA was removed by three non-stringency washes and three stringency washes. The supernatant of the third wash was saved each time and stored at -20°C. The non-stringency washes were performed by adding 400µl TEN_{1000} {10mM Tris-HCl (pH 8), 1mM EDTA (pH 8), 1M NaCl, pH 8} to the beads, gentle mixing and leaving the mixture at room temperature for 5 minutes after which the beads were magnetically removed and the aqueous phase discarded. The stringency washes were performed by adding 400µl of 0.2x SSC, 0.1% SDS to the

Stellenbosch University http://scholar.sun.ac.za

beads, gentle mixing and leaving the mixture at room temperature for 5 minutes after which the beads were magnetically removed and the aqueous phase discarded.

DNA was separated from the beads-probe complex by two denaturation steps. The first denaturation step was performed by adding 50μl TE {10mM Tris-HCl (pH 8), 1mM EDTA (pH 8), pH 8} to the beads and it was incubated for 5 minutes at 95°C. The supernatant was removed and quickly stored at -20°C. The second denaturation step was performed by adding 15μl 0.15M NaOH to the beads. The supernatant was neutralised by adding 1μl 0.1667M CH_3COOH (acetic acid) and then stored at -20°C.

DNA was precipitated by adding 1x volume isopropanol and 0.15M Sodium Acetate to the eppendorf tubes containing the supernatants of the four different washes. The tubes were left overnight at -20°C. The DNA was pelleted by centrifugation at 16.1rcf for 30 minutes. The pellet was dried and resuspended in 50μl ddH_2O.

2.3.5) 2ND AFLP AMPLIFICATION

Two micro litres of the DNA obtained from the washes and denaturation steps were amplified using the same conditions as used during the 1st AFLP amplification (see section 2.3.2; p44), the only difference being the use of 30 cycles with the 2nd amplification. The products were analysed on an agarose gel (2%, 1xTBE; Ethidium Bromide) and visualised by Beta 4.0.2 of Scion Image while being exposed to UV-light.

2.3.6) CLONING

The pCR®4-TOPO® vector (*Invitrogen™ Life Technologies*) was used to clone the desired PCR-product. Luria Bertani (LB) –medium (1% Bacto-Tryptone, 0.5% Bacto-Yeast, 0.5% NaCl, 1.2% Bacterial Agar, pH 7.5) containing

Stellenbosch University http://scholar.sun.ac.za

100µg/ml (100mg/ml stock solution) Ampicillin (*Roche Applied Science*) was prepared and poured into petri dishes.

A transformation reaction containing 4µl fresh PCR-product, 1µl salt solution and 1µl TOPO® vector was incubated at room temperature for 30 minutes, after which 2µl of the reaction was added to competent cells and mixed gently. The competent cells were incubated on ice for 30 minutes, heat-shocked at 42°C for 30 seconds and immediately put on ice again. A volume of 250µl of the transformed cells was added to SOC-medium (2% Bacto-tryptone, 0.5% Yeast Extract, 0.05% NaCl; g/ml; 10ml 250mM KCl, 18ml 20% sterile Glucose, 5ml 2M sterile $MgCl_2$, pH7), at room temperature and shaken at 175rpm for 1 hour. The SOC-medium containing the transformed cells were evenly divided among 5 LB-medium plates, spread with a glass hockey stick and incubated overnight at 37°C.

2.3.7) SCREENING OF COLONIES

Single colonies were selected from the plates and a colony-PCR was performed. The M13 forward (5'-GGTTTTCCCAGTCACGAC-3') and reverse (5'-GGAAACAGCTATGACCATG-3') vector-specific primers were used to screen the colonies in a 50µl PCR-reaction {1x $MgCl_2$ free buffer, 1.5mM $MgCl_2$, 200µM of each dNTP, 5pmol M13 forward and reverse primer, 0.2 µl Taq polymerase (*Promega*), DNA from picked colony}. The PCR-program was set for 94°C 10min; 25 cycles of 94°C 1min, 55°C 1min, 72°C 1min; 72°C 10min. The PCR-products were analysed on an agarose gel (1%, 1xTBE, Ethidium Bromide) along with a 100bp size marker and visualised by Beta 4.0.2 of Scion Image while being exposed to UV-light.

2.3.8) SEQUENCING OF CLONES

The clones were sequenced in a 10µl sequencing-reaction containing 2µl BigDye v3 terminator (*Applied Biosystems*), 1.6pmol primer and 3.3 – 6.6ng/

Stellenbosch University http://scholar.sun.ac.za

µl DNA. The sequencing reactions were analysed with the ABI PRISM® 3100 DNA automated sequencer.

2.3.9) DESIGNING OF MICROSATELLITE PRIMERS

Sequence data obtained from the ABI PRISM® 3100 DNA automated sequencer was edited using Chromas 1.45 (Conor McCarthy, 1996 – 1998). The edited sequence data was checked against the NCBI database to ensure that the DNA were not contaminated by other organisms. Sequences of different clones were aligned against each other using BioEdit version 5.0.9 in order to detect any possible duplication of flanking regions. The microsatellite primers were then designed and analysed by using Oligo 4.1. Analysis of the primers included checking for dimer formation and annealing temperature. All primers were designed to have an annealing temperature of approximately 60°C. Primers were designed to be some distance away from the beginning and end of the repeat units (Koorey *et al*, 1993).

2.4) OPTIMIZATION OF PCR WITH UNLABELLED MICROSATELLITE PRIMERS

Microsatellite primers were dissolved in SABAX water to make a stock solution of 100µM and stored at -20°C.

Different annealing temperatures, different $MgCl_2$ concentrations and different primer concentrations were tested (Rahman *et al*, 2000) to optimise the primers.

A touch-down PCR strategy (Rahman *et al*, 2000) was used during optimisation. A 10µl PCR-reaction (1x $MgCl_2$ free buffer, 2mM $MgCl_2$, 50µM of each dNTP, 0.2U Taq polymerase, 2pmol/µl primer and 20ng DNA) was performed. The PCR-program was set for an initial denaturation step of 94°C 5min followed by 2 cycles of 94°C 30s, 65°C 30s, 72°C 30s. Hereafter the annealing temperature was lowered by 1°C in each two consecutive cycles,

Stellenbosch University http://scholar.sun.ac.za

until an annealing temperature of 55°C was reached and maintained for 30 cycles of 94°C 1min, 55°C 1min, 72°C 1min, with a final elongation step at 72°C 7min. The only parameter that was changed during optimisation was the primer concentration.

2.5) CHOICE OF PRIMERS TO BE LABELLED

The level of polymorphism of the primers that successfully amplified, were tested using poly-acrylamide gel electrophoresis. A 12% acrylamide:bisacrylamide {37.5:1 (30%)} Mighty Small (*Hoefer*) poly-acrylamide (PAA) gel was used (Table 2.9).

Table 2.9: Poly-acrylamide gel mix.

	STOCK	[FINAL]
AA	**37.5:1 (30%)**	**12% (v/v)**
TBE	10x	1x
APS	1mg/ml	0.8% (v/v)
Temed		0.16% (v/v)

2.5.1) PCR-REACTIONS

Test PCR-reactions were performed on 8 different individuals (Cohort 16 Individual 7; Brood Stock Individuals 1, 24, 140; Saldanha Bay Individuals 7, 11, 20, 23) using the touch-down protocol (TD-PCR). The success of the reactions were checked with a 100bp size marker, by agarose gel electrophoresis (2%, 1xTBE, Ethidium Bromide) and visualised by Beta 4.0.2 of Scion Image while being exposed to UV-light. The reactions were then loaded along with a 100bp size marker on the PAA-gels. The PAA-gels were visualised by silver staining (see section 2.5.2; p50).

Stellenbosch University http://scholar.sun.ac.za

2.5.2) SILVER STAINING

The gel was removed from the electrophoresis plates, placed in the fixing solution (10% Ethanol, 0.5% Acetic Acid) and shaken on The Belly Dancer (*Stovall*) for 10 minutes. The fixing solution was removed and the gel was rinsed twice for 1 minute each with ddH_2O. The staining solution {0.1% $Ag(NO_3)_2$} was added and shaken for 10 minutes. The staining solution was removed and the gel quickly rinsed with ddH_2O again. The developing solution {1.5% NaOH, 0.15% Formaldehyde (added fresh)} was added and the gel was shaken until all the bands could be seen, usually between 10 and 20 minutes. The developing solution was removed; the gel allowed to dry and then sealed in order to preserve it.

2.6) LABELLING OF PRIMERS

Primers were labelled with the fluorescent labels VIC, NED, FAM and PET (*Whitehead Scientific*). The sequences of the labelled primers were submitted to GenBank (Bester *et al*, *in press*).

2.7) OPTIMISATION OF PCR WITH LABELLED PRIMERS

Labelled primers were tested using the same PCR-conditions as for the unlabelled primers. Different PCR-programs were used for further optimisation of the primers. Ten microlitre reactions were used for the amplification of products and the only change from the unlabelled primer PCR-conditions was the use of 0.1U Taq polymerase instead of 0.2U. Another PCR-program was also used for two of the primers to increase amplification and resolution during the analysis of the labelled primers (see section 2.8; p51). The PCR-machine was set for 94°C 5min; 5 cycles of 94°C 30s, 60°C 40s; 30 cycles of 94°C 30s, 55°C 40s and 72°C 10min.

Stellenbosch University http://scholar.sun.ac.za

2.8) GENOTYPING USING THE ABI PRISM® 3100 DNA AUTO SEQUENCER

PCR-products were first quantified by agarose gel electrophoresis and diluted accordingly (Fernando *et al*, 2001); dilutions were optimised by trial and error (approximately 5 – 10ng of DNA). Samples were run with the Genescan 500LIZ (*Applied Biosystems*) size standard on an ABI-PRISM® 3100 DNA Auto Sequencer and analysed and scored with the Genotyper® software (*Applied Biosystems*). The Black Rock samples were genotyped by Aletta Bester (Department of Genetics, University of Stellenbosch, Aquaculture Division).

2.9) STATISTICAL ANALYSIS

2.9.1) SAMPLES ANALYSED

Rows 2 to 4 were analysed as individual populations. The populations being studied are shown in Table 2.10. The brood stock of Sea Plant Products consists of an admixture of animals from both the West (Saldanha Bay) and East (Port Elizabeth) coast. This composite brood stock was sub-divided into different subgroups. The subgroups were called Rows.

Table 2.10: The individual populations being studied.

Population name	Population composition	Number of Individuals
Brood2	Row 2 brood stock	37
Brood3	Row 3 brood stock	24
Brood4	Row 4 brood stock	27
Prog2	Row 2 commercial stock	31
Prog3	Row 3 commercial stock	33
Prog4	Row 4 commercial stock	34
Saldanha Bay	Saldanha Bay natural population	32
Black Rock	Black Rock natural population	24

The data of the rows and natural populations was pooled together to construct three new populations, which were used to investigate the effect of population

Stellenbosch University http://scholar.sun.ac.za

subdivisions. The pooled populations are shown in Table 2.11. The samples were also treated separately.

Table 2.11: The pooled populations being studied.

Population name	Population composition
AllBrood	Brood2 + Brood3 + Brood4
AllProg	Prog2 + Prog3 + Prog4
AllWild	Saldanha Bay + Black Rock

2.9.2) GENETIC VARIATION

2.9.2.1) Number of alleles

The number of alleles per locus and per population was calculated using GenePop 3.4 (Raymond & Rousset, 1995) and Genetix 4.03 (Belkhir *et al*, 2000).

2.9.2.2) Allele frequency

The frequency of each allele for each locus was calculated using Genetix 4.03 (Belkhir *et al*, 2000).

Estimate of allelic frequency in codominant multiple-allele system:

$$p = \frac{N_{11} + \frac{1}{2}\sum N_{12}}{N} \qquad \text{(Eq 1)}$$

where

N_{11} = number of homozygotes

N_{12} = number of heterozygotes

N = number of individuals in population

2.9.2.3) Effective number of alleles

The effective number of alleles (a_e; Kimura & Crow, 1964) were calculated using Microsoft® Excel 2000:

Stellenbosch University http://scholar.sun.ac.za

$$a_e = \frac{1}{\sum x_i^2} \qquad \text{(Eq 2)}$$

where

x_i is the frequency of the *i*th allele.

2.9.2.4) Hardy-Weinberg equilibrium

Deviations from Hardy-Weinberg equilibrium were calculated by the exact Hardy-Weinberg test (Haldane, 1954; Guo & Thompson, 1992) for each locus within each population using GenePop 3.4 (Raymond & Rousset, 1995). The Markov Chain method (10 000 dememorization, 100 batches and 5000 iterations; alleles per locus > 5) was used to calculate the *P*-values (H_o: random union of gametes).

2.9.2.5) Heterozygosity

Expected heterozygosities for each locus for the pooled as well as the individual populations were calculated using Genetix 4.03.

$$H_e = 1 - \sum p^2 \qquad \text{(Eq 3)}$$

where

p = allele frequency

H_e = expected heterozygosity

2.9.2.6) Heterozygote deficiency index

The heterozygote deficiency or excess index (*D*) was calculated using the heterozygosity values previously calculated by Genetix 4.0.3.

$$D = \frac{H_o - H_e}{H_e} \qquad \text{(Eq 4)}$$

(Selander, 1970)

Stellenbosch University http://scholar.sun.ac.za

where

H_o = observed heterozygosity

H_e = expected heterozygosity

2.9.2.7) Null allele frequency

The frequency of null alleles per locus for each population and the average frequency of null alleles per population for each locus were calculated using Microsoft® Excel 2000.

$$r = \frac{H_e - H_o}{1 + H_e} \qquad \text{(Eq 5)}$$

(Brookfield, 1996)

where

H_o = observed heterozygosity

H_e = expected heterozygosity

2.9.2.8) Mann-Whitney test (*U*) for distribution calculations for average number of alleles and average expected heterozygosity

The non-parametric, two sample Mann-Whitney (*U*) test (Eason *et al*, 1989) was used to calculate the significance of observations (also used by Li *et al*, 2004). The *U*-value was calculated using Microsoft® Excel 2000 while the *P*-value (H_0: *X* and *Y* have the same distribution) was calculated by SISA (Simple Interactive Statistical Analysis; Uitenbroek, 1997).

$$U = S_x - \frac{n_x(n_x + 1)}{2} \qquad \text{(Eq 6)}$$

where

S_x = sum of *X* ranks

n_x = number of observations for *X*

Stellenbosch University http://scholar.sun.ac.za

2.9.3) GENETIC DIFFERENTIATION

2.9.3.1) F_{st}

Population differentiation was calculated with the use of the infinite allele model using the *F*-statistics estimator of Weir & Cockerham (1984). The software program Genetix 4.03 (Belkhir *et al*, 2000) was used to calculate the Fst between all the populations. The significance levels were calculated using 1000 permutations.

2.9.3.2) F_{is}

F_{is} calculates the correlation of loci within individuals within populations by taking the probability of identity in state of pairs of loci into account.
F_{is} estimates were calculated for each locus per population and over all loci using Genetix 4.03. The calculations followed Weir & Cockerham (1984). The significance levels were calculated using 1000 permutations

2.10) PARENTAGE ASSIGNMENTS

2.10.1) SAMPLES USED FOR ANALYSIS

The animals from Row 2, Row 3 and Row 4 were used for the parentage analysis {Tables 2.1 – 2.6 (individuals used in genetic variation and differentiation analysis) and 2.11 – 2.14 (individuals used in QTL-analysis)}. A total of 88 potential parents and 215 offspring were genotyped. The total of the parents and F1 of each row are shown in Table 2.11. No prior information of a known parent for the offspring was available, due to the mass-spawning strategy employed by the farm. No samples were received from the farm for some parents. This was discovered (too late) when records of the brood stock layout of the farm was received. The missing parental samples are shown in Table 2.13.

Stellenbosch University http://scholar.sun.ac.za

Table 2.12: The number of individuals genotyped.

POPULATION	NUMBER OF INDIVIDUALS	TABLE
Parents: Row 2	37	2.2
F1: Row2	31	2.5
Parents: Row 3	24	2.3
F1: Row 3	33	2.6
Parents: Row 4	27	2.4
F1: Row 4	151	2.7, 2.12 – 2.15

Table 2.13: Parents not sampled.

PARENTS NOT SAMPLED					
ROW 2		ROW 3		ROW 4	
Animal ID	Gender	Animal ID	Gender	Animal ID	Gender
347	Male	48	Female	116	Male
348	Male	437	Female	360	Male
349	Male			447	Female
351	Male				
352	Male				
405	Female				
406	Female				
448	Female				

2.10.2) CHOICE OF LOCI FOR GENOTYPING

Loci for parentage analysis were chosen using the following criteria:

1) The absence of null alleles (Pemberton *et al*, 1995).
   - Brookfield's null allele frequencies (*r*; Brookfield, 1996) were calculated per locus. All loci with $r > 0.10$ were discarded for the analysis.
2) The avoidance of size homoplasy (Estoup *et al*, 1995).
   - Loci with composite repeats were rejected, because the origin of the alleles would be uncertain.
3) Allele frequency distribution, number of alleles and heterozygosity.
   - Loci with skewed allele frequency distributions, few number of alleles and low heterozygosity levels could score too many homozygous individuals within parents or offspring individuals. Homozygous loci could result in ambiguous parentage assignment.

Stellenbosch University http://scholar.sun.ac.za

4) Exclusion probabilities of the loci.
   - Example: If the loci met the above-mentioned criteria, but a researcher wished to use a minimum number of loci needed for parentage assignments, the individual exclusion probabilities of a locus could be used to make the choice. If a researcher wished to use only 2 loci for the assignments, the 2 loci with the highest individual exclusion probabilities should be chosen.

All calculations were done using the pooled data of all the parents and all the F1 of all rows.

Allele frequencies were calculated as in section 2.9.2.2; p52.
The number of observed alleles was calculated as in section 2.9.2.1; p51.
Expected heterozygosity was calculated as in section 2.9.2.5; p53.
The null allele frequencies were calculated as in section 2.9.2.7; p54.
Exclusion probabilities were calculated using EPC v.1.0, Slabbert 2004 (Appendix B) by the following two formulae:

Exclusion of first parent:

$$P = 1 - 4\sum_{i=1}^{n} p_i^2 + 2\left(\sum_{i=1}^{n} p_i^2\right)^2 + 4\sum_{i=1}^{n} p_i^3 - 3\sum_{i=1}^{n} p_i^4 \qquad \text{(Eq 7)}$$

where

$(p_1 \neq p_2 \ldots p_i)$

p = allele frequency

Exclusion of second parent:

$$P = 1 - 2\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} p_i^3 + 2\sum_{i=1}^{n} p_i^4 - 3\sum_{i=1}^{n} p_i^{5^2} - 2\left(\sum_{i=1}^{n} p_i^2\right)^2 + 3\sum_{i=1}^{n} p_i^2 \sum_{i=1}^{n} p_i^3$$

(Eq 8)

where

$(p_1 \neq p_2 \ldots p_i)$

p = allele frequency

Stellenbosch University http://scholar.sun.ac.za

Combined Exclusion Probability for different number of loci (k)

$$P = 1-(1-P_1)(1-P_2)(1-P_3)\Lambda\,(1-P_k) \qquad \text{(Eq 9)}$$

(Jamieson & Taylor, 1997)

PCR-reactions were performed as described in sections 2.7 and 2.8; p50 – 51.

2.10.3) ASSIGNMENT OF PARENTS

The software program Cervus 2.0 (Marshall *et al*, 1998; Slate *et al*, 2000) was used to assign both male and female parents to the offspring of the three different cohorts.

The program used exclusion probabilities to indicate the reliability of a marker or a combination of markers. Exclusion probabilities for individual markers and marker combinations were calculated using the same two formulae as shown in section 2.10.2; p57 – 58.

Cervus 2.0 used a likelihood-based paternity inference method to indicate the most likely parent or parental pair. The method calculates the natural logarithm of the likelihood ratio (LOD-score) of a particular parent in relation to the likelihood of an arbitrary parent (Marshall *et al*, 1998). The likelihood ratio is given by:

$$L(H_1, H_2 \mid D) = \frac{P(D \mid H_1)}{P(D \mid H_2)} \qquad \text{(Eq 10)}$$

where

$P(D \mid H_1)$ = probability of obtaining data D under hypothesis H_i

D = genotypes of parents and offspring

H_1 = hypothesis whether alleged father is the true father

Stellenbosch University http://scholar.sun.ac.za

H_2 = hypothesis whether alleged father is an unrelated individual

(Marshall *et al*, 1998)

An offspring was assigned to a parent or parents of one gender. The assigned parent or parents were then used as input for Cervus 2.0 as the known parent in order to assign a second parent of the opposite gender. A likelihood error rate of 0.01 was used for the assignment of both the first and second parent, while the threshold for the second parent was set at strict (95%). All parents with positive LOD-scores (see below) were selected as output for Cervus 2.0. Each offspring's individual genotype was then compared to the genotype of the Cervus-assigned parents to check the accuracy of the assignments. This was done using Microsoft® Excel 2000.

2.10.3.1) LOD-scores

The LOD score is the log (to base e) of the product of the likelihood ratios at each locus, or equivalently the sum of the log-likelihood ratios at each locus. A LOD score was calculated for each candidate parent, based on the genotypes of the candidate parent, offspring and other parent (if known).

A negative LOD score implied that the candidate parent was less likely to be the true parent than an arbitrary randomly-chosen individual: generally this happened when the candidate parent mismatched at one or more loci. If likelihoods were calculated without taking typing errors into account, any mismatch led to a likelihood ratio of zero, meaning that the LOD score was undefined. Negative LOD scores could also occur when the candidate parent and offspring shared very common alleles or was homozygous at many of the studied loci.

A LOD score of zero implied that the candidate parent was equally likely to be the true parent as an arbitrary randomly-chosen individual.

Stellenbosch University http://scholar.sun.ac.za

A positive LOD score implied that the candidate parent was more likely to be the true parent than an arbitrary randomly-chosen individual. Unless typing errors were very frequent, the true parent would almost always have a positive LOD score.

The most likely candidate parent was the candidate parent with the highest positive LOD score (Cervus-help file).

2.10.4) LOD-SCORE AND CONTRIBUTIONS OF PARENTS

The LOD-scores, the contributions of the parents and possible dominating genotypes were analysed using Microsoft® Excel 2000.

2.11) QUANTITATIVE TRAIT LOCI (QTL) DISCOVERY

2.11.1) ALLELE SEGREGATION

The analysis used bulked segregant analysis developed by Micklemore *et al* (1991) by making two different phenotypic pools namely fast growing (large) and slow growing (small) individuals of *Haliotis midae*. Robinson *et al*, 2003, described the method adapted for the QTL discovery.

The method investigated different parameters to search and test for significant segregation between an allele and a specific trait (fast and slow growth). The differences in allele peak heights (this is only for the pooled DNA PCR) and observed number of alleles between the large and the small samples were tested following a conventional χ^2-test.

The DNA concentration of all the fast and slow growers was standardized and a preliminary analysis (done by Robinson *et al*, 2003) was performed by constructing two DNA pools: one pool contained the DNA of the fast growers (60 individuals) and the other contained DNA from the slow growers (59 individuals). The pools were genotyped and an electropherogram for each

Stellenbosch University http://scholar.sun.ac.za

locus was obtained. The peak heights of the electropherograms (Figure 2.3; differences are indicated with red oval) were compared between the two pools and the significance of the observations were calculated following a conventional χ^2-test.

All loci showing significant differences for allele peak heights between the two pools were further used for the individual genotyping of the large and small animals. Significant differences in the observed number of alleles between the large individuals and the small individuals were used to identify segregating alleles. The significance of the observations were calculated following a conventional χ^2-test. It was hypothesised that alleles that showed significant allele number differences, is segregating with either the slow or fast growth trait.

The segregation of the alleles must be confirmed by doing backcrosses between the brood stock and the F1. This was not done during this study and is seen as future work.

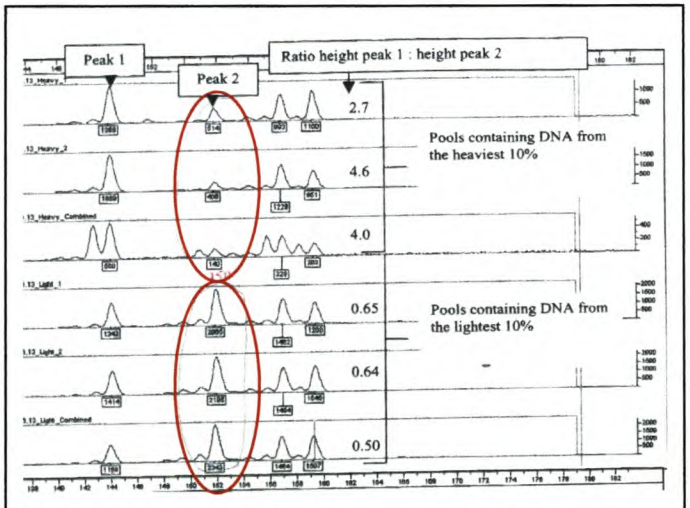


Figure 2.3: A visual difference are detected in the peak heights (red) between the two pools.

(Robinson *et al*, 2003)
Difference is indicated by red oval

Stellenbosch University http://scholar.sun.ac.za

2.11.2) GENOTYPE SEGREGATION

A bulked segregant analysis (developed by Michelmore *et al*, 1991) method was adopted from the one used by Lee *et al* (2003) to determine whether or not certain genotypes were linked to the fast growing (large) or slow growing (small) trait. Only F1 individuals were used in the analysis, because no information of the growth rate of the parents was available. The different genotypes per locus were compared between the large and the small individuals and the significant differences in the number of occurrences of each genotype were calculated following a conventional χ^2-test. It was hypothesised that a genotype is segregating with either the fast or slow growth rate trait if significant differences in the number of observations were found.

The segregation of the alleles must be confirmed by doing backcrosses between the brood stock and the F1. This was not done during this study and is seen as future work (see section 6.1.2; p130)

2.11.3) SAMPLE COLLECTION

The trait being studied was growth rate. Samples that contained fast growing (large) animals (10mm and more in length) and slow growing (small) animals (7mm and less in length) were obtained from Batches 238/03 and 266/03 from the Sea Plant Products hatchery. Each batch corresponded to an independent spawning event from Row 4 (Table 2.4). The size difference between the large and the small animals are shown in Figure 2.4.

Sampling was done in the destructive manner. The details of the samples are shown in Tables 2.14 – 2.17.

The animals from batch 266/03 (Tables 2.14 & 2.15) were sampled at the age of 147 days. The animals from batch 238/03 (Table 2.16 & 2.17) were sampled at the age of 175 days.

Stellenbosch University http://scholar.sun.ac.za

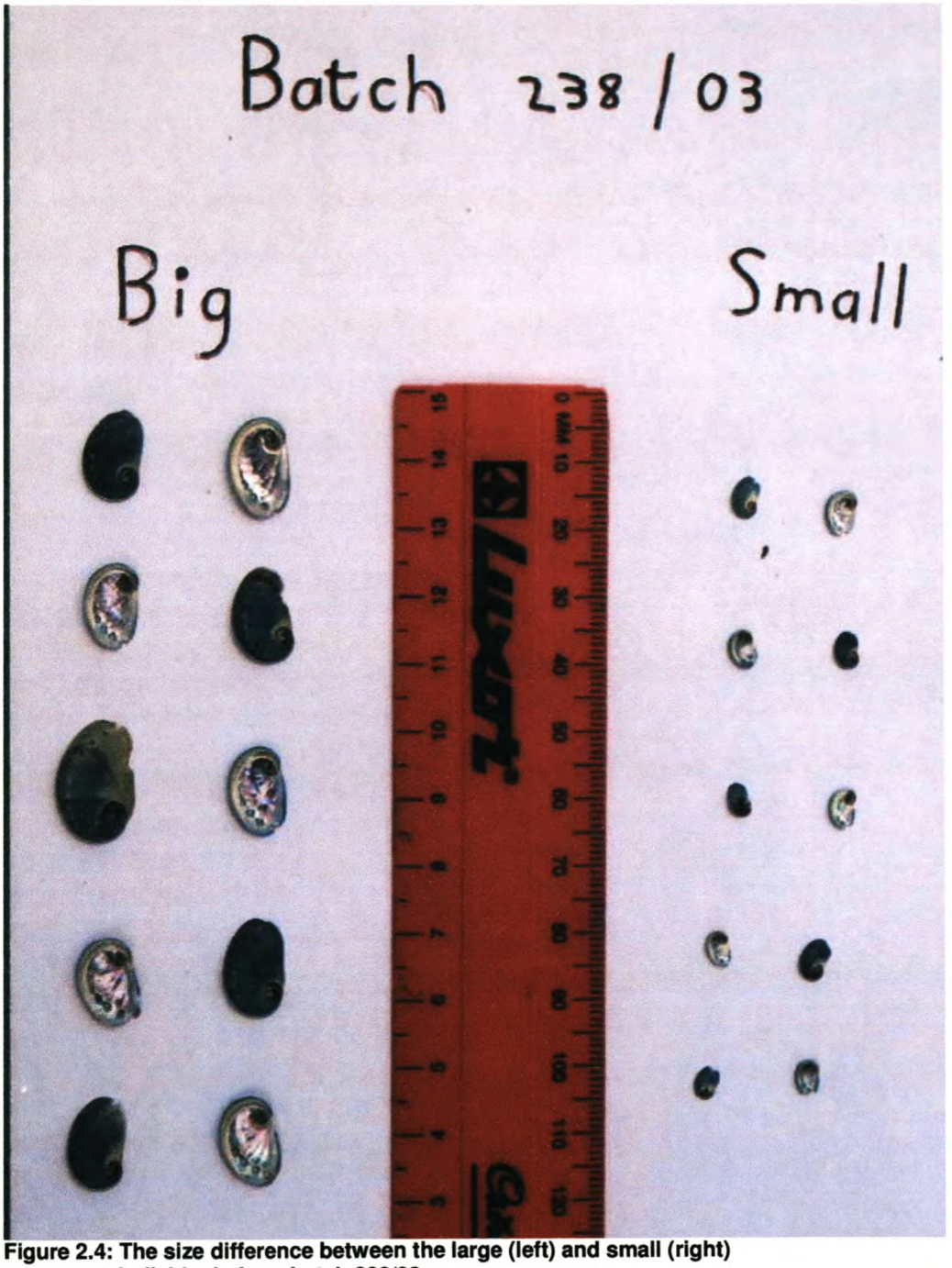


Figure 2.4: The size difference between the large (left) and small (right) individuals from batch 238/03.
(Photo taken by author)

Stellenbosch University http://scholar.sun.ac.za

Table 2.14: Batch 266/03 Large (266/03 – 48/04).

Animal ID	Length (mm)	Width (mm)	Preserved Sample	Extraction Sample
1	12	8	Whole	Muscle
2	11	7	Whole	Muscle
3	12	8	Whole	Muscle
4	11	7	Whole	Muscle
5	11	7	Whole	Muscle
6	12	8	Whole	Muscle
7	12	8	Whole	Muscle
8	11	7	Whole	Muscle
9	12	8	Whole	Muscle
10	11	7	Whole	Muscle
11	12	8	Whole	Muscle
12	11	7	Whole	Muscle
13	12	8	Whole	Muscle
14	12	8	Whole	Muscle
15	12	8	Whole	Muscle
16	12	8	Whole	Muscle
17	11	8	Whole	Muscle
18	12	7	Whole	Muscle
19	11	7	Whole	Muscle
20	11	7	Whole	Muscle
21	11	8	Whole	Muscle
22	12	8	Whole	Muscle
23	12	8	Whole	Muscle
24	11	7	Whole	Muscle
25	12	8	Whole	Muscle
26	12	8	Whole	Muscle
27	12	8	Whole	Muscle
28	11	7	Whole	Muscle
29	11	8	Whole	Muscle
30	11	7	Whole	Muscle

Stellenbosch University http://scholar.sun.ac.za

Table 2.15: Batch 266/03 Small (266/03 – 48/04).

Animal ID	Length (mm)	Width (mm)	Preserved Sample	Extraction Sample
1	5	3	N/A	Muscle
2	6	4	N/A	Muscle
3	5	3	N/A	Muscle
4	4	3	N/A	Muscle
5	5	4	N/A	Muscle
6	5	4	N/A	Muscle
7	5	3	N/A	Muscle
8	6	4	N/A	Muscle
9	5	4	N/A	Muscle
10	4	3	N/A	Muscle
11	5	4	N/A	Muscle
12	5	3	N/A	Muscle
13	5	4	N/A	Muscle
14	4	3	N/A	Muscle
15	5	4	N/A	Muscle
16	4	3	N/A	Muscle
17	6	4	N/A	Muscle
18	5	4	N/A	Muscle
19	4	3	N/A	Muscle
20	5	3	N/A	Muscle
21	5	4	N/A	Muscle
22	5	4	N/A	Muscle
23	5	4	N/A	Muscle
24	5	4	N/A	Muscle
25	4	2	N/A	Muscle
26	5	4	N/A	Muscle
27	5	4	N/A	Muscle
28	4	3	N/A	Muscle
29	4	3	N/A	Muscle
30	4	3	N/A	Muscle

Stellenbosch University http://scholar.sun.ac.za

Table 2.16: Batch 238/03 Large (238/03 – 48/04).

Animal ID	Length (mm)	Width (mm)	Preserved Sample	Extraction Sample
1	14	9	Whole	Muscle
2	14	9	Whole	Muscle
3	13	8	Whole	Muscle
4	15	10	Whole	Muscle
5	18	12	Whole	Muscle
6	15	10	Whole	Muscle
7	15	9	Whole	Muscle
8	15	9	Whole	Muscle
9	13	8	Whole	Muscle
10	13	8	Whole	Muscle
11	13	9	Whole	Muscle
12	13	9	Whole	Muscle
13	13	8	Whole	Muscle
14	14	9	Whole	Muscle
15	14	9	Whole	Muscle
16	14	9	Whole	Muscle
17	13	8	Whole	Muscle
18	14	9	Whole	Muscle
19	14	9	Whole	Muscle
20	13	8	Whole	Muscle
21	13	9	Whole	Muscle
22	14	9	Whole	Muscle
23	13	8	Whole	Muscle
24	14	8	Whole	Muscle
25	14	9	Whole	Muscle
26	13	8	Whole	Muscle
27	13	8	Whole	Muscle
28	13	8	Whole	Muscle
29	13	8	Whole	Muscle
30	13	8	Whole	Muscle

Stellenbosch University http://scholar.sun.ac.za

Table 2.17: Batch 238/03 Small (238/03 – 48/04).

Animal ID	Length (mm)	Width (mm)	Preserved Sample	Extraction Sample
1	6	4	N/A	Muscle
2	6	4	N/A	Muscle
3	5	4	N/A	Muscle
4	6	4	N/A	Muscle
5	7	4	N/A	Muscle
6	4	3	N/A	Muscle
7	4	2	N/A	Muscle
8	5	3	N/A	Muscle
9	6	4	N/A	Muscle
10	6	4	N/A	Muscle
11	6	4	N/A	Muscle
12	5	4	N/A	Muscle
13	7	5	N/A	Muscle
14	6	4	N/A	Muscle
15	6	4	N/A	Muscle
16	5	4	N/A	Muscle
17	5	4	N/A	Muscle
18	5	4	N/A	Muscle
19	5	4	N/A	Muscle
20	5	4	N/A	Muscle
21	5	4	N/A	Muscle
22	5	4	N/A	Muscle
23	5	4	N/A	Muscle
24	6	4	N/A	Muscle
25	4	2	N/A	Muscle
26	4	3	N/A	Muscle
27	6	4	N/A	Muscle
28	5	4	N/A	Muscle
29	4	2	N/A	Muscle

Stellenbosch University http://scholar.sun.ac.za

2.11.4) DNA QUANTIFICATION AND POOLING

2.11.4.1) Quantification

See section 2.2; p43.

2.11.4.2) Pools

Two DNA-pools (Table 2.18) were created after standardizing the DNA concentration ([DNA] ≈ 40ng/μl) using agarose gel electrophoresis. The pools were created when 5μl of each individual was added together. The ALL_Large pool consisted of the DNA of 60 large individuals and the ALL_Small pool consisted of 59 small individuals.

Tabel 2.18: DNA pools.

