Microsatellite genotyping of contributing broodstock and selected offspring of *Haliotis midae* submitted to a growth performance recording scheme

Nicola Ribeiro Ruivo

Thesis presented in partial fulfilment of the requirements of the degree of Master of Science (M.Sc.) in Genetics at the University of Stellenbosch.

Supervisor: Dr Rouvay Roodt-Wilding

December 2007

DECLARATION

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

Signed: _____ **Date:** _____

Nicola Ruivo

SUMMARY

The indigenous abalone *Haliotis midae* is one of the most remarkable and highly exploited species of marine molluscs in South Africa. It is the only species of southern African Haliotidae to be commercially reared and has been successfully cultured for almost two decades. Its short history of domestication along with market demands and the need to develop efficiency in the production process has resulted in an increased interest in the possible genetic improvement of this species. The unhurried growth rate associated with *H. midae* is a cause of particular concern to the industry, predominantly with regards to profitability and competitiveness in the market place. A modest amount of work has so far been directed at establishing a means of enhancement for selective breeding on the commercial level. Genetics plays a key role in the establishment of successful improvement programmes in various aquaculture species. The aim of this study was to develop species-specific microsatellite markers for the abalone and subsequently perform parentage assignment on farm produced animals entered into a growth performance recording scheme. Animals were obtained from the hatcheries of three commercial abalone farms situated in the Walker Bay region in the Western Cape.

Microsatellites were isolated using the enrichment-based FIASCO method, and characterised into perfect, imperfect and compound repeats according to the structural nature of their repetitive units. From the partial gDNA libraries obtained and 365 screened colonies, a total of 54 loci were located. PCR primers were designed for 36 markers and the 15 primer pairs that displayed loci with the highest level of polymorphism were subsequently chosen for fluorescent labelling. The markers were tested on a subset of 32 wild *H. midae* individuals to determine their usefulness and efficiency in genotyping. Five markers, along with five others that were previously designed, were chosen for assigning parentage to the animals submitted to the performance recording scheme. Three thousand offspring from each of the three participating farms were equally divided and reared at five different locations. From each location 20 fast growing and 20 slow growing juveniles, as well as the broodstocks, were sampled and genotyped using the ten chosen microsatellite loci. Two farms had 60% of offspring unambiguously assigned to a single parental couple. Assignments showed patterns of dominant male and female brooders, but no trend in brooders specifically contributing to fast or slow growing offspring. Parentage assignment for the third farm was, however, unsuccessful due to lack of broodstock data. In future, screening of all available broodstock will ensure acquisition of relevant pedigree information. The results obtained in this study are an initial step in the development of a genetic improvement programme for commercial *Haliotis midae*.

OPSOMMING

Die inheemse skulpvis *Haliotis midae* is een van die mees merkwaardige en hoogs oorbenutte mariene slakspesies in Suid-Afrika. Dit is die enigste suidelike Afrika Haliotidae spesie wat kommersieel benut word en dit word al meer as twee dekades suksesvol geteel. Die spesie se kort domestiseringsgeskiedenis, toenemende mark aanvraag en die behoefte om meer effektiewe produksie daar te stel, het gelei tot toenemende belangstelling in die moontlike genetiese verbetering van die spesie. Die stadige groeitempo geassosieer met *H. midae* is veral ‘n punt van kommer vir die industrie, veral in terme van winsgewendheid en kompetering in die markplek. Minimale werk is sover gedoen in die daarstelling van verbetering deur selektiewe teling op ‘n kommersiële skaal. Genetika speel ’n sleutelrol in die daarstelling van suksesvolle verbeteringsprogramme van verskeie akwakultuur spesies. Die doel van hierdie studie was om spesie-spesifieke mikrosatelliet merkers vir perlemoen te ontwikkel en vervolgens ouerskapsbepaling van kommersiële diere, wat deelneem aan ‘n groeiprestasie aantekenstelsel, uit te voer. Diere is voorsien deur die teelstasies van drie kommersiële perlemoenplase geleë in die Walker Bay omgewing in die Wes-Kaap.

Mikrosatelliete is geïsoleer deur die verrykings-gebaseerde FIASCO metode, en gekarakteriseer as perfekte, onderbroke of saamgestelde herhalings gebaseer op die strukturele aard van die herhalings eenhede. Vanaf die gedeeltelik gDNA biblioteke wat bekom is en 365 gesifte kolonies, is ‘n totaal van 54 loki opgespoor. PKR inleiers is ontwerp vir 36 merkers en die 15 inleierpare, wat loki met die hoogste polimorfisme geamplifiseer het, is vervolgens geselekteer vir fluoreserende merking. Die merkers is getoets op ’n kleiner groep van 32 natuurlike *H. midae* individue om hulle bruikbaarheid en genotiperingseffektiwiteit te bepaal. Vyf merkers is saam met vyf reeds ontwikkelde merkers gekies vir ouerskapsbepaling van die diere in die prestasie aantekenstelsel. Drieduisend nageslag diere vanaf elkeen van die drie deelnemde plase is gelykop verdeel en grootgemaak op die vyf verskillende lokaliteite. ‘n Monster van 20 vinnig groeiende en 20 stadig groeiende jong perlemoen, sowel as broeidiere, is vanaf elke lokaliteit geneem en gegenotipeer deur middel van die 10 geselekteerde mikrosatelliet loki. Sestig persent van twee van die plase se nageslag is onteenseglik toegesê aan ‘n enkele ouerpaar. Ouerskapstoekenning het patrone van dominante vroulike en manlike broeidiere getoon, maar geen tendens in terme van bydrae tot vinnig en stadig groeiende nageslag kon gevind word nie. Ouerskapstoekenning vir die derde plaas was onsuksesvol as gevolg van ’n gebrek aan data vir die broeidiere. In die toekoms sal genotipering van alle beskikbare broeidiere die daarstelling van relevante stamboominligting verseker. Die resultate verkry in hierdie studie verteenwoordig ‘n eerste stap in die ontwikkeling van ’n genetiese verbeteringsprogram vir kommersiële *Haliotis midae*.

For my mom & dad, who gave me the courage to pursue my dreams.

ACKNOWLEDGEMENTS

Firstly, many thanks due to Dr Rouvay Roodt-Wilding, my supervisor, for giving me the opportunity to be involved in this project, for motivating me to stay focused, for challenging me and never allowing me to give up, especially when I felt like there was no hope. I thank her for being an exceptional mentor and for expanding my mind to corners I thought were unreachable. I am truly grateful for all the support and guidance she has given over the last few years, and deeply indebted to her for her invaluable contribution to the format and contents of this thesis.

To my fellow laboratory colleagues who went through the daily dramas with me, I thank them all for putting up with me especially when I was at my most unbearable. The groups of honours students that came and went during my time in the department also deserve a special mention. Many thanks are due to Dr Helen Whitaker and Aletta van der Merwe for all their friendly technical and administrative assistance, and for being patient with my never-ending curiosity. I would like to acknowledge Dr Michael-John Freeborough for all the technical support and guidance he gave me whenever I felt I was in trouble. I could not have asked for better help than that I received from him.

I would like to extend my thanks to friends in the department, for all the times we shared together. Your support made the past three years an unforgettable and very colourful experience for me. Each one played their part in making the daily grind that little bit more fun and interesting. Thanks have to go to those girls who were there with me from day one, Aletta, Daleen, Jenn and Nicol. Very special thanks are sincerely due to go to my closest ally, Nicola Panton. An amazing person with an insurmountable amount of patience and compassion without whose constant reassurance and support I could not have gotten through this degree.

To my grandparents, aunts, uncle and cousins, I thank them all for being great influences in my life, and for supporting and encouraging me. Claudia, Monica and Nadia, who each played their role in me becoming who I am today, and baby Tiago who brought the much needed brightness and joy to an often serious person. To Natalie, my sister and my best friend, for never letting me down when I needed someone to lean on, for constantly challenging me and most of all, for helping me through all my ups and downs. Thank you, I love you dearly.

Last, but certainly not least, to the most important people and role models in my life. My mom and dad without whose encouragement, emotional support, and financial support, I would not have

written this. There are no words to describe the love and affection I feel for the both of them. I thank God everyday for having the life they have given to me. I wish to thank them for never letting me doubt myself and for bringing me up with the confidence to be whatever I can be. Their patience, pride and love have been the fuel that drove me in life. I love you both very much.

"Love and work are the cornerstones of our humanness."