POOLS	COMPOSITION
ALL_Large	Large individuals of batch 238/03 and 266/03
ALL_Small	Small individuals of batch 238/03 and 266/03

2.11.5) GENOTYPING

Loci HmD11, HmD30, HmSP5, HmD55, HmD59, HmD14 and HmT35 were used in this analysis. The loci are shown in Table 2.19.

PCR-reactions for the pooled samples and the 119 individuals were performed and analysed as described in sections 2.7, 2.8 and 2.9; p50 – 55.

Stellenbosch University http://scholar.sun.ac.za

Table 2.19: Labelled microsatellite loci used in QTL-discovery.

Locus	Repeat Sequence	Primer (5'-3')	Label	Acc Nr
HmD11	$(TCTG)_8$	F-AGCTCAGAAAAGTGGTGTACG R-TTACCTAGCTAAAGTTGACAACG - #	VIC	AY303341
HmD30	$(AGTC)_2GGTC(AGTC)_{11}$	F-TGATGTTGCTGGAATATTGC R-CAATTTCATTTTCAACAGTTCA - #	VIC	AY303342
HmSP5	$(AC)_{13}$	F-TTCGGCAAGTGAATGTCTAG - # R-ATGCGACACTTACTACACCG	FAM	AY303344
HmD55	$(GTGA)_{12}$	F-ATCAAGATAAAACGAGGCG R-ACCACTGTGAAAACGTCCA - #	VIC	AY303337
HmD59	$(CA)_{15}$	F-TATACTGCCATTTCCGTCTG - # R-TCTGTATTCTGGTCCTGTCG	FAM	AY303338
HmD14	$(CA)_{10}$	F-TAAGGCAAGTGAATGTCTAG - # R-ATTGCAAGAATCACAACTGC	NED	AY303333
HmT35	$(TGAC)_5(TCAC)_3$	F-TTAGATATCGCTGTTTATGCTG - # R-AGTTGATTGTGTCTGAGAGGG	NED	unpublished

- indicates labelled primer

2.11.6) PROCEDURE FOR QTL-DISCOVERY

2.11.6.1) Conventional χ^2test

H_o: The peak height, allele frequencies and observed number of alleles are equally distributed between groups with an expected segregation of 1:1 per allele.

H_o would be rejected if a significant χ^2-value for one degree of freedom was obtained. See Appendix E for the χ^2-table.

$$\chi_{LS}^{\ 2} = \frac{(o_L - e_L)^2}{e_L} + \frac{(o_S - e_S)^2}{e_S} \qquad \text{(Eq 11)}$$

where

χ_{LS}^2 = χ^2-value per allele

Stellenbosch University http://scholar.sun.ac.za

o_L = observed value for large sample
e_L = expected value for large sample
o_S = observed value for small sample
e_S = expected value for small sample

(adapted from Liu, 1997)

2.11.6.2) Allele segregation

2.11.6.2.1) Genotyping of DNA pools

The DNA of 60 large individuals and 59 small individuals was pooled together and genotyped using 7 microsatellite DNA markers (Table 2.19).

The peaks (from electropherograms) for each allele per locus were compared between the ALL_Large pool and the ALL_Small pool to look for differences in the peak heights. The significance of the observed peak heights was calculated per allele following a conventional χ^2-test.

2.11.6.2.2) Genotyping of individual samples

One hundred and nineteen individuals (Tables 2.14 – 2.17) were genotyped using 7 microsatellite DNA markers (Table 2.19).

The observed number of alleles was calculated by using Microsoft® Excel 2000. Differences in the number of observed alleles per locus between the large individuals and the small were investigated. The significance of the observed differences were calculated following a conventional χ^2-test.

2.11.6.2.3) Selection of potential segregating alleles

The criteria for selecting alleles potentially segregating with the large or small trait was (Robinson *et al*, 2003):

Stellenbosch University http://scholar.sun.ac.za

1. Significant differences in peak heights between the ALL_Large and ALL_Small pools were used as estimators of allele number, but not as conclusive proof of segregation.
2. Significant differences in the observed number of alleles between the large individuals and the small individuals were used as possible proof of segregation.

2.11.6.3) Genotype segregation

The number of observations of the different genotypes per locus for the large and small individuals were calculated using Microsoft® Excel 2000. The number of observations was compared between the large individuals and the small individuals. The significance of the observed differences were calculated following a conventional χ^2-test.

A genotype was considered as a segregant if significant differences were detected in the number of observations.

2.11.7) DIFFERENTIAL CONTRIBUTIONS OF THE PARENTS

The male / female combinations of 25 fast growing (large) individuals and 24 slow growing (small) individuals that were 100% assigned during parentage analysis were compared between the fast growing and the slow growing trait. A histogram was drawn with Microsoft® Excel 2000 to visualise the contributions of each pair to either the fast growers (large) or the slow growers (small).

Stellenbosch University http://scholar.sun.ac.za

2.12) REFERENCES

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L., Bonhomme, F., 2000. *GENETIX*, logiciel sous Window TM pour la génétique des populations. Laboratoire Génome et Populations, Montpellier.

Bester, A.E., Slabbert, R., D'Amato, M.E. Isolation and characterisation of microsatellite markers in South African abalone (*Haliotis midae*). In press. *Molecular Ecology Notes*. (Appendix A)

Brookfield, J.F., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**, 453 – 455.

Eason, G., Coles, C.W., Gettinby, G., 1989. Mathematics and statistics for the bio-sciences. Ellis Horwood Limited Publishers, Chichester, pp. 487 – 492.

Estoup, A., Tailliez, C., Cornuet, J., Solignac, M., 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, Apis mellifera and Bombus terrestris (Apidae). *Molecular and Biological Evolution* **12**, 1074 – 1084.

Fernando, P., Evans, B.J., Morales, J.C., Melnick, D.J., 2001. Electrophoresis artefacts – a previously unrecognised cause of error in microsatellite analysis. *Molecular Ecology* **1**, 325 – 328.

Fishtech, Inc. Facts about abalone [Online]. Available: http://www.fishtech.com/facts.html. [2004, August]

Guo, S.W., Thompson, E.A., 1992. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics* **48**, 361 – 372.

Stellenbosch University http://scholar.sun.ac.za

Haldane, J.B.S., 1954. An exact test for randomness of mating. Journal of *Genetics* **52**, 631 – 635.

Jamieson, A., Taylor, St.C.S., 1997. Comparisons of three probability formulae for parentage exclusion. *Animal Genetics* **28**, 397 – 400.

Kimura, M., Crow, J.F., 1964. The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725 – 738.

Koorey, D.J., Bishop, G.A., McCaughan, G.W., 1993. Allele non-amplification: a source of confusion in linkage studies employing microsatellite polymorphisms. *Human Molecular Genetics* **2**, 289 – 291.

Lee, B.-Y., Penman, D.J., Kocher, T.D., 2003. Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. *Animal Genetics* **34**, 1 – 5.

Li, Q., Park, C., Endo, T., Kijima, A., 2004. Loss of genetic variation at microsatellite loci in hatchery strains of the Pacific abalone (*Haliotis discus hannai*). *Aquaculture* **237**, 207 – 222.

Liu, B.-H., 1997. Chapter 6: Two-locus models: The controlled crosses. In Liu, B.-H. *Statistical Genomics: Linkage, Mapping, and QTL Analysis*. Boca Raton, London, New York, Washington, D.C.: CRC Press, 161 – 166.

Marshall, T.C., Slate, J., Kruuk, L.E.B., Pemberton, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639 – 655.

Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using

Stellenbosch University http://scholar.sun.ac.za

segregating populations. *Proceedings of the National Academy of Science USA* **88**, 9828 – 9832.

Pemberton, J.M., Slate, J., Bancroft, D.R., Barrett, J.A., 1995. Nonamplification alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**, 249 – 252.

Rahman, M.H., Jaquish, B., Khasa, P.D., 2000. Optimization of PCR protocol in microsatellite analysis with silver and SYBR stains. *Plant Molecular Biology Reporter* **18**, 339 – 348.

Raymond, M., Rousset, F., 1995. Population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248 – 249.

Robinson, N., Baranski, M., Rourke, M., Hayes, B., 2003. Application of advanced genetic techniques to abalone breeding.10th Abalone Aquaculture Workshop.

Saghai-Maroof, M.A., Solima, K.M., Jorgenson, R.A., Allard, R.W., 1984. Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences (USA)* **81**, 8014 – 8018.

Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin Ver. 2.000: A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Selander, R.K.,1970. Behaviour and genetic variation in natural populations. *American Zoologist* **10**, 53 – 66.

Slate, j., Marshall, T.C., Pemberton, J.M., 2000. A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology* **9**, 801 – 808.

Stellenbosch University http://scholar.sun.ac.za

Uitenbroek, D.G., 1997. *SISA Binomial*. Online: http://home.clara.net/sisa/binomial.htm. [1 Jan 2004].

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P.F., 2003. Micro-Checker. Department of Biological Sciences, Department of Computer Science, The University of Hull.

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

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1 – 16.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 3 – ISOLATION OF MICROSATELLITE LOCI

3.1) ISOLATION OF GENOMIC DNA

Genomic DNA was successfully isolated by using abalone muscle tissue. Yields of between 50 ng and 200 ng were achieved.

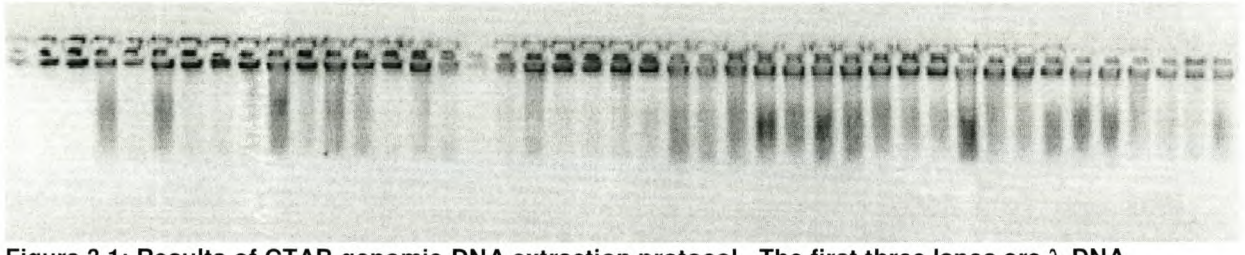


Figure 3.1: Results of CTAB genomic-DNA extraction protocol. The first three lanes are λ-DNA standards (50, 100, 200 ng).

3.2) MICROSATELLITE ISOLATION

3.2.1) 1ST AFLP AMPLIFICATION

Optimal PCR-products were obtained (Figure 3.2). Optimal PCR-products were described as smears containing no distinct bands (multi copy sequences) and all products were expected to be above 200bp (Zane *et al*, 2002). The smears contained different sizes of bands flanked by *Mse* sites. The number of cycles had an influence on the amplification of the bands. The higher number (26) of cycles showed fewer amplification products at the lower and upper size ranges, but the lower number (17) of cycles showed a greater size range for products (Figure 3.2).

Stellenbosch University http://scholar.sun.ac.za

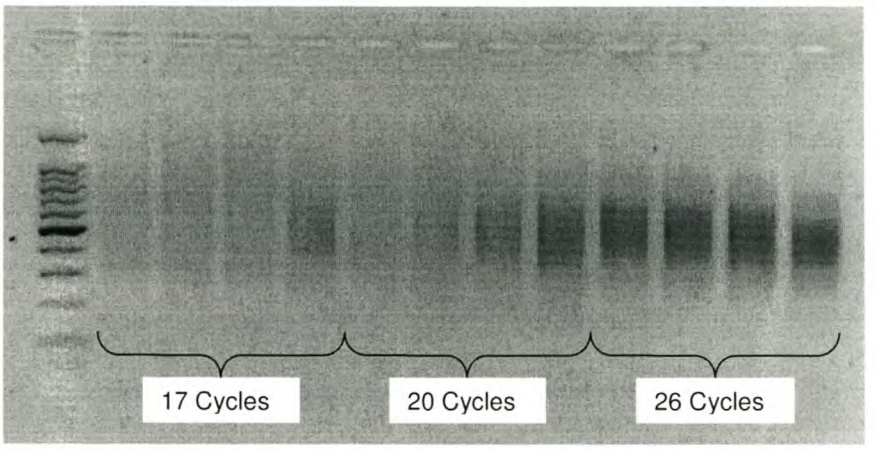


Figure 3.2: 1st AFLP amplification showing amplification products with 17, 20 and 26 cycles.

3.2.2) HYBRIDISATION AND SELECTIVE CAPTURE OF HYBRIDISED DNA

Amplification products obtained from the PCR-program using 17 cycles were selected for the hybridisation and capturing steps.

3.2.3) 2ND AFLP AMPLIFICATION

The elusions from the four washes were expected to contain DNA fragments with *Mse*I-N priming sites at the ends. Higher concentrations of these DNA fragments were expected from the first denaturation step with decreasing concentrations in the stringency and non-stringency washes. The second denaturation step should contain little or no DNA fragments with the *Mse*I-N priming sites.

The DNA fragments obtained from the washes were amplified. The amplification products (smears) are shown in Figure 3.3. As expected, the concentration of the amplification products varied in the three washes, being the highest for the first denaturation step, followed by the stringency and non-stringency washes and the second denaturation step. The second denaturation step was expected to have no amplification products, but it would have no effect on further results. The smears were all above 200bp, which were also expected, because the amplification products used in section

Stellenbosch University http://scholar.sun.ac.za

3.2.2; p77 should only have contained amplifiable fragments of more than 200bp (Zane *et al*, 2002).

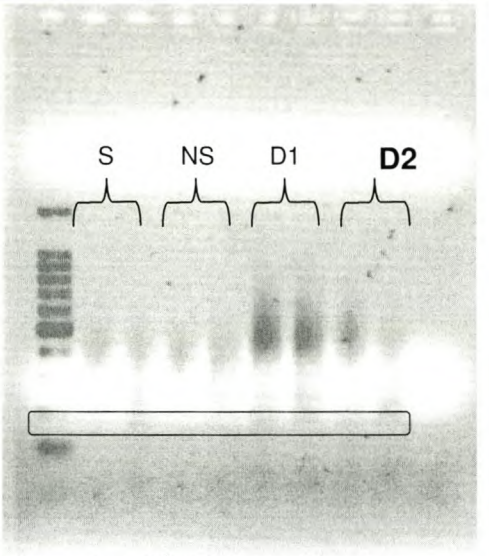


Figure 3.3: 2nd AFLP amplification showing amplification products above 200bp (indicated by block).
S = stringency washes
NS = non-stringency washes
D1 = first denaturation step
D2 = second denaturation step

3.2.4) CLONING

The amplification product of the first denaturation step from individual 16-20 was cloned into the TOPO-TA vector (*Invitrogen*) in order to construct a highly enriched microsatellite library.

3.2.5) COLONY SCREENING AND SEQUENCING

A total of 188 colonies were screened. Colonies were selected for sequencing according to size (> 500bp) and whether or not a single product was observed on the gel (Figure 3.4). The products used for sequencing were not cleaned, because it was deemed that they contained minimal primer

Stellenbosch University http://scholar.sun.ac.za

dimers that could impair the sequencing reaction. A total of 38 colonies were sequenced. Twenty-one of these clones contained microsatellite repeats, including dinucleotide (Figure 3.5), tetranucleotide (Figure 3.6) and pentanucleotide (Figure 3.7) repeats. The names of the clones and repeats are shown in Table 3.1.

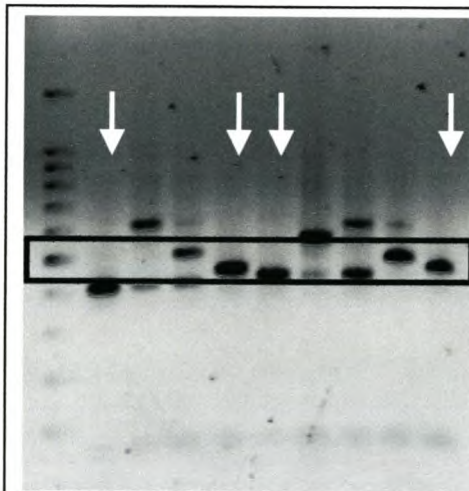


Figure 3.4: Results of screened colonies. Only clones above 500bp (indicated by block) are used for sequencing. Arrows indicate single products.

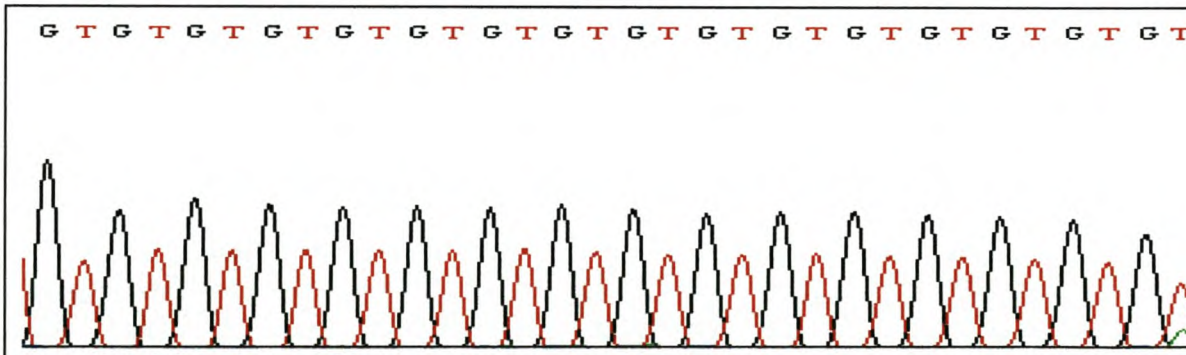


Figure 3.5: A GT – dinucleotide repeat.

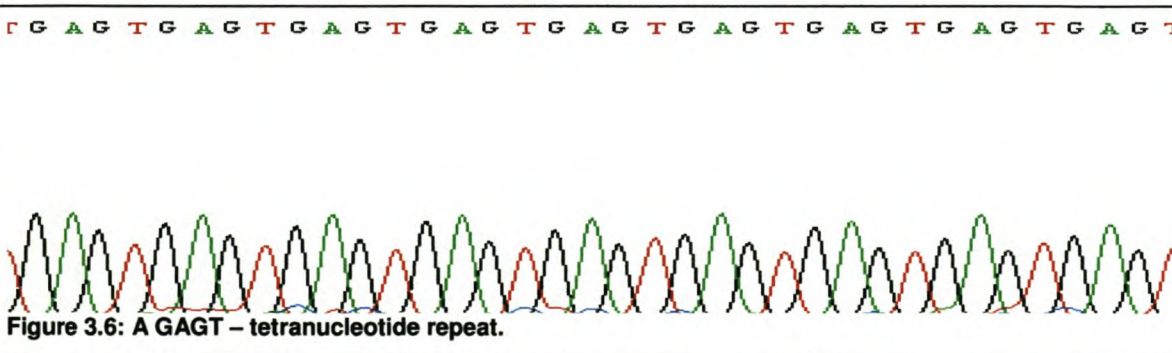


Figure 3.6: A GAGT – tetranucleotide repeat.

Stellenbosch University http://scholar.sun.ac.za

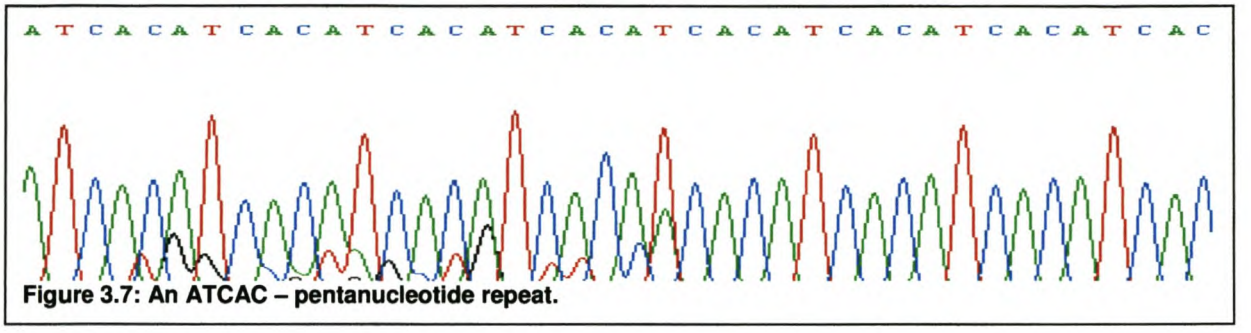


Figure 3.7: An ATCAC – pentanucleotide repeat.

Table 3.1: Names of clones containing repeats and their respective repeats.

Clone Name	Repeat
AD5	$(GT)_3$ $(AT)_3$ $(GT)_8$ $(GA)_{14}$
AD9	$(GT)_{12}$ $(ATGT)_4$ $(GT)_5$ $(TG)_4$
AD19	$(ATCAC)_{14}$ $(ATAC)_{19}$ $(AC)_{17}$ $(AG)_2$ $(AC)_3$ (AG) $(AC)_4$ $(AGAC)_3$ $(AC)_6$
AD24	$(AG)_5$ AT $(AG)_{21}$ $(GT)_{33}$ $(GT)_8$
AD25	(TG)12 TA $(TG)_{22}$
AD28	$(GAGT)_3$ $(GT)_2$ $(GAGT)_{14}$ $(GTGA)_5$
AD31	$(CACG)_3$ (AC)9 GC $(GCAC)_6$
AD33	$(CA)_4$ $(GACA)_2$ $(CA)_{10}$ $(CGCA)_2$ $(CA)_4$
AD35	$(TG)_{17}$
AD44	$(GAGT)_{20}$
AD46	$(GTGC)_4$ $(GT)_{11}$
AD56	(AT)4 GTAT $(GTGC)_3$ $(GT)_4$ $(GC)_3$ (CT) (GT) (GC) (GT) $(CT)_3$ (GT)2 CT $(GT)_3$ $(GTGC)_4$ $(TG)_{11}$
AD63	(GTGGT)3 AT $(GGTGT)_4$ $(AGTGT)_4$ $(GCGT)_6$
AD65	(TG)2 CG $(TG)_4$ $(CG)_4$ $(TG)_3$
AD69	$(GT)_9$ $(GC)_5$ $(GTGC)_6$ $(GT)_4$
AD96	$(GT)_3$ $(TG)_4$ $(CG)_3$ $(TG)_{13}$ $(CGTG)_3$
AD102	$(CACT)_{15}$
AD133	$(AC)_{13}$
AD139	$(GT)_{13}$
AD142	$(GAGT)_{19}$ $(GAGT)_4$
ATE	$(GTCA)_4$

Stellenbosch University http://scholar.sun.ac.za

The most common of the dinucleotide repeats was GT – AC (Figure 3.5), making up 49% of all the observed dinucleotide repeats. The two most common tetranucleotide repeats were GAGT (Figure 3.6) and GTGC, making up 23% and 18% of all observed tetranucleotide repeats, respectively. Many of the repeats were compound repeats (Figure 3.8), for example AD5 that contained four dinucleotide repeats and AD46 that contained one di- and one tetranucleotide repeat. It was possible to design primer pairs for these compound repeats, but it would be impossible to explain the origin of any mutations observed in allele sizes (Estoup *et al*, 1995). Observed polymorphisms of the AD5 compound repeat could be due to a loss of one or more GT, AT or GA repeat units for example $(GT)_3$ $(AT)_3$ $(GT)_8$ **$(GA)_{14}$** versus $(GT)_3$ $(AT)_3$ $(GT)_8$ **$(GA)_{12}$** (where two GA repeats were lost) or even a loss of a combination of a GT and an AT repeat unit, for example $(GT)_3$ **$(AT)_3$** **$(GT)_8$** $(GA)_{14}$ versus $(GT)_3$ **$(AT)_2$** **$(GT)_7$** $(GA)_{14}$ (where one GT and one AT repeat were lost). Both afore mentioned alleles had the same size identity, but their ancestral origins differed. This occurrence, called homoplasy, would influence analysis such as parental assignments and the search for linked marker systems.

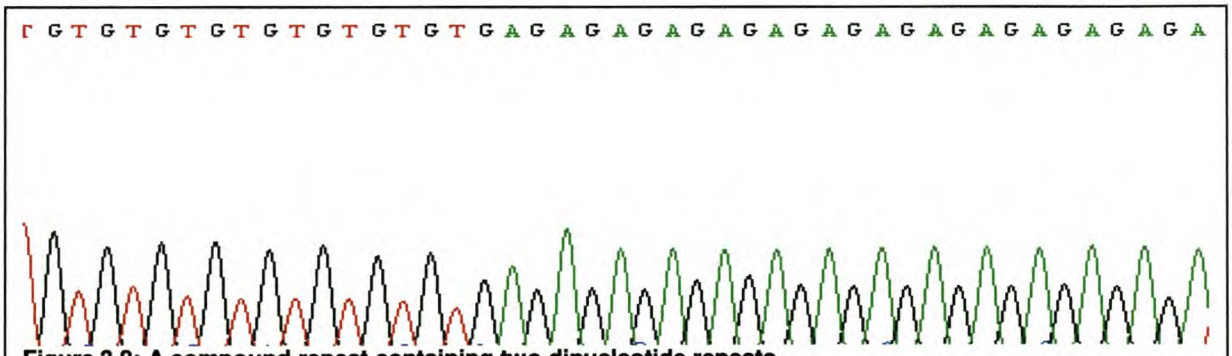


Figure 3.8: A compound repeat containing two dinucleotide repeats.

Sequence alignment of the repeat containing clones detected one duplicate locus within clones D26 and D30 (cloned by Aletta Bester, Department of Genetics, University of Stellenbosch, Aquaculture Division), most probably caused by the recombination between clones. No further overlapping regions between the various flanking regions were detected. It was safe to assume

Stellenbosch University http://scholar.sun.ac.za

that no other recombination events took place between any of the clones to cause nested priming sites. **Blastn** (available: www.ncbi.nlm.nih.gov/BLAST) were performed to verify whether or not the cloned inserts showed major homology with unrelated taxa. The results showed that the flanking regions of the repeat were mostly related to other marine species (for example: *Penaeus monodon*; Acc. Nr.: AY654002) or to other *Haliotis*-species (for example: *Haliotis rubra*; Acc. Nr: AF302830).

A total of 11 primer pairs were designed. Clones containing adequate and unambiguous flanking sequences were used to design the primer pairs. The primers and their respective repeats being amplified are shown in Table 3.2. Primer sizes were kept at a minimum of 20 nucleotides in order to obtain an annealing temperature of approximately 60°C.

Table 3.2: Primers designed from clones containing repeats.

Primer Name	Repeat Type	Primer (5' – 3')
AD9	$(GT)_{12}$ $(ATGT)_4$ __ $(GT)_5$	F-TTCCATTGGAGTCAGAGTCG R-TTAGCTCACACACAGACGTATTG
AD24A	$(AG)_5$ AT $(AG)_{21}$	F-TGTGGACTGCCTGTGTTGTA R-CATTATACTCTCTAAATTCTGTGTGTG
AD24B	$(GT)_{33}$	F-AGAGAGTATAATGAGGTAAAATGGT R-TGATTCCTGCTCAATTTTGT
AD28	$(GAGT)_3$ $(GT)_2$ __ $(GAGT)_{14}$ __ $(GTGA)_5$	F-ATAAAAGGTCTTTCCAAGTGG R-AAGACTCAGGTTTGATTCACC
AD33	$(CA)_{10}$ $(CGCA)_2$ $(CA)_4$	F-ATAGTGGTCATACAGTCATCACCT R-TAGGCATGTTTGAGTTCGTGT
AD35D	$(TG)_{17}$	F-ATGGTTATACTCGCGAAATG R-TGAAAATGAAGTGTGGTTCC
AD44T	$(GAGT)_{20}$	F-CAAACAGTAGAGTGAAACTCGC R-CCAGTGTATATTGCGCTAGC
AD56	(AT)4 GTAT $(GTGC)_3$ __ $(GT)_4$ $(GC)_3$ (CT) (GT) (GC) (GT) $(CT)_3$ (GT)2 CT $(GT)_3$ __ $(GTGC)_4$ __ $(TG)_{11}$	F-TGTTAACTTTGCGTTCAAGAT R-CAAGGCTGTACTGAAGGCTA
AD65	(TG)2 CG $(TG)_4$ $(CG)_4$ $(TG)_3$	F-GATCGTGAACTGAACCATGA R-GTCAAATTCGAGGGATGATT
AD133D	$(AC)_{13}$	F-TTCGGCAAGTGAATGTCTAG R-ATGCGACACTTACTACACCG
AD142T	$(GAGT)_{19}$	F-ATATTTTAGGCACAAAAGGTTT R-ACTCACATGATATTGCTGGAA

Stellenbosch University http://scholar.sun.ac.za

Of the 11 primer pairs, 5 should amplify perfect repeats and the other 6 pairs should amplify compound repeats. AD24A-reverse contained a TG-repeat sequence at the 3'-priming end, that could cause problems. This problem was caused, because the flanking regions were too short to design a 20-nucleotide long primer without including repeat units. Any mutation in this short repeat unit would influence the binding of the 3'-priming end and no elongation will take place. A high probability for the occurrence of null alleles would exist at this locus.

3.3) OPTIMIZATION OF PCRs USING UNLABELED MICROSATELLITE PRIMERS

The PCR conditions for the two primer pairs designed by Aletta Bester (Department of Genetics, University of Stellenbosch, Aquaculture Division) namely D11 and D30 (see section 3.2.5; p81 for note on D30), were also optimised. DNA from 3 different individuals (16-7, S4 and S27) was used to test for amplification products. Of the 13 primer pairs optimised for PCR-reactions, only 3 pairs failed to amplify at all. The 3 unsuccessful primer pairs were AD9, AD24A and AD56. PCR-conditions for the different primers are shown in Table 3.3.

Table 3.3: Optimised PCR-conditions of primers.

Primer Name	[DNA] (ng)	[$MgCl_2$] (mM)	[Primer] (pmol/*ul*)	Annealing Temperature
D11	20	2	2	Touch-Down
D30	20	2	2	Touch-Down
AD24B	20	2	10	Touch-Down
AD28	20	2	5	Touch-Down
AD33	20	2	2	Touch-Down
AD35D	20	2	10	Touch-Down
AD44T	20	2	10	Touch-Down
AD65	20	2	10	Touch-Down
AD133D	20	2	2	Touch-Down
AD142T	20	2	10	Touch-Down

Stellenbosch University http://scholar.sun.ac.za

3.4) CHOICE OF PRIMERS TO BE LABELLED AND LABELLING OF PRIMERS

Mighty Small (*Hoefer*) poly-acrylamide gels were used to test the polymorphism of the products amplified by the 10 primer pairs (Table 3.3). Before large-scale analysis was done, a test run was done to establish whether there was a difference in resolution between a 6% and a 12% poly-acrylamide gel. After the test run of identical samples on a 6% and a 12% gel, it was concluded that the 12% poly-acrylamide gel had a better resolution and an increased ability to separate the PCR-products.

The results of the PAA-gel analysis are shown in Table 3.4. A total of 5 out of the 10 primer pairs were found to be polymorphic. The choice of primers to be labelled is also shown in Table 3.4. Loci AD35D, AD44T and AD142T were labelled, even though they showed no levels of polymorphism between individuals (see Table 3.2). It could also be argued that the PAA-gel analysis was not sensitive enough to detect the existence of small mutations between alleles. The primer pairs of loci AD24B and AD28 were not labelled and the sequences were not submitted to GenBank. The loci names and GenBank accession numbers are also shown in Table 3.4 and Appendix A.

Table 3.4: Results after PAA-gel analysis and labelling.

Loci Name	Polymorphic	LABEL	Primer Name	Accession Number
D11	✔	VIC	HmD11	AY303341
D30	✔	VIC	HmD30	AY303342
AD24B	✘	-	-	Not submitted
AD28	✘	-	-	Not submitted
AD33	✔	PET	HmSP1	AY303346
AD35D	✘	NED	HmSP2	AY303347
AD44T	✘	FAM	HmSP3	AY303343
AD65	✔	NED	HmSP4	AY303348
AD133D	✔	FAM	HmSP5	AY303344
AD142T	✘	FAM	HmSP6	AY303345

Stellenbosch University http://scholar.sun.ac.za

3.5) OPTIMISATION AND ANALYSIS OF LABELLED PRIMERS

Four additional labelled primer pairs were received from Aletta Bester (Department of Genetics, University of Stellenbosch, Aquaculture Division), and are shown in Table 3.5 and Appendix A (Bester et al, in press). These primer pairs amplified successfully. Their respective PCR conditions are summarised in Table 3.6.

Table 3.5: Additional labelled primers received from Aletta Bester*.

Primer Name	Repeat Type	Primer (5' – 3')	Label	Accession Number
HmD55	$(GTGA)_{12}$	F-ATCAAGATAAAACGAGGCG R-ACCACTGTGAAAACGTCCA - #	VIC	AY303337
HmD59	$(CA)_{15}$	F-TATACTGCCATTTCCGTCTG - # R-TCTGTATTCTGGTCCTGTCG	FAM	AY303338
HmD14	$(CA)_{10}$	F-TAAGGCAAGTGAATGTCTAG - # R-ATTGCAAGAATCACAACTGC	NED	AY303333
HmT35	$(TGAC)_5(TCAC)_3$	F-TTAGATATCGCTGTTTATGCTG - # R-AGTTGATTGTGTCTGAGAGGG	NED	unpublished

- indicates labelled primer
* Department of Genetics, University of Stellenbosch, Aquaculture Division

Table 3.6: Optimised PCR-conditions of labelled primers.

Primer Name	[DNA] (ng)	[$MgCl_2$] (mM)	[Primer] (pmol/*ul*)	Annealing Temperature
HmD11	20	2	2	Touch-Down / 60°C
HmD30	20	2	5	Touch-Down 2
HmSP1	20	2	5	Touch-Down 2
HmSP3	20	2	5	Touch-Down / 60°C
HmSP5	20	2	2	Touch-Down
HmD55	20	2	2	Touch-Down
HmD59	20	2	2	Touch-Down
HmD14	20	2	2	Touch-Down
HmT35	20	2	5	Touch-Down

Five of the 8 labelled primers (Table 3.4) amplified successfully and their respective PCR-conditions are shown in Table 3.6. HmSP3 failed to show

Stellenbosch University http://scholar.sun.ac.za

any polymorphism between different individuals after analysis with the Genotyper® software (*Applied Biosystems*). HmSP3 had only one allele at a size of 380bp. HmT35 also showed little polymorphism by only amplifying 3 alleles, two of which occurred at very low frequencies.

Alleles were scored using the following criteria to classify a peak as an allele: The two peaks shouldn't differ more than 2/3 in height from each other, especially if the first peak correponds to the short allele. The first of the two peaks shown in Figure 3.9 was too small to be scored as an allele. The second peak was usually expected to have a smaller peak height, because of allele dropout (Gagneux *et al*, 1997). Allele dropout occurs when amplification of the smaller allele was favoured, resulting in a lower peak height for the larger allele.

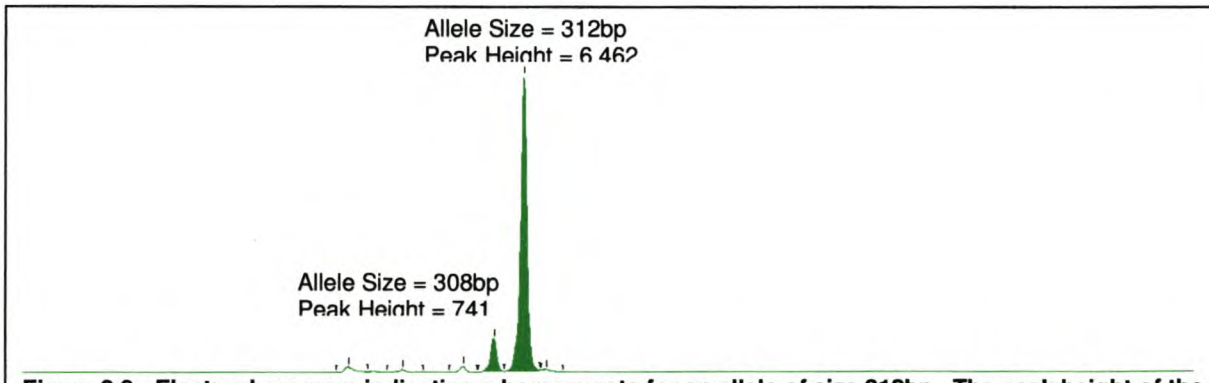


Figure 3.9 : Electropherogram indicating a homozygote for an allele of size 312bp. The peak height of the allele at 308bp is too short to be an allele.