Sigmund Freud

TABLE OF CONTENTS

CHAPTER 2: MICROSATELLITE ISOLATION & CHARACTERISATION FOR THE SOUTH ARFICAN ABALONE, *HALIOTIS MIDAE*

CHAPTER 3: ASSIGNMENT OF PARENTAGE TO HATCHERY-REARED ABALONE SUBMITTED TO A GROWTH PERFORMANCE RECORDING SCHEME

CHAPTER 4: CONCLUSIONS, IMPLICATIONS & FUTURE PROSPECTS

APPENDICES

LIST OF FIGURES

CHAPTER 1: General Introduction

CHAPTER 2: Microsatellite Isolation & Characterisation for the South African Abalone, *Haliotis midae*

LIST OF TABLES

CHAPTER 2: Microsatellite Isolation & Characterisation for the South African Abalone, *Haliotis midae*

CHAPTER 3: Assignment of Parentage to Hatchery-Reared Abalone Submitted to a Growth Performance Recording Scheme

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

≈	Approximately
β	Beta
χ^2	Chi-square
ºC	Degrees Celsius
>	Greater than
<	Less than
µ	Micro (10^{-6})
µg	Microgram
µl	Microlitre
µM	Micromolar
µg/µl	Microgram per microlitre
µg/ml	Microgram per millilitre
%	Percentage
®	Registered trademark
$	American dollars
™	Trademark
3’	3-prime
5’	5-prime

ABBREVIATIONS

A	Adenosine
AFASA	Abalone Farmers Association of South Africa
AFLP	Amplified fragment length polymorphism
APS	Ammonium persulphate: $(NH_4)_2S_2O_8$
ATP	Adenosine-5’-triphosphate
BLAST	Basic local alignment search tool
bp	Base pair
BSA	Bovine serum albumin
C	Cytosine
cM	Centimorgan
CTAB	N-cetyl-N,N,N-trimethylammoniumbromide
dATP	2’-deoxyadenosine-5’-triphosphate

dCTP	2'-deoxycytosine-5'-triphosphate
DEAT	Department of Environmental Affairs and Tourism
df	Degrees of freedom
dGTP	2'-deoxyguanine-5'-triphosphate
DNA	Deoxyribonucleic acid
dNTP	2'-deoxy-nucleotide-5'-triphosphate
DTT	Dithiothreitol
dTTP	2'-deoxythymine-5'-triphosphate
eg.	For example
EDTA	Ethylenediaminetetraacetate: $C_{10}H_{16}N_2O_8$
EST	Expressed sequence tag
et al.	And others
etc.	Etcetera
EtBr	Ethidium bromide: $C_{21}H_{20}BrN_3$
EtOH	Ethanol: CH_3CH_2OH
F	Forward primer
F_1	First generation of offspring
FAP	Family assignment program
FIASCO	Fast isolation by AFLP sequence containing repeats
g	Gram
G	Guanine
gDNA	Genomic deoxyribonucleic acid
HCl	Hydrogen chloride
H_e	Expected heterozygosity
H_o	Observed heterozygosity
hr	Hours
HWE	Hardy-Weinberg equilibrium
i.e.	That is / such as
IAM	Infinite allele model
IF	Innovation Fund
I&J	Irvin and Johnson
Inc.	Incorporated
kb	Kilo base
kg	Kilogram
LB	Luria-Bertani medium

LOD	Natural logarithm of the likelihood ratio
m	Milli (10^{-3})
M	Molar (moles per litre)
MAS	Marker-assisted selection
MCM	Marine and Coastal Management
$MgCl_2$	Magnesium Chloride
min	Minutes
ml	Millilitre
MLS	Minimum legal size
mM	Millimolar
mm	Millimetre
mRNA	Messenger ribonucleic acid
Mse I	*Micrococcus roseus*; 1st enzyme
mtDNA	Mitochondrial deoxyribonucleic acid
N	Indicative of one of A, C, G or T base pair
n	Nano (10^{-9})
N_a	Number of alleles
NaCl	Sodium chloride
NCBI	National Centre for Biotechnology information
ng	Nanogram
ng/µl	Nanogram per microlitre
NOR	Nucleolus organising region
NS	Non-stringency
P	Probability
PAA	Poly-acrylamide; C_3H_5NO
PCR	Polymerase chain reaction
pH	Concentration of hydrogen ions in a solution
PIC	Polymorphic information content
PRS	Performance recording scheme
QTL	Quantitative trait loci
r	Null allele frequency
R	Reverse primer
R	South African rand
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism

rpm	Rotations per minute
S	Stringency
SDS	Sodium-dodecyl sulphate
sec	Seconds
SMM	Stepwise mutation model
SNP	Single nucleotide polymorphism
SSC	Standard saline citrate
SSR	Simple sequence repeat
STR	Short tandem repeat
T	Thymine
TAC	Total allowable catch
Taq	*Thermus aquaticus*
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
TEMED	N,N,N',N'-tetramethylethylenediamine
TEN	Tris-ETDA-NaCl buffer
Tris-HCl	Tris hydrochloride
U	Units
UV	Ultraviolet
V	Volts
v/v	Volume per volume
VNTR	Variable number tandem repeat
w/v	Weight per volume
X-gal	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside: $C_{14}H_{15}BrClNO_6$

Chapter 1

General Introduction

1.1 GENETICS IN AQUACULTURE

1.1.1 HISTORY

Aquaculture, a practise that has existed for many centuries, is found globally among many diverse cultures. It began in the east, China specifically, in the year 2500BC with the formation of artificial lakes for storing fish during the flood seasons. The Romans were also proficient in breeding fish in mock ponds, which grew in popularity in European monasteries during the middle ages because fish was scarce and therefore expensive (Dunham *et al*. 2001). Aquaculture genetics was born when the first fish culturists began changing gene frequencies and altering the performance of domesticated stocks, without realising it. When farmers noticed changes in colour and body conformation, i.e. mutations, those fish with more attractive phenotypic traits were chosen as broodstock. This concept formed the basis of selective breeding. Direct breeding programmes were first developed in the 1800's but only became more established in the 1900's with the knowledge of Mendelian principles (Figure 1.1; Wilkins 1981). Aquaculture genetics became popular in the 1960's with the emergence of genetic enhancement programmes for finfish, followed by the advent of molecular genetics in the 1980's. Today, genetic programmes are well established for domesticated finfish and are being used increasingly in the development of aquatic invertebrate domestication. Genetics has proven to significantly contribute to production efficiency and enhancement as well as to increased sustainability for the aquaculture industry (Dunham 2004).

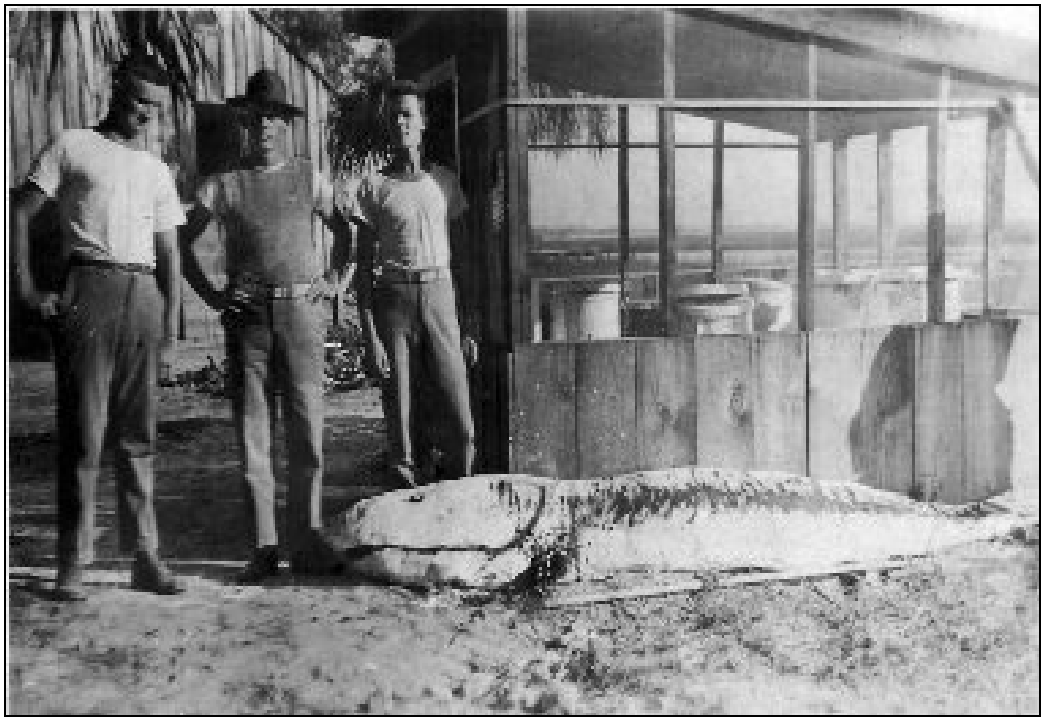


Figure 1.1 Mystery fish! A postcard dated between 1904 and 1918 indicating an abnormally large fish of unknown origin. On closer inspection one does not see fins or even a tail. Could this fish have undergone evolutionary genetic modification? (www.cryptomundo.com/cryptozoo-news/name-the-mystery-fish/).

1.1.2 BACKGROUND

Globally, there is an increasingly high demand for high-quality aquatic protein and other fish products, especially from marine resources. Thus, an increase in aquaculture management and production is clearly needed to meet this demand. Major advances have been achieved through enhanced management, nutrition and the genetic improvement of production traits in aquaculture (Dunham *et al*. 2001). Genetics has been of great importance to these advances. A review of world aquaculture resources by the Food and Agricultural Organisation (FAO 1995), identified two essential areas in which genetics is particularly significant in the commercial development of aquaculture species: 1) the genetic implications and importance of the translocation of species for restocking or conservation programmes, and 2) the genetic enhancement of commercially and economically important production traits.

1.1.3 GENETICS & CONSERVATION

A variety of species are becoming extinct at a rapid rate, one that can be compared to "the mass extinctions of geological time" (Beaumont & Bruford 1999). This is due to species being forced into small and isolated populations as a result of destruction or fragmentation of their habitats. These populations are at greater risk of being affected by environmental variation as well as being drastically reduced in numbers (Knaepkens *et al.* 2004). Management of the conservation of these species needs to be seriously considered when populations decrease to perturbing shortages and species are threatened with extinction.

In addition to the ecological factors associated with species decline, altered genetic processes (such as the loss of genetic diversity and species bottlenecks) must also be acknowledged. Since it is well established that genetic variation promotes adaptation to changing environmental conditions, conservation of genetic resources is an essential component of many species management programmes (Salgueiro *et al.* 2003). A number of direct environmental factors (water temperature and quality, nutrition and disease) have an effect on the genetic quality of cultured stocks. When farm-reared animals are released in to the natural environment (wild), such effects can ultimately create changes (such as loss of genetic diversity) in the genetic structure of the wild population and thus result in a loss of population identity. Genetic studies have a major role in these implications by assisting in the conservation of genetic resources among aquaculture species.

1.1.4 GENETIC ENHANCEMENT IN AQUACULTURE

Genetic improvement occurs when a variant of a particular trait existing in a population is made to change in frequency in the next generation so that the new generation performs better than the

parental stock. According to Wilkins (1981), genetic improvement is based on the existence of variation for a chosen character within or between available stocks of a species. It also includes the design of breeding schemes which selectively alter the expression of said character in subsequent generations. The extent of improvement that will be realised depends on the features of the breeding scheme, the heritability of the character and the interaction of that character with biological aspects of the genome and physical aspects of the environment.

Genetic improvement involves 3 broad features, namely the detection, propagation and assessment of the cultured species (Wilkins 1981). The initial step in any enhancement programme is *detection* of the character that is to be superior followed by the detection of deviations of that particular characteristic existing in the available population. Adequate detection implies measurement of the frequency distribution of discontinuous characters or the frequency density of continuous variables. *Propagation* implies various genetic manipulations used to modify the frequency of one or more of the detected and measured characters, in a generation. It may involve positive selection of individuals with the desired phenotype as broodstock, or the removal of undesirable phenotypes before breeding takes place. Propagation results in the alteration of the offspring generation only. The process of *assessment* is just as important as the features of the improvement process. The degree of alteration after propagation can be assessed after comparing offspring mean with parental mean. Whether or not the modification represents a genetic improvement requires a comparison of the overall performance of the broodstock and the changed offspring under normal commercial cultivation conditions.

Despite having a high reproductive capacity and therefore the potential for genetic improvement, most aquaculture species have not yet benefited from contemporary developments in selective breeding. Selective breeding could result in increased productivity, as demonstrated in livestock (Lymbery *et al*. 2000). Genetic enhancement of farmed aquaculture species has the capacity to deliver growing and continuous improvements in various production commodities such as product efficiency, product quality and ultimately financial profitability of the industry (Davis & Hetzel 2000). As aquaculture has been the world's fastest growing food production system for the past 20 years, the industry would benefit from the introduction of genetic improvement programmes. Genetic improvement is seen by many as offering the greatest potential to meet future food demands of a growing world population (Lymbery *et al*. 2000).

The improvement that has been achieved in aquaculture species thus far can be attributed, to a large extent, to good genetic management (Figure 1.2). Successful implementation of an amalgamation

of various genetic enhancement programmes is likely to result in a combination of the best suitable genotypes for the ever growing aquaculture industry (Dunham 2004). Genetic improvement of aquaculture species will therefore enhance competitiveness in relation to other food sources in the supermarket, which will in turn expand, stabilise and ultimately sustain the aquaculture industry.

Figure 1.2 Genetically enhanced Nile Tilapia, *Oreochromis niloticus*. This species was chosen as the first species for improving productivity of aquaculture to increase the supply of fish for the poorer sections of society and improve the earnings of fish farmers and related workers. It is a species of great importance to developing countries because of its many desirable traits and suitability for culture in low/high input farms (www.worldfishcenter.org/reshigh01_3.htm).

1.2 MOLECULAR MARKERS

Genetic markers are heritable characters, referred to as loci, with multiple states, or alleles, at each character (Strachan & Read 2004). In a diploid organism, as aquatic animals are, each individual can have one or two different alleles per locus. Multilocus approaches, in which more that one inherited trait is involved, are technically convenient but have distinct limitations. A substantial proportion of the variation they detect is non-heritable and inheritance occurs in a dominant manner – loci can only be scored by the presence or absence of one allele. In spite of their limitations, multilocus techniques can have powerful applications in linkage mapping and QTL analysis. In contrast, single-locus markers are more informative. Inheritance is co-dominant where both alleles at a particular locus can be identified in an individual and thus analysed more accurately. Using many markers simultaneously can yield greater sensitivity (Sunnucks 2000).

The various marker systems currently available show marked differences in their characteristics and in the amount of variability they display. There are also species-specific differences in levels of variation and mode of inheritance (Féral 2002). Markers are either dominantly or co-dominantly inherited. With dominant markers only the genotype of individuals of a population and not the

allele frequencies can be determined. This is because banding patterns of homozygotes and heterozygotes are indistinguishable. With co-dominant markers, banding patterns can easily be distinguished between homozygotes and heterozygotes, allowing for the estimation of allele frequencies in populations (Strachan & Read 2004). Markers can also be classified based on mode of inheritance: maternal (mitochondrial markers) versus biparental (nuclear markers) modes of inheritance (Ferguson & Danzmann 1998).

Molecular markers are divided into two types (O'Brien 1991). Initially, aquaculture genetic studies were based on type I allozyme and restriction fragment length polymorphism (RFLP) markers which are based on their association with genes of known functions, as well as mitochondrial marker systems. More recently, type II markers associated with anonymous genomic fragments, have been used. This switch was driven as a result of the low level of information of type I and mtDNA markers in relation to type II markers, and the fact that with type I markers assumptions on phenotypes were based merely on visual examination (Davis & Hetzel 2000). Type II markers include random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), single nucleotide polymorphism (SNP) and microsatellite DNA markers. At present, a combination of the strengths of all types of marker systems is the best solution for an overall understanding of the genomics of any aquaculture species as they offer fundamental and essential biological information (Liu & Cordes 2004).