Sometimes a sequential series of peaks would occur with increasing (Figure 3.10) or decreasing peak heights that will differ in one repeat size from each other. In dinucleotide repeats the peaks showed a difference of 2 basepairs ($n \pm 2$ peaks; Figure 3.10) and in tetranucleotide repeats showed a difference of 4 basepairs ($n \pm 4$ peaks). The $n \pm x$ peaks were caused by in vitro slippage during DNA amplification. A slippage event was caused when the nascent DNA strand dissociated from the template DNA strand. When microsatellite repeat-units occurred, the nascent strand may have reannealed out of phase and formed loop structures in the repeat sequence. The loop

Stellenbosch University http://scholar.sun.ac.za

structures consisted of the unpaired repeat units. The Taq polymerase being used for PCR-amplification had no proofreading ability and would not correct the misannealed replication complexes which contained these loop structures. The $n+x$ peaks were caused by the occurrence of loop structures in the nascent DNA strand (Figure 3.11a), while the $n-x$ peaks were caused by the occurrence of loop structures in the template DNA strand (Figure 3.11b; Hancock, 1999).

The shorter $n \pm x$ peak heights observed in the chromatograms could be explained as follows. It was assumed that the amplification of the correct sized allele would occur more frequently than the slippage events. Slippage events usually caused the loss of one repeat unit, but it could lead to the loss of more repeat units. With dinucleotide repeats it was expected that the peak heights of $n \pm 2$ peaks would be higher than $n \pm 4$ etc. as shown in Figure 3.10. The allele with the greatest peak height would be scored.

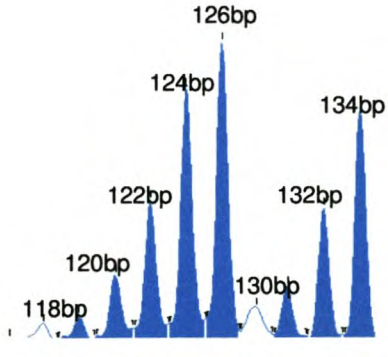


Figure 3.10 : Electropherogram showing $n-x$ peaks (118bp – 124bp and 130bp – 132bp).

Allele = 126bp
$n-2$ = 124bp
$n-4$ = 122bp
$n-6$ = 120bp
$n-8$ = 118bp

Allele = 134bp
$n-2$ = 132bp
$n-4$ = 130bp

Stellenbosch University http://scholar.sun.ac.za

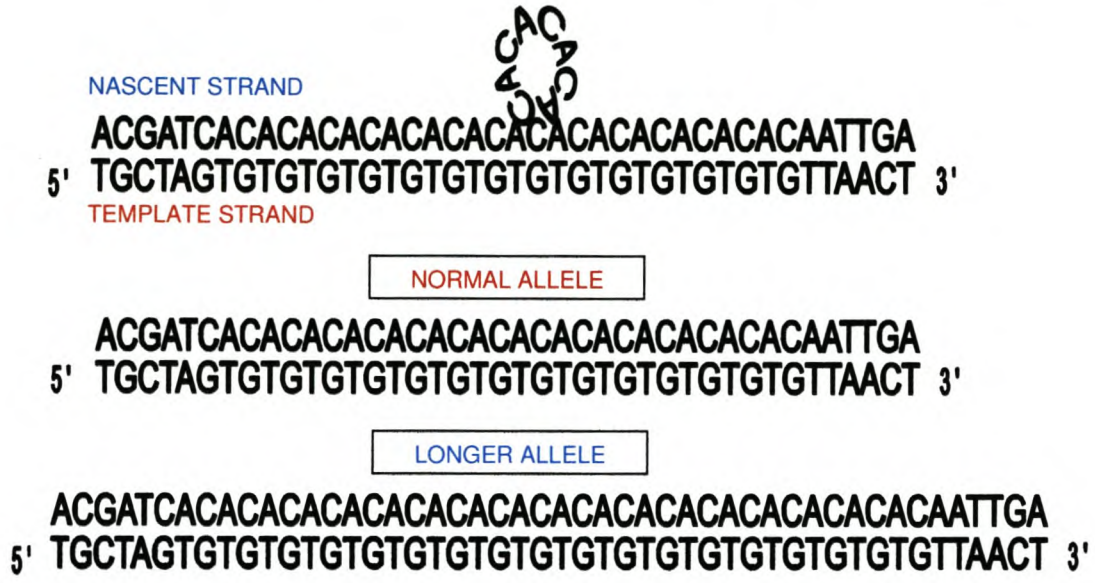


Figure 3.11a: The $n+x$ peaks (longer allele) are caused by the occurrence of loop structures in the nascent DNA strand.

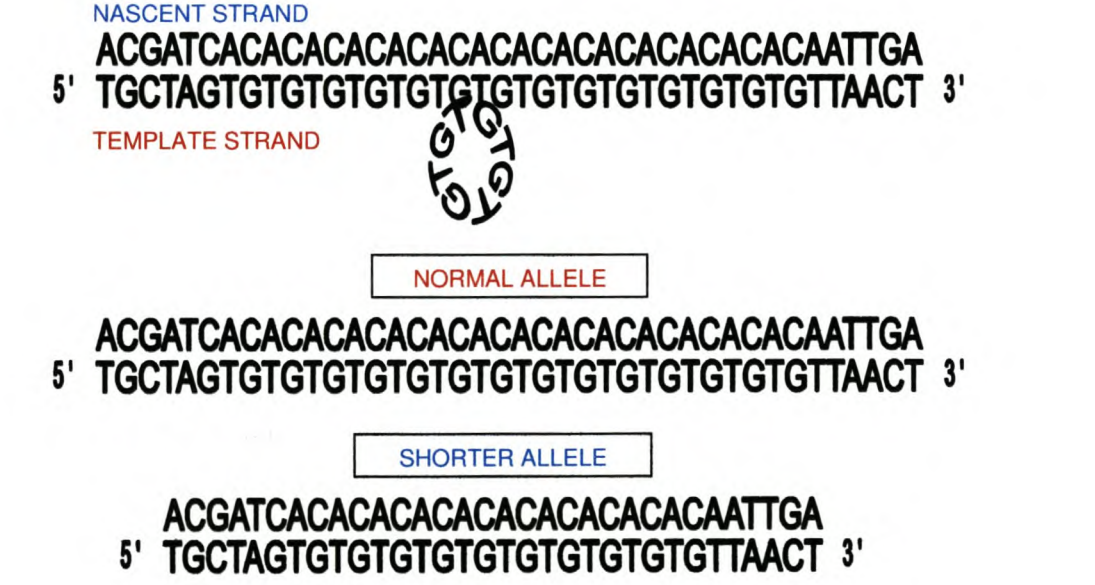


Figure 3.11b: The $n-x$ peaks (shorter allele) are caused by the occurrence of loop structures in the template DNA strand.

Stellenbosch University http://scholar.sun.ac.za

3.6) MICROSATELLITES TO BE USED FOR GENETIC DIVERSITY, POPULATION STRUCTURE, PARENTAGE AND QTL ANALYSIS

The labelled microsatellite DNA markers used for further analysis in Chapters 4 – 6 are shown in Table 3.7.

Table 3.7: Labelled primers used in further analysis in Chapters 4 – 6.

Locus	Repeat Sequence	Primer (5'-3')	Acc Nr	Label
HmSP1	$(CA)_{10}CGCA(CA)_4$	F-ATAGTGGTCATACAGTCATCACCT R-TAGGCATGTTTGAGTTCGTGT	AY303346	PET
HmSP5	$(AC)_{13}$	F-TTCGGCAAGTGAATGTCTAG R-ATGCGACACTTACTACACCG	AY303344	FAM
HmD55	$(GTGA)_{12}$	F-ATCAAGATAAAACGAGGCG R-ACCACTGTGAAAACGTCCA	AY303337	VIC
HmD59	$(CA)_{15}$	F-TATACTGCCATTTCCGTCTG R-TCTGTATTCTGGTCCTGTCG	AY303338	FAM
HmD11	$(TCTG)_8$	F-AGCTCAGAAAAGTGGTGTACG R-TTACCTAGCTAAAGTTGACAACG	AY303341	VIC
HmD30	$(AGTC)_2GGTC(AGTC)_{11}$	F-TGATGTTGCTGGAATATTGC R-CAATTTCATTTTCAACAGTTCA	AY303342	VIC

Stellenbosch University http://scholar.sun.ac.za

3.7) REFERENCES

Estoup, A., Tailliez, C., Cornuet, J., Solignac, M., 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Molecular and Biological Evolution* **12**, 1074 – 1084.

Gagneux, P., Boesch, C., Woodruff, D.S., 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* **6**, 861 – 868.

Hancock, J.M., 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In Goldstein, D.B., Schlötterer, C. *Microsatellites, Evolution and Applications*. Oxford: Oxford University Press, 1 – 9.

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1 – 16.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 4 – GENETIC DIVERSITY AND DIFFERENTIATION

4.1) ALLELE FREQUENCY

Allele frequencies for each locus over all the pooled and subdivided populations are shown in Tables 4.1 – 4.6.

The allele frequencies of the brood stock were representative of the wild populations with a few exceptions, for example allele 208 of locus HmSP1 (Table 4.2), where the allele had a frequency of more than 0.2 in the three brood populations, but only a frequency of 0.0476 in the Black Rock population. Alleles occurring at very low frequencies were also present and are known as rare alleles.

Private alleles (alleles that occur in only one population) were observed in the natural population {for example: alleles 258 – 276 of locus HmSP1 (Table 4.2) in the Black Rock population}, in the brood populations {for example: allele 217 of locus HmD55 (Table 4.1) in the Brood2 population} and in the commercial populations {for example: allele 231 of locus HmD55 (Table 4.1) in the Prog3 and Prog4 populations}. The observation of the private alleles could be due to typing errors or to small sample sizes where allelic richness were under represented.

Stellenbosch University http://scholar.sun.ac.za

Table 4.1: Allele frequencies of locus HmD55.

Locus	AllBrood	Brood2	Brood3	Brood4	AllProg	Prog2	Prog3	Prog4	AllWild	Saldanha	Black Rock
HmD55											
179	0.0432	0.0571	0.0455	0.0208	0.0158	0.0167	0.0313	—	0.0208	0.0385	—
183	0.1173	0.1286	0.0682	0.1458	0.1632	0.0667	0.1406	0.2727	0.0938	0.0769	0.1136
185	—	—	—	—	0.0053	—	—	0.0152	—	—	—
187	0.2407	0.1857	0.2955	0.2708	0.1526	0.2500	0.1563	0.0606	0.1146	0.1731	0.0455
191	0.1420	0.1143	0.1591	0.1667	0.1789	—	0.3281	0.1970	0.2708	0.2115	0.3409
193	—	—	—	—	—	—	—	—	0.0104	0.0192	—
195	0.1914	0.2143	0.1136	0.2292	0.1842	0.3333	0.0625	0.1667	0.1771	0.1346	0.2273
197	0.0062	—	—	0.0208	—	—	—	—	—	—	—
199	0.0617	0.0571	0.0682	0.0625	0.1053	0.0167	0.1094	0.1818	0.1250	0.1731	0.0682
201	0.0864	0.0857	0.1818	—	0.0526	—	0.0781	0.0758	0.0313	—	0.0682
203	0.0309	0.0286	0.0227	0.0417	0.0579	0.1333	0.0469	—	0.0625	0.0577	0.0682
205	0.0062	0.0143	—	—	—	—	—	—	—	—	—
207	0.0432	0.0571	0.0227	0.0417	—	—	—	—	0.0208	0.0192	0.0227
209	—	—	—	—	—	—	—	—	0.0104	0.0192	—
211	0.0062	0.0143	—	—	0.0053	—	0.0156	—	0.0417	0.0385	0.0455
215	—	—	—	—	—	—	—	—	0.0104	0.0192	—
217	0.0062	0.0143	—	—	—	—	—	—	—	—	—
219	0.0185	0.0286	0.0227	—	0.0579	0.1833	—	—	0.0104	0.0192	—
231	—	—	—	—	0.0211	—	0.0313	0.0303	—	—	—

Stellenbosch University http://scholar.sun.ac.za

Table 4.2: Allele frequencies of locus HmSP1.

Locus	AllBrood	Brood2	Brood3	Brood4	AllProg	Prog2	Prog3	Prog4	AllWild	Saldanha	Black Rock
HmSP1											
192	0.0347	0.0455	—	0.0476	0.0188	—	—	0.0577	0.0233	—	0.0476
194	0.0069	—	—	0.0238	0.0437	—	—	0.1346	—	—	—
196	0.0139	0.0303	—	—	0.0125	—	—	0.0385	0.0349	0.0682	—
198	0.1042	0.1364	0.0556	0.0952	0.1187	0.3103	0.0200	—	0.0349	0.0682	—
200	0.0417	0.0606	0.0556	—	0.0125	—	0.0400	—	0.0233	—	0.0476
202	0.0139	0.0303	—	—	0.0063	—	—	0.0192	0.0465	0.0227	0.0714
204	0.0694	0.0455	0.0278	0.1429	0.0375	0.0172	0.0600	0.0385	0.0930	0.0682	0.1190
206	0.0764	0.0606	0.1389	0.0476	0.1063	0.1379	0.1200	0.0577	0.0465	0.0455	0.0476
208	0.2292	0.2273	0.2500	0.2143	0.2438	0.2069	0.2200	0.3077	0.1163	0.1818	0.0476
210	0.0486	0.0455	0.0278	0.0714	0.0437	—	0.0400	0.0962	0.0930	0.1364	0.0476
212	0.0972	0.0758	0.0278	0.1905	0.0250	—	0.0200	0.0577	0.1047	0.0682	0.1429
214	0.0486	0.0455	0.1111	—	0.0125	—	0.0400	—	0.0233	0.0227	0.0238
216	0.0694	0.0909	0.0278	0.0714	0.0938	0.1724	0.0800	0.0192	0.1047	0.1591	0.0476
218	0.0139	0.0303	—	—	0.0250	—	0.0400	0.0385	0.0116	—	0.0238
220	0.0069	—	0.0278	—	0.0812	—	0.1600	0.0962	—	—	—
222	0.0278	0.0152	0.0278	0.0476	0.0188	—	0.0600	—	0.0233	—	0.0476
224	0.0208	0.0152	0.0278	0.0238	—	—	—	—	0.0116	—	0.0238
226	0.0069	—	—	0.0238	0.0188	—	0.0200	0.0385	—	—	—
228	—	—	—	—	0.0063	—	0.0200	—	0.0116	0.0227	—
232	0.0069	0.0152	—	—	—	—	—	—	—	—	—
236	—	—	—	—	—	—	—	—	0.0116	0.0227	—
238	—	—	—	—	—	—	—	—	0.0116	0.0227	—
240	0.0486	0.0303	0.1389	—	0.0750	0.1552	0.0600	—	0.0465	0.0682	0.0238
242	0.0139	—	0.0556	—	—	—	—	—	—	—	—
244	—	—	—	—	—	—	—	—	0.0116	—	0.0238
246	—	—	—	—	—	—	—	—	0.0116	0.0227	—
258	—	—	—	—	—	—	—	—	0.0233	—	0.0476
264	—	—	—	—	—	—	—	—	0.0116	—	0.0238
268	—	—	—	—	—	—	—	—	0.0233	—	0.0476
272	—	—	—	—	—	—	—	—	0.0233	—	0.0476
276	—	—	—	—	—	—	—	—	0.0233	—	0.0476

Stellenbosch University http://scholar.sun.ac.za

Table 4.3: Allele frequencies of locus HmSP5.

Locus	AllBrood	Brood2	Brood3	Brood4	AllProg	Prog2	Prog3	Prog4	AllWild	Saldanha	Black Rock
HmSP5											
179	—	—	—	—	0.0052	—	—	0.0147	—	—	—
181	—	—	—	—	0.0052	—	—	0.0147	—	—	—
183	0.0244	0.0139	0.0625	—	0.0361	0.0833	—	0.0294	—	—	—
185	0.6463	0.7222	0.6458	0.5227	0.7938	0.8500	0.8636	0.6765	0.6250	0.7500	0.4792
187	0.0427	0.0417	0.0417	0.0455	0.0052	—	—	0.0147	0.0096	—	0.0208
189	0.0671	0.0694	0.0625	0.0682	0.0309	0.0667	0.0152	0.0147	0.0577	0.0179	0.1042
191	0.0061	—	—	0.0227	—	—	—	—	0.0096	0.0179	—
193	0.0305	0.0417	—	0.0455	—	—	—	—	0.0481	0.0714	0.0208
195	0.0122	—	0.0208	0.0227	0.0103	—	0.0303	—	0.0673	0.0714	0.0625
197	0.0122	0.0139	—	0.0227	—	—	—	—	—	—	—
199	0.0366	0.0139	0.0208	0.0909	—	—	—	—	0.0385	—	0.0833
201	0.0061	0.0139	—	—	0.0052	—	—	0.0147	0.0385	0.0179	0.0625
203	0.0122	0.0139	—	0.0227	0.0155	—	0.0152	0.0294	0.0096	—	0.0208
205	0.0183	0.0278	—	0.0227	0.0155	—	0.0303	0.0147	0.0385	0.0536	0.0208
207	0.0305	—	0.1042	—	0.0052	—	0.0152	—	—	—	—
209	—	—	—	—	—	—	—	—	0.0096	—	0.0208
211	0.0305	0.0278	0.0417	0.0227	0.0412	—	0.0152	0.1029	0.0096	—	0.0208
213	—	—	—	—	—	—	—	—	0.0096	—	0.0208
215	0.0122	—	—	0.0455	0.0155	—	—	0.0441	—	—	—
217	0.0061	—	—	0.0227	0.0103	—	—	0.0294	0.0192	—	0.0417
219	—	—	—	—	—	—	—	—	0.0096	—	0.0208
221	0.0061	—	—	0.0227	0.0052	—	0.0152	—	—	—	—

Table 4.4: Allele frequencies of locus HmD59.

Locus	AllBrood	Brood2	Brood3	Brood4	AllProg	Prog2	Prog3	Prog4	AllWild	Saldanha	Black Rock
HmD59											
106	0.0172	0.0139	0.0208	0.0185	0.0567	0.1500	0.0303	—	0.0204	0.0192	0.0217
108	0.0172	0.0278	0.0208	—	—	—	—	—	0.0102	—	0.0217
110	0.0460	0.0556	0.0417	0.0370	0.1237	0.2167	0.0455	0.1176	0.0204	0.0192	0.0217
112	0.0690	0.0556	0.0625	0.0926	0.0722	—	0.0606	0.1471	0.0714	0.0769	0.0652
114	0.0575	0.0417	0.0833	0.0556	0.0464	—	0.1061	0.0294	0.0510	0.0192	0.0870
116	0.0632	0.0417	0.0833	0.0741	0.0722	—	0.1061	0.1029	0.0306	—	0.0652
118	0.0287	0.0417	0.0208	0.0185	0.0464	0.0167	0.0152	0.1029	0.0510	0.0962	—
120	0.1149	0.1111	0.1042	0.1296	0.0464	0.0500	0.0455	0.0441	0.0918	0.1538	0.0217
122	0.1322	0.1111	0.1667	0.1296	0.1598	0.0500	0.1970	0.2206	0.2041	0.1154	0.3043
124	0.1034	0.0972	0.0833	0.1296	0.0515	0.0333	0.0455	0.0735	0.1939	0.1923	0.1957
126	0.0575	0.0556	0.0417	0.0741	0.1082	0.2167	0.1061	0.0147	0.0204	—	0.0435
128	0.0977	0.1250	0.0833	0.0741	0.0773	0.1500	0.0606	0.0294	0.0714	0.0962	0.0435
130	0.0575	0.0694	0.0417	0.0556	0.0412	0.1000	0.0303	—	0.0714	0.0962	0.0435
132	0.0172	0.0139	0.0208	0.0185	0.0052	0.0167	—	—	—	—	—
134	0.0345	0.0278	0.0625	0.0185	0.0258	—	0.0606	0.0147	—	—	—
136	0.0287	0.0278	—	0.0556	0.0155	—	0.0152	0.0294	0.0408	0.0577	0.0217
138	—	—	—	—	—	—	—	—	0.0102	0.0192	—
140	0.0115	0.0139	0.0208	—	0.0052	—	0.0152	—	0.0204	0.0385	—
144	0.0287	0.0556	0.0208	—	0.0052	—	0.0152	—	0.0102	—	0.0217
146	0.0057	0.0139	—	—	—	—	—	—	—	—	—
148	0.0057	—	—	0.0185	0.0412	—	0.0455	0.0735	—	—	—
150	0.0057	—	0.0208	—	—	—	—	—	0.0102	—	0.0217

Stellenbosch University http://scholar.sun.ac.za

Table 4.5: Allele frequencies of locus HmD11.

Locus	AllBrood	Brood2	Brood3	Brood4	AllProg	Prog2	Prog3	Prog4	AllWild	Saldanha	Black Rock
HmD11											
288	—	—	—	—	0.0053	—	—	0.0147	0.0278	0.0469	—
292	0.0120	0.0143	—	0.0200	0.0368	0.0517	0.0313	0.0294	0.0463	—	0.1136
296	0.0060	—	—	0.0200	—	—	—	—	0.0093	0.0156	—
300	—	—	—	—	—	—	—	—	0.0185	0.0313	—
304	0.3614	0.3714	0.3478	0.3600	0.4053	0.4138	0.3750	0.4265	0.4259	0.4688	0.3636
308	0.0241	—	0.0870	—	0.0737	—	0.1406	0.0735	0.0278	0.0469	—
312	0.5422	0.5571	0.4783	0.5800	0.4684	0.5345	0.4219	0.4559	0.4074	0.3906	0.4318
316	0.0482	0.0571	0.0652	0.0200	0.0105	—	0.0313	—	0.0278	—	0.0682
320	0.0060	—	0.0217	—	—	—	—	—	—	—	—
352	—	—	—	—	—	—	—	—	0.0093	—	0.0227

Table 4.6: Allele frequencies of locus HmD30.

Locus	AllBrood	Brood2	Brood3	Brood4	AllProg	Prog2	Prog3	Prog4	AllWild	Saldanha	Black Rock
HmD30											
124	0.0139	—	—	0.0455	—	—	—	—	0.0408	0.0192	0.0652
126	0.0069	0.0156	—	—	—	—	—	—	—	—	—
128	0.0208	0.0313	—	0.0227	0.0309	0.1136	—	—	0.0102	—	0.0217
132	0.1458	0.1250	0.1389	0.1818	0.1667	0.1591	0.2308	0.1212	0.1122	0.0962	0.1304
136	0.0069	0.0156	—	—	0.0309	—	0.0577	0.0303	0.0306	0.0192	0.0435
138	0.4375	0.4688	0.4167	0.4091	0.4198	0.5682	0.3846	0.3485	0.4796	0.5577	0.3913
140	0.0486	0.0469	0.0278	0.0682	0.0494	—	—	0.1212	0.0714	0.0769	0.0652
142	0.0556	0.0781	0.0278	0.0455	0.0802	—	0.0385	0.1667	0.0408	0.0192	0.0652
144	0.1458	0.1250	0.1944	0.1364	0.1173	0.0909	0.1538	0.1061	0.1327	0.2115	0.0435
146	—	—	—	—	0.0556	0.0682	0.0962	0.0152	0.0204	—	0.0435
148	0.0347	0.0469	0.0556	—	0.0062	—	—	0.0152	0.0510	—	0.1087
150	0.0486	0.0469	0.0556	0.0455	—	—	—	—	0.0102	—	0.0217
154	0.0278	—	0.0556	0.0455	—	—	—	—	—	—	—
158	0.0069	—	0.0278	—	0.0432	—	0.0385	0.0758	—	—	—

4.2) NUMBER OF ALLELES AND EFFECTIVE NUMBER OF ALLELES

The number of observed alleles and effective number of alleles for each pooled and individual population are shown in Table 4.7. The averages of observed and effective number of alleles are also shown in Table 4.7. The number of observed alleles shown in Table 4.7 was a further indication of the high level of polymorphism of the microsatellite DNA markers. The locus with

Stellenbosch University http://scholar.sun.ac.za

the highest number of observed alleles in the hatchery and wild populations was HmSP1 with 24 and 19 alleles within the AllBrood and AllProg populations and 26 alleles within the AllWild population. The locus with the least number of observed alleles was HmD11 with 7 and 6 alleles within the AllBrood and AllProg populations and 9 alleles within the AllWild population.

Table 4.7: The number of observed alleles and effective number of alleles for pooled and individual populations.

Locus	Hatchery Populations								Wild Populations		
	AllBrood (88)	Brood2 (37)	Brood3 (24)	Brood4 (27)	AllProg (98)	Prog2 (31)	Prog3 (33)	Prog4(34)	AllWild (56)	Saldanha (32)	Black Rock (24)
HmD11											
N	**7**	4	5	5	**6**	3	5	5	**9**	6	5
A_e	**2.3**	2.2	2.8	2.1	**2.6**	2.2	2.9	2.5	**2.8**	2.6	3.0
HmD30											
N	**13**	10	9	9	**10**	5	7	9	**11**	7	11
A_e	**4.1**	3.8	4.1	4.3	**4.3**	2.7	4.2	5.1	**3.7**	2.7	5.0
HmSP1											
N	**24**	17	14	12	**19**	6	15	13	**26**	15	20
A_e	**9.9**	9.6	7.8	7.7	**8.8**	4.7	8.8	6.8	**15.0**	9.5	14.7
HmSP5											
N	**17**	11	8	14	**15**	3	8	12	**15**	7	14
A_e	**2.3**	1.9	2.3	3.4	**1.6**	1.4	1.3	2.1	**2.5**	1.7	3.8
HmD55											
N	**14**	13	10	9	**12**	7	10	8	**14**	13	9
A_e	**6.9**	7.7	5.8	5.5	**7.3**	4.3	5.6	5.4	**6.7**	7.3	5.0
HmD59											
N	**21**	19	18	16	**18**	10	17	13	**18**	13	15
A_e	**13.1**	13.3	12.1	11.4	**11.7**	6.4	10.6	8.3	**8.9**	8.8	6.4
Average											
N	**16**	12	11	11	**13**	6	10	10	**15**	10	12
A_e	**6.4**	6.4	5.8	5.7	**6.0**	3.6	5.6	5.0	**6.6**	5.4	6.4

N = Number of Observed Alleles
A_e = Effective Number of Alleles

The effective number of alleles was lower than the observed number of alleles for each locus. This indicated an uneven distribution of allele frequencies ($p_1 > p_2 > p_3 > p_4 \mathrm{K}\ etc.$, where p = allele frequency; for example locus HmSP1: in Brood2 N = 17 and A_e = 9.6; Table 4.7) across samples. The effective number of alleles would have been numerically closer to the observed number of alleles if $p_1 \approx p_2 \approx p_3 \approx p_4 \mathrm{K}\ p_i$, (where p = allele

Stellenbosch University http://scholar.sun.ac.za

frequency). The low effective number of alleles could also indicate the presence of rare alleles: the greater the difference between the effective and observed number alleles, the higher the number of rare alleles present.

This study showed no significant loss in the average number of observed alleles from the natural populations to the hatchery populations (Mann-Whitney test, $P > 0.05$, Table 4.8), the only exception was Prog2 (Mann-Whitney test, $P < 0.05$, Table 4.8). In contrast, other farmed aquatic species such as *Haliotis discus hannai* (Li *et al*, 2004), *Haliotis midae* (Evans *et al*, 2004), Japanese flounder, *Paralichthys olivaceus* (Sekino *et al*, 2002) and Atlantic salmon, *Salmo salar* (Norris *et al*, 1999) showed the loss of alleles. No significant loss of alleles was detected between the brood stock populations and the progeny populations (Mann-Whitney test, $P > 0.05$; Table 4.9), with Prog2 once again being the exception (Mann-Whitney test, $P < 0.05$; Table 4.9). During the analyses of the influence of subdivision of the brood stock (Table 4.10), significant allele losses were observed for Brood3, Brood4 and Prog3 at the 10% confidence level and for Prog2 and Prog4 at the 5% confidence level when compared to AllBrood. Brood2 was also very close to the 10% cut-off. The loss of alleles during the sub-division of the brood stock is also shown in Figure 4.1. The decrease in the average number of observed alleles from AllBrood to Brood2 (25%), Brood3 (31.3%) and Brood4 (31.3%) and from AllProg to Prog2 (62.5%), Prog3 (37.5%) and Prog4 (37.5%) is clear.

Table 4.8: Comparison of average number of alleles between hatchery populations and wild populations using the Mann-Whitney-test.

Comparison	*U*-value (Mann-Whitney Test)	*P*-Value
AllBrood x AllWild	39.5	0.532
AllProg x AllWild	37	0.374
Brood2 x Saldanha Bay	43	0.739
Brood2 x Black Rock	38.5	0.468
Brood3 x Saldanha Bay	41	0.626
Brood3 x Black Rock	34.5	0.236
Brood4 x Saldanha Bay	41	0.626
Brood4 x Black Rock	35	0.261
Prog2 x Saldanha Bay	26.5	**
Prog2 x Black Rock	25.5	**
Prog3 x Saldanha Bay	40.5	0.595
Prog3 x Black Rock	35	0.261
Prog4 x Saldanha Bay	38	0.436
Prog4 x Black Rock	33	0.168

** = $P < 0.05$

Stellenbosch University http://scholar.sun.ac.za

Table 4.9: Comparison of average number of alleles between hatchery populations using the Mann-Whitney-test.

Comparison	*U*-value (Mann-Whitney Test)	*P*-Value
AllBrood x AllProg	34	0.212
Brood2 x Prog2	25.5	**
Brood3 x Prog3	40.5	0.595
Brood4 x Prog4	42	0.685

** = $P < 0.05$

Table 4.10: Comparison of average number of alleles between pooled and subdivided hatchery populations using the Mann-Whitney-test.

Comparison	*U*-value (Mann-Whitney Test)	*P*-Value
AllBrood x Brood2	32	0.131
AllBrood x Brood3	30.5	*
AllBrood x Brood4	29.5	*
AllBrood x Prog2	22.5	**
AllBrood x Prog3	30	*
AllBrood x Prog4	27	**
AllProg x Prog2	24	**
AllProg x Prog3	32	0.131
AllProg x Prog4	31.5	0.115

* = $P < 0.1$
** = $P < 0.05$

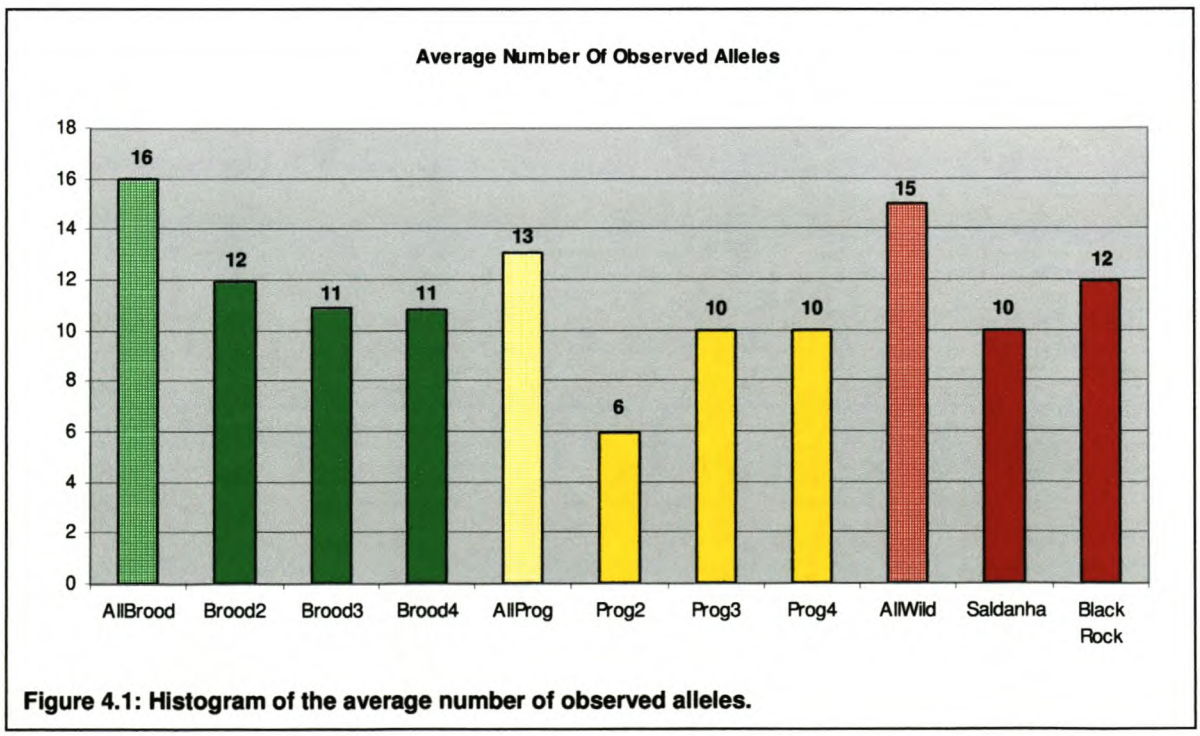


Figure 4.1: Histogram of the average number of observed alleles.

Stellenbosch University http://scholar.sun.ac.za

The loss of the alleles was a direct result of the division of the original brood stock into sub-populations (Lacy & Lindenmayer, 1995). The differential contributions of parents might also cause the loss of genetic diversity within the progeny (Hedgecock, 1994; see section 5.4; p124 – 128). The original brood stock was a combined population in itself (see section 2.9.1; p51). All alleles weren't represented in all the populations, as was suggested by the presence of private alleles and the loss of the average number of observed alleles (see section 4.1 & Figure 4.1). Three new artificial populations (brood stocks) were then created from the subdivision of the combined population. As a result, a loss of the average number of observed alleles would be inflated. Figure 4.1 shows the decline in average number of alleles between the commercial populations (yellow) and the brood stock (green).

The rare alleles in a population are the most likely to be lost due to genetic drift. The allele frequency data shown in Tables 4.1 – 4.6 showed the presence of rare alleles. Besides, the large difference between the observed and effective number of alleles also suggested the presence of rare alleles and possibly null alleles as well. From the allele frequency data it was clear that only alleles occurring at very low frequencies ($p < 0.05$, Tables 4.1 – 4.6) were lost. If the F1 were used to replenish the brood stock, genetic diversity would decrease due to the backcrossing of the F1 to the brood stock.

4.3) HARDY-WEINBERG EQUILIBRIUM

The *P*-values for Hardy-Weinberg equilibrium for each locus per population is shown in Table 4.11. The average of the *P*-values was taken over the subdivided populations. The pooled populations showed significant deviation ($P < 0.05$ after sequential Bonferroni correction) from Hardy-Weinberg equilibrium in 14 of the 18 tests. The subdivided populations showed significant deviation ($P < 0.05$ after sequential Bonferroni correction) from Hardy-Weinberg equilibrium in 17 of the 48 tests. The deviations from Hardy-Weinberg could be ascribed to the presence of null alleles (Callen *et al*, 1993), which was confirmed by Brookfield's null allele frequencies (*r*, Brookfield,

Stellenbosch University http://scholar.sun.ac.za

1996; Table 4.16). The null alleles were responsible for heterozygote deficiencies, which would affect the Hardy-Weinberg equilibrium of the populations. Heterozygote deficiencies were confirmed by Selander's heterozygote deficiency index (*D*; Selander, 1970; Table 4.16).

Table 4.11: Probability of Hardy-Weinberg equilibrium for each locus per population.