1.2.1 TYPE I MARKERS

With the progression of aquaculture genetics, type I markers have been shown to be more useful and more important then previously anticipated. In addition to their effectiveness in comparative genomics, genome evolution and candidate gene identification, these markers have been shown to play crucial roles in linkage and QTL mapping. They serve as a bridge for comparing and transferring genomic information from a map-rich species to a relatively map-poor species, and also serve as genetic anchors when comparing sequence conservation among species (Liu & Cordes 2004).

i) Allozymes

Allozyme markers are allelic variants of proteins and are co-dominant, suggesting that heterozygotes are phenotypically different from homozygotes. They represent gene products, visualised on a gel matrix, that actually affect performance and are therefore of interest as markers because polymorphism exists (Dunham 2004). The number of polymorphic loci and alleles per locus are, however, often too low to characterise all genetic patterns or to assign parentage with

confidence (Féral 2002). Methods detecting allozyme variation were first developed in the 1960's and allozymes were amongst the earliest markers used for aquaculture genetics (Beaumont & Hoare 2003). They have previously been used in tracking inbreeding, stock identification and in pedigree analysis, as well as population and evolutionary genetics studies (Liu & Cordes 2002).

ii) Restriction Fragment Length Polymorphisms (RFLPs)

Restriction endonucleases are specific enzymes that cut DNA at highly conserved recognition sequences, usually four to eight base pairs in length (Féral 2002). The resulting fragments are visualised as bands after separation by electrophoresis. Restriction enzyme recognition sites are gained or lost through mutation (i.e. insertions, deletions or nucleotide substitutions) and in this way polymorphisms can be identified based on patterns of size fragments (Beaumont & Hoare 2003). RFLPs are co-dominant markers indicating that both alleles in a heterozygous individual can be observed, and because the size difference is often large, scoring is relatively easy. RFLPs have a relatively low level of polymorphism, thus decreasing their potential power in revealing genetic variation (Dunham 2004). This marker system can be classified as a type I or type II marker (Liu & Cordes 2004).

iii) Expressed Sequence Tags (ESTs)

Expressed sequence tags are sequences generated from the random sequencing of complementary DNA (cDNA) library sequences reverse-transcribed from messenger RNAs (mRNA; Strachan & Read 2004). They provide an efficient way to identify genes and analyse their expression in specific tissues types, under specific physiological and/or environmental conditions or during specific developmental stages (Dunham 2004). ESTs are powerful in their analyses of gene identification and the expression of those genes. They are most useful for functional studies as well as linkage and physical mapping in aquaculture genomics (Liu & Cordes 2004).

1.2.2 TYPE II MARKERS

In recent years type II markers have proven to be most appropriate in the genetic management of aquaculture species (Dunham 2004). These markers are amplified from random non-coding regions of genomic DNA *via* the polymerase chain reaction (PCR) and are therefore selectively neutral. The usefulness of such markers is based on their polymorphic information content (PIC) which depends on the number of detectable alleles and the distribution of their frequencies within a population (Liu & Cordes 2004).

i) Random Amplified Polymorphic DNA (RAPDs)

Short arbitrary sequences of nucleotides (primers) about eight to ten base pairs in length are used to amplify random fragments of DNA. This technique produces a large number of fragments which are individual-specific (Féral 2002). Variation between individuals is scored by the presence or absence of bands after gel electrophoresis. This is a reflection of mutations that occur at primer binding sites (Beaumont & Hoare 2003). The presence or absence of bands is an indication that RAPD markers are inherited in a Mendelian, dominant fashion. A band produced by homozygotes will, therefore, be identical to that seen for heterozygotes. Additionally, it is also difficult to determine whether bands represent different loci or different alleles of a single locus (Liu & Cordes 2004).

ii) Amplified Fragment Length Polymorphisms (AFLPs)

AFLP analysis is a multilocus fingerprinting technique that combines the power and overcomes the limitations of RFLP and RAPD methods (Liu & Cordes 2004). AFLP technology involves both restriction digestion and PCR, and thus, is rapidly becoming very popular among geneticists for the detection of a larger number of polymorphic DNA markers in a reproducible manner (Féral 2002). AFLP markers are inherited as dominant markers and their analysis involves the detection of the presence or absence of restriction fragments rather than differences in lengths (Beaumont & Hoare 2003). Bands are considered to be bi-allelic and therefore have relatively low PIC (Liu & Cordes 2004).

iii) Single Nucleotide Polymorphisms (SNPs)

Single nucleotide polymorphisms (SNPs) are variations caused by point mutations at a single nucleotide position. SNPs give rise to different alleles containing alternative bases and are the most abundant polymorphism in any given organism (Dunham 2004). SNPs are regarded as being bi-allelic because of their restriction to one of two alleles, either purine bases (A or T) or pyrimidine bases (C or G). SNPs have become one of the most popular markers for analyses over the past decade because it is a fast, high-throughput method (Liu & Cordes 2004).

iv) Microsatellites

Microsatellites have gained popularity with aquaculture geneticists over the last decade. They have proven to be exceptionally useful in a wide range of applications which helped in ensuring the sustainable development of various species of economic importance. Microsatellites form the basis of the current project and are discussed in greater detail in section 1.3.

1.2.3 MITOCHONDRIAL DNA MARKERS

Mitochondrial DNA (mtDNA) occurs in high copy numbers and is stored within the energy-generating organelles of all cells, i.e. the mitochondria (Strachan & Read 2004). It is therefore referred to as extra-chromosomal DNA as it does occur in genomic areas of cells. Mitochondrial DNA is said to be maternally inherited, suggesting that it passed on to the next generation originating entirely from the maternal parent (Dunham 2004). Maternal inheritance limits the ability to provide information on the male component of a population (Ferguson & Danzmann 1998), but it does offer insight into sex-biased population structure (Féral 2002). Mitochondrial markers have been popular among aquaculture geneticists as they have been used extensively in the investigation of stock structure and the identification of broodstock animals. In earlier research, mtDNA markers were preferred over allozymes because of their relatively higher level of polymorphism. Although mtDNA loci can exhibit larger numbers of alleles per locus, the limited number of markers available however, makes them less powerful than highly variable nuclear DNA markers (Liu & Cordes 2004). Due to the non-Mendelian mode of inheritance, the mtDNA molecule is considered a single locus and the mutation rate of these markers has been shown to be higher than that of nuclear DNA (Avise 1994).

1.2.4 ROLE OF MOLECULAR MARKERS IN AQUACULTURE

The role molecular markers are beginning to play in the genetics of aquaculture species is a central one particularly with regards to species produced for commercial purposes. The main areas for the use of DNA marker technologies include their application in interspecies, intraspecies and hybrid identification, genetic variability and inbreeding, parentage assignment, genetic selection, mapping economically valuable quantitative traits and genes associated with said traits, and marker-assisted selection (MAS) with the aid of the construction of linkage maps. With the incorporation of marker technologies there has been a significant improvement in the understanding of the molecular biology of aquaculture species (Liu & Cordes 2004). By detecting genetic variation, DNA markers may provide useful information at different levels including: population structure, levels of gene flow, phylogenetic relationships and patterns of biogeography (Féral 2002). Marker technologies can also be applied with the aim of enhancing health, production, and ultimately the profitability of the aquaculture industry. Much progress has been made in the development of DNA marker technologies in aquaculture species including salmon, trout, catfish, tilapia, shrimp and marine molluscs. Microsatellites, AFLPs and SNPs have, over the last ten years, shown to have the largest impact in aquaculture genetics, particularly in research involving various species that are important for commercial purposes (Liu & Cordes 2004).

1.3 MICROSATELLITES

Microsatellites are short and simple, tandem repetitive regions of DNA abundantly distributed across both the coding and non-coding areas in the genomes of all organisms. They consist of specific sequences of nucleotides (nitrogenous bases) which contain, for example, di-, tri- and tetranucleotides, etc (Figure 1.3a). Microsatellites fall under the group of repetitive sequences known as variable number tandem repeats (VNTR) and are also referred to as simple sequence repeats (SSR) or short tandem repeats (STR) (Chambers & MacAvoy 2000). Typically, microsatellites are found in the form of dinucleotides, with the most common motif of vertebrates and invertebrates being $(AC)_n$, followed by tetra-, tri- and mononucleotides, with hexa- and pentanucleotides occurring relatively in much lower frequencies (Christiakov *et al*. 2006). Microsatellites are inherited in a Mendelian manner and are highly variable. Alleles at a specific locus differ in the number of repeat units. For example, at a particular locus one allele may have 10 repetitive units of a CT motif, while another may have 12 repeats, and because of the diploid nature (two sets of chromosomes) of most organisms, alleles may be either homozygous for a locus or heterozygous (Figure 1.3b).

ACACACACACAC – a dinucleotide referred to as $(AC)_6$

CTGCTGCTGCTGCTG – a trinucleotide referred to as $(CTG)_5$

ACTGACTGACTGACTG – a tetranucleotide referred to as $(ACTG)_4$

Figure 1.3a Examples of various microsatellite sequences.

I) Homozygous: both strands have 10 CT repeats

…CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG…

…CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG…

5' Flanking Region **Microsatellite Locus** **3' Flanking Region**

II) Heterozygous: one strand has 10 CT repeats while the other has 12 CT repeats

…CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG…

…CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTCTCTATCGGTACTAC…

Figure 1.3b Representation of a microsatellite locus in two individuals. I) a homozygote has the same allele on homologous chromosomes. II) a heterozygote has two different alleles on the same locus on homologous chromosomes.

In any population, many alleles of a single microsatellite locus exist, with each allele having a different length. The different alleles are the result of chromosomal mutations that occur from one generation to the next. The flanking regions surrounding the locus for all individuals of a given

populations will, however, comprise of conserved sequences. This is important for the application of PCR when amplifying microsatellite loci in various individuals of a species.