Locus	Hatchery Populations								Wild Populations			
	AllBrood (88)	Brood2 (37)	Brood3 (24)	Brood4 (27)	AllProg (98)	Prog2 (31)	Prog3 (33)	Prog4 (34)	AllWild (56)	Saldanha (32)	Black Rock (24)	Average (*P*)
HmD11												
P	**	**	**	0.2217	**	0.0235	**	**	**	0.2970	*	**0.0679**
HmD30												
P	*	0.2201	0.0079	0.4311	**	*	*	**	**	**	0.0992	**0.0950**
HmSP1												
P	**	**	0.0054	**	**	**	0.0668	0.0911	**	**	**	**0.0204**
HmSP5												
P	0.0354	0.1095	0.8447	0.0468	0.2069	0.2793	1.0000	0.1442	*	0.0257	0.0406	**0.3114**
HmD55												
P	**	0.0109	0.0311	0.0273	**	0.1822	**	**	**	**	0.3180	**0.0712**
HmD59												
P	0.8282	0.1812	0.3991	0.9886	**	0.3999	0.0549	0.0140	0.1371	0.3734	0.4231	**0.3543**

P = *P*-Value for Hardy-Weinberg Equilibrium
* = $P < 0.05$
** = $P < 0.01$

4.4) HETEROZYGOSITY

The observed and expected heterozygosities per locus per sample, along with averages across loci are shown in Table 4.12. Figure 4.2 show a histogram of the averages of the observed and expected heterozygosity per population over all loci. This study found no significant loss of expected heterozygosity between any of the populations (Mann-Whitney test, $P > 0.05$; Tables 4.13 – 4.15). Figure 4.2 also shows no significant differences between the expected heterozygosities of the pooled or subdivided populations. Similar results had been found in Japanese flounder, *Paralichthys olivaceus* (Sekino *et al*, 2002) and *Haliotis midae* (Evans *et al*, 2004). In other studies a loss of expected heterozygosity was reported for the abalone *Haliotis discus hannai* (Li *et al*, 2004) and Atlantic salmon, *Salmo salar* (Norris *et al*, 1999). This contrast in observations could be explained by the loss of low-frequency alleles that had

Stellenbosch University http://scholar.sun.ac.za

little effect on heterozygosity (Allendorf, 1986) and if a strain was found with heterozygous parents, the heterozygosity would be inflated (Sekino *et al*, 2002).

The average observed heterozygosities (Figure 4.2) followed the same pattern as the average expected heterozygosities. The average observed heterozygosities were lower, which might suggest the presence of inbreeding (see section 4.6.2).

Table 4.12: Observed and expected heterozygosities.

Locus	Hatchery Populations								Wild Populations		
	AllBrood (88)	Brood2 (37)	Brood3 (24)	Brood4 (27)	AllProg (98)	Prog2 (31)	Prog3 (33)	Prog4 (34)	AllWild (56)	Saldanha (32)	Black Rock (24)
HmD11											
H_o	0.2289	0.0857	0.2609	0.4000	0.3053	0.3793	0.2813	0.2647	0.4074	0.4688	0.3182
H_e	0.5723	0.5482	0.6380	0.5328	0.6094	0.5404	0.6597	0.6038	0.6476	0.6221	0.6632
HmD30											
H_o	0.6111	0.6563	0.4444	0.6818	0.6296	0.4091	0.5385	0.8485	0.5306	0.3846	0.6957
H_e	0.7555	0.7349	0.7577	0.7676	0.7665	0.6260	0.7596	0.8030	0.7272	0.6280	0.7987
HmSP1											
H_o	0.5000	0.4848	0.7222	0.3333	0.6000	0.2414	0.7600	0.8462	0.4884	0.5000	0.4762
H_e	0.8995	0.8958	0.8719	0.8696	0.8861	0.7878	0.8864	0.8521	0.9329	0.8946	0.9320
HmSP5											
H_o	0.4878	0.3611	0.5833	0.5909	0.3505	0.2333	0.2727	0.5294	0.4423	0.2857	0.6250
H_e	0.5701	0.4676	0.5599	0.7035	0.3648	0.2661	0.2511	0.5260	0.5938	0.4235	0.7396
HmD55											
H_o	0.7160	0.7429	0.7273	0.6667	0.5789	0.9000	0.4375	0.4242	0.5833	0.5000	0.6818
H_e	0.8551	0.8706	0.8285	0.8168	0.8628	0.7700	0.8218	0.8154	0.8498	0.8632	0.8006
HmD59											
H_o	0.9310	0.8889	0.9167	1.0000	0.8660	0.9000	0.9091	0.7941	0.8367	0.8846	0.7826
H_e	0.9235	0.9248	0.9175	0.9122	0.9143	0.8444	0.9054	0.8789	0.8871	0.8861	0.8440
Average											
H_o	0.5791	0.5366	0.6091	0.6121	0.5551	0.5105	0.5332	0.6179	0.5481	0.5040	0.5966
H_e	0.7627	0.7403	0.7623	0.7671	0.7340	0.6391	0.7140	0.7465	0.7731	0.7196	0.7964

H_o = Observed Heterozygosity
H_e = Expected Heterozygosity

Stellenbosch University http://scholar.sun.ac.za

Table 4.13: Comparison of average expected heterozygosity between hatchery populations and wild populations using the Mann-Whitney-test.

Comparison	*U*-value (Mann-Whitney Test)	*P*-Value
AllBrood x AllWild	38	0.436
AllProg x AllWild	38	0.436
Brood2 x Saldanha Bay	36	0.315
Brood2 x Black Rock	37	0.374
Brood3 x Saldanha Bay	37	0.374
Brood3 x Black Rock	37	0.374
Brood4 x Saldanha Bay	37	0.374
Brood4 x Black Rock	38	0.436
Prog2 x Saldanha Bay	33	0.168
Prog2 x Black Rock	30	0.075
Prog3 x Saldanha Bay	41	0.626
Prog3 x Black Rock	37	0.374
Prog4 x Saldanha Bay	42	0.685
Prog4 x Black Rock	39	0.500

Table 4.14: Comparison of average expected heterozygosity between hatchery populations using the Mann-Whitney-test.

Comparison	*U*-value (Mann-Whitney Test)	*P*-Value
AllBrood x AllProg	39	0.500
Brood2 x Prog2	33	0.168
Brood3 x Prog3	39	0.500
Brood4 x Prog4	37	0.374

Table 4.15: Comparison of average expected heterozygosity between pooled and subdivided hatchery populations using the Mann-Whitney-test.

Comparison	*U*-value (Mann-Whitney Test)	*P*-Value
AllBrood x Brood2	37	0.374
AllBrood x Brood3	38	0.436
AllBrood x Brood4	38	0.436
AllBrood x Prog2	32	0.131
AllBrood x Prog3	38	0.436
AllBrood x Prog4	36	0.315
AllProg x Prog2	33	0.168
AllProg x Prog3	38	0.436
AllProg x Prog4	42	0.685

Stellenbosch University http://scholar.sun.ac.za

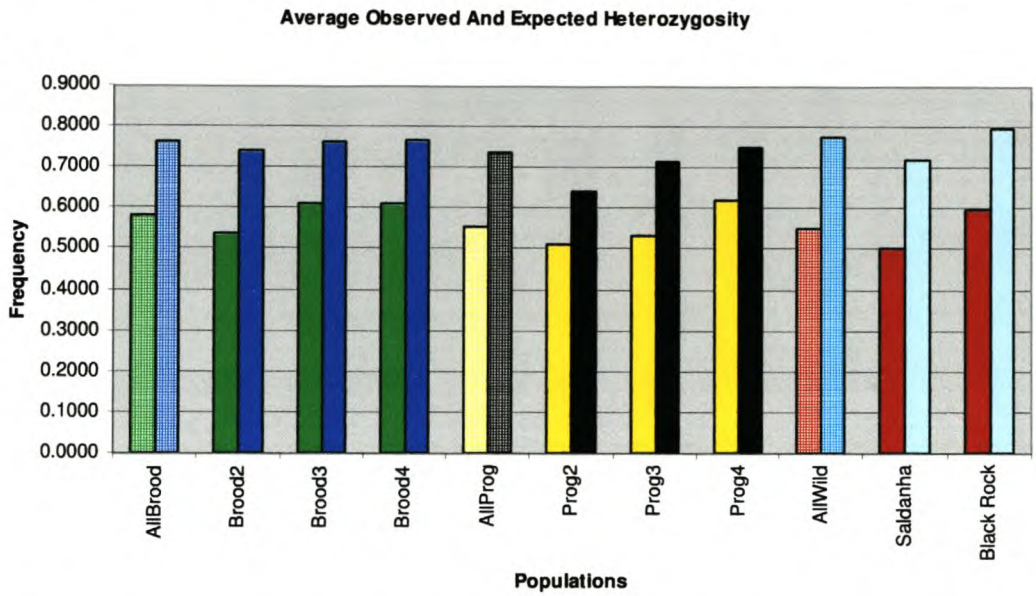


Figure 4.2: Histogram of the observed and expected heterozygosities for each population.

Observed heterozygosity = green, yellow, red
Expected heterozygosity = dark blue, black, light blue

4.5) HETEROZYGOTE DEFICIENCY INDEX AND NULL ALLELE FREQUENCY

Selander's heterozygote deficiency (*D*; Selander, 1970) and Brookfield's null allele frequencies (*r*, Brookfield, 1996) and the average of the null allele frequency for each locus for the eight subdivided populations, are shown in Table 4.16. The negative values of *D* indicated a deficiency of heterozygotes within a population. This deficiency could be ascribed to the presence of null alleles, population subdivision, inbreeding or non-random mating (Brookfield, 1996). In the case of the presence of null alleles, the genotype of a heterozygous individual would be mistakenly considered a homozygote instead. The total number of heterozygotes would thus be underestimated (Pemberton *et al*, 1995). Null alleles could be caused by a mutation within the complementary DNA sequence where the primer should bind, causing the reduction or absence of an allele (Callen *et al*, 1993). This scenario could be remedied by redesigning the effected microsatellite primers (Peatkau &

Stellenbosch University http://scholar.sun.ac.za

Strobeck, 1995). Allele dropout also cause the underscoring of heterozygotes, because the smaller allele was missed due to its low peak intensity.

Table 4.16: Null allele frequencies and heterozygote deficiency indexes.

Locus	Hatchery Populations								Wild Populations			
	AllBrood (88)	Brood2 (37)	Brood3 (24)	Brood4 (27)	AllProg (98)	Prog2 (31)	Prog3 (33)	Prog4 (34)	AllWild (56)	Saldanha (32)	Black Rock (24)	Average (*r*)
HmD11												
r	0.2184	0.2987	0.2302	0.0866	0.1890	0.1046	0.2280	0.2114	0.1458	0.0945	0.2074	**0.1827**
D	-0.6000	-0.8437	-0.5911	-0.2492	-0.4990	-0.2981	-0.5736	-0.5616	-0.3709	-0.2464	-0.5202	
HmD30												
r	0.0823	0.0453	0.1782	0.0485	0.0775	0.1334	0.1257	-0.0252	0.1138	0.1495	0.0573	**0.0891**
D	-0.1911	-0.1070	-0.4135	-0.1118	-0.1786	-0.3465	-0.2911	0.0567	-0.2704	-0.3876	-0.1290	
HmSP1												
r	0.2103	0.2168	0.0800	0.2869	0.1517	0.3056	0.0670	0.0032	0.2300	0.2083	0.2359	**0.1755**
D	-0.4441	-0.4588	-0.1717	-0.6167	-0.3229	-0.6936	-0.1426	-0.0069	-0.4765	-0.4411	-0.4891	
HmSP5												
r	0.0524	0.0726	-0.0150	0.0661	0.0105	0.0259	-0.0173	-0.0022	0.0951	0.0968	0.0659	**0.0366**
D	-0.1444	-0.2278	0.0418	-0.1601	-0.0392	-0.1233	0.0860	0.0065	-0.2551	-0.3254	-0.1549	
HmD55												
r	0.0750	0.0683	0.0553	0.0826	0.1524	-0.0734	0.2109	0.2155	0.1441	0.1949	0.0660	**0.1025**
D	-0.1627	-0.1467	-0.1221	-0.1838	-0.3290	0.1688	-0.4676	-0.4798	-0.3136	-0.4208	-0.1484	
HmD59												
r	-0.0039	0.0187	0.0004	-0.0459	0.0252	-0.0301	-0.0019	0.0451	0.0267	0.0008	0.0333	**0.0025**
D	0.0081	-0.0388	-0.0009	0.0963	-0.0528	0.0658	0.0041	-0.0965	-0.0568	-0.0017	-0.0727	

r = Brookfield's Null Allele Frequency
D = Selander's Heterozygote Deficiency Index

4.6) GENETIC DIFFERENTIATION

4.6.1) F_{st}

The infinite allele model of mutation (IAM) was preferred over the stepwise mutation model (SMM) for the genetic differentiation analysis. The uneven allele frequency distributions observed for some loci would be better described by using this model. The allele frequency distributions are shown in Figures 4.3 – 4.8. The differential allele distributions could be caused by

Stellenbosch University http://scholar.sun.ac.za

typing errors (a possibility in locus HmD55; Figure 4.7) of even the presence of null alleles (a possibility for locus HmD11; Figure 4.3).

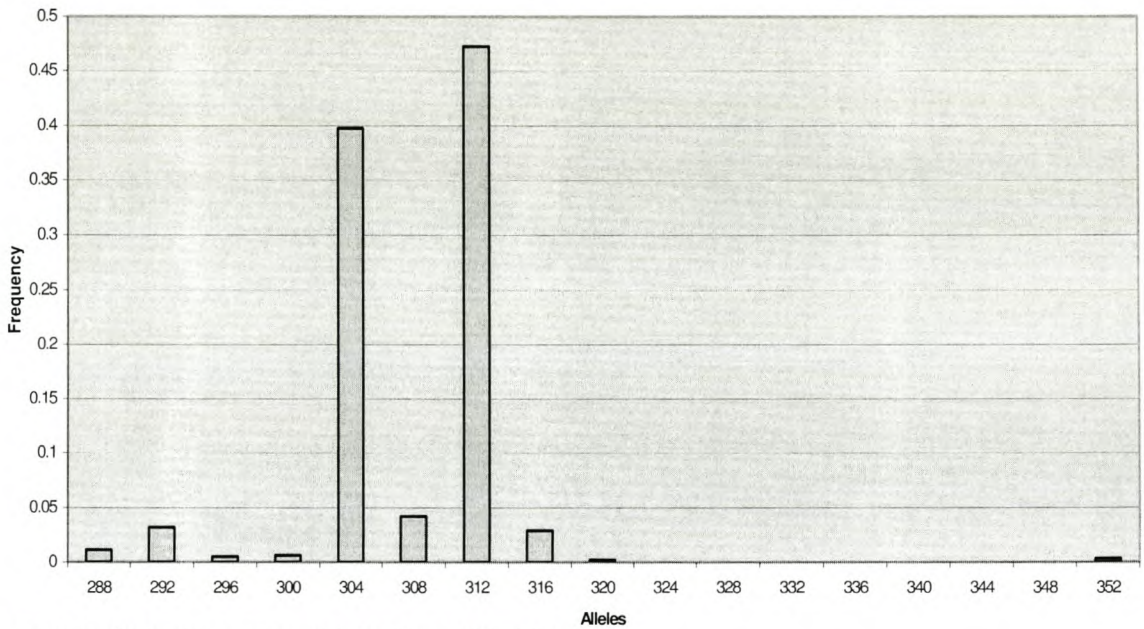


Figure 4.3: Allele frequency distribution of HmD11.

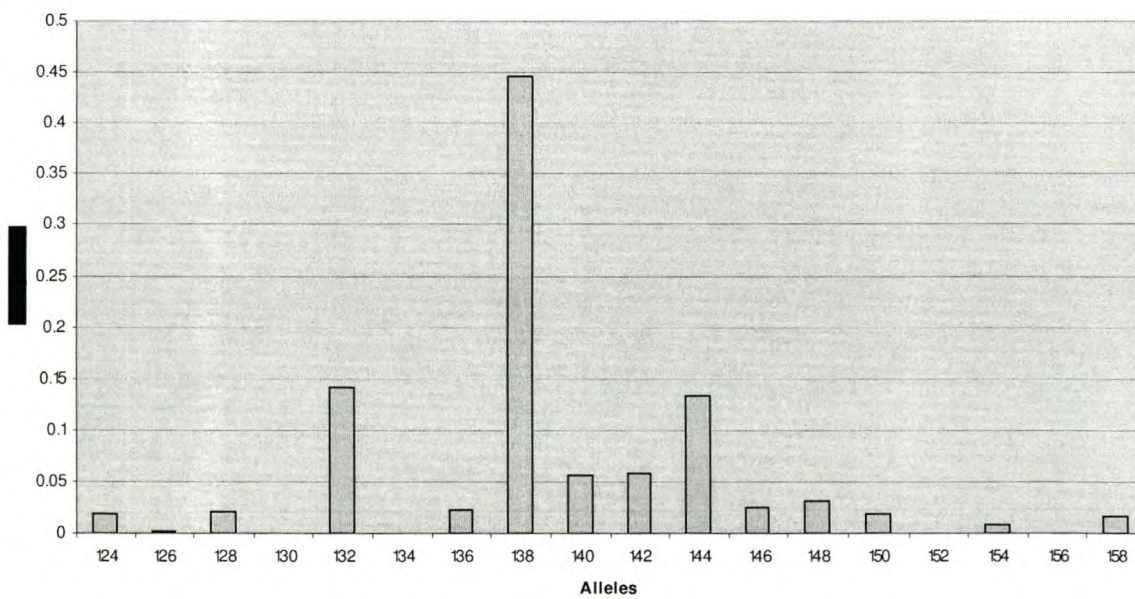


Figure 4.4: Allele frequency distribution of HmD30.

Stellenbosch University http://scholar.sun.ac.za

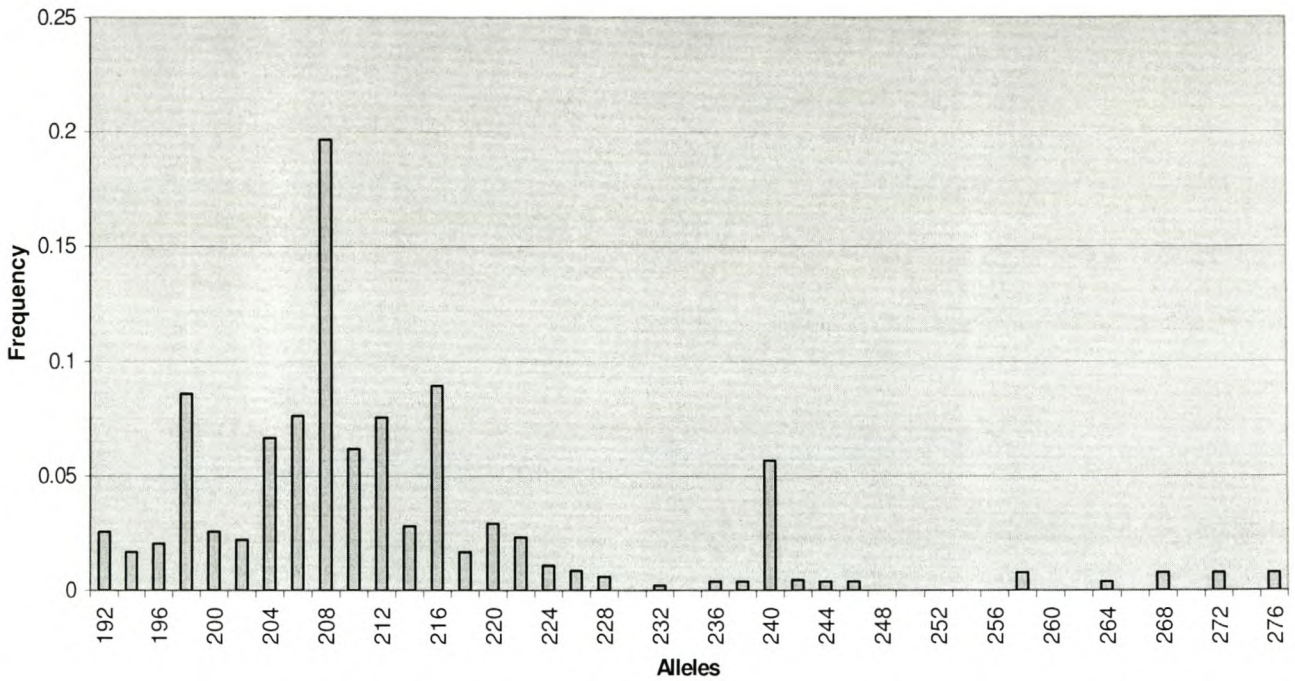


Figure 4.5: Allele frequency distribution of HmSP1.

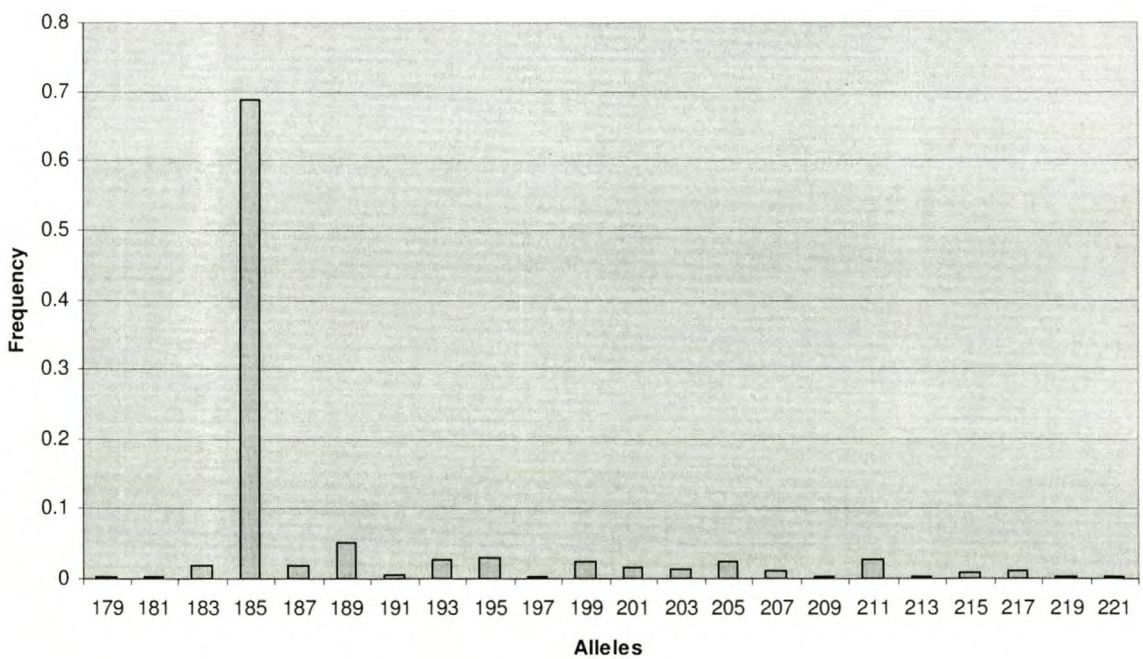


Figure 4.6: Allele frequency distribution of HmSP5.

Stellenbosch University http://scholar.sun.ac.za

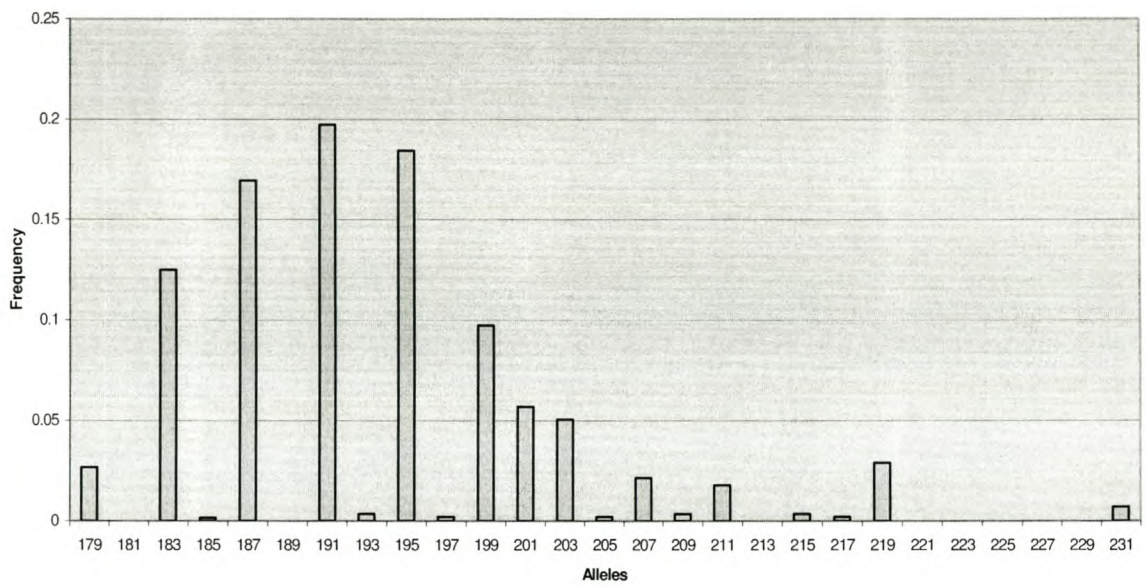


Figure 4.7: Allele frequency distribution of HmD55.

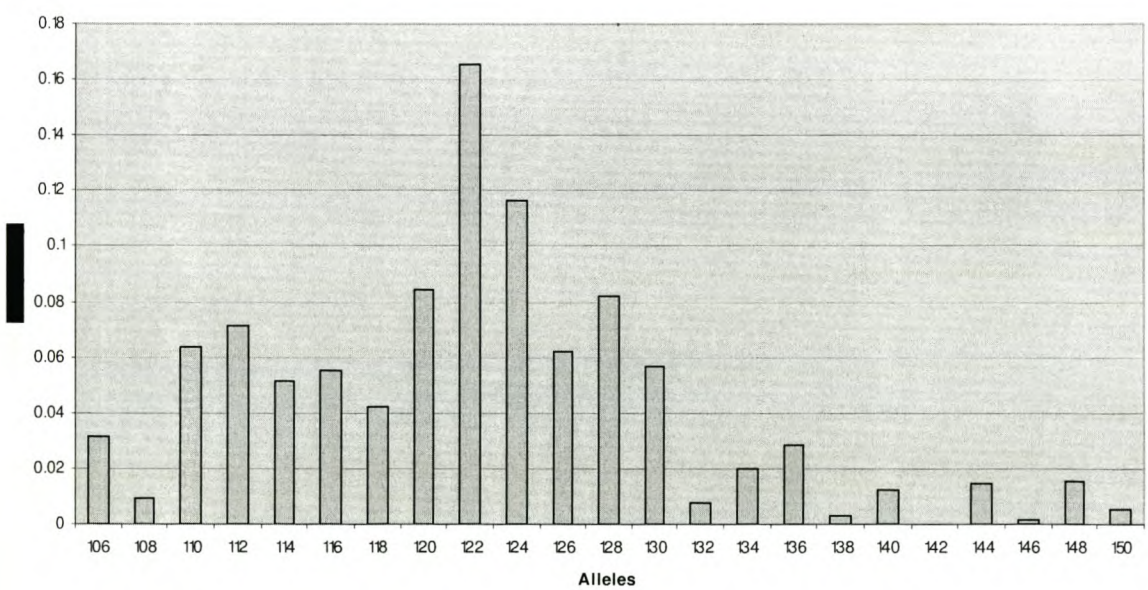


Figure 4.8: Allele frequency distribution of HmD59.

Stellenbosch University http://scholar.sun.ac.za

The F_{st} values for the hypothetical pooled populations are shown in Table 4.17. F_{st} values ranged from 0.004 – 0.011.

Table 4.17: F_{st} values for the hypothetical pooled populations.

	AllBrood	AllProg	AllWild
AllBrood	**0**	0.005	0.004
AllProg		**0**	0.011
AllWild			**0**

There was no differentiation between the AllBrood population and the AllWild population. This was an expected observation, because the brood stock originated directly from the wild. It indicated that a representative brood stock was collected from the wild (by sheer luck). The F_{st} value between the AllBrood and the AllProg populations showed no differentiation. This was once again expected, because the AllProg are the F1 of the AllBrood population. There was no significant differentiation between the AllProg and AllWild populations.

The F_{st} values for the subdivided populations are shown in Table 4.18. F_{st} values ranged from -0.008 – 0.073.

No significant differentiation ($P > 0.1$) was detected between any of the subpopulations, but the only exception was Prog2. Prog2 showed significant differentiation with all the populations except with Brood2 and Brood3. The fact that there was no genetic differentiation detected between the farm and wild populations suggested that the release or even escape of farm animals into the wild would have no negative effect on the natural populations.

Stellenbosch University http://scholar.sun.ac.za

Table 4.18: F_{st} values for the subdivided populations.

	Farm Populations						Wild Populations	
	Brood2	**Brood3**	**Brood4**	**Prog2**	**Prog3**	**Prog4**	**Black Rock**	**Saldanha**
Brood2	**0**	-0.008	-0.007	0.018	0.011	0.012	0.014	-0.000
Brood3		**0**	-0.002	0.035	0.001	0.015	0.014	0.005
Brood4			**0**	**0.039***	0.020	0.013	0.002	0.006
Prog2				**0**	**0.051***	**0.068***	**0.073***	**0.039***
Prog3					**0**	0.011	0.023	0.012
Prog4						**0**	0.019	0.019
Black Rock							**0**	0.020
Saldanha								**0**

* = significant differentiation ($P < 0.05$)

4.6.2) F_{is}

The F_{is} values for the hypothetical pooled populations per locus and over all loci are shown in Table 4.19. Loci HmD11, HmD30, HmSP1 and HmD55 showed significant levels of inbreeding. Locus HmSP5 showed significant inbreeding in only the natural populations, but this could be a sampling error (related animals were sampled by chance). Locus HmD59 showed no significant levels of inbreeding. Significant levels of inbreeding were found in all populations when the cumulative effects of the loci were studied.

Table4.19: F_{is} values per locus for the hypothetical pooled populations.

LOCUS	POPULATIONS		
	AllBrood	**AllProg**	**AllWild**
HmD11	0.60386*	0.50305*	0.37896*
HmD30	0.19784*	0.18457*	0.27986*
HmSP1	0.44973*	0.32849*	0.48556*
HmSP5	0.15039	0.04436	0.26412*
HmD55	0.16861*	0.33372*	0.32305*
HmD59	-0.00237	0.05805	0.06708
All Loci	0.24665*	0.24908*	0.30053*

* = significant for $P < 0.05$

Stellenbosch University http://scholar.sun.ac.za

The F_{is} values for the subdivided populations per locus are shown in Table 4.20. Fifteen out of 48 observations had significant levels of inbreeding and the levels of inbreeding over all loci were significant.

Table 4.20: F_{is} values per locus for the subdivided populations.

Locus	POPULATIONS							
	Brood2	**Brood3**	**Brood4**	**Prog2**	**Prog3**	**Prog4**	**Black Rock**	**Saldanha**
HmD11	0.84776*	0.60538*	0.26829	0.31403	0.58420*	0.57174*	0.53701*	0.26132
HmD30	0.12264	0.43685*	0.13462	0.36683	0.30898	-0.04126	0.15078	0.40405*
HmSP1	0.47080*	0.19928	0.63158*	0.70258*	0.16253	0.02655	0.50739*	0.45965*
HmSP5	0.24103	-0.02060	0.18263	0.13983	-0.07063	0.00835	0.17563	0.34146
HmD55	0.16089	0.14504	0.20432	-0.15232	0.47993*	0.49149*	0.17105	0.43674*
HmD59	0.05285	0.02222	-0.07751	-0.04889	0.01133	0.11122	0.09486	0.02128
All Loci	0.28875*	0.22435*	0.22323*	0.21879*	0.26940*	0.18772*	0.27223*	0.31756*

* = significant for $P < 0.05$

The high levels of inbreeding seen for some loci (for example: 0.84776 for locus HmD11 under Brood2; Table 4.20) indicated that some genetic diversity is lost between the individuals within a population. The expected heterozygosity values observed in section 4.3 might be too high, because the descent of an allele was accounted for when heterozygosities were calculated.

The loss of genetic variation (Figure 4.1) from the brood stock to the commercial stock should be seen in increasing F_{is} values from the brood stock to the commercial stock. Table 4.21 gives a summary of the comparisons in F_{is} values between the corresponding brood stock and commercial stock per locus. Only 6 observations out of 18 showed an increase in the F_{is} from brood to commercial stock (Table 4.21). The reason for the unexpected observations was the higher number of homozygous animals at each locus within the brood stock when compared to the commercial stock. A high number of homozygotes will inflate the F_{is} value of a locus. The number of homozygotes per sample was then compiled in Table 2.21. The only exceptions were locus HmSP5 (in all Rows) and HmD59 (for Row2). The HmSP5 observations could be due to the low frequencies of all but one allele (Figure 4.6; allele 185).

Stellenbosch University http://scholar.sun.ac.za

Table 2.21: Summary of the comparisons in F_{is} values between the corresponding brood stock and commercial stock per locus.

ROW	POPULATION	HmD11		HmD30		HmSP1	
		OBSERVATION	FREQUENCY OF HOMOZYGOTES	OBSERVATION	FREQUENCY OF HOMOZYGOTES	OBSERVATION	FREQUENCY OF HOMOZYGOTES
Row2	Brood2	>	0.914	<	0.344	<	0.515
	Prog2		0.586		0.591		0.724
Row3	Brood3	>	0.773	>	0.625	>	0.294
	Prog3		0.719		0.462		0.240
Row4	Brood4	<	0.600	>	0.350	>	0.737
	Prog4		0.735		0.152		0.160

continued...

ROW	POPULATION	HmSP5		HmD55		HmD59	
		OBSERVATION	FREQUENCY OF HOMOZYGOTES	OBSERVATION	FREQUENCY OF HOMOZYGOTES	OBSERVATION	FREQUENCY OF HOMOZYGOTES
Row2	Brood2	>	0.639	>	0.257	>	0.083
	Prog2		0.767		0.100		0.100
Row3	Brood3	>	0.435	<	0.286	>	0.087
	Prog3		0.750		0.563		0.091
Row4	Brood4	>	0.450	<	0.364	<	0.000
	Prog4		0.471		0.576		0.206

> = F_{is}-brood stock > F_{is}-F1

> = F_{is}-brood stock > F_{is}-F1

Stellenbosch University http://scholar.sun.ac.za

4.7) REFERENCES

Allendorf, F.W., 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology* **5**, 181 – 190.

Brookfield, F.F.Y., 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**, 453 – 455.

Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C., Sutherland, G.R., 1993. Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics* **52**, 922 – 927.

Evans, B., Bartlett, J., Sweijd, N., Cook, P., Elliott, N.G., 2004. Loss of genetic variation at microsatellite loci in hatchery produced abalone in Australia (*Haliotis rubra*) and South Africa (*Haliotis midae*). *Aquaculture* **233**, 109 – 127.

Hedgecock, D., 1994. Does variance in reproductive success limit effective population size of marine organisms? In Beaumont, A.R. *Genetics and Evolution of Aquatic Organisms*. London: Chapman & Hall, 122 – 134.

Lacy, R.C., Lindenmayer, D.B., 1995. A simulation of the impacts of population subdivision on the mountain brushtail possum *Trichosurus caninus* ogilby (Phalangeridae: Marsupialia), in South-Eastern Australia. II. Loss of genetic variation within and between subpopulations. *Biological Conservation* **73**, 131 – 142.

Li, Q., Park, C., Endo, T., Kijima, A., 2004. Loss of genetic variation at microsatellite loci in hatchery strains of the Pacific abalone (*Haliotis discus hannai*). *Aquaculture* **237**, 207 – 222.

Stellenbosch University http://scholar.sun.ac.za

Norris, A.T., Bradley, D.G., Cunningham, E.P., 1999. Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture* **180**, 247 – 264.

Peatkau, D., Strobeck, C., 1995. The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology* **4**, 519 – 520.

Pemberton, J.M., Slate, J., Bancroft, D.R., Barrett, J.A., 1995. Nonamplification alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* **4**, 249 – 252.

Sekino, M., Hara, M., Taniguchi, N., 2002. Loss of microsatellite and mitochondrial DNA variation in hatchery strains of Japanese flounder *Paralichthys olivaceus*. *Aquaculture* **213**, 101 – 122.

Selander, R.K.,1970. Behaviour and genetic variation in natural populations. *American Zoologist* **10**, 53 – 66.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 5 – PARENTAGE ANALYSIS

1. A major factor that influenced the parentage assignment was the incomplete parent data sets. The missing parents are shown in section 2.10.1, Table 2.12.
2. All the samples that were used for parentage analysis was used in Chapter 4 as well. The additional Row 4 offspring was used for the QTL analysis to be discussed in Chapter 6.

5.1) LABELLED MICROSATELLITE PRIMERS USED FOR ANALYSIS

The available labelled microsatellite markers and their respective number of alleles, average expected heterozygosity and null allele frequency, are shown in Table 5.1. Calculations were performed for the pooled data of the parents and F1. Allele frequencies are shown in Appendix C.

Table 5.1: Available labelled microsatellite loci.

Loci Names	Repeat Sequence	Number of Alleles	H_e	r	Exclusion Probability For One Parent	Exclusion Probability For Second Parent
HmD11	$(TCTG)_8$	8	0.6361	0.21	0.22	0.37
HmD30	$(AGTC)_2GGTC(AGTC)_{11}$	14	0.7686	0.08	0.40	0.58
HmSP1	$(CA)_{10}CGCA(CA)_4$	22	0.8939	0.16	0.66	0.79
HmSP5	**$(AC)_{13}$**	**19**	**0.4656**	**0.06**	**0.13**	**0.31**
HmD55	**$(GTGA)_{12}$**	**16**	**0.8327**	**0.09**	**0.51**	**0.68**
HmD59	**$(CA)_{15}$**	**21**	**0.9230**	**0.02**	**0.73**	**0.84**

r = Brookfield's Null Allele Frequency
H_e = expected heterozyosity

Of the six markers only three could be used for parentage analysis. HmD11 was discarded, because of a high null allele frequency (21%). A high null allele frequency is the cause of false homozygotes and lowers the

Stellenbosch University http://scholar.sun.ac.za

exclusionary power of the marker (compare HmSP5 with lower H_e, but with almost identical exclusion probabilities, Table 5.1). HmD30 and HmSP1 were discarded because they selected for composite repeats. Composite repeats can originate homoplasic alleles (not identical by descent), that was the product of independent mutational events and (Estoup *et al*, 1995; see section 3.2.5; p78 – 83)

The 3 remaining microsatellite markers, HmSP5, HmD55 and HmD59 (typed in bold in Table 5.1) were chosen for the parentage analysis. Potential problems with HmSP5 were, (1) lower level of heterozygosity (0.4656), which may indicate too many homozygous parents or offspring. This could have an effect on the accurate estimation of the number of potential parents, given that more false parents could be assigned to an offspring. This problem may occur in HmD55 and HmD59 as well, but the higher level of heterozygosity (0.8327 & 0.9230) rules this out. (2) The allele frequency distribution within the parental animals and offspring (Figure 5.1) indicates one dominating allele. Allele 185 is present in 73% of individual animals and this is responsible for the low level of heterozygosity for HmSP5. The abundance of this allele could also have an effect on the assignment of false parents as true parents (as seen in low exclusion probabilities (13% and 31%; Table 5.1).

A problem with HmD55 is the null allele frequency of 9%, suggesting that false homozygous parents or offspring may be present. This suggested that the null alleles might have a major influence on the analysis, because the null alleles may be responsible for false homozygous parents and offspring.

HmD55 and HmD59 have an even allele frequency distribution within the parental animals and the offspring, with no single allele clearly dominating the others (Figures 5.2 – 5.3).

The potential pitfalls encountered with the individual markers can be overcome by combining the three markers in order to achieve more accurate results. This is suggested by comparing the individual exclusion probabilities to the combined exclusion probabilities of the three markers. The combined

Stellenbosch University http://scholar.sun.ac.za

exclusion probability of HmSP5, HmD55 and HmD59 together was 0.88 for the first parent and 0.97 for the second parent, which is much higher than their individual exclusion probabilities shown in Table 5.1.

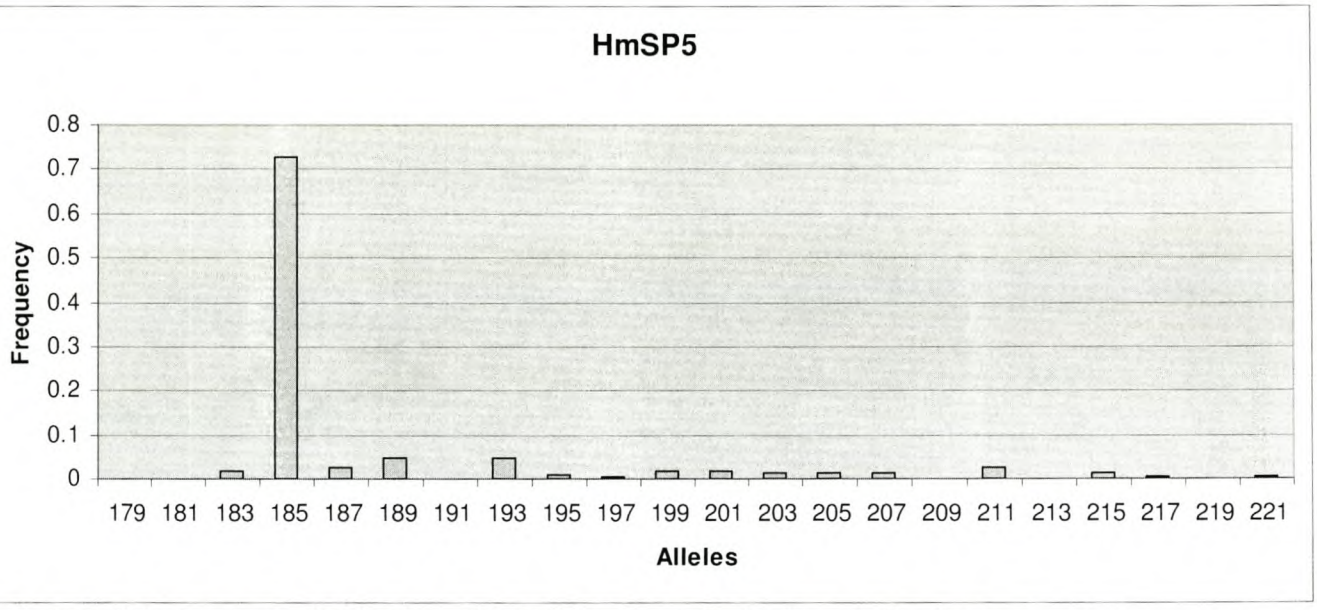


Figure 5.1: Allele frequency distribution of HmSP1 for parental animals and offspring.

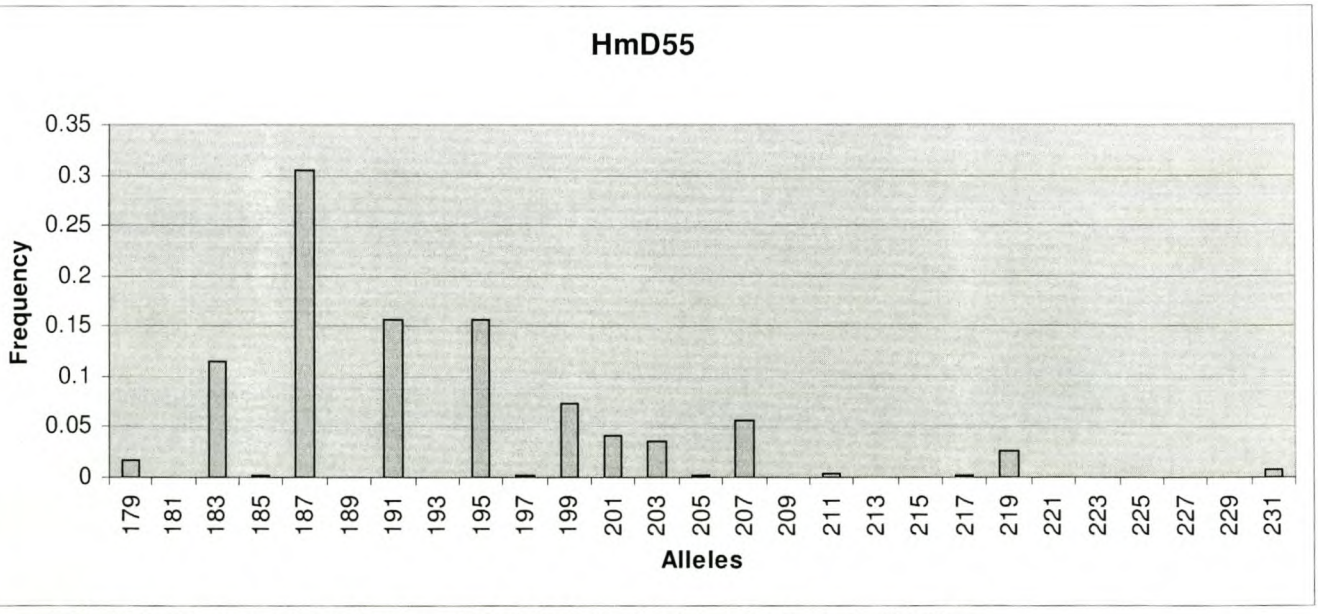


Figure 5.2: Allele frequency distribution of HmD55 for parental animals and offspring.

Stellenbosch University http://scholar.sun.ac.za

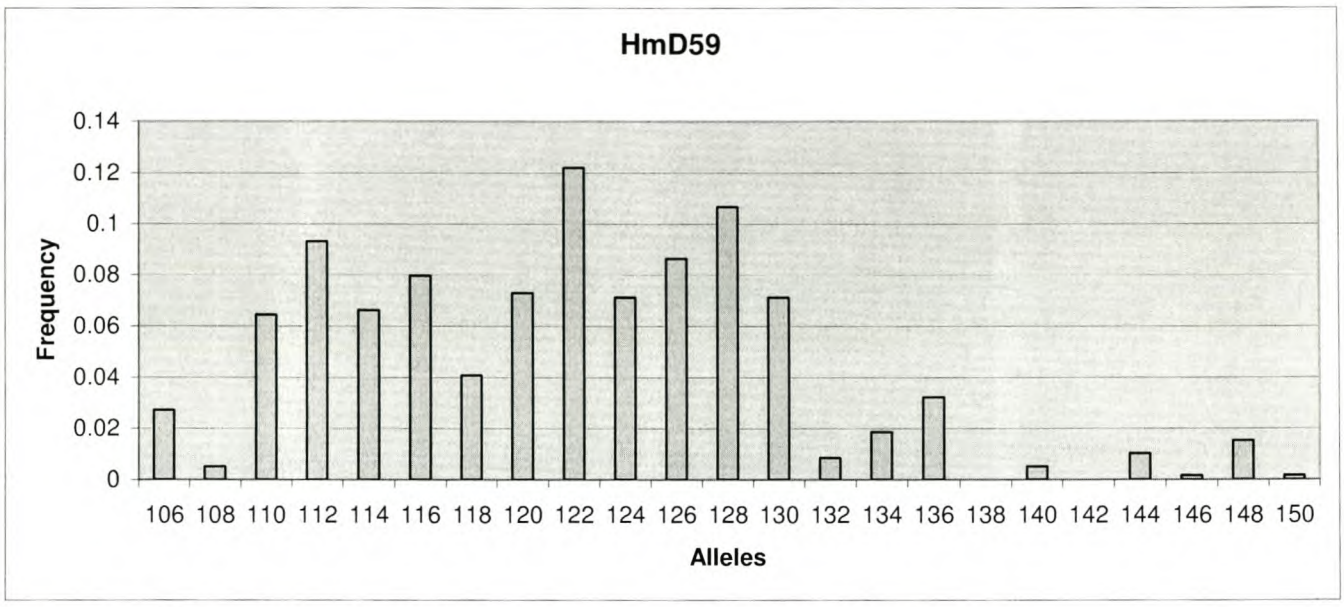


Figure 5.3: Allele frequency distribution of HmD59 for parental animals and offspring.

5.2.1) FACTORS INFLUENCING PARENTAL ASSIGNMENTS

(1) Parental assignments were influenced by incomplete data sets, because samples of all parental animals could not be obtained. This may explain some of the unassigned offspring as well as the substantial number of putative parents observed after the analysis.

(2) Locus HmD59 might have been responsible for some of the mismatches between the offspring and the potential parent, because of the difficulty experienced during genotyping. Many typing errors might occur between alleles 106 and 130 (Figure 5.3).

(3) Another possible reason for the high number of observed putative parents could be the presence of two or three homozygous loci in the same individual. The frequencies of homozygous parents and offspring for two or three loci are shown in Table 5.2. Exclusion power is lowered by the presence of homozygote genotypes. This problem could be overcome by analysing more loci.

Table 5.2: Frequencies of parents and offspring that is homozygous for two or three loci.

ROW 2		ROW 3		ROW 4	
Parents	**Offspring**	**Parents**	**Offspring**	**Parents**	**Offspring**
0.16	0.19	0.21	0.39	0.11	0.32

Stellenbosch University http://scholar.sun.ac.za

5.2.2) ROW 2

5.2.2.1) Assignment of first parent

The frequencies of the offspring for which one, 2 to 5, and more than 5 parents of the same gender were assigned, are shown in Table 5.3. All offspring were assigned a parent. Only 3% of offspring were assigned one father, but none were assigned only one mother. The different reasons for the high frequencies of multiple paternity (≤ 5 = 0.81 and > 5 = 0.16) or maternity (≤ 5 = 0.29 and > 5 = 0.68) assignments were discussed in 5.2.1.

Table 5.3: The frequency of the offspring assigned one, 2 – 5 or more than 5 parents of the same gender.

Offspring assigned to:	Frequency of Offspring Assigned Potential ♂	Frequency of Offspring Assigned Potential ♀
One parent	0.03	0
≤ 5 parents	0.81	0.29
> 5 parents	0.16	0.68
Unassigned	0	0

5.2.2.2) Assignment of second parent

The percentages of offspring assigned both a father and mother is shown in Table 5.4. Complete assignment means that the genotypes of the offspring matched the genotypes of both the father and mother. Putative assignment means that the genotype of one parent corresponds to the genotype of the offspring, while the genotype of the other parent differs at one of the loci. In some instances both parents may differ at one locus from the genotype of the offspring.

In some cases when more than one breeding pair was assigned to an offspring, the pair with the highest LOD-score was taken as the correct one. All offspring were assigned to a breeding pair. About 45.16% of all offspring were confidently assigned to a breeding pair, while the remaining 54.84% had

Stellenbosch University http://scholar.sun.ac.za

one or two putative parents. The reasons for the large number of putative parents were discussed in 5.2.1.

More accurate assignments would have been obtained if all parental animals were sampled (Table 2.12) and more loci were used for the analysis.

Table 5.4: The percentages of offspring assigned both a father and mother.

Assignment	% Offspring assigned
Complete	45.16
Putative	54.84
Unassigned	0

Complete = no loci mismatches
Putative = loci mismatch in one or both parents

5.2.3) ROW 3

5.2.3.1) Assignment of first parent

The frequencies of the offspring for which one, 2 to 5, and more than 5 parents of the same gender were assigned, are shown in Table 5.5. About 9% of offspring were not assigned a father, but all were assigned a mother. About 3% of the offspring were assigned one father and 9% one mother. The different reasons for the high frequencies of multiple paternity (≤ 5 = 0.88) or maternity (≤ 5 = 0.82) assignments were discussed in 5.2.1.

Table 5.5: The frequency of the offspring assigned one, 2 – 5 or more than 5 parents of the same gender.

Offspring assigned to:	Frequency of Offspring Assigned Potential ♂	Frequency of Offspring Assigned Potential ♀
One parent	0.03	0.09
≤ 5 parents	0.88	0.82
> 5 parents	0	0.09
Unassigned	0.09	0

5.2.3.2) Assignment of second parent

The percentages of offspring assigned both a father and mother are shown in Table 5.6. Complete assignment means that the genotypes of the offspring matched the genotypes of both the father and mother. Putative assignment

Stellenbosch University http://scholar.sun.ac.za

means that the genotype of one parent corresponds to the genotype of the offspring, while the genotype of the other parent differs at one of the loci. In some instances both parents may differ at one locus from the genotype of the offspring. Offspring that remained unassigned had putative parents differing at 2 loci with them.

In some cases when more than one breeding pair was assigned to an offspring, the pair with the highest LOD-score was taken to be the correct one. Nine percent of the offspring were not assigned to a breeding pair. One third of all offspring were confidently assigned to a breeding pair, while the remaining 57.58% had one or two putative parents within the breeding pair. The reasons for the large number of putative parents were discussed in 5.2.1.

More accurate assignments would have been obtained if all parental animals were sampled (see Table 2.13; p56) and more loci were used for the analysis.

Table 5.6: The percentages of offspring assigned both a father and mother.

Assignment	% Offspring assigned
Complete	33.33
Putative	57.58
Unassigned	9.09

Complete = no loci mismatches
Putative = loci mismatch in one or both parents

5.2.4) ROW 4

5.2.4.1) Assignment of first parent

The frequencies of the offspring for which one, 2 to 5, and more than 5 parents of the same gender were assigned, are shown in Table 5.7. Only 1% of offspring were not assigned a father, but all were assigned a mother. About 11% of the offspring were assigned one father and 3% one mother. The different reasons for the high frequencies of multiple paternity ($\leq 5 = 0.88$) or maternity ($\leq 5 = 0.76$ and $> 5 = 0.18$) assignments were discussed in 5.2.1.

Stellenbosch University http://scholar.sun.ac.za

Table 5.7: The frequency of the offspring assigned one, 2 – 5 or more than 5 parents of the same gender.

Offspring assigned to:	Frequency of Offspring Assigned Potential ♂	Frequency of Offspring Assigned Potential ♀
One parent	0.11	0.03
≤ 5 parents	0.88	0.77
> 5 parents	0	0.20
Unassigned	0.01	0

5.2.4.2) Assignment of second parent

The percentages of offspring assigned both a father and a mother is shown in Table 5.8. Complete assignment means that the genotypes of the offspring matched the genotypes of both the father and mother. Putative assignment means that the genotype of one parent corresponds to the genotype of the offspring, while the genotype of the other parent differs at one of the loci. In some instances both parents may differ at one locus from the genotype of the offspring. Unassigned offspring was either not assigned a father or mother in the previous analysis or the genotypes of the potential parental animals differ at 2 loci from the genotype of the offspring.

In some cases when more than one breeding pair was assigned to an offspring, the pair with the highest LOD-score was taken to be the correct one. Nearly 6% of the offspring were not assigned to a breeding pair. About 34.44% of all offspring were confidently assigned to a breeding pair, while the remaining 60.26% had one or two putative parents within the breeding pair. The reasons for the large number of putative parents were discussed in 5.2.1.

More accurate assignments would have been obtained if all parental animals were sampled and more loci were used for the analysis.

Table 5.8: The percentages of offspring assigned both a father and mother.

Assignment	% Offspring assigned
Complete	34.44
Putative	60.26
Unassigned	5.30

Complete = no loci mismatches
Putative = loci mismatch in one or both parents

Stellenbosch University http://scholar.sun.ac.za

5.3) LOD-SCORE

The histograms of the pooled frequencies of the observed LOD-scores for each assignment with two putative parents (both parents had one mismatch at a locus), with one putative parent (one parent had one mismatch at a locus) and complete assignments (confirmed assignment; no mismatches) are shown in Figures 5.4 – 5.6. The data was checked for possible trends in the LOD-scores to assess their viability as discriminatory factors for the three loci used, in assigning true breeding pairs to the correct offspring.

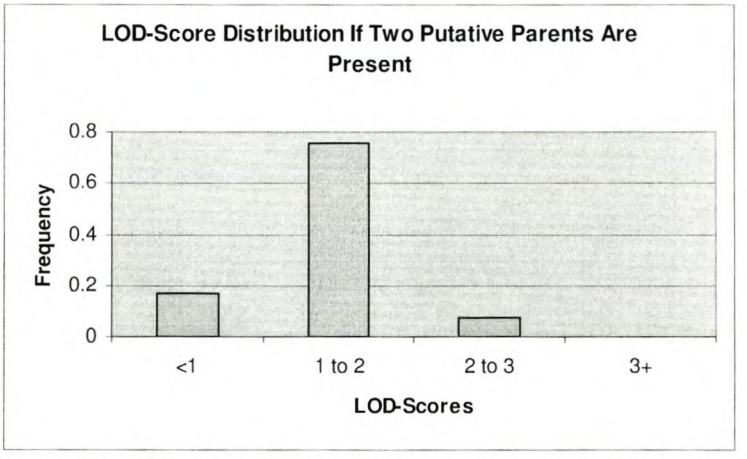


Figure 5.4: Histogram of LOD-scores for two putative parents.

- <1 = assignments with LOD-scores < 1
- 1 to 2 = assignments with 1 ≤ LOD-score < 2
- 2 to 3 = assignments with 2 ≤ LOD-score < 3
- 3+ = assignments with LOD-scores > 3

Stellenbosch University http://scholar.sun.ac.za

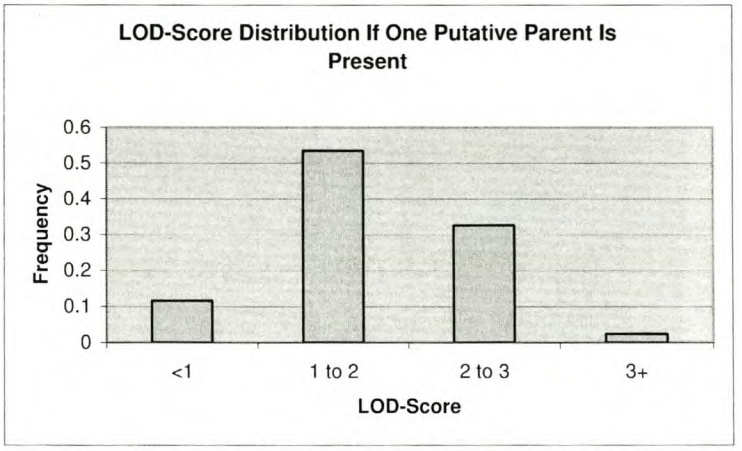


Figure 5.5: Histogram of LOD-scores for one putative parent.

- <1 = assignments with LOD-scores < 1
- 1 to 2 = assignments with 1 ≤ LOD-score < 2
- 2 to 3 = assignments with 2 ≤ LOD-score < 3
- 3+ = assignments with LOD-scores > 3

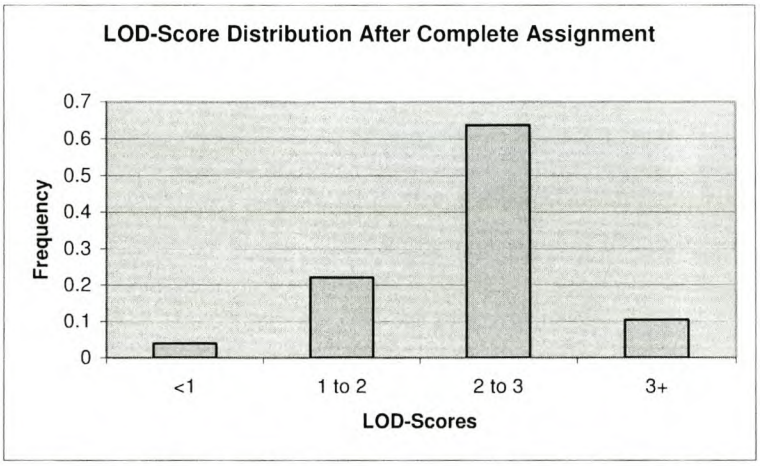


Figure 5.6: Histogram of LOD-scores for complete assignments.

- <1 = assignments with LOD-scores < 1
- 1 to 2 = assignments with 1 ≤ LOD-score < 2
- 2 to 3 = assignments with 2 ≤ LOD-score < 3
- 3+ = assignments with LOD-scores > 3

Slate *et al* (2000) used a LOD-score of 3 and higher to confirm paternity. Figure 5.6 (confirmed assignments) was used to confirm this, but as could be seen, most LOD-scores (90%) were below 3. Just using LOD-scores to assign parents were thus an unreliable factor for in this study.

Stellenbosch University http://scholar.sun.ac.za

The histograms in Figures 5.4 – 5.6 showed that the LOD-scores increased as the number of confidently assigned parents increased. Approximately 93% of the LOD-scores where two putative parents were assigned (Figure 5.4) were below 2. Approximately 65% of the LOD-scores where one putative parent were assigned (Figure 5.5) were below 2, while only 26% of the LOD-scores where parents were confidently assigned (Figure 5.6) were below 2. This increase in LOD-scores was expected – the LOD-score increases as the likelihood of assigning the correct parent increases (Marshall *et al*, 1998).

5.4) DIFFERENTIAL CONTRIBUTION OF PARENTS TO F1

5.4.1) ROW 2

The contributions of the males are shown in Figure 5.7 and the contributions of the females are shown in Figure 5.8.

Males 307 and 324 (Table 5.7) are the main contributors to the offspring, each having fathered 26% of the F1. The contributions of the other active males are close to 10%. There are four main contributors among the females, which are females 316 (19%), 318 (23%), 321 (32%) and N (16%).

The main observation of these results was the variation in the reproductive success of the brood stock animals – only 50% of males and 30% of females were contributing to the F1. This drastic reduction in effective population size might explain the loss of the observed number of alleles (Figure 4.1) and the high level of inbreeding as shown by F_{is} values seen in section 4.6.2; p109 – 110. Genetic variation was lost and inbreeding increased (as predicted, Frankham, 1995). The observation also suggested that the other brood animals in Row 2 is less fit for breeding in captivity, because they may not be able to handle the stresses of the captive environment.

High variation in reproductive success of marine organisms and its consequences on the genetic variation in populations was suggested by

Stellenbosch University http://scholar.sun.ac.za

Hedgecock, (1994) and recently discussed and proven by Arnason (2004, and references therein).

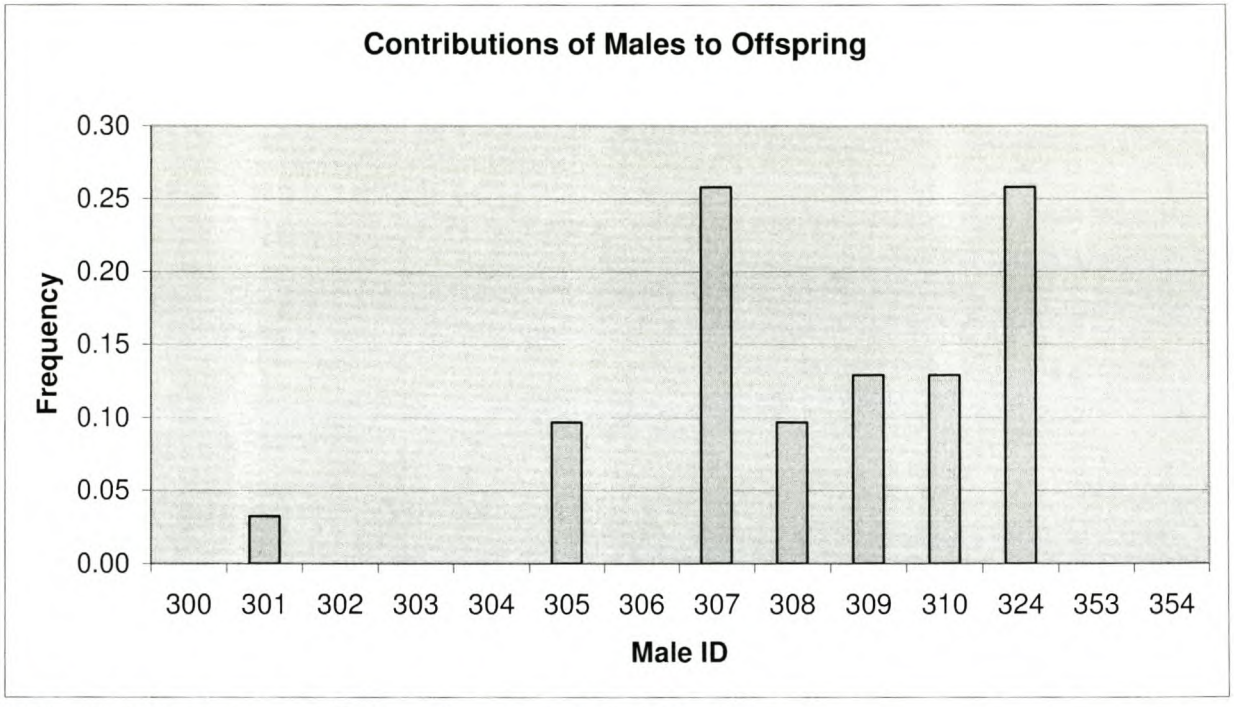


Figure 5.7: Differential contributions of males to Row 2 offspring.

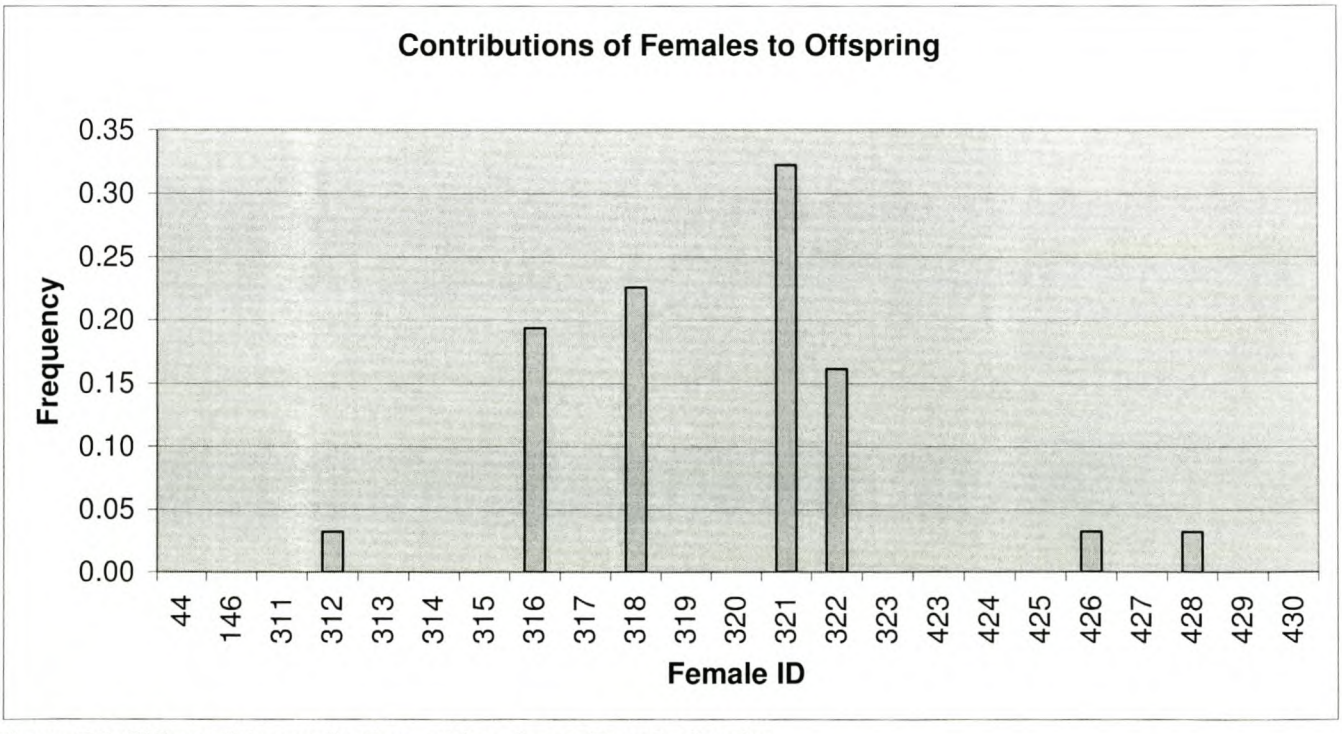


Figure 5.8: Differential contributions of females to Row 2 offspring.

Stellenbosch University http://scholar.sun.ac.za

5.4.2) ROW 3

The contributions of the males are shown in Figure 5.9 and the contributions of the females are shown in Figure 5.10.

The contributions of the males are spread more or less evenly, with the highest contributor being male 197 (15%). The contributions of the females are also evenly spread with the exception of the main contributor, female 435 (27%).

In contrast to Row2, 75% of all male animals and 83% of all female animals contributed to the offspring. This will ensure a more even distribution of the original genetic material as was seen by the smaller loss of genetic material in comparison to Row 2 (Figure 4.1). This also suggested that the brood stock in this group in comparison to the Row 2 brood stock, could be more suited to the captive environment, or they were under less physiological stresses or they were younger.

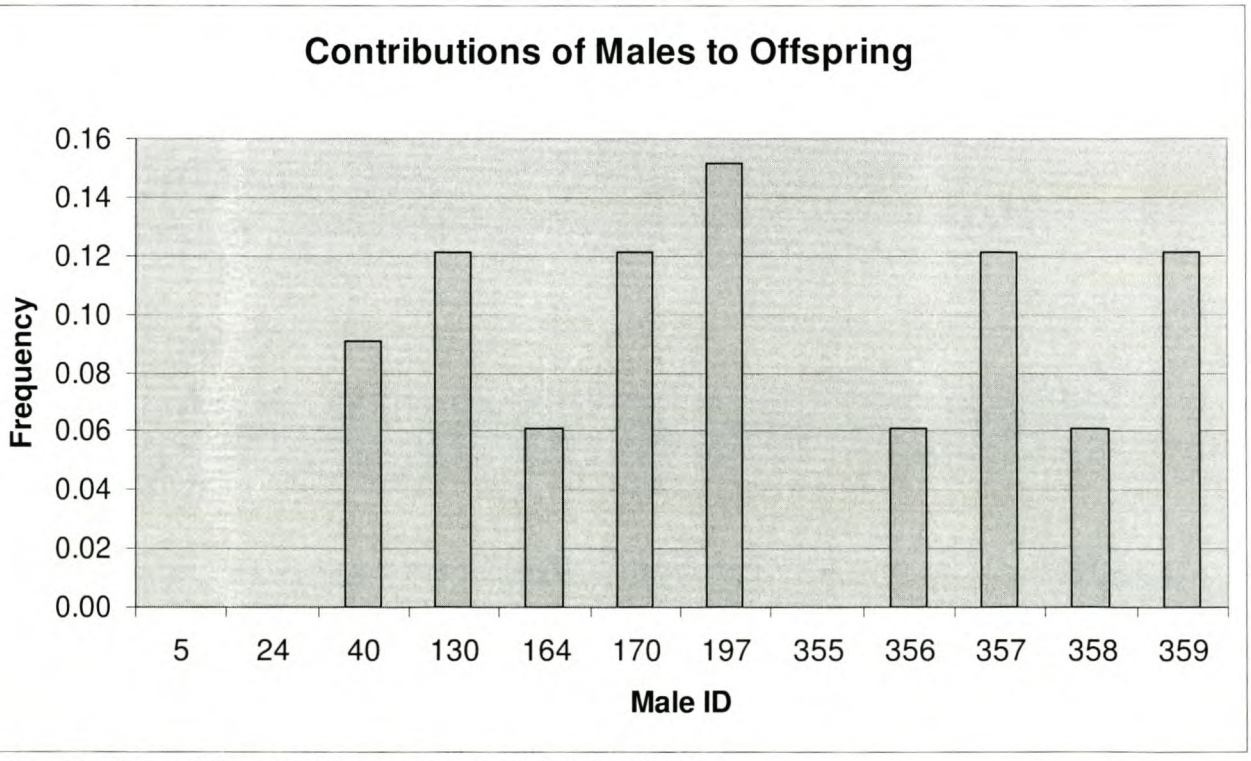


Figure 5.9: Differential contributions of males to Row 3 offspring.