1.3.1 EVOLUTIONARY DYNAMICS OF MICROSATELLITES

The main contributing factor to the keen interest in using microsatellites is their hypermutability, and therefore their hypervariability (Christiakov *et al.* 2006). For DNA sequences' evolving neutrally, the amount of polymorphism is expected to be directly proportional to the underlying mutation rate (Ellegren 2000). Non-repetitive eukaryotic sequences generally mutate at a rate of approximately 10^{-9} per nucleotide per generation, which is less than those of repetitive sequences. Microsatellite mutation rates have been estimated to range between 10^{-6} and 10^{-2} per locus per generation (Ellegren 2000; Schlötterer 2000; Christiakov *et al.* 2006). It is important to keep in mind however that mutation rates vary between species and more importantly between loci. Variation between mutation rates at different loci have been attributed to several influencing factors, i.e.: the length (number or repeats) of individual microsatellites, the sequence of the repeat motif, the length of the repeat unit, the flanking sequences, as well as interruptions within the microsatellites and recombination rates (Schlötterer 2000). Interestingly, microsatellites are said to be highly unstable when situated in expressed regions and therefore undergo a much higher rate of mutation than those found in the non-coding areas. This may be due to the expressed DNA having a higher tolerance to mutations (Hancock 1999).

Mutations are the resulting factor in the development of the various alleles of a microsatellite locus and therefore result in their polymorphic nature. There are two focal theories with regards to how these mutations occur within a given genome. The predominant theory was first described by Fresco and Alberts (1960) and is referred to as DNA replication slippage. The second mutational mechanism suggested by Smith (1976), is termed non-reciprocal recombination, a theory which is becoming less popular with researchers as a result of increasing evidence against its importance.

i) Replication/DNA Slippage

Slippage during replication takes place when the nascent strand of DNA dissociates from the template strand and re-anneals "out-of-phase" with the template strand (Hancock 1999). DNA polymerase is the enzyme used to replicate DNA molecules. By using both strands of the double helix as templates, nascent (new) strands are created (Figure 1.4). The polymerase elongates the nascent strands by attaching corresponding nitrogenous bases to those of the template strand. Sometimes during the replication event the polymerase slips, causing the new strand to mispair from its template, resulting in repeat units being left out or too many repetitive units being added.

Decreases in length at a microsatellite locus occur when misannealing gives rise to looped-out bases in the template strand, resulting in the nascent strand slipping down in its binding to the template strand. Length increases at a locus occur as a result of the nascent strand forming a misannealing loop, causing the polymerase to add extra repeats in order to fill the gap (Moxon & Wills 1999). Thus, the new strand will have a different allele from the parent strand.

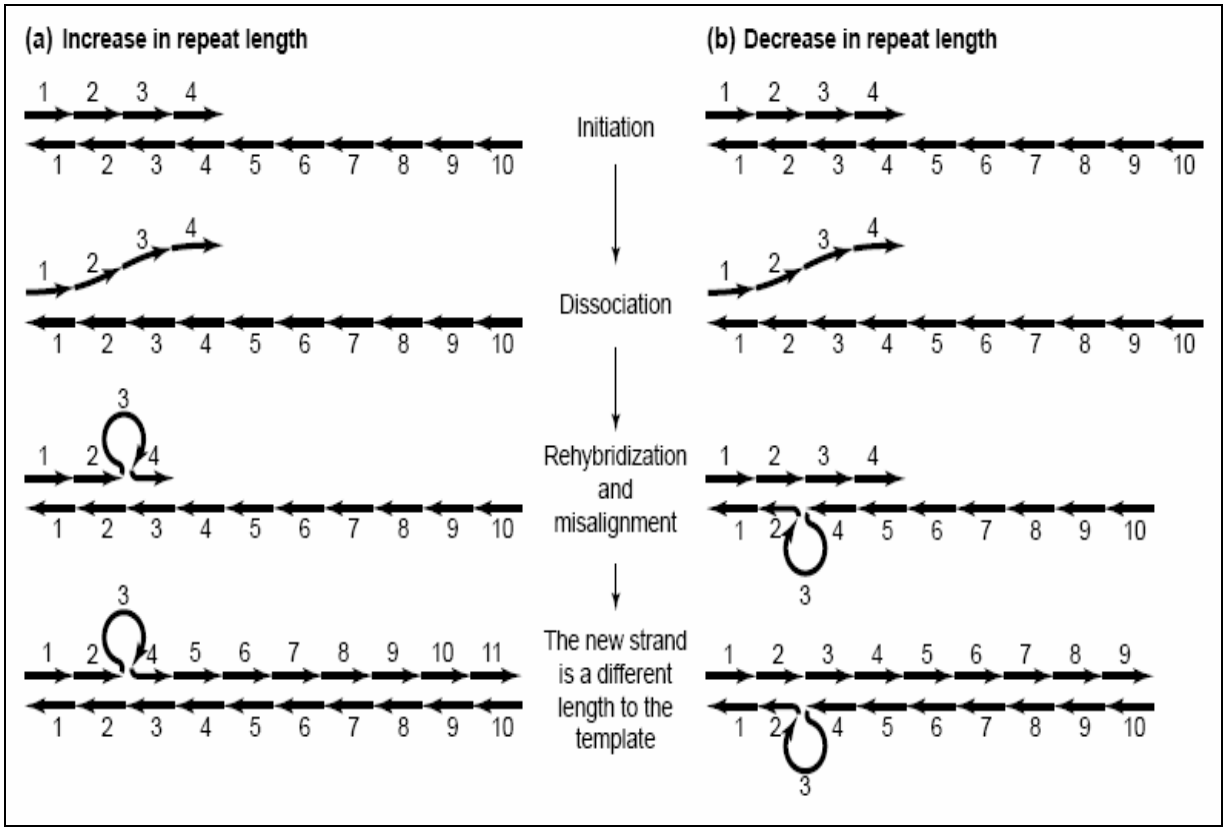


Figure 1.4 An illustration of replication slippage resulting in variations of a microsatellite locus; repeat units are denoted by arrows (Ellegren 2000).

ii) Non-reciprocal Recombination

During meiosis, homologous chromosomes align, enabling alleles of the same locus to switch places between the paired chromosomes. Unequal crossing-over involves the cross-over between either the chomatids of the same chromosome or between homologous chromosomes which are misaligned (Figure 1.5). This is particularly common with repetitive sequences where recombination machinery cannot easily determine the correct alignment between the two strands (Hancock 199). Instead of whole alleles switching places, parts of them switch over leaving behind part of the original microsatellite resulting in an expanded or contracted allele. Therefore, non-reciprocal recombination gives rise to a deletion in one DNA molecule and an insertion in the other, and is said to cause more drastic changes in numbers of repeats than those associated with DNA slippage (Moxon & Wills 1999).

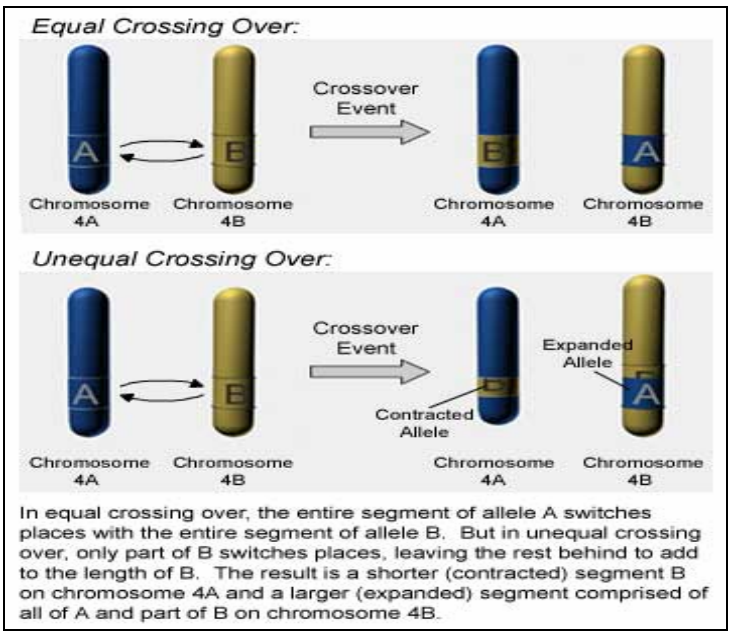


Figure 1.5 Unequal crossing over may result in microsatellite variation (http://www.stanford.edu/group/hopes/causes/mutation/q4.html).