Stellenbosch University http://scholar.sun.ac.za

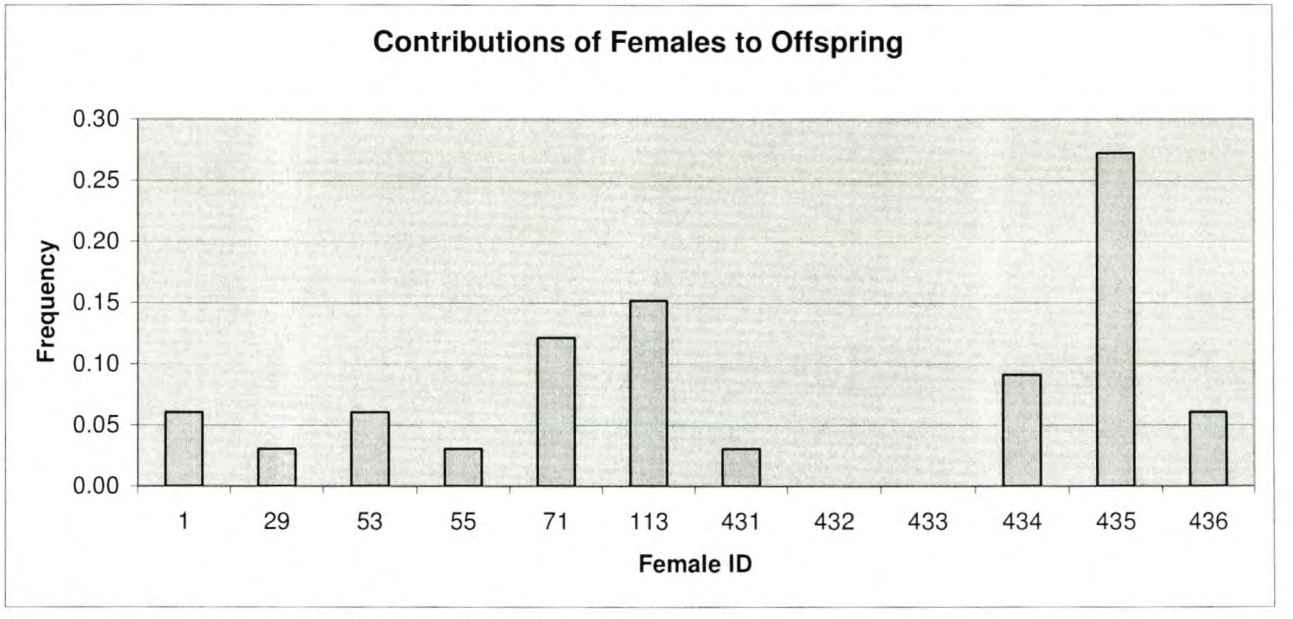


Figure 5.10: Differential contributions of females to Row 3 offspring.

5.4.3) ROW 4

The contributions of the males are shown in Figure 5.11 and the contributions of the females are shown in Figure 5.12.

Male 142 was the main contributor, being responsible for 35% of all offspring. Males 152 and 190 were the only other males that contributed to more than 10% of the offspring. The contributions of the females were unequal with female 140 contributing to 17% and female 131 to 15% of all offspring.

About 90% of the male animals contributed to the offspring. A reason for concern was that even though 88% of female animals contributed to the offspring, almost 40% of all females contributed to less than 2% of all offspring. The effective population size was thus still small, and this could be seen by the loss in alleles (Figure 4.1) and the high F_{is} values seen in section 4.6.2; p109 – 110. It was thus suggested that the female animals didn't adapt

Stellenbosch University http://scholar.sun.ac.za

too well to captivity or they were in a poor physiological condition or they were too old or the spawning conditions in the tank (environmental effect) was not ideal (for example: water temperature).

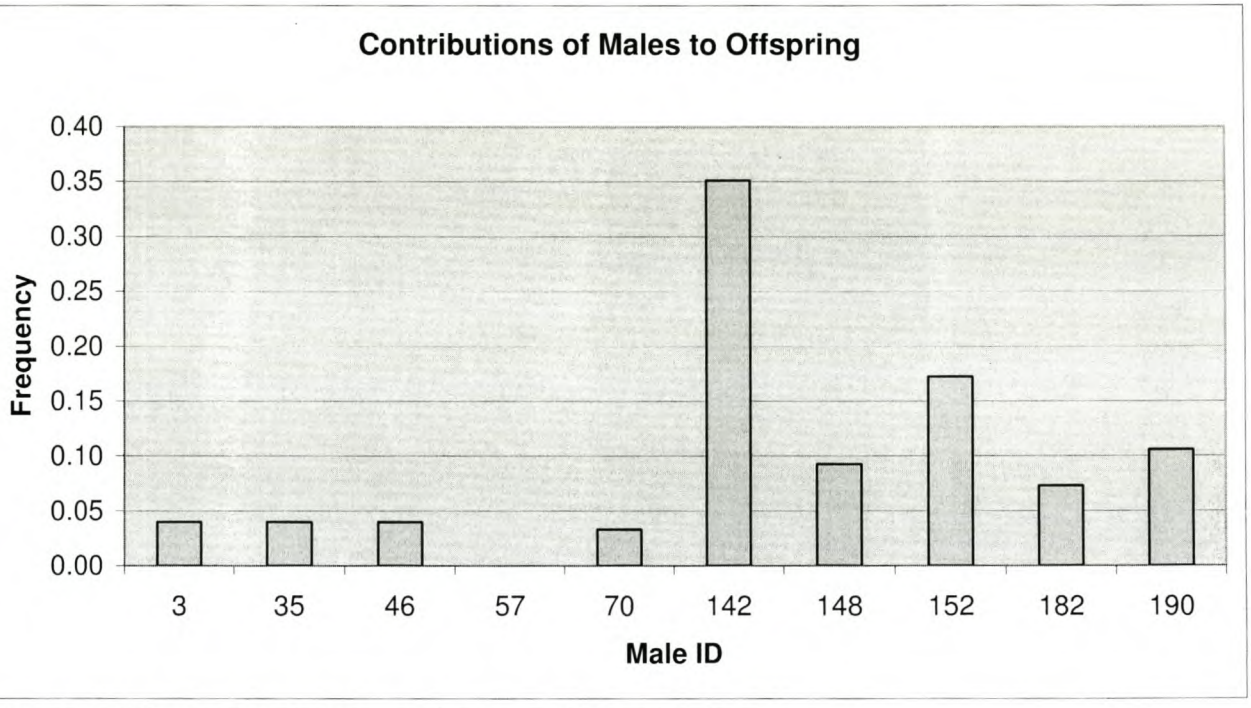


Figure 5.11: Differential contributions of males to Row 4 offspring.

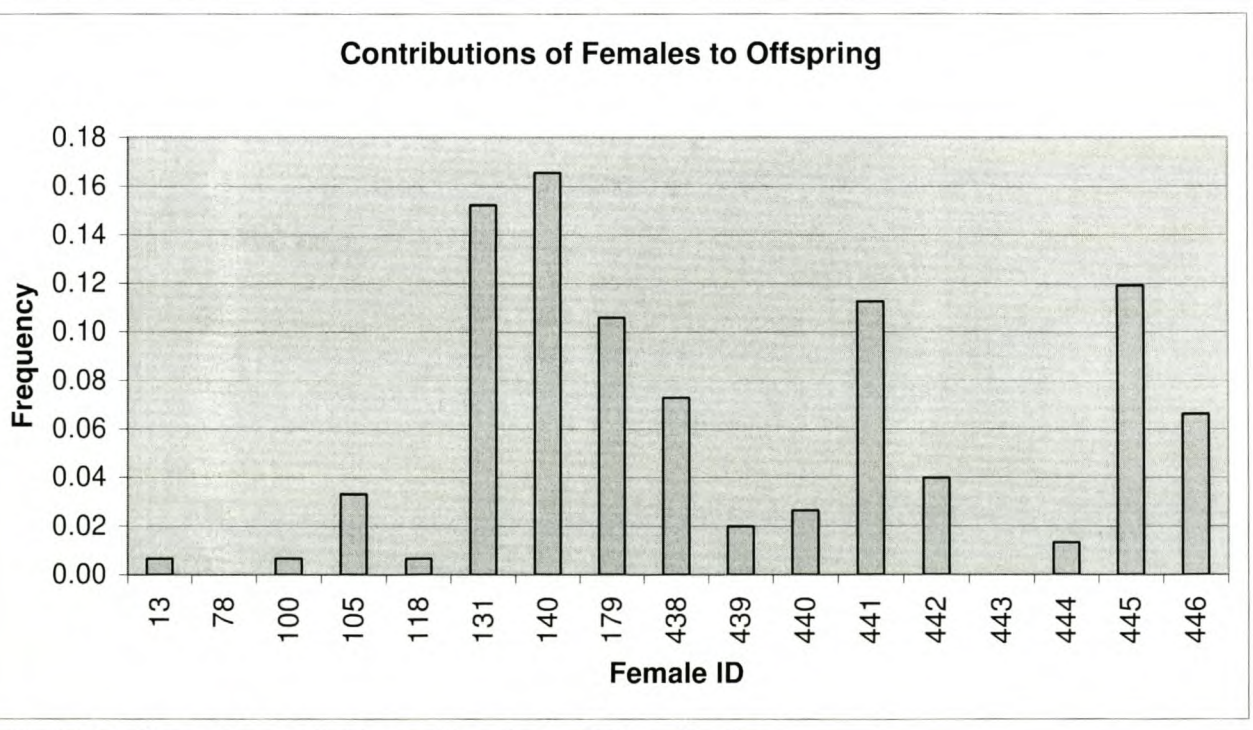


Figure 5.12: Differential contributions of females to Row 4 offspring.

Stellenbosch University http://scholar.sun.ac.za

5.5) REFERENCES

Arnason, E., 2004. Mitochondrial cytochrome *b* variation in the high-fecundity atlantic cod: Trans Atlantic clines and shallow gene genealogy. *Genetics* **166**, 1871 – 1885.

Estoup, A., Tailliez, C., Cornuet, J., Solignac, M., 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, Apis mellifera and Bombus terrestris (Apidae). Molecular and Biological *Evolution* **12**, 1074 – 1084.

Frankham, R., 1995. Inbreeding and extinction: a threshold effect. *Conservation Biology* **9**, 792 – 799.

Hedgecock, D., 1994. Does variance in reproductive success limit effective population size of marine organisms? In Beaumont, A.R. *Genetics and Evolution of Aquatic Organisms*. London: Chapman & Hall, 122 – 134.

Marshall, T.C., Slate, J., Kruuk, L.E.B., Pemberton, J.M., 1998. Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* **7**, 639 – 655.

Slate, J., Marshall, T., Pemberton, J., 2000. A retrospective assessment of the accuracy of the paternity inference program CERVUS. *Molecular Ecology* **9**, 801 – 808.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 6 – QUANTITATIVE TRAIT LOCI (QTL) DISCOVERY

6.1) PROCEDURE FOR QTL-DISCOVERY

6.1.1) OBJECTIVE

The objective of this part of the work was the identification of alleles and genotypes that segregated with growth rate (fast or slow). Animals carrying alleles or genotypes segregating with fast growth trait would further be used in selective breeding programs.

6.1.2) ACCURACY OF RESULTS

Due to the small sample sizes and few loci that were used in this study, any results should be taken as preliminary. According to Robinson *et al* (2003) more than 75 animals for each extremity of the studied trait should be used to confirm the segregation of an allele with a specific trait. The sample sizes used in this study were 60 per group (large or small) and this would give lower experimental power to the analysis. The likelihood of detecting segregating alleles would increase if more loci were studied. Robinson *et al* (2003) studied 102 loci, while this study used 7 loci.

Another hurdle for future work in QTL discovery is imposed by the reproductive biology of this species. Sexual maturity is reached at the age of 7 years (see section 1.1.3; p2). The potential segregation of the allele / genotype with the traits (fast and slow growth) studied here could only be confirmed in the F2 by backcrosses between the F1 and the brood stock. The segregating allele / genotype could then be used to select fast growers from the F2 to establish a new brood stock. A control brood stock consisting of slow growers from the F2 should also be established. The segregation of the allele / genotype should be studied using the two F3's and the growth rate of

Stellenbosch University http://scholar.sun.ac.za

the two F3's should then be compared to assess the success of the selection process of the fast growers from the F2. The confirmation of the segregation of an allele / genotype will take a minimum of 14 years to complete – 7 years for the F1 to reach sexual maturity plus a further 7 years for the F2 to reach sexual maturity.

6.1.3) PARAMETERS AND SIGNIFICANCE OF OBSERVATIONS

To achieve the objective of this study, different parameters were used to search and test for significant segregation between a genotype and / or allele and a specific trait (fast and slow growth). The differences in allele peak heights and observed number of alleles / genotypes between the large and the small samples were tested following a conventional χ^2 test.

The allele peak heights should only be used as rough estimators of allele numbers within a DNA pool and not as a reliable parameter for segregation, because: (1) the height could be influenced by the effectiveness of the amplification of product during a PCR reaction; (2) the concentration of the fluorescent label present within the sample might increase the peak area and cause the peak height to decrease; (3) the peak might "flat-top": an event that occurs if the allele reached its maximum peak height and this differs between reactions and alleles. Peak heights will thus not give an accurate assessment of the number of alleles and would need to be verified by genotyping individuals.

Differences were considered significant for $\chi^2 > 3.84$ ($P < 0.05$). Differences were considered putative for $2.706 < \chi^2 < 3.84$ ($0.05 < P < 0.1$). Significant differences were indicated with a ( ) in the electropherograms and graphs (for example Figure 6.5 shows significant differences at peak 5), and were highlighted in red within tables (for example Table 6.1 shows significant differences at peak 5).

Stellenbosch University http://scholar.sun.ac.za

6.1.4) ALLELE SEGREGATION

The bulked segregant analysis (Micklemore *et al*, 1991) procedure followed here was based on the work of Robinson *et al* (2003).

A preliminary analysis (done by Robinson *et al*, 2003) was performed by constructing two DNA pools: one containing the DNA of the fast growers (60 individuals) and the other containing DNA from the slow growers (59 individuals). Allele peak heights were used to detect differences among samples. All loci showing significant differences in allele peak heights between the two pools were further used for the individual genotyping of the large and small animals. Significant differences in the observed number of alleles between the large individuals and the small individuals were used to identify segregating alleles. The segregation of the alleles should then be confirmed as described above in section 6.1.2.

For this study individual genotyping was performed even if the results of the preliminary analysis showed no significant peak height differences – this was done because only 7 loci were studied. Results per locus are shown in sections 6.4.1 – 6.4.7. A summary of segregating alleles is given at the end of the chapter in Table 6.15; p154

6.1.5) GENOTYPE SEGREGATION

A bulked segregant analysis (developed by Michelmore *et al*, 1991) method used here, was adopted from Lee *et al* (2003) to determine whether certain genotypes were linked to the fast growing (large) or slow growing (small) trait. Only F1 individuals were used in the analysis, because no information of the growth rate of the parents was available. The genotypes per locus were compared between the large and the small individuals and the significant differences in the number of occurrences of each genotype were calculated by a conventional χ^2 test. Any segregation found for a genotype should be confirmed as described above in section 6.1.2.

Stellenbosch University http://scholar.sun.ac.za

Results per locus are shown in section 6.5. A summary of the genotypes (Tables 6.16 & 6.17) is shown at the end of this section; p155 – 156.

6.2) DNA QUANTIFICATION AND POOLING

DNA was successfully extracted from the mantle tissue of the large (60 animals) and small (59 animals) offspring of batches 266/03 and 238/03 (shown in Figures 6.1 – 6.2) and the DNA concentrations were standardized (see sections 2.11.4.2; p68).

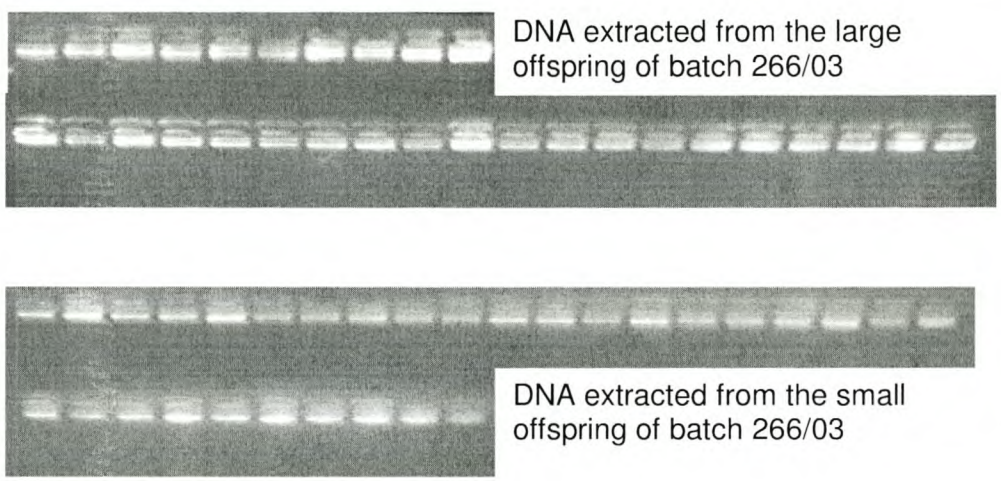


Figure 6.1: DNA extractions of batch 266/03.

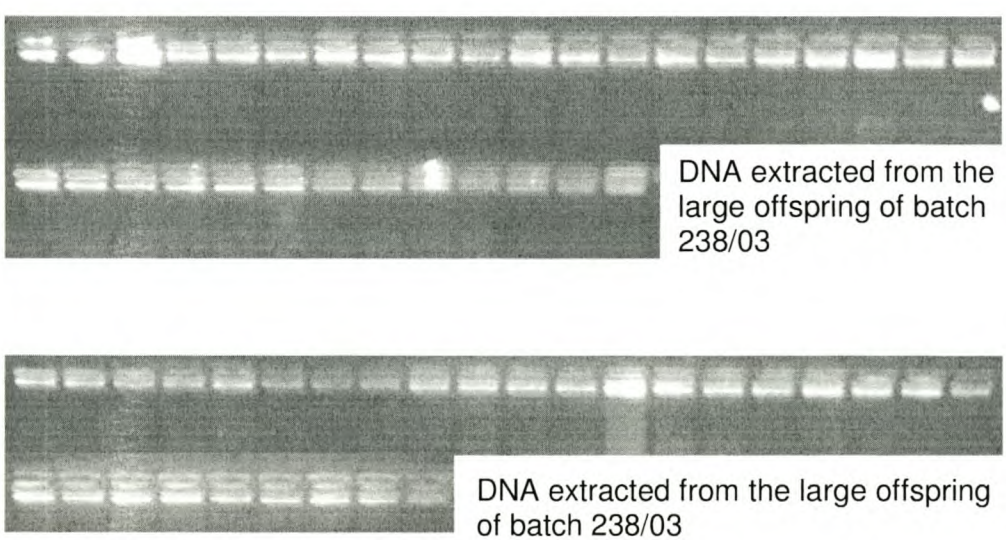


Figure 6.2: DNA extractions of batch 238/03.

Stellenbosch University http://scholar.sun.ac.za

6.3) GENOTYPING

6.3.1) PCR OF POOLS

The results of the PCR-reactions of pooled DNA using labelled microsatellite markers were tested in agarose gels (Figure 6.3) and later resolved in an ABI automated sequencer

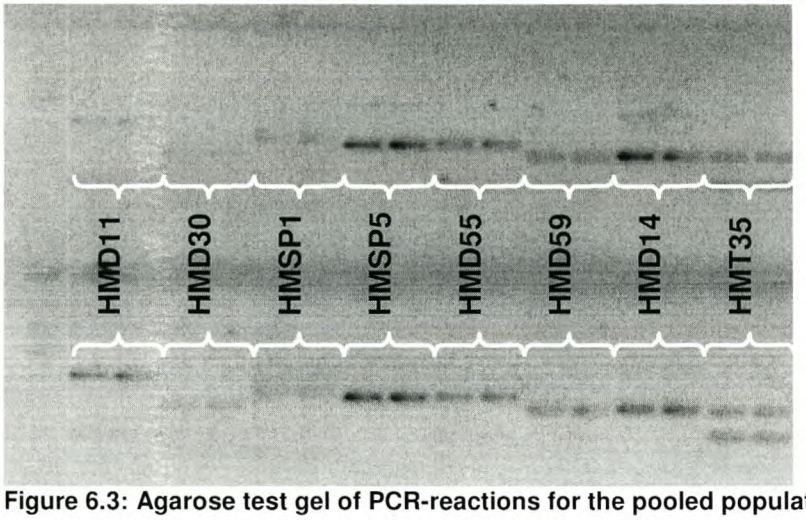


Figure 6.3: Agarose test gel of PCR-reactions for the pooled populations.

6.3.2) PCR OF INDIVIDUALS

The results of the PCR-reactions using locus HmT35, of 8 of the 60 large individuals and 8 of the 59 small individuals were tested in an agarose gel (Figure 6.4) and later resolved in an ABI automated sequencer. Allele frequencies of the pooled data are shown in Appendix D.

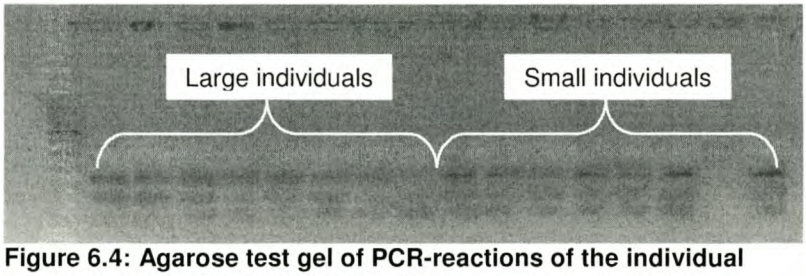


Figure 6.4: Agarose test gel of PCR-reactions of the individual samples for locus HmT35.

Stellenbosch University http://scholar.sun.ac.za

6.4) ANALYSIS

6.4.1) LOCUS HmD11

6.4.1.1) Genotyping of pools

The electropherogram of locus HmD11 for the ALL_Large and ALL_Small pools are shown in Figure 6.5 and the peak heights along with the results of the χ^2-test are shown in Table 6.1. The peak numbers in Figure 6.5 correspond to the peak numbers in Table 6.1.

Significant differences were found at alleles 292 (peak 1; Figure 6.5 & Table 6.1), 308 (peak 4; Figure 6.5 & Table 6.1) and 312 (peak 5; Figure 6.5 & Table 6.1)(Table 6.1; $P < 0.05$). The other two alleles showed no significant differences in peak heights between the two samples (Table 6.1; $P > 0.1$).

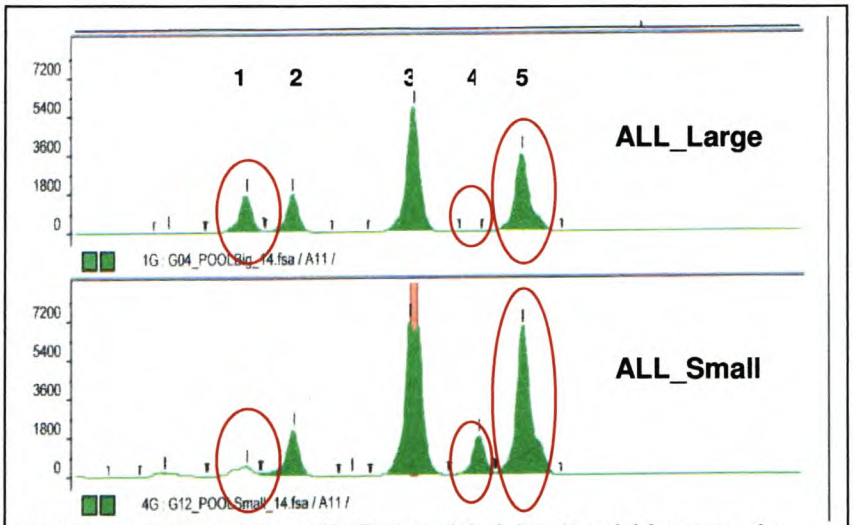


Figure 6.5: Comparison of HmD11 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 5.

Significant differences are shown by red oval

Table 6.1: Comparison of HmD11 peak heights between the ALL_Large and ALL_Small pools.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	292	1802	0	18.02
2	296	1822	2228	0.41
3	304	5863	7253	1.47
4	308	0	1850	18.50
5	312	3670	6987	10.78

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

Stellenbosch University http://scholar.sun.ac.za

6.4.1.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.2.

A significant difference was found at allele 308 (Table 6.2; $P < 0.05$). The differences in the observed numbers of the other alleles were non-significant (Table 6.2; $P > 0.1$).

Table 6.2: Comparison of observed allele numbers between large and small individuals.

HmD11			
Allele	Number of Alleles		χ^2
	Large	Small	
292	11	7	0.89
296	13	7	1.80
304	42	45	0.10
308	0	7	7.00
312	36	33	0.13

$\chi^2 > 3.84$ ($P < 0.05$)): significant for difference between large and small
Significant differences are highlighted in red

6.4.1.3) Possible segregating alleles

The results suggest that allele 308 could be segregating with the slow growing (small) individuals. The low allele frequency (Figure 6.6 & Table 6.2; $p = 0.0686$), sample size and the low number of observations (4 out of 59 small individuals carried allele 308) were the limiting factors in the positive identification of the allele as a segregating allele. The use of a bigger sample size would be able to test the hypothesis of allele 308 segregating with slow growth.

Significant differences for alleles 292 and 312 were only detected by peak height data (Figure 6.5 & Table 6.1) and since peak height is not a reliable parameter for segregation (see section 6.1.3; p131), further proof was needed in allele number differences. The data for the observed number of alleles (Table 6.2) failed to show significant differences between the large and small

Stellenbosch University http://scholar.sun.ac.za

individuals. Alleles 292 and 312 were therefore not considered as segregating alleles.

6.4.2) LOCUS HmD30

6.4.2.1) Genotyping of pools

The electropherogram of locus HmD30 for the ALL_Large and ALL_Small pools are shown in Figure 6.6 and the peak heights and the results of the χ^2-test are shown in Table 6.3. The peak numbers in Figure 6.6 correspond to the peak numbers in Table 6.4.

No significant differences were found at any alleles in locus HmD30 (Table 6.3; $P > 0.1$).

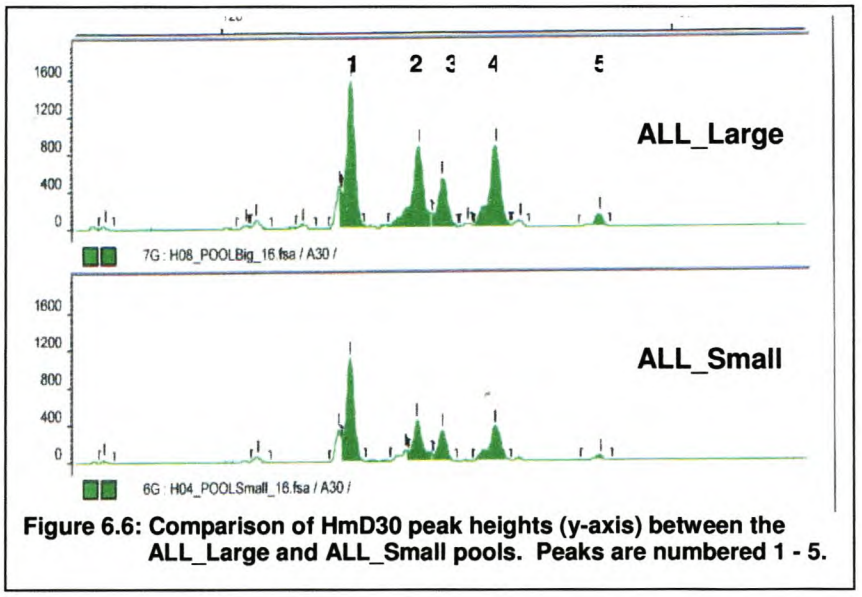


Figure 6.6: Comparison of HmD30 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 5.

Table 6.3: Comparison of HmD30 peak heights between the ALL_Large and ALL_Small pools.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	132	1602	1129	0.82
2	138	889	452	1.42
3	140	552	335	0.53
4	144	895	386	2.02
5	154	150	61	0.38

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small

Stellenbosch University http://scholar.sun.ac.za

6.4.2.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.4.

Significant differences were found at alleles 132 and 140 (Table 6.4; $P < 0.05$). The differences in the observed numbers of the other alleles were non-significant (Table 6.4; $P > 0.1$).

Table 6.4: Comparison of observed allele numbers between large and small individuals.

HmD30			
Allele	Number of Alleles		χ^2
	Large	Small	
124	0	1	1.00
128	2	0	2.00
132	18	33	4.41
138	28	22	0.72
140	2	10	5.33
142	3	2	0.20
144	14	13	0.04
148	1	0	1.00
154	0	1	1.00

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

6.4.2.3) Possible segregating alleles

The results suggest that alleles 132 and 140 segregate with the slow growing (small) individuals.

The data for alleles 132 and 140 was, however, not very reliable: even though the observed number of alleles data (Table 6.4) found significant differences, the observations could be ascribed to missing data in the individual genotypes, as might be suggested by the similar peak heights between the two samples for alleles 132 and 140 (Figure 6.6 & Table 6.3) – of a possible 60 fast growers (large) animals only 35 was typed at locus HmD30, while 42 out of 59 slow growers (small) were typed.

Stellenbosch University http://scholar.sun.ac.za

6.4.3) LOCUS HmSP5

6.4.3.1) Genotyping of pools

The electropherogram of locus HmSP5 for the ALL_Large and ALL_Small pools are shown in Figure 6.7 and the peak heights and the results of the χ^2-test are shown in Table 6.5. The peak numbers in Figure 6.7 correspond to the peak numbers in Table 6.5.

Significant differences were found at alleles 185 (peak1; Figure 6.7 & Table 6.5) and 189 (peak 2; Figure 6.7 & Table 6.5)(Table 6.5; $P < 0.05$). No significant differences were found at allele 193 (Table 6.5; $P > 0.1$).

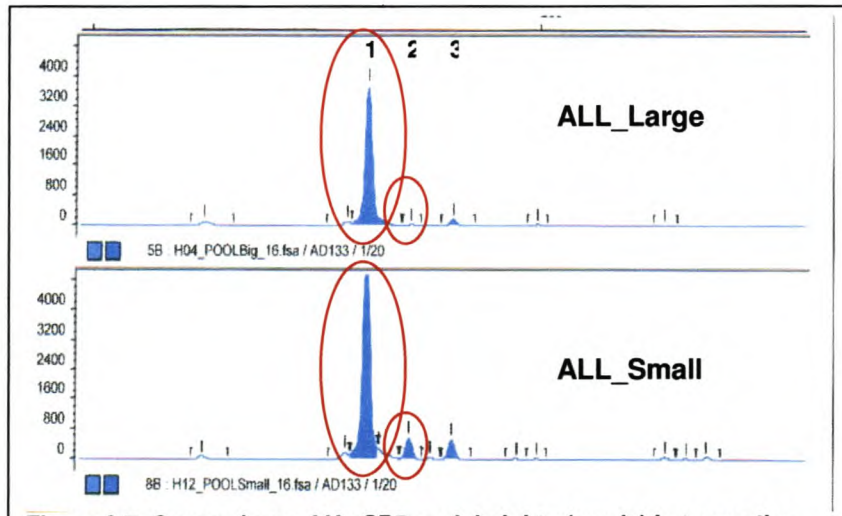


Figure 6.7: Comparison of HmSP5 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 3.

Significant differences are shown by red oval

Table 6.5: Comparison of peak heights between the large and small pools for HmSP5.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	185	3806	6247	5.93
2	189	57	597	4.46
3	193	198	572	1.82

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

Stellenbosch University http://scholar.sun.ac.za

6.4.3.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.6.

A significant difference was found at allele 189 (Table 6.6; $P < 0.05$). The observed numbers of the other alleles were non-significant (Table 6.6; $P > 0.1$).

Table 6.6: Comparison of observed allele numbers between large and small individuals.

HmSP5			
Alleles	Number of Alleles		χ^2
	Large	Small	
185	88	76	0.88
187	3	3	0
189	1	9	6.40
193	9	13	0.73
Other	11	13	0.17

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

6.4.3.3) Possible segregating alleles

Significant differences for allele 185 was only detected by peak height data (Figure 6.7 & Table 6.5) and since peak height is not a reliable parameter for segregation (see section 6.1.3; p131), further proof was needed in allele number differences. The data for the observed number of alleles (Table 6.6) failed to show significant differences between the large and small individuals. Allele 185 was therefore not considered as a segregating allele.

The results suggest that allele 189 is segregating with the slow growing (small) individuals. The use of a bigger sample size would be more accurate for testing the hypothesis of allele 189 segregating with slow growth.

Stellenbosch University http://scholar.sun.ac.za

6,4,4) LOCUS HmD55

6.4.4.1) Genotyping of pools

The electropherogram of locus HmD55 for the ALL_Large and ALL_Small pools are shown in Figure 6.8 and the peak heights and the results of the χ^2-test are shown in Table 6.7. The peak numbers in Figure 6.8 correspond to the peak numbers in Table 6.7.

A significant difference was found at allele 191 (peak 3; Figure 6.8 & Table 6.7; $P < 0.05$). The other alleles showed no significant differences (Table 6.7; $P > 0.1$).

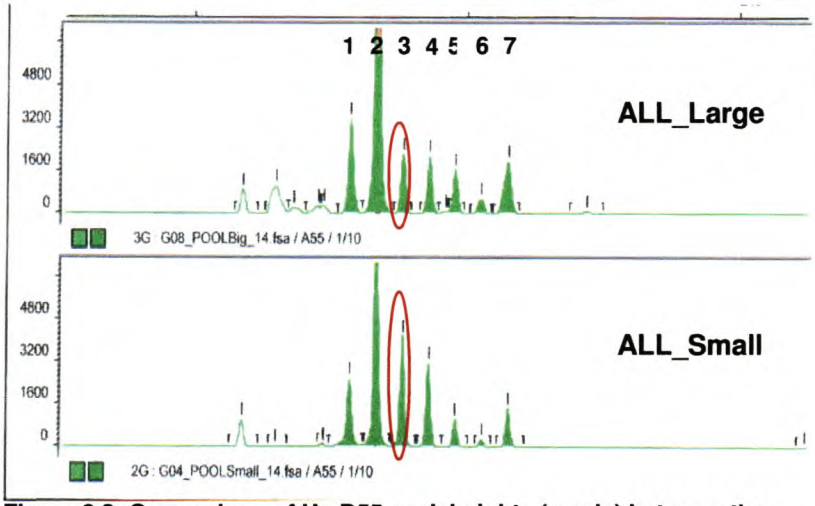


Figure 6.8: Comparison of HmD55 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 7.

Significant differences are shown by red oval

Table 6.7: Comparison of peak heights between the large and small pools for HmD55.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	183	3596	2603	1.59
2	187	7223	7426	0.03
3	191	2275	4222	5.83
4	195	2191	3217	1.95
5	199	1688	1077	1.35
6	203	591	307	0.90
7	207	1973	1480	0.70

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

Stellenbosch University http://scholar.sun.ac.za

6.4.4.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.8.

No significant differences were found at the alleles of locus HmD55 (Table 6.12; $P > 0.1$). A putative difference was found at allele 191 (Table 6.8; $0.05 < P < 0.1$).

Table 6.8: Comparison of observed allele numbers between large and small individuals.

HmD55			
Alleles	**Number of Alleles**		χ^2
	Large	**Small**	
183	9	8	0.06
187	60	49	1.11
191	**13**	**23**	**2.78**
195	9	15	1.50
199	7	6	0.08
203	3	2	0.20
207	16	10	1.38
219	1	0	1.00

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Putative difference is highlighted in yellow

6.4.4.3) Possible segregating alleles

Results suggest the possible segregation of allele 191 with the slow growing (small) individuals. The observed number of allele data found putative differences (Table 6.8; $0.05 < P < 0.1$) between the two samples. A bigger sample size might be more informative when testing the hypothesis on the potential segregation of allele 191.

Stellenbosch University http://scholar.sun.ac.za

6.4.5) LOCUS HmD59

6.4.5.1) Genotyping of pools

The electropherogram of locus HmD59 for the ALL_Large and ALL_Small pools are shown in Figure 6.9 and the peak heights and the results of the χ^2-test are shown in Table 6.9. The peak numbers in Figure 6.9 correspond to the peak numbers in Table 6.9.

Significant differences were found at alleles 112 (peak 5; Figure 6.9 & Table 6.9), 114 (peak 6; Figure 6.9 & Table 6.9) and 116 (peak 7; Figure 6.9 & Table 6.9)(Table 6.9; $P < 0.05$). Alleles 106 (peak 2; Figure 6.9 & Table 6.9), 110 (peak 4; Figure 6.9 & Table 6.9), 118 (peak 8; Figure 6.9 & Table 6.9), 130 (peak 14; Figure 6.9 & Table 6.9) and 132 (peak 15; Figure 6.9 & Table 6.9) showed putative differences between the two samples (Table 6.9; $0.05 < P < 0.1$). The other alleles showed no significant differences (Table 6.9; $P > 0.1$).