There is little knowledge on the actual details of how mutations lead to varying allele numbers. Subsequently, two different models of microsatellite mutation have been hypothesised with regards to the process:

1) In the classical stepwise mutation model (SMM; Kimura & Ohta 1973), one repeat unit is either added or deleted upon mutation. This implies that two alleles that differ by one repeat are more closely related than alleles that differ by numerous repeats, and follows the changes in length, allowing a step up (or down) at each mutation. Its simplest form is one-step symmetry, which assumes only one step per mutation. Furthermore, probabilities of increasing and decreasing are equal and constant across alleles (Ellegren 2000; Dieringer & Schlötterer 2003).

2) Mutation rates are, however, not constant across different allele sizes, and it can therefore be assumed that mutations do not always constitute single-step changes (Ellegren 2000). In the infinite allele model (IAM; Kimura & Crow 1964), each mutation can create any allele randomly. Thus, there are an infinite number of states each allele can mutate to; hence each mutation is assumed to be unique.

1.3.2 FUNCTIONS OF MICROSATELLITES

Although microsatellites represent selectively neutral DNA markers, various studies have shown that they also have a functional significance when it comes to various aspects of DNA such as structure, replication, recombination and gene expression (Christiakov *et al.* 2006).

DNA Structure - Microsatellites situated at the telomeric and centromeric regions of chromosomes of assorted fish species have been shown to play various roles in the organisation of

chromosome structure: AT-rich motifs found in centromeric regions of chromosomes of gobiid species are considered important for the control of chromatin compactness (Canapa *et al.* 2002), whereas telomeric-associated repeats have been related to nucleolus organising regions (NORs) in rainbow trout (Abuin *et al.* 1996) and Nile tilapia (Foresti *et al.* 1993).

DNA Replication - Microsatellite repeats can influence DNA replication. Observations suggest that they can affect enzymes controlling mutation rates and cell cycles, and therefore initiate tumorigenesis in humans (Chang *et al.* 2001).

Recombination - Microsatellites are considered hot spots for recombination events (Jeffreys *et al.* 1998). Dinucleotide motifs are particularly significant sites due to their high affinity for recombination enzymes and their direct influence on recombination through their structural influences (Christiakov *et al.* 2006).

Gene Expression - A range of effects of microsatellites with regard to gene expression have also been recorded. Edwards *et al.* (1998) showed that simple repetitive sequences located in the promoter regions of genes can influence expression in puffer fish. In Nile tilapia it was seen that microsatellites found in the promoter of the prolactin 1 gene significantly influenced the expression of the gene (Streelman & Kocher 2002). Intronic microsatellite sequences also affect gene transcription by serving as binding sites for expression-regulating proteins (Epplen *et al.* 1996).

1.3.3 APPLICATIONS OF MICROSATELLITES IN AQUACULTURE

Due to their multi-allelic nature, co-dominant inheritance, small length, wide genome distribution and abundance, microsatellites have been applied in a variety of research fields within the aquaculture industry (Christiakov *et al.* 2006). These include genetic mapping, QTL mapping and MAS and, individual identification and parentage assignment. As parentage assignment is one of the main focuses of this study, it will be discussed in greater detail from the other applications, further on in this chapter.

i) Genetic-Linkage Mapping

Genetic mapping is an illustration of the order or position of, and distances between genes and molecular markers, relative to each other, on the chromosomes of an organism. These maps are based on the frequencies of recombination between markers during crossover of homologous chromosomes, and aid researchers in locating markers and/or genes by testing for genetic linkage of already known markers (Beaumont & Hoare 2003). Single families, whether they are full-sib or half-sib, are used in the construction of such maps. This is due to the fact that single families are heterozygous for almost all genetic markers (Dunham 2004). Linkage maps are usually constructed in the framework of quantitative genomic programmes, including long term studies directed at

locating genomic regions underlying phenotypic variation associated with economically important traits within family structures (Saavedra & Bachère 2006). Genetic mapping represents a key research field of aquaculture and has made remarkable progress over the last ten years. The construction of framework genetic linkage maps for several aquatic animals has set in motion an increased trend in the development of commercially important aquaculture species including catfish (Waldbieser *et al.* 2001), tilapia (Agresti *et al.* 2000), salmon (Hoyheim *et al.* 1998), trout (Sakamoto *et al.* 2000), shrimp (Li *et al.* 2000) and oysters (Hubert *et al.* 2000).

Over the last few years a large number of molecular markers have been developed and evaluated. SSRs remain the markers of choice for the construction of linkage maps because they are highly polymorphic and informative, and require small amounts of DNA for testing (Christiakov *et al.* 2006). Microsatellites and AFLP markers have demonstrated reliability, efficiency and reproducibility for genetic linkage mapping. They are extremely useful in establishing primary frameworks which could be further enriched (Dunham 2004). Microsatellites in particular have played an important role in the creation of linkage maps for a variety of fish and shellfish species. A linkage map of the Pacific oyster is based on over 100 microsatellites (Hubert & Hedgecock 2004), while 191 microsatellite markers were used for the linkage map established for rainbow trout (Sakamoto *et al.* 2000) and 200 markers were used to map the tiger pufferfish genome (Kai *et al.* 2005).

In the case of abalone, there are only two known genetic maps and both are based on microsatellite DNA markers. Baranski *et al.* (2006) were the first to develop a linkage map for any *Haliotis* species. They used 122 microsatellite markers to represent a first generation linkage map of *Haliotis rubra*. The second and most recent published haliotid linkage map was developed by Sekino and Hara (2007; Figure 1.6) who used three F_1 outbred families in which 180 microsatellite markers were mapped to 18 linkage groups for the Pacific abalone *Haliotis discus hannai*.

ii) QTL Mapping

A quantitative trait locus is a region of DNA associated with a particular phenotypic trait, and many QTL are used to identify candidate genes underlying said trait. Such traits can vary in degree and can be accredited to polygenic interactions and influences from the environment (Beaumont & Hoare 2003). According to Mackay (2001), the principal of QTL mapping was noted over 80 years ago in a study done on *Phaseolus vulgaris* executed by Sax in 1923. The mapping of quantitative trait loci is the first step towards the identification of genes and causal polymorphisms for important traits (Seaton *et al.* 2002). It requires the statistical study of the alleles that occur at a locus and the

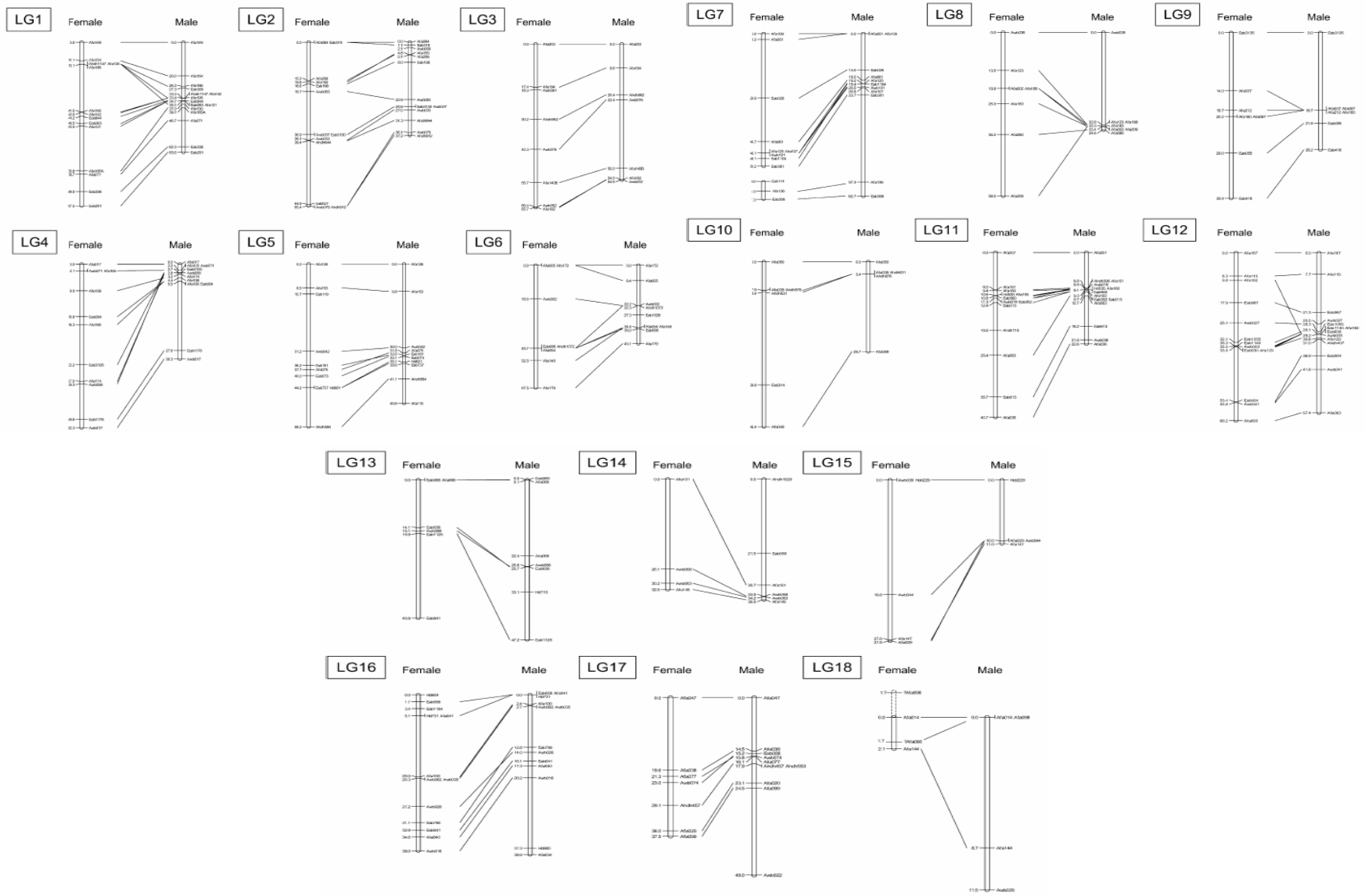


Figure 1.6 Genetic linkage map of the Pacific abalone (*Haliotis discus hannai*), based on 180 microsatellite DNA markers arranged into 18 linkage groups, for male and female, spanning approximately 702.4cM and 888.1cM of the genome, respectively (Sekino & Hara 2007).

phenotypes they are coupled to and produce.

Genetic linkage maps are required to map QTL. The markers used need to be polymorphic, have good genome coverage and be suitable for automated genotyping, and require families in which the markers segregate (Martinez 2006). Microsatellite-based strategies represent appropriate techniques to identify QTL, particularly those associated with economically important complex traits. For a given QTL, the probable success and mapping resolution is dependent on the number of loci screened and the magnitude of their effect on the trait of interest. Coarse mapping of QTL detects a rough location of between 10 and 30 centimorgan (cM), while fine mapping narrows the putative location of a QTL to less than 5cM (Chistiakov *et al.* 2006).

In aquaculture, most performance and production traits are controlled by multiple genes and are therefore inherited as quantitative traits. Analysis of their related QTL promises to be a very significant part of aquaculture genetics (Liu & Cordes 2004). The first mapping of an economically important trait in aquaculture was reported by Jackson *et al.* (1998) who identified a QTL linked to temperature tolerance in rainbow trout (*Oncorhynchus mykiss*). Microsatellite-based QTL screening for growth, feed-conversion efficiency and temperature tolerance have also been performed in other salmonid species (Danzmann *et al.* 1999; Ozaki *et al.* 2001; Robinson *et al.* 2001; Kause *et al.* 2002; Somorjai *et al.* 2003; Araneda *et al.* 2005; Hayes *et al.* 2006), however to date, no QTL genes have been identified. In the case of shell fish, little work has been done with regards to QTL mapping, and those available have been based on AFLP markers. These include mapping of a QTL linked to shell colour in bay scallop, *Agropectin irradians irradians* (Qin *et al.* 2006), and disease resistance in the eastern oyster, *Crassostrea virginica* (Yu & Guo 2006). Recently, the successful identification of QTL for growth-related traits was achieved in the Pacific abalone, *Haliotis discus hannai* (Liu *et al.* 2007). Researches made use of a selection of molecular markers including ALFPs, RAPDs as well as microsatellites. With the availability of family structures and DNA markers such as microsatellites, it is expected that greater achievements will be realised in QTL mapping in aquaculture species, which will eventually lead to successful selective breeding programmes, particularly MAS.

iii) Marker-assisted Selection

Marker-assisted selection refers to the concept that it is possible to deduce the presence of a gene from the presence of a marker closely linked to that gene. It enables breeders to select desirable traits based on genotype rather than phenotype using molecular markers (Young 1999). Molecular markers have several advantages over phenotypic markers in that they offer a greater scope for

improving the efficiency of conventional breeding because they are not environmentally regulated (Mohan *et al.* 1997). According to Hulata (2001), knowledge of linkage between markers and QTL alleles for MAS increase the rate of genetic progress above that for selective breeding alone. To implement MAS, it is important to have high-density and high-resolution genetic maps, which are saturated by markers in close proximity to the target gene that will be selected (Chritiakov *et al.* 2006). The main focus for the future of aquaculture genetics is QTL and comparative mapping in order to increase skills in identifying candidate genes controlling performance and production traits through MAS.

Marker-assisted selection has been successfully implemented in various agricultural industries such as crop plants including tomatoes, potatoes, beans, maize, barley and rice (Mohan *et al.* 1997), and livestock including pigs and dairy cattle (Soller 1994). To date, in aquaculture, most genetic improvement has focussed on the use of traditional selective breeding techniques such as selection, crossbreeding and hybridisation. Although few, genetic maps have been generated for various fish and shellfish, providing the genetic framework and necessary tools for developing MAS programmes (Dunham 2004). A few examples of MAS actually exist for fish: improved growth rate and feed-conversion efficiency as well as heat tolerance have been demonstrated using MAS in rainbow trout and salmon (Ferguson & Danzmann 1998; Danzmann *et al.* 1999; Davis & Hetzel 2000). MAS has the potential to greatly maximise genetic improvement in aquaculture species.

1.4 PARENTAGE ASSIGNMENT

For a long time, a major concern for farmers has been the identification of their stocks. In order for them to make economically sound choices regarding stock production and to enhance the value of those stocks, knowledge of pedigree is required. From the times of antiquity to the 20th century various methods of traceability have been employed for livestock, including descriptive documents and certificates distinguishing prized animals, branding animals directly on their bodies, and/or attaching removable exterior markers such as collars or rings (Blancou 2001). In aquaculture, fish have undergone fin clipping, external and internal tagging with various devices, as well as thermal and chemical branding (Håstein *et al.* 2001). With the new millennium came the development and establishment of biotechnology, making biological identification systems a certainty. Within the specialised area of pedigree analysis, DNA testing is widely used to confirm parentage, and molecular verification is progressively replacing old and outdated methods of identification.

1.4.1 THE FUNDAMENTALS OF GENETIC IDENTIFICATION

With the exception of monozygotic twins and clones, organisms differ from each other at the most basic level, i.e. at their DNA. Certain DNA sequences code for the production of particular proteins associated with certain traits that are specific to individuals, thereby resulting in variations in physical characteristics such as size, shape and colour. Since the genome of every living organism contains approximately three billions base pairs, the capacity for individual variation, particularly at the DNA level, is vast (Cunningham & Meghen 2001). Each individual has a unique genotype, which represents that individuals genetic make up. Due to recombination of parental chromosomes during meiosis, one can say the individual acquired half of its genomic data from the maternal chromosomes and the other half from the paternal chromosomes.

i) Meiosis & Sexual Reproduction

Meiosis is critical to the successful sexual reproduction of all diploid organisms. Sexual reproduction physically mixes chromosomes through recombination, at every generation, providing an almost unlimited assortment of genetic combinations unique to each new generation (Klug & Cummings 2000). During prophase of meiosis I homologous chromosomes form a synapsis whereby segments of one maternal and one paternal chromatid are randomly exchanged *via* recombination (crossing over; Figure 1.7).

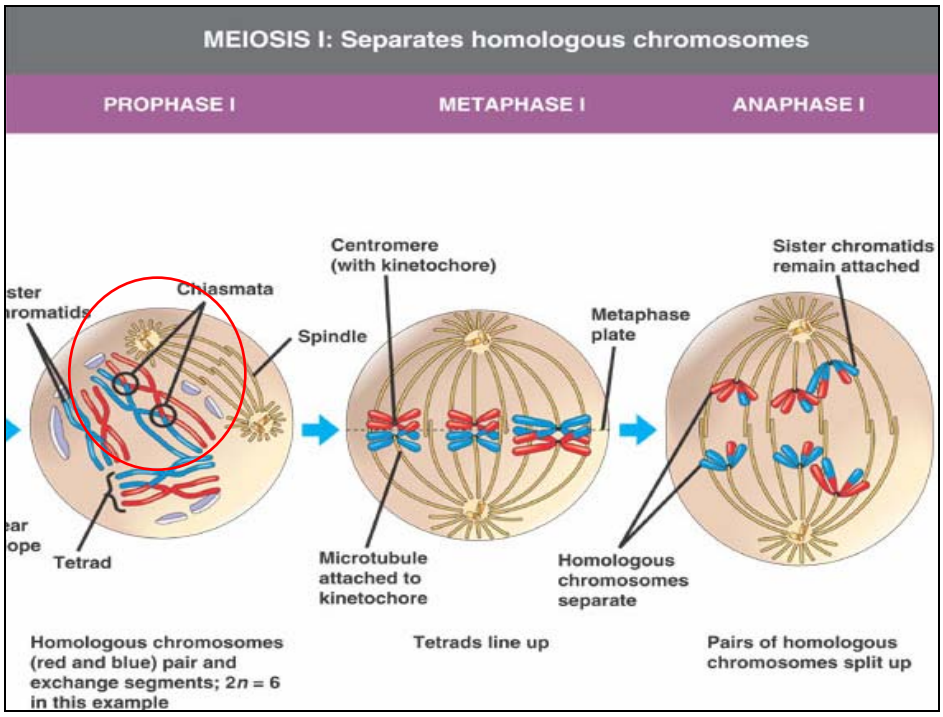


Figure 1.7 An illustration of meiosis which demonstrates the recombination (crossing over) of homologous chromosomes, at the chiasmata (points of crossover indicated in the red circle) which result in the exchange of genetic information to be inherited by the next generation of individuals (kvhs.nbed.nb.ca/gallant/biology/biology.html).

ii) Inheritance of Genetic Loci

On each set (two) of haploid chromosomes in the cells of diploid organisms, there are various genes and DNA markers. Every genetic locus situated on every chromosome from one parent will have a reciprocal and corresponding locus on the homologous chromosome from the other parent. These molecular loci will be the in the same position on homologous chromosomes, but may vary in structure. This forms the basis for distinct genotype profiles that give each individual organism its uniqueness. Therefore with a set of markers and enough pedigree information, one can trace specific genetic loci to a specific parent and even a grandparent, resulting in the ability to trace the segregation of said loci.

1.4.2 THE ROLE OF MICROSATELLITES

An appropriate methodology should be chosen for the accurate analysis of genotyping data to reconstruct parentage and pedigree structures. Microsatellite loci have provided a very popular means of learning about the behaviour of breeding and the structure of populations. The pattern of relationships between individuals can be determined using these popular molecular markers. They are the best tools for this application because they are such highly discriminating and bi-parentally inherited co-dominant markers, which have consequently found extensive approval in parentage and pedigree analysis (Chambers & MacAvoy 2000).

Using a panel of several markers, a unique genotype profile can be produced for each individual. This profile is highly discriminating, suggesting a random individual would have a low probability of matching any other given genotype (Christiakov *et al.* 2006). For parentage assignment, microsatellites provide the basis for an affordable and reliable genetic profiling system that can be productively used for identification and parentage control. They also provide more flexibility as to sample type and their robustness has been demonstrated in large scale efforts (Bowling 2001).

1.4.3 PARENTAGE ASSIGNMENT IN AQUACULTURE

Knowledge of pedigree structure or co-ancestries between the individuals which belong to a population is important to many areas of the biological and genetic background of that population. This information is needed to estimate genetic parameters such as heritability and genetic correlations, breeding values, loss of diversity and an increase in inbreeding. It also allows for the study of mating systems, parental care and dispersal in ecological and behavioural genetics (Fernández and Toro 2006). The genetic evaluation of animals relies on the collection of pedigree information on the population under consideration, as well as trait information of individuals and/or their relatives. Traditional pedigree records for livestock breeding were based on the observation of

mating between dams and sires as well as birthing in order to tag individual offspring accordingly to maintain lineage information (Dodds *et al.* 2005). Fish have complex mating systems, making the traceability of pedigrees fairly difficult. It is therefore not possible to simply attach identifiers to the offspring at birth. Effective methods of traceability are required for basic research within different types of aquaculture operations.

Using DNA marker technology to identify parents overcomes the management constraints of visual recording. With the introduction of microsatellite markers and emerging programmes to assign parentage, it is now possible to analyse heredity and inheritance. Microsatellites provide the best results because of their high genetic variation among individuals (Liu & Cordes 2004). In the majority of studies performed, genotypes are simply compared to reconstruct parentage. Such parentage matching is typically very reliable and effective for an offspring with few possible parents. However in aquaculture, it is rather difficult to unambiguously assign parents for all progeny because of mass spawning methods. Many other factors could also affect the amount of genetic progress achieved using molecular markers. These include the number of microsatellites used, missing animals and genotypes, and the heritability of the markers under test (Dodds *et al.* 2005).

Over the last decade, aquaculture has made use of parentage assignment, with microsatellites, for various applications. Herbinger *et al.* (1995) used relatedness testing in order to estimate the parental effects of growth and survival on progeny for a group of rainbow trout broodstock. Microsatellites have been used to track parents and their reproductive contribution, to offspring of various economically important species including: bluegill sunfish, *Lepomis macrochirus* (Neff 2001); turbot, *Scophthalmus maximus* (Borrell *et al.* 2004); Nile tilapia, *Oreochromis niloticus* (Fessehaye *et al.* 2006); Atlantic salmon, *Salmo salar* (Norris *et al.* 2000); Pacific oyster, *Crassostrea gigas* (Li & Kijuma 2006) and shrimp *Penaeus japonicus* and *Penaeus monodon* (Jerry *et al.* 2004, 2006). Accurate parental assignment levels varied from 47% in shrimp to 98% in Atlantic halibut, with the use of a set of microsatellite markers ranging from four to fifteen (Jackson *et al.* 2003; Jerry *et al.* 2004). Parentage assignment has also been used in the abalone industry as a means to improve commercial stocks by tracing growth rates through the aid of genetic enhancement (Selvamani *et al.* 2001; Lucas *et al.* 2006). The current project focuses on assigning parentage to cultured South African abalone using newly developed species-specific microsatellites. The information obtained will then allow the determination of contributions among broodstock to offspring.

1.5 ABALONE

1.5.1 CLASSIFICATION

According to the Concise Oxford Dictionary (Oxford University Press), the word abalone originated from the American-Spanish term abulónes, the plural of abulón, meaning an edible mollusc of warm seas with a shallow ear-shaped shell lined with mother-of-pearl. Abalone belong to the phylum Mollusca, a group that includes mussels, scallops and clams, and fall under the class Gastropoda, part of the clade Vetigastropoda, whose members have no shell or one shell. In common marine snails shells are visibly spiral; in abalone, alternatively, they are flattened. Abalone are members of the family Haliotidae, which includes a number of recognised species in the genus *Haliotis*. The name *Haliotis* also refers to the shallow ear-shaped shell lined with mother-of-pearl (Geiger 2000).

1.5.2 GENERAL ANATOMY

Molluscs are predominantly marine creatures with soft bodies surrounded by a mantle, a frontal head and a large muscular foot. They are best known for their brilliantly coloured calcareous shells which are secreted by the mantle. Abalone are univalve (one shell), oceanic molluscs whose most prominent characteristic is their shell with its row of respiratory pores scientifically termed tremata (Taylor & Ragg 2005). The larger of the pores are situated at the anterior end of the shell while the smaller pores are located towards the posterior end (Figure 1.8a). These openings play an important role in respiration, the release of gametes, as well as defecation (Fallu 1991). Underlying the shell is an iridescent layer of mother-of-pearl. The shell, which makes up approximately 30% of the total body mass, is attached to the body of the animal by a muscular column known as the adductor muscle (Baldwin *et al.* 2007). The adductor muscle extends into a pedal musculature that has a strong suction power, enabling the abalone to secure itself firmly to rock surfaces. This muscular foot is the source of the commercially valuable food that is so greatly sought after. The body can account for more than a third of the total weight of the animal (Nash 1991). The mantle, extending from the shell, covers the foot that is circled by a lip known as the epipodium. The epipodium is a sensory extension of the foot that bears sensitive epipodial tentacles. The muscular foot, situated at the head of the abalone, consists of a pair of eyes, a mouth and an enlarged pair of tentacles. Inside the mouth is a long tongue called the radula, which scrapes in algae, to be ingested (Fallu 1991).

The internal organs of the abalone are arranged around the foot and under the shell (Figure 1.8b). The crescent-shaped gonad extends around the right side of the body, situated opposite the tremata and towards the posterior. It is the most prominent organ that distinguishes male from female

(Fallu 1991). The male testis is cream or light brown in colour whereas the female ovary is grey-brown, green or violet. The gonad constitutes 15-20% of the soft body mass during the breeding season and remains so until spawning, after which is rapidly decreases in size (Henry