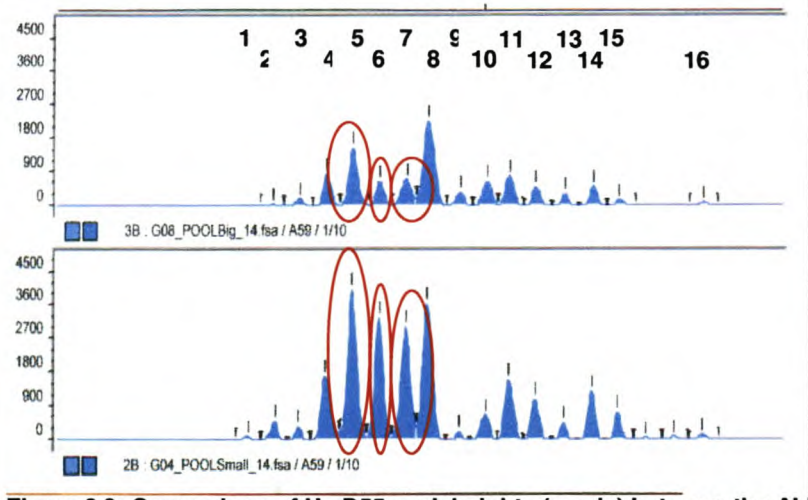


Figure 6.9: Comparison of HmD55 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 16.

Significant differences are shown by red oval

Stellenbosch University http://scholar.sun.ac.za

Table 6.9: Comparison of peak heights between the large and small pools for HmD59.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	104	0	112	1.12
2	106	64	514	3.50
3	108	199	389	0.61
4	110	871	1706	2.71
5	112	1588	4062	10.83
6	114	676	3307	17.38
7	116	723	3058	14.42
8	118	2278	3669	3.25
9	120	366	224	0.34
10	122	662	690	0.01
11	124	825	1609	2.53
12	126	517	1099	2.10
13	128	333	462	0.21
14	130	535	1305	3.22
15	132	163	747	3.75
16	136	106	161	0.11

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red
Putative differences ($0.05 < P < 0.1$) are highlighted in yellow

6.4.5.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.10.

A significant difference was found at allele 126 (Table6.10; $P < 0.05$). A putative difference was found at allele 114 (Table 6.10; $0.05 < P < 0.1$). The other alleles showed no differences (Table 6.10; $P > 0.1$).

Table 6.10: Comparison of observed allele numbers between large and small individuals.

HmD59			
Alleles	Number of Alleles		χ^2
	Large	Small	
106	0	2	2.00
110	3	2	0.20
112	14	15	0.03
114	6	14	3.20
116	9	13	0.73
118	7	3	1.60
120	8	6	0.29
122	9	9	0
124	4	10	2.57
126	14	5	4.26
128	16	15	0.03
130	10	14	0.67
132	1	0	1.00
136	7	4	0.82

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red
Putative differences ($0.05 < P < 0.1$) are highlighted in yellow

Stellenbosch University http://scholar.sun.ac.za

6.4.5.3) Possible segregating alleles

The results suggest that four possible segregating alleles are present for locus HmD59.

Alleles 106, 110, 112, 116, 118, 130 and 132 were not considered as segregants, because allele peak height is not a reliable parameter for segregation (see section 6.1.3; p131) and the observed number of allele data (Table 6.10; $P > 0.1$) showed no differences between the samples.

Allele 114 segregates with slow growers. Allele 114 only showed a significant difference for peak height data (Figure 6.9 & Table 6.9; $P < 0.05$) and a putative difference for the observed number of allele data (Table 6.10; $0.05 < P < 0.1$). The use of a bigger sample size would be able to test the hypothesis of allele 114 segregating with slow growth.

Results suggest that allele 126 segregates with the fast growing (large) individuals. Peak height data (Figure6.14 & Table 6.9; $P > 0.1$) showed no significant difference between the two samples, but the observed number of allele data (Table 6.10; $P < 0.05$) showed significant differences between the two samples. The use of a bigger sample size would be able to test the hypothesis of allele 126 segregating with fast growth.

Stellenbosch University http://scholar.sun.ac.za

6.4.6) LOCUS HmD14

6.4.6.1) Genotyping of pools

The electropherogram of locus HmD14 for the ALL_Large and ALL_Small pools are shown in Figure 6.10 and the peak heights and the results of the χ^2-test are shown in Table 6.11. The peak numbers in Figure 6.10 correspond to the peak numbers in Table 6.11.

A significant difference was found at allele 142 (peak 1; Figure 6.10 & Table 6.11; $P < 0.05$). The other alleles of locus HmD14 showed no significant differences (Table 6.16; $P > 0.1$).

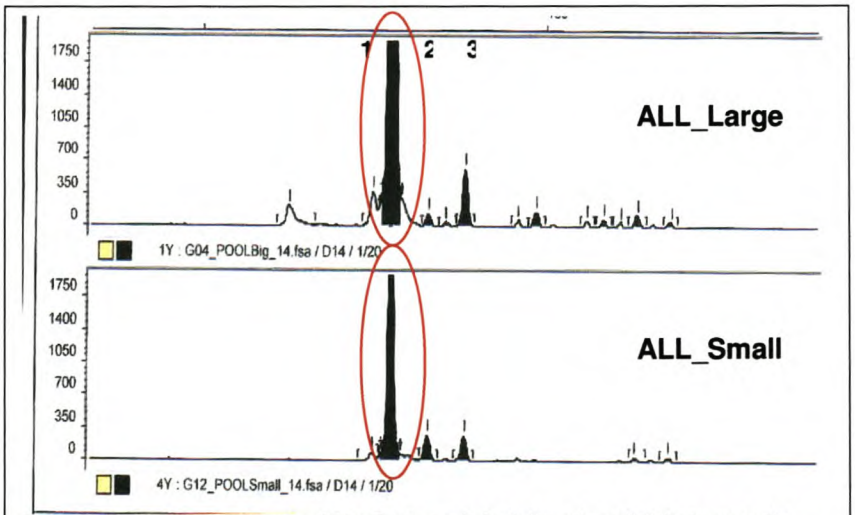


Figure 6.10: Comparison of HmD14 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 3.

Significant differences are shown by red oval

Table 6.11: Comparison of peak heights between the large and small pools for HmD14.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	142	7088	2784	18.76
2	146	147	277	0.40
3	150	623	277	1.33

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

Stellenbosch University http://scholar.sun.ac.za

6.4.6.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.12.

A significant difference was found at allele 146 (Table 6.12; $P < 0.05$). Putative differences were found at alleles 142, 144, 156 and 160 (Table 6.12; $0.05 < P < 0.1$). No significant differences were detected for the other alleles (Table 6.12; $P > 0.1$).

Table 6.12: Comparison of observed allele numbers between large and small individuals.

HmD14			
Alleles	Number of Alleles		χ^2
	Large	Small	
142	98	74	3.35
144	0	3	3.00
146	0	9	9.00
150	12	14	0.15
154	0	1	1.00
156	0	3	3.00
158	4	3	0.14
160	0	3	3.00

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red
Putative differences ($0.05 < P < 0.1$) are highlighted in yellow

6.4.6.3) Possible segregating alleles

The data suggests that allele 142 segregates with the fast (large) growing individuals. The allele peak height data were significant (Table 6.11; $P < 0.05$), while the observed number of allele data showed putative differences (Table 6.12; $0.05 < P < 0.1$).

The data suggests allele 146 segregates with the slow growing (small) individuals. The few observations (6 individuals out of 59 carried the allele) was a concern – the allele might have segregated by chance. The use of a bigger sample size would be able to test the hypothesis of allele 146 segregating with slow growth.

Stellenbosch University http://scholar.sun.ac.za

Alleles 144, 156 and 160 need bigger sample sizes to test the hypothesis of their segregation with growth rate. They were ignored as segregants, because no significant differences in allele peak heights (Figure 6.10 and Table 6.11; $P > 0.1$) were found. The observed number of allele differences was only putative (Table 6.12; $0.05 < P < 0.1$) and the number of observations was very low.

Stellenbosch University http://scholar.sun.ac.za

6.4.7) LOCUS HmT35

6.4.7.1) Genotyping of pools

The electropherogram of locus HmT35 for the ALL_Large and ALL_Small pools are shown in Figure 6.11 and the peak heights and the results for the χ^2-test are shown in Table 6.13. The peak numbers in Figure 6.11 correspond to the peak numbers in Table 6.13.

No significant differences were found at the alleles of locus HmT35 (Table 6.13; $P > 0.1$).

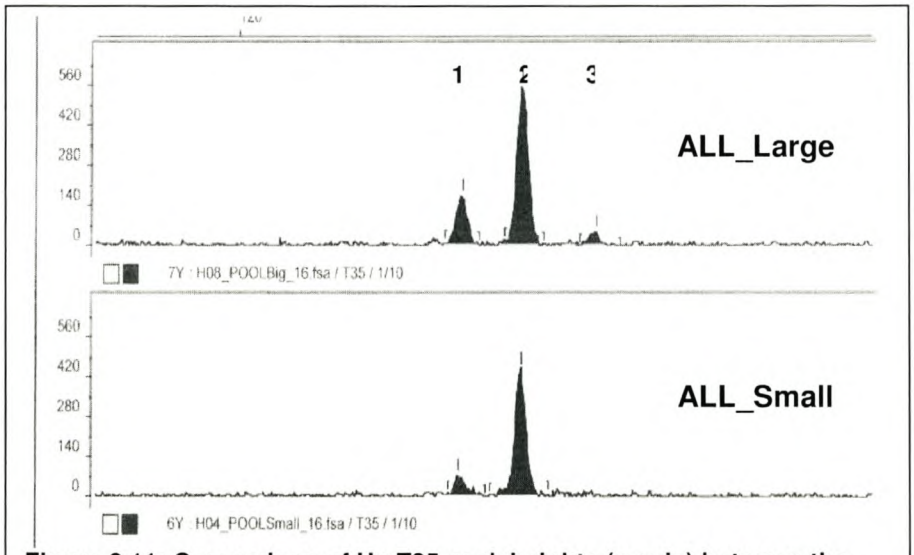


Figure 6.11: Comparison of HmT35 peak heights (y-axis) between the ALL_Large and ALL_Small pools. Peaks are numbered 1 - 3.

Table 6.13: Comparison of peak heights between the large and small pools for HmT35.

Peak Nr	Allele	Peak Height (ALL_Large)	Peak Height (ALL_Small)	χ^2
1	135	181	78	0.41
2	139	574	455	0.14
3	143	54	0	0.54

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small

Stellenbosch University http://scholar.sun.ac.za

6.4.7.2) Genotyping of individuals

The observed number of alleles and the results of the χ^2-test are shown in Table 6.14.

A significant difference was found at allele 135 (Table 6.14; $P < 0.05$). The other alleles showed no significant difference between the two samples (Table 6.14; $P > 0.1$).

Table 6.14: Comparison of observed allele numbers between large and small individuals.

HmT35			
Allele	Number of Alleles		χ^2
	Large	Small	
135	23	11	4.24
139	86	100	1.05
143	5	1	2.67

$\chi^2 > 3.84$ ($P < 0.05$): significant for difference between large and small
Significant differences are highlighted in red

6.4.7.3) Possible segregating alleles

The data suggests that allele 135 is segregating with fast growing (large) individuals. The differences of the observed number of allele data (Figure 6.11 & Table 6.14; $P < 0.05$) were significant. The use of a bigger sample size would be able to test the hypothesis of allele 135 segregating with fast growth.

6.4.8) CONCLUDING REMARKS

Table 6.15 gives a summary for the potential segregating alleles. The potential segregation of two alleles with the fast growing (large) individuals provides valuable information to continue exploring these loci and the genotypes of parents producing fast-growing offspring. All the potential segregating alleles that were identified need to be verified by using larger sample sizes. Their potential segregation needs to be tested within the F2 of backcrosses between the original parents and the F1.

Stellenbosch University http://scholar.sun.ac.za

6.5) SEGREGATING GENOTYPES

Table 6.16 shows the loci with their various genotypes, the number of observations for each genotype per locus and the result of the χ^2-test per genotype.

Five loci had one segregating genotype each: locus HmD11 – 292 / 304, locus HmSP5 – 185 / 189, HmD59 – 114 / 128, locus HmD14 – 142 / 142 and HmT35 – 135 / 139.

The genotype of locus HmD11 is potentially segregating with fast growing (large) individuals.

The genotype of locus HmSP5 is potentially segregating with slow growing (small) individuals. The genotype carried allele 189 that is suggested to segregate with slow growers (see above). This could be seen as further evidence for the segregation of allele 189.

The genotype of locus HmSP5 is potentially segregating with slow growing (small) individuals. The genotype carried allele 114 that is suggested to segregate with slow growers (see above). This could be seen as further evidence for the segregation of allele 114.

The genotype of locus HmD14 is potentially segregating with fast growing (large) individuals. The genotype carries allele 142 and this could be seen as further evidence for the segregation of the allele.

The genotype of locus HmT35 is potentially segregating with fast growing (large) individuals. The genotype carried allele 135 that is suggested to segregate with fast growers (see above). This could be seen as further evidence for the segregation of allele 135.

Stellenbosch University http://scholar.sun.ac.za

To confirm the segregation of certain genotypes with certain traits, it is necessary to use larger sample sizes and use families to track the genotypes.

6.6) DIFFERENTIAL CONTRIBUTION OF PARENTS

Row4 parents from offspring that was assigned 100% (see Chapter 5; p114) was used to draw the histogram of differential contributions shown in Figure 6.12.

The data obtained from Figure 6.12 could be useful when animals are selected for future QTL-experiments where linked markers for the fast growing trait will be sought. The combination of male 152 and female 131 will be extremely useful during the testing for segregation of potential markers selecting for fast growers. The data also suggest the possibility of designing early selective breeding strategies. By only choosing the combinations of parents that spawned fast growers and discarding the combinations that spawned only slow growers, it will be theoretically possible to maximize the number of fast growing individuals within the F1. Examples of parental combinations that spawned only fast growers: male 46 + female 140; male 46 + female 446; male 142 and female 105; male 152 + female 131; male 152 + female 100 and male 190 + female 140 (Figure 6.12). The results for the differential contributions of the Row4 parents (see section 5.4.3; p127 – 128) suggests that the best breeding pairs may be male 152 + female 140 and male 190 + female 140. These four animals each contributed to 15% or more of the offspring, which make them reproductively fit. Offspring would be guaranteed for these crossings.

Examples of parental combinations that spawned only slow growers: male 35 + female 13; male 46 + female 118; male 142 + female 438; male 142 + female 439; male 152 + female 445; male 152 + female 446; male 190 + female 438 and male 190 + female 445.

Stellenbosch University http://scholar.sun.ac.za

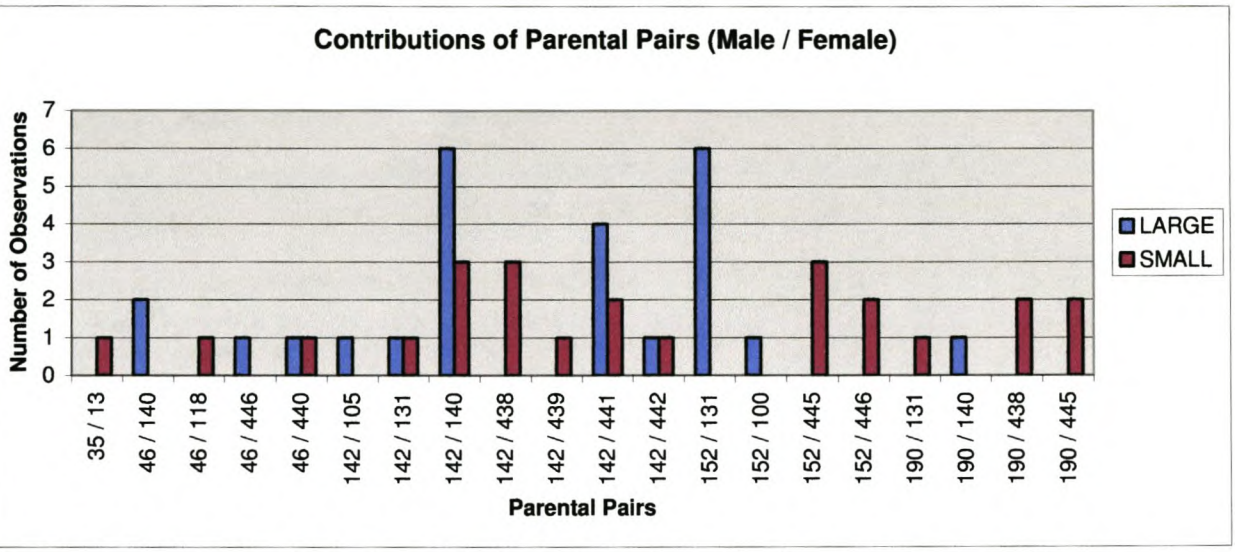


Figure 6.12: Histogram of differential contributions of parents to either the large or the small offspring.

Parental Pairs = Male / Female

Table 6.17 shows the genotypes over 2 loci that showed putative differences between the large and the small animals following a conventional χ^2-test. The genotypes were 185 / 193 + 187 / 191 and 185 / 193 + 187 / 195 (locus HmSP5 + HmD55) and both segregated with slow growth. The use of a bigger sample size would be able to test the hypothesis of these genotypes segregating with slow growth.

Stellenbosch University http://scholar.sun.ac.za

Table 6.15: Summary of significance and problems of the potential segregating alleles.

Potential Segregant			Significance of Observed Differences			Problems			
Locus	**Allele**	**Trait**	**Peak Height**	**Observed Number of Alleles**	**Small Sample Sizes**	**Low Allele Frequency ($p < 0.1$)**	**Few Observations**	**Missing Genotype Data**	**Usefulness In Selective Breeding Programs Selecting For Fast Growers**
HmD11	**308**	**Small**	**	**	Yes	Yes	Yes	No	No
HmD30	**132**	**Small**	NS	**	Yes	No	No	Yes	No
	140	**Small**	NS	**	Yes	No	No	Yes	No
HmSP5	**189**	**Small**	**	**	Yes	Yes	Yes	No	No
HmD55	**191**	**Small**	**	*	Yes	No	No	No	No
HmD59	**114**	**Small**	**	*	Yes	No	No	No	No
	126	**Large**	NS	**	Yes	No	No	No	Yes
HmD14	**142**	**Large**	**	*	Yes	No	No	No	Yes
	146	**Small**	NS	**	Yes	Yes	Yes	No	No
HmT35	**135**	**Large**	NS	**	Yes	No	No	No	Yes

Significant: $\chi^2 > 3.84$ = **
Putative: $2.706 < \chi^2 < 3.84$ = *
Non-significant $\chi^2 < 2.706$ = NS
Small Sample Sizes: Indicates that the small sample size used in this study could influence the results.
Few Observations: Indicates whether the allele is carried by less than 15% of the individuals from the trait that the allele segregates with.
Missing Genotype Data: Indicates whether some individual genotypic data is missing for the locus.

Stellenbosch University http://scholar.sun.ac.za

Table 6.16: Summary of potentially linked genotypes per locus.

LOCUS	GENOTYPE	LARGE	SMALL	χ^2	LOCUS	GENOTYPE	LARGE	SMALL	χ^2	LOCUS	GENOTYPE	LARGE	SMALL	χ^2
HmD11	292 / 292	1	3	1.0		185 / 207	1	0	1.0		110 / 122	1	0	1.0
	292 / 304	5	0	*		185 / 211	1	1	0.0		110 / 130	1	0	1.0
	292 / 312	4	1	1.8		185 / 215	0	1	1.0		112 / 112	3	1	1.0
	296 / 296	5	2	1.3		189 / 189	0	2	2.0		112 / 114	2	1	0.3
	296 / 304	3	5	0.5		189 / 193	1	0	1.0		112 / 116	1	3	1.0
	304 / 304	14	16	0.1		189 / 215	0	1	1.0		112 / 124	0	2	2.0
	304 / 308	0	1	1.0		193 / 193	1	0	1.0		112 / 126	2	1	0.3
	304 / 312	6	8	0.3		193 / 197	0	1	1.0		112 / 128	2	4	0.7
	308 / 308	0	3	3.0		199 / 199	0	1	1.0		112 / 130	1	2	0.3
	312 / 312	13	12	0.0		199 / 203	1	0	1.0		114 / 120	0	1	1.0
						199 / 215	0	1	1.0		114 / 122	0	2	2.0
HmD30	124 / 138	0	1	1.0		201 / 201	1	1	0.0		114 / 124	1	2	0.3
	128 / 128	1	0	1.0							114 / 126	2	1	0.3
	132 / 132	3	7	1.6	HmD55	183 / 183	1	0	1.0		114 / 128	0	5	*
	132 / 138	2	6	2.0		183 / 187	4	5	0.1		114 / 130	1	2	0.3
	132 / 140	2	7	2.8		183 / 191	1	2	0.3		116 / 116	1	0	1.0
	132 / 144	8	6	0.3		183 / 195	0	1	1.0		116 / 118	0	1	1.0
	138 / 138	9	4	1.9		183 / 207	2	0	2.0		116 / 122	0	2	2.0
	138 / 142	2	2	0.0		187 / 187	16	9	2.0		116 / 126	1	2	0.3
	138 / 144	4	5	0.1		187 / 191	5	6	0.1		116 / 128	4	2	0.7
	138 / 146	1	0	1.0		187 / 195	3	7	1.6		116 / 130	0	3	3.0
	138 / 148	1	0	1.0		187 / 199	7	6	0.1		116 / 136	1	0	1.0
	140 / 140	0	1	1.0		187 / 207	9	7	0.3		118 / 118	1	0	1.0
	140 / 154	0	1	1.0		191 / 191	2	5	1.3		118 / 126	1	0	1.0
	142 / 146	1	0	1.0		191 / 195	1	2	0.3		118 / 130	4	2	0.7
	144 / 144	1	1	0.0		191 / 203	1	1	0.0		120 / 124	1	1	0.0
	146 / 146	0	1	1.0		191 / 207	1	2	0.3		120 / 128	5	2	1.3
						195 / 195	1	2	0.3		120 / 130	2	2	0.0
HmSP5	185 / 185	36	26	1.6		195 / 203	1	1	0.0		122 / 122	2	1	0.3
	185 / 187	3	3	0.0		195 / 207	1	1	0.0		122 / 124	1	2	0.3
	185 / 189	0	4	*		195 / 219	1	0	1.0		122 / 126	1	0	1.0
	185 / 193	6	12	2.0		203 / 207	1	0	1.0		122 / 128	1	0	1.0
	185 / 199	0	1	1.0		207 / 207	1	0	1.0		122 / 130	1	2	0.3
	185 / 201	3	1	1.0							124 / 128	1	0	1.0
	185 / 203	0	1	1.0	HmD59	106 / 106	0	1	1.0		124 / 136	0	2	2.0
	185 / 205	2	0	2.0		110 / 110	1	1	0.0		126 / 126	2	0	2.0

Stellenbosch University http://scholar.sun.ac.za

LOCUS	GENOTYPE	LARGE	SMALL	χ^2	LOCUS	GENOTYPE	LARGE	SMALL	χ^2	LOCUS	GENOTYPE	LARGE	SMALL	χ^2
Continued...														
HmD59	126 / 128	1	1	0.0		142 / 144	0	1	1.0		150 / 154	0	1	1.0
	126 / 136	3	0	3.0		142 / 146	0	3	3.0		156 / 156	0	1	1.0
	128 / 128	1	0	1.0		142 / 150	4	9	1.9		158 / 158	1	1	0.0
	128 / 136	0	1	1.0		142 / 156	0	1	1.0		160 / 160	0	1	1.0
	130 / 136	0	1	1.0		142 / 158	2	1	0.3					
	132 / 136	1	0	1.0		142 / 160	0	1	1.0	HmT35	135 / 139	23	11	*
	136 / 136	1	0	1.0		144 / 144	0	1	1.0		139 / 139	29	44	3.1
						146 / 146	0	3	3.0		139 / 143	5	1	2.7
HmD14	142 / 142	46	29	*		150 / 150	4	2	0.7					

* = significant for difference between two samples ($P < 0.05$)

Table 6.17: Summary of potentially linked genotypes over two loci.

LOCI	GENOTYPE	LARGE	SMALL	χ^2
HmSP5 + HmD55	185 / 193 + **187** / **191**	0	3	*
HmSP5 + HmD55	185 / 193 + **187** / **195**	0	3	*

* = putative significance for difference between two samples ($0.05 < P < 0.1$)
Genotype = allele1 – locusA / allele2 – locusA + allele1 – locusB / allele2 – locusB

Stellenbosch University http://scholar.sun.ac.za

6.7) REFERENCES

Lee, B.-Y., Penman, D.J., Kocher, T.D., 2003. Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. *Animal Genetics* **34**, 1 – 5.

Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Science USA* **88**, 9828 – 9832.

Robinson, N., Baranski, M., Rourke, M., Hayes, B., 2003. Application of advanced genetic techniques to abalone breeding.10th Abalone Aquaculture Workshop.

Stellenbosch University http://scholar.sun.ac.za

CHAPTER 7 – SUMMARIES AND CONCLUSIONS

7.1) MICROSATELLITE DESIGN

The FIASCO method (Zane *et al*, 2002) for microsatellite enrichment was successfully used in the South African abalone species, *Haliotis midae*. Several species-specific microsatellite markers (the first for *Haliotis midae*) were designed, of which eight were used for this study. Bester *et al* (in press; Appendix A) published the newly designed microsatellite markers.

7.1.1) Future

Further microsatellite enrichments will be done in and more microsatellite markers will be designed and characterised for *Haliotis midae*.

7.2) GENETIC DIVERSITY AND DIFFERENTIATION

The current brood stock of Sea Plant Products was representative of the natural population in terms of average number of observed alleles, heterozygosity levels and genetic differentiation (see sections 4.2; p95 – 99; 4.3; p99 – 100; 4.6; p104 – 110).

This study detected a loss of genetic variation within the F1 of Sea Plant Products (see section 4.2, 95 – 99). The most likely causes for the decline were the subdivision of the original brood stock (see section 2.9.1 p51), the high levels of inbreeding (F_{is}; section 4.6.2; p109 – 110) and the differential contributions of parental animals (see section 5.4; p124 – 128). A bottleneck effect was thus present caused by a founder effect.

No differentiation was detected between the farm populations and the wild populations or between the different farm populations or between the Saldanha Bay and Black Rock populations. The only population that show

Stellenbosch University http://scholar.sun.ac.za

differentiation from all the other populations was Prog2 (F_{st}; see section 4.6; p104 – 109), but this was due to the differential contributions of the parental animals from Row 2 (see section 5.4.1;p 124 – 125).

7.2.1) FUTURE GENETIC MANAGEMENT

The current brood stock does not need to be replenished or added to. If the brood stock is replenished or added to from wild populations or F1, it will be important to study the genetic profiles of the newly introduced animals. The profiles will be used to assess impact on the present levels of genetic variation within the farm.

The problem of the differential contributions of parental animals can be solved by revising the number of animals in each row to ensure a spawn ratio of 2:3 (males:females).

A rotational mating system can be used to increase the overall genetic variation within the F1 by moving the male animals from one row to the next on a yearly basis. Care should be taken if the F1 is considered to replenish or enlarge the existing brood stock. The backcross between the less genetic variable F1 and its corresponding brood stock will increase the loss of genetic variation. A strategy for the incorporation of the F1 into the brood stock is shown in Figure 7.1. The F1 of Row 2 must be incorporated into the brood stock of Row 3 and so on, and thus avoid backcrossing (pers. comm.: Dr Danie Brink, University of Stellenbosch, Department of Genetics, Aquaculture Division).

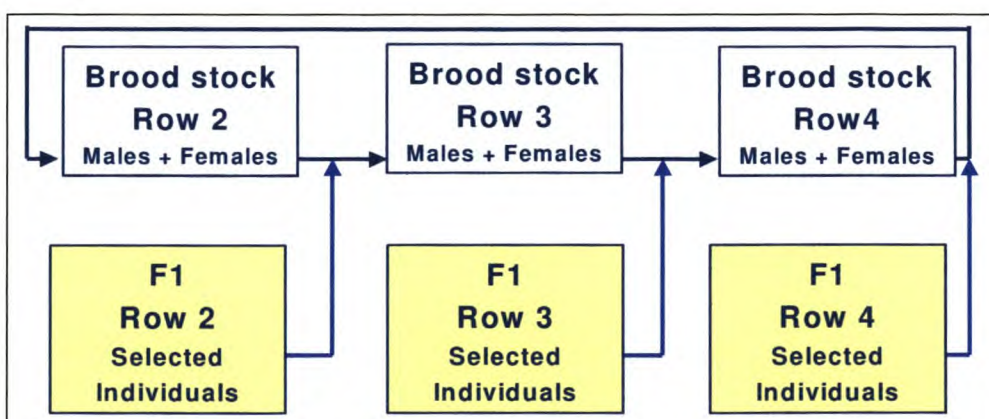


Figure 7.1: A strategy for incorporating the F1 into the brood stock.
Dr. D.Brink (University of Stellenbosch, Department of Genetics, Aquaculture Division)

Stellenbosch University http://scholar.sun.ac.za

The brood stock and F1 of the farm poses no environmental danger in terms of genetic composition, because no differentiation was found between the farm and the natural populations (see above). Results also suggest that the F1 can be used for future reseeding or restocking projects of the natural populations. It will however, be essential to study the genetic profiles of the animals before they are reintroduced. The profiles will be used to determine whether or not the chosen animals show the desired level of genetic variation and whether or not they are genetically similar to the local natural population (see sections 1.2.2 & 1.2.3; p6 – 10 for more information).

7.3) PARENTAGE

Parentage was inferred for the first time in *Haliotis midae* using three microsatellite loci (see Chapter 5; p114). Thirty-eight percent of all studied offspring (151 individuals) were confidently assigned to a male and female pair. Null alleles, missing data sets, scoring difficulties with locus HmD59 and individuals that carried more than one homozygous locus hampered the success of the assignments.

Parentage analysis was useful in genetic diversity studies as well. The differential contributions of the parents to the offspring (see section 5.4, 124 – 128) could also be analysed and used to explain the loss of genetic diversity within the farm population.

Parentage analysis was also used in QTL discovery. Parental pairs that spawned only fast growing offspring was identified. These animals could be used in future experiments for QTL-mapping in abalone.

A computer program, EPC v1.0 (Slabbert, 2004; Appendix B) was written to calculate the exclusion probabilities of individual loci as well as combinations of loci.

Stellenbosch University http://scholar.sun.ac.za

7.3.1) FUTURE WORK

The accuracy of parental assignments can be optimised by testing more microsatellite loci.

7.4) QTL-DISCOVERY

Alleles and genotypes segregating with growth rate was the target of this analysis. All the results of this QTL search (see Chapter 6; p130) must be seen as **PRELIMINARY**. The accuracy of the study was influenced by the small sample sizes and the few microsatellite loci that were used.

A number of alleles and genotypes were found that segregated with growth rate. Summaries can be seen in Tables 6.15 – 6.17, 154 – 156.

7.4.1) FUTURE

Future QTL-studies must be planned carefully (use Robinson *et al* 2003 as guideline). The appropriate number of individuals must be sampled (75+; Robinson *et al*, 2003) to increase the experimental power of the study. Backcrossing family groups that will be used to study the segregation of potential QTL-linked loci must be created and maintained under the correct physiological, biological and environmental conditions. More microsatellite loci must also be used to increase the likelihood of finding segregating alleles.

Stellenbosch University http://scholar.sun.ac.za

7.5) REFERENCES

Bester, A.E., Slabbert, R., D'Amato, M.E. Isolation and characterisation of microsatellite markers in South African abalone (*Haliotis midae*). In press. *Molecular Ecology Notes*. (Appendix A)

Robinson, N., Baranski, M., Rourke, M., Hayes, B., 2003. Application of advanced genetic techniques to abalone breeding.10th Abalone Aquaculture Workshop.

Zane, L., Bargelloni, L., Patarnello, T., 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1 – 16.

Stellenbosch University http://scholar.sun.ac.za

APPENDIX A

Bester *et al*, in press

Stellenbosch University http://scholar.sun.ac.za

Molecular Ecology Notes (2004) doi: 10.1111/j.1471-8286.2004.00755.x

PRIMER NOTE

Isolation and characterization of microsatellite markers in the South African abalone (*Haliotis midae*)

ALETTA E. BESTER, RUHAN SLABBERT and MARÍA EUGENIA D'AMATO
Department of Genetics, University of Stellenbosch, Private BagXI, 7602 Matieland, Stellenbosch, South Africa

Abstract

We report the isolation and characterization of 11 polymorphic microsatellite loci in the South African abalone *Haliotis midae*. These loci showed a range of five to 21 alleles per locus and observed heterozygosities ranging from 0.14 to 0.93 in a wild population of 32 individuals. All loci except four conformed to Hardy–Weinberg expectations and did not show linkage disequilibrium. The polymorphism exhibited at these loci indicate that they would be useful in determining levels of genetic variability in natural and commercial *Haliotis midae* populations as well as in parentage and Quantitative Trait Loci (QTL) analysis in hatchery reared abalone.

Keywords: abalone, *Haliotis*, microsatellites

Received 18 May 2004; revision received 23 June 2004; accepted 23 June 2004

Haliotidae (Gastropoda) includes about 90 species of abalone distributed in coastal waters of all continents, of which 15 are subjected to commercial exploitation in Australia, Japan, South Africa, North America and Southeast Asia. In South Africa, *Haliotis midae* is the only one of the six South African species with commercial potential.

A combined effect of over-exploitation and illegal harvesting has caused the decline of this species, to the extent that local extinctions have been forecasted to take place in the next 5 years (http://eces.org/articles/000263.php). Research has been conducted to evaluate the feasibility of stock enhancement or ranching of *Haliotis midae* in South Africa (Sweijd *et al.* 1998; De Waal *et al.* 2003). Both restocking programmes and genetic management in aquaculture practices require the application of molecular markers. Here we report a set of 11 polymorphic microsatellites used to investigate the genetic variability of wild and commercial *Haliotis midae*.

Genomic DNA was isolated from mantle tissue following a standard CTAB extraction method (Saghai Maroof *et al.* 1984). Tissue was homogenized in 700 µL of CTAB lysis buffer containing 0.5 mg/mL Proteinase K and incubated at 60 °C. Following phenol-chloroform: isoamyl alcohol (25 : 24 : 1) extractions, the supernatant was precipitated with two volumes of 100% cold ethanol. DNA was redissolved in 100 µL of distilled water and stored at –20 °C.

In this study, microsatellite repeat sequences were isolated using an enrichment technique (FIASCO) described by Zane *et al.* 2002. An enriched partial genomic library was constructed using DNA from a single individual. For this, 250 ng DNA was simultaneously digested with *Mse*I and ligated to *Mse*I AFLP adaptors. DNA was selectively amplified using a mixture of four adaptor specific primers (*Mse*I-N) and hybridized independently with a biotinylated $(AC)_{12}$ and a $(GATC)_6$ probe. Repeat-containing fragments were recovered by streptavidin magnetic particles and cloned into a TOPO-TA cloning vector (Invitrogen) in order to produce a highly enriched microsatellite library.

Approximately 1200 recombinant clones were obtained, of which 250 were sequenced on an ABI 3100 Automated Sequencer to verify the presence of repeat sequences. Forty-five percent of the clones contained repeat sequences but only a small number had sufficient flanking regions for primer design. Oligonucleotide primers were designed for 20 loci using the program OLIGO™. version 4.0 (National Biosciences Inc.). Using the template DNA from which these loci were isolated, amplification products of expected size were obtained for 16 of the 20 microsatellite loci isolated.

A total of 32 individuals from Black Rock, on the east coast of South Africa, were genotyped to test the polymorphism of the markers. For each primer pair, one of the primers was labelled with FAM, NED, VIC or PET dyes. All polymerase chain reactions (PCR) were conducted in a Geneamp 2700 thermo cycler (Applied Biosystems) in 10 µL reactions containing 20 ng DNA, 0.3 µM of each primer,

Correspondence: A. E. Bester. Fax: ; E-mail: aeb@sun.ac.za

© 2004 Blackwell Publishing Ltd

Stellenbosch University http://scholar.sun.ac.za

Table 1 Primer sequences and characteristics of 11 *Haliotis midae* microsatellite loci

Locus	Repeat Sequence	Primer Sequence (5′–3′)	n	Size Range	T_a °C	H_O	H_E	Allele no.	Acc Nr
HmD14	$(CA)_{10}$	F TAAGGCAAGTGAATGTCTAG R ATTGCAAGAATCACAACTGC	27	142–180	60	0.67	0.76	16	AY303333
HmD33	$(GAGT)_{12}AAGT(GAGT)_6$	F TTGAAAGTGAACCAAAATCTG R CATGGGTACAATGTGTAAAGC	22	129–205	59	0.32	0.87	11	AY303334
HmD36	$(GTGA)_{14}$	F AGATCGAATGACATCAGCTTC R CATATAGCAAGCCTGAAACC	23	220–304	60	0.43	0.89	15	AY303335
HmD55	$(GTGA)_{12}$	F ATCAAGATAAAACGAGGCG R ACCACTGTGAAAACGTCCA	32	183–211	60	0.68	0.8	9	AY303337
HmD59	$(CA)_{15}$	F TATACTGCCATTTCCGTCTG R TCTGTATTCTGGTCCTGTCG	32	106–150	60	0.78	0.84	15	AY303338
HmD60	$(CA)_{16}$	F AAGTTGTTCTCCATAAAGTCGTA R GAAGATCCGGGTTAGAACTG	14	155–171	60	0.14	0.86	8	AY303339
HmD61	$(CA)_{24}$	F GATATCCAACCCCTGATCAC R GAACATCAACATCTCCATGG	28	234–298	60	0.61	0.82	11	AY303340
HmD11	$(TCTG)_8$	F AGCTCAGAAAAGTGGTGTACG R TTACCTAGCTAAAGTTGACAACG	30	292–352	61	0.32	0.66	5	AY303341
HmD30	$(AGTC)_2GGTC(AGTC)_{11}$	F TGATGTTGCTGGAATATTGC R CAATTTCATTTTCAACAGTTCA	27	124–150	60	0.7	0.8	11	AY303342
HmSP1	$(CA)_{10}(CGCA)_2(CA)_4$	F ATAGTGGTCATACAGTCATCACCT R TAGGCATGTTTGAGTTCGTGT	21	192–276	61	0.48	0.93	21	AY303346
HmSP5	$(AC)_{13}$	F TTCGGCAAGTGAATGTCTAG R ATGCGACACTTACTACACCG	31	185–219	60	0.63	0.74	14	AY303344

T_a, optimum annealing temperature; H_O, observed heterozygosity; H_E, expected heterozygosity.

200 μM dNTP's, 0.1 unit of *Taq* polymerase (Promega), 1 × PCR Buffer A (Promega) and 2 mM $MgCl_2$. PCR consisted of an initial denaturing step at 94 °C for 5 min followed by 35 cycles of 30 s at 94 °C, 30 s at 59–61 °C, and 1 min at 72 °C, and a final extension for 10 min at 70 °C. PCR products were separated on a ABI 3100 Automated Sequencer and analysed using the GENESCAN software program (Applied Biosystems).

Eleven loci showed polymorphism. PCR primer sequences, optimal annealing temperature, repeat motif and allele size ranges are shown in Table 1. Observed and expected heterozygosities and probability of Hardy–Weinberg equilibrium (PHWE) were calculated using GENETIX version 4.02 (Belkhir *et al.* 2000). Allelic number ranged from five to 21. However, more alleles are likely to be present for loci HmD14 and HmSP5, due to an apparent allele dropout effect. An extensive PCR optimization (variation in DNA, primer and $MgCl_2$ concentration), following Goossens *et al.* (1998) could not overcome the problem. Presence of null alleles is suspected in loci HmD33, HmD36, HmD60 and HmSP1 because of their departure from HWE ($P < 0.01$). All other loci conformed to HWE expectations and no linkage disequilibrium was detected. In summary, these loci exhibited high levels of polymorphism and heterozygosity and provide an invaluable tool for analysing genetic diversity of wild and commercial *Haliotis midae* populations. In addition, these markers could be used for parentage analysis and possible QTL identification in hatchery reared abalone.

Acknowledgements

We are indebted to Brian Godfrey for collecting samples of the test population in Black Rock. This work was partially funded by a Claude Harris Leon Foundation research grant.

References

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (2000) *GENETIX, logiciel sous Window TM pour la génétique des populations*. Laboratoire Génome et Populations, Montpellier, France.

De Waal SWP, Branch GM, Navarro R (2003) Interpreting evidence of dispersal by *Haliotis midae* juveniles seeded in the wild. *Aquaculture*, **221**, 299–310.

Goossens B, Waits LP, Taberlet P (1998) Plucked hair samples as a source of DNA: reliability of dinucleotide microsatellite genotyping. *Molecular Ecology*, **7**, 1237–1241.

Saghai Maroof MA, Solima KM, Jorgenson RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences (USA)*, **81**, 8014–8018.

Sweijd N, Snethlage Q, Harvey D, Cook P (1998) Experimental abalone (*H. midae*) seeding in South Africa. *Journal of Shellfish Research*, **17**, 897–904.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.

Stellenbosch University http://scholar.sun.ac.za

APPENDIX B

EPC V1.0

Stellenbosch University http://scholar.sun.ac.za

EPC V1.0

TABLE OF CONTENTS

Stellenbosch University http://scholar.sun.ac.za

EPC V1.0 (EXCLUSION PROBABILITY CALCULATOR)

B.1) INTRODUCTION

EPC v1.0 is used to calculate the exclusion probability using the allele frequencies of the loci used for parentage assignments. The formulae that is used, was taken from Jamieson & Taylor, 1997.

The exclusion probabilities for assigning the first parent, the second parent or a couple to the offspring can be calculated. Exclusion probabilities can be calculated over individual loci or over different combinations of loci.

The output is given in a file with extension *.exp and is saved in the C-drive.

The program is given on a stiffy attached to the manuscript.

B.2) INPUT

Unfortunately all data must be input manually, but this will be rectified in the next version.
Input required: (1) Loci names, (2) allele number and (3) allele frequencies.
An example of how to input data follows at the end.

B.3) OPTIONS AVAILABLE

"Paternity": Calculates the exclusion probability for second parent.
"One Genotype Missing": Calculates the exclusion probability for first parent.
"Exclude Both Parents": Calculates the exclusion probability for a couple.
"General": Calculates the exclusion probability if $p_1 \neq p_2 \neq p_3 \ldots p_x$ (p = allele frequency).
"Pmax": Calculates the exclusion probability if $p_1 = p_2 = p_3 \ldots p_x$ (p = allele frequency).

Stellenbosch University http://scholar.sun.ac.za

B.4) FORMULAE USED

Exclusion of first parent (One Genotype Missing):

$$P = 1 - 4\sum_{i=1}^{n} p_i^2 + 2\left(\sum_{i=1}^{n} p_i^2\right)^2 + 4\sum_{i=1}^{n} p_i^3 - 3\sum_{i=1}^{n} p_i^4$$

where

$(p_1 \neq p_2 \ldots p_i)$

p = allele frequency

Exclusion of second parent (Parenity):

$$P = 1 - 2\sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} p_i^3 + 2\sum_{i=1}^{n} p_i^4 - 3\sum_{i=1}^{n} p_i^{5^2} - 2\left(\sum_{i=1}^{n} p_i^2\right)^2 + 3\sum_{i=1}^{n} p_i^2 \sum_{i=1}^{n} p_i^3$$

where

$(p_1 \neq p_2 \ldots p_i)$

p = allele frequency

Exclusion of couple (Exclude Both Parents):

$$P = 1 + 4\sum_{i=1}^{n} p_i^4 - 4\sum_{i=1}^{n} p_i^5 - 3\sum_{i=1}^{n} p_i^6 - 8\left(\sum_{i=1}^{n} p_i^2\right)^2 + 8\left(\sum_{i=1}^{n} p_i^2\right)\left(\sum_{i=1}^{n} p_i^3\right) + 2\left(\sum_{i=1}^{n} p_i^3\right)^2$$

where

$(p_1 \neq p_2 \ldots p_i)$

p = allele frequency

Combined Exclusion Probability for combination of loci (k)

$$P = 1 - (1 - P_1)(1 - P_2)(1 - P_3)\cdots(1 - P_k)$$

(Jamieson & Taylor, 1997)

Stellenbosch University http://scholar.sun.ac.za

B.5) FUTURE

The future version will have a more user-friendly input interface, as well as other features.

B.6) EXAMPLE OF INPUT

B.6.1) STEP 1

Choose a file name for the output by clicking NEW (new file) or MODIFY (existing file) – Figure A. It will be saved with extension *.exp under the C-drive.

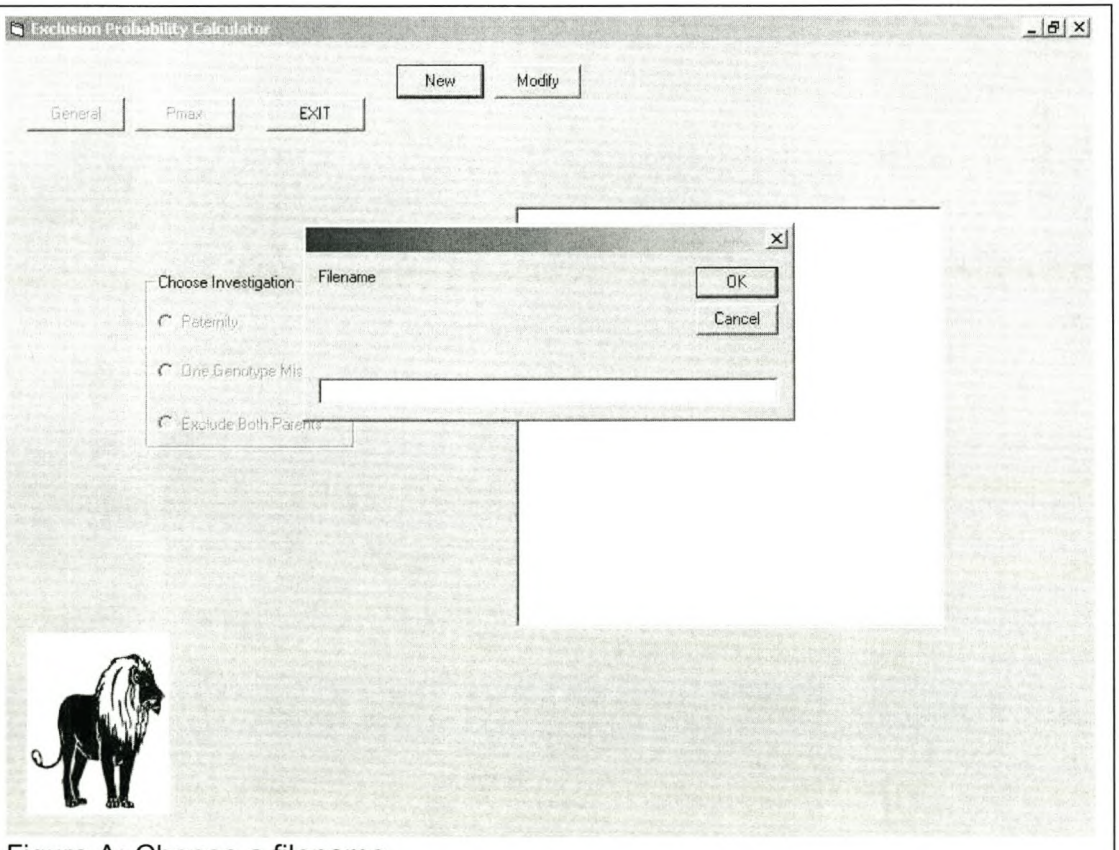


Figure A: Choose a filename.

Stellenbosch University http://scholar.sun.ac.za

B.6.2) STEP 2

Choose one of the options indicated by arrow (Figure B).

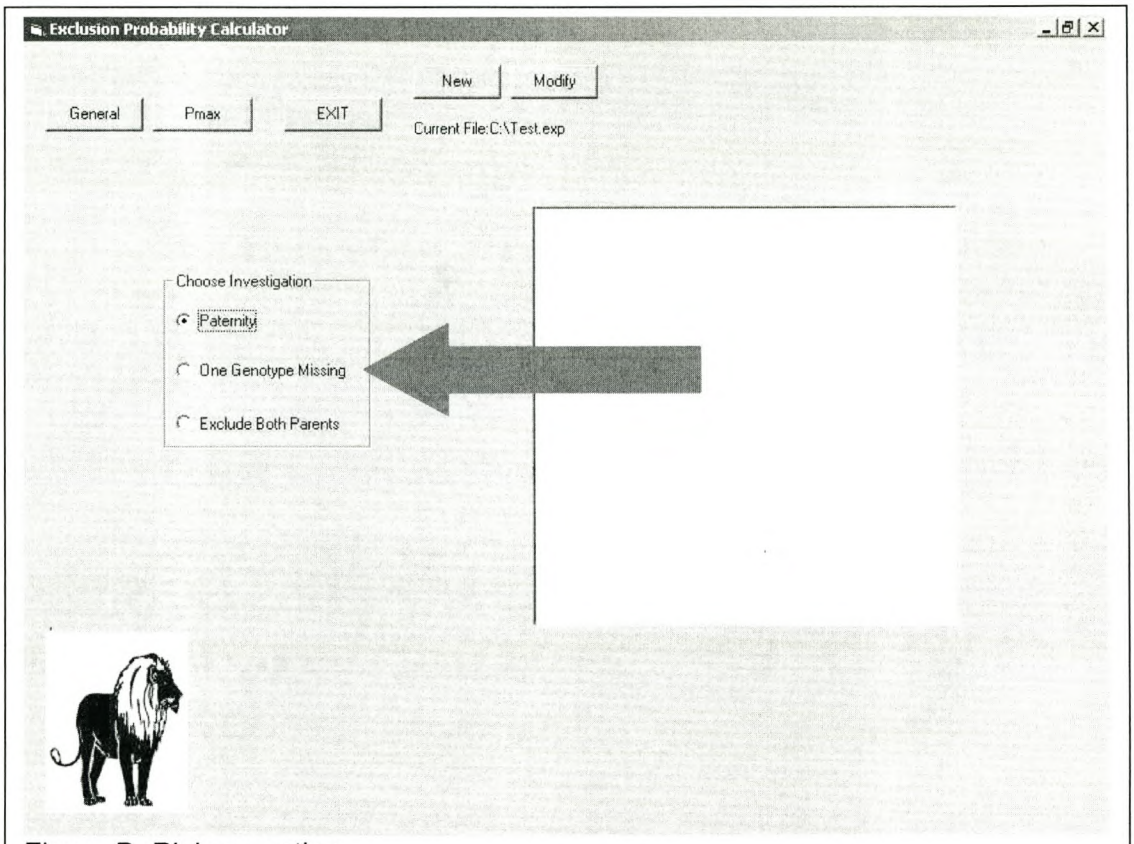


Figure B: Pick an option.

Stellenbosch University http://scholar.sun.ac.za

B.6.3) STEP 3

Choose one of the options General or Pmax (Figure C).

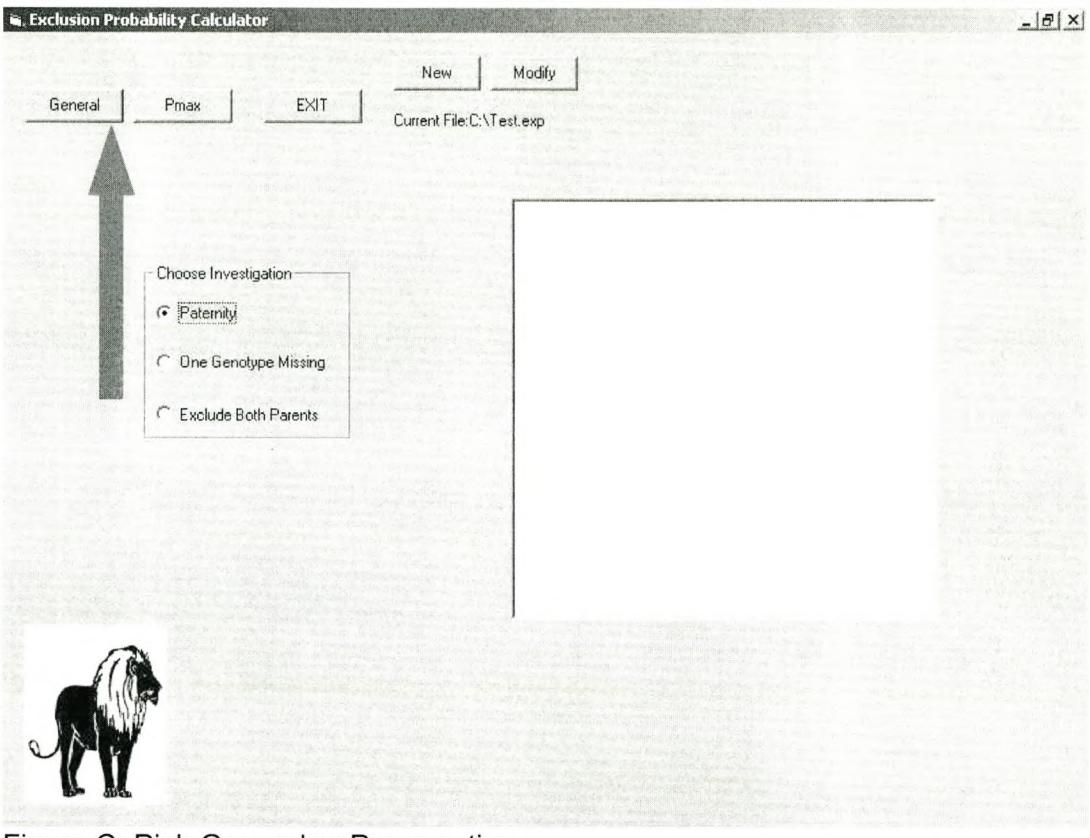


Figure C: Pick General or Pmax option.

Stellenbosch University http://scholar.sun.ac.za

B.6.4) STEP 4

Give the number of loci you wish to analyse (Figure D).

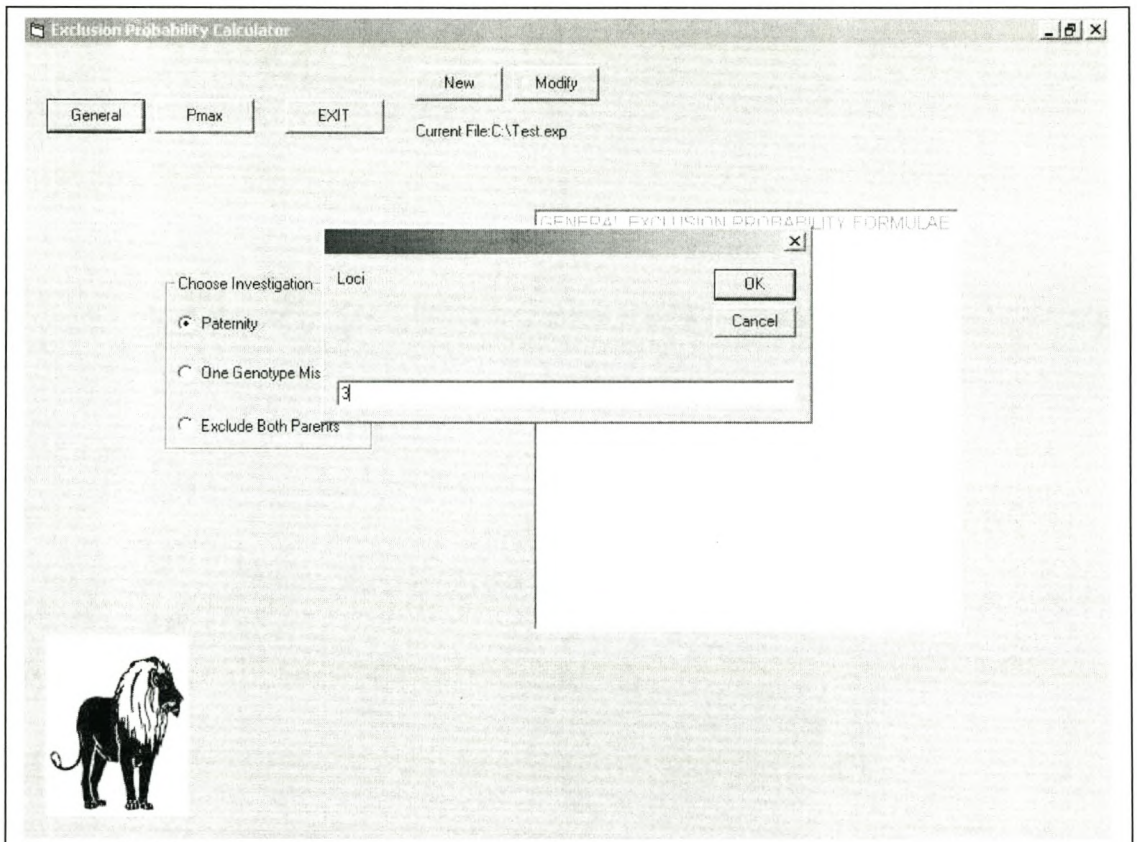


Figure D: Number of loci input.

Stellenbosch University http://scholar.sun.ac.za

B.6.5) STEP 5

Give the name of the first locus (Figure E).

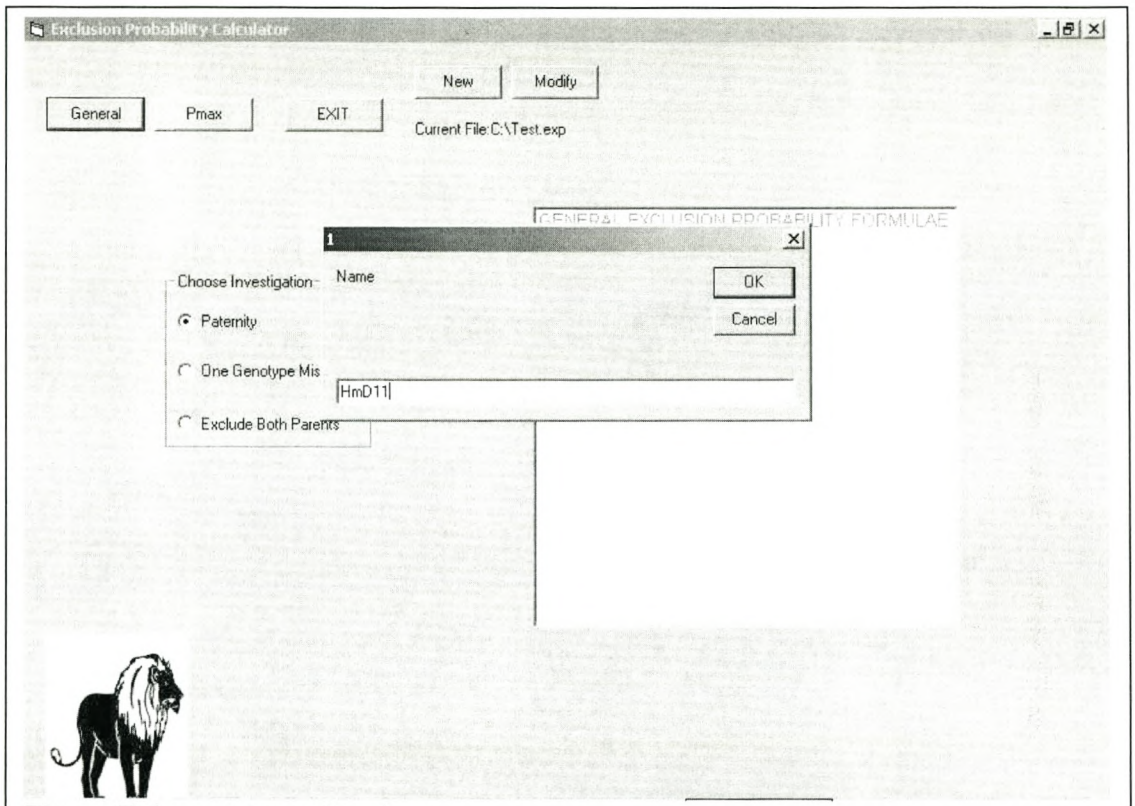


Figure E: Locus name input.

Stellenbosch University http://scholar.sun.ac.za

B.6.6) STEP 6

Give the number of alleles of the locus (Figure F).

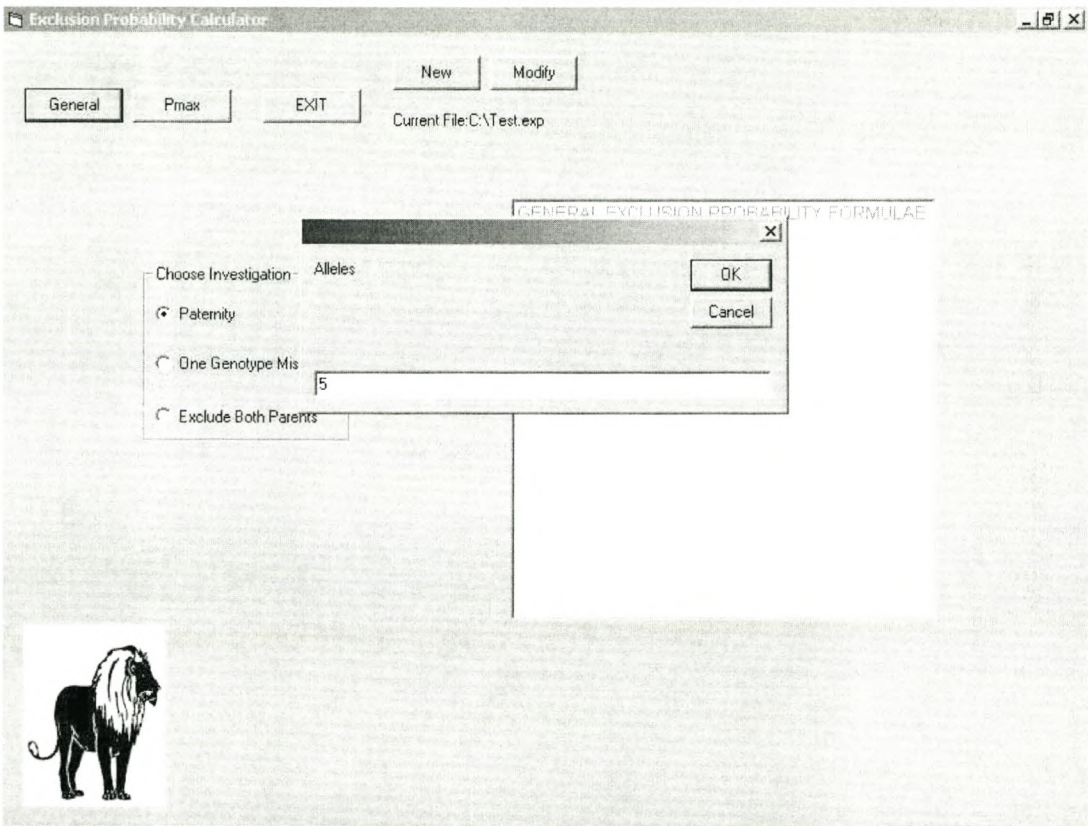


Figure F: Number of alleles input.

Stellenbosch University http://scholar.sun.ac.za

B.6.7) STEP 7

Give the allele frequency of the first allele (Figure G). Repeat this until the frequencies of all the alleles were input.

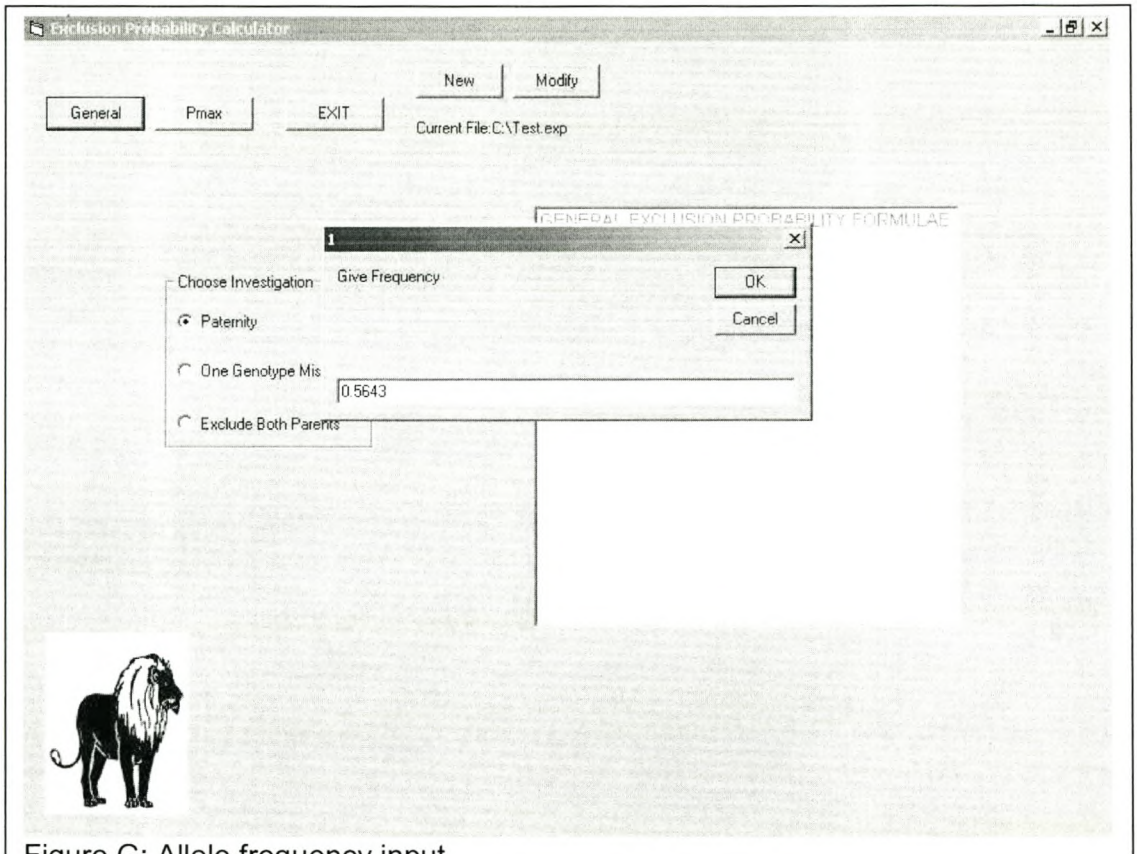


Figure G: Allele frequency input.

Stellenbosch University http://scholar.sun.ac.za

B.6.8) STEP 8

If more than one locus is being analysed, repeat STEP 5 – 7 for each locus.

B.6.9) STEP 9

View output file by opening it with Notepad.

B.7) REFERENCES

Jamieson, A., Taylor, St.C.S., 1997. Comparisons of three probability formulae for parentage exclusion. *Animal Genetics* **28**, 397 – 400.

Stellenbosch University http://scholar.sun.ac.za

APPENDIX C

p: Parents + Offspring

Stellenbosch University http://scholar.sun.ac.za

APPENDIX C – ALLELE FREQUENCIES (PARENTS + OFFSPRING)

Locus	Parents + Offspring: Allele Frequencies	Locus	Parents + Offspring: Allele Frequencies	Locus	Parents + Offspring: Allele Frequencies	Locus	Parents + Offspring: Allele Frequencies	Locus	Parents + Offspring: Allele Frequencies
A11		continued…		continued…		continued…		continued…	
288	0.0018	**AD33**		**AD133**		**185**	0.0017	**122**	0.1220
292	0.0482	**192**	0.0175	**179**	0.0017	**187**	0.3041	**124**	0.0712
296	0.0411	**194**	0.0461	**181**	0.0017	**191**	0.1564	**126**	0.0864
304	0.4018	**196**	0.0088	**183**	0.0188	**195**	0.1564	**128**	0.1068
308	0.0446	**198**	0.0746	**185**	0.7260	**197**	0.0017	**130**	0.0712
312	0.4429	**200**	0.0175	**187**	0.0240	**199**	0.0739	**132**	0.0085
316	0.0179	**202**	0.0154	**189**	0.0462	**201**	0.0412	**134**	0.0186
320	0.0018	**204**	0.0482	**191**	0.0017	**203**	0.0361	**136**	0.0322
A30		**206**	0.0614	**193**	0.0462	**205**	0.0017	**140**	0.0051
124	0.0065	**208**	0.2171	**195**	0.0068	**207**	0.0567	**144**	0.0102
126	0.0022	**210**	0.0614	**197**	0.0051	**211**	0.0034	**146**	0.0017
128	0.0217	**212**	0.1667	**199**	0.0188	**217**	0.0017	**148**	0.0153
132	0.2152	**214**	0.0439	**201**	0.0171	**219**	0.0258	**150**	0.0017
136	0.0130	**216**	0.0811	**203**	0.0120	**231**	0.0069		
138	0.3935	**218**	0.0175	**205**	0.0137	**A59**			
140	0.0587	**220**	0.0417	**207**	0.0120	**106**	0.0271		
142	0.0565	**222**	0.0154	**211**	0.0257	**108**	0.0051		
144	0.1457	**224**	0.0066	**215**	0.0137	**110**	0.0644		
146	0.0283	**226**	0.0088	**217**	0.0051	**112**	0.0932		
148	0.0152	**228**	0.0022	**221**	0.0034	**114**	0.0661		
150	0.0152	**232**	0.0022	**A55**		**116**	0.0797		
154	0.0109	**240**	0.0417	**179**	0.0172	**118**	0.0407		
158	0.0174	**242**	0.0044	**183**	0.1151	**120**	0.0729		
	continues…		continues…		continues…		continues…		

Stellenbosch University http://scholar.sun.ac.za

APPENDIX D

p: Large + Small

Stellenbosch University http://scholar.sun.ac.za

APPENDIX D – ALLELE FREQUENCIES (LARGE + SMALL INDIVIDUALS)

LOCUS	LARGE INDIVIDUALS	SMALL INDIVIDUALS
HmD11		
292	0.1078	0.0686
296	0.1275	0.0882
304	0.4118	0.4510
308	-	0.0686
312	0.3529	0.3235
HmD30		
124	-	0.0119
128	0.0286	-
132	0.2571	0.3929
138	0.4000	0.2619
140	0.0286	0.1190
142	0.0429	0.0238
144	0.2000	0.1548
146	0.0286	0.0238
148	0.0143	-
154	-	0.0119
HmSP5		
185	0.7857	0.6667
187	0.0268	0.0263
189	0.0089	0.0789
193	0.0804	0.1140
197	-	0.0088
199	0.0089	0.0351
201	0.0446	0.0263
203	0.0089	0.0088
205	0.0179	-
207	0.0089	-
211	0.0089	0.0088
215	-	0.0263
HmD55		
183	0.0763	0.0714
187	0.5085	0.4375
191	0.1102	0.1875
195	0.0763	0.1429
199	0.0593	0.0536
203	0.0254	0.0179
207	0.1356	0.0893
219	0.0085	-

continues...

LOCUS	LARGE INDIVIDUALS	SMALL INDIVIDUALS
continued...		
HmD59		
106	-	0.0179
110	0.0364	0.0179
112	0.1273	0.1339
114	0.0545	0.1250
116	0.0818	0.1161
118	0.0636	0.0268
120	0.0727	0.0536
122	0.0818	0.0804
124	0.0364	0.0893
126	0.1364	0.0446
128	0.1455	0.1339
130	0.0909	0.125
132	0.0091	-
136	0.0636	0.0357
HmD14		
142	0.8596	0.6727
144	-	0.0273
146	-	0.0818
150	0.1053	0.1273
154	-	0.0091
156	-	0.0273
158	0.0351	0.0273
160	-	0.0273
HmT35		
135	0.2018	0.0982
139	0.7544	0.8929
143	0.0439	0.0089

Stellenbosch University http://scholar.sun.ac.za

APPENDIX E

χ^2-TABLE

Stellenbosch University http://scholar.sun.ac.za

APPENDIX E - χ^2-TABLE

Table D1: Table showing significant values for χ^2-test.

df	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.01	0.02	0.051	0.103	0.211	4.605	5.991	7.378	9.21	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.86
5	0.412	0.554	0.831	1.145	1.61	9.236	11.07	12.833	15.086	16.75
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.69	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.18	2.733	3.49	13.362	15.507	17.535	20.09	21.955
9	1.735	2.088	2.7	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.94	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.92	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.3
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.66	5.629	6.571	7.79	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.39	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.26	9.591	10.851	12.443	28.412	31.41	34.17	37.566	39.997
21	8.034	8.897	10.283	11.591	13.24	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.26	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.98	45.559
25	10.52	11.524	13.12	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.16	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.29
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.42	76.154	79.49
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.54	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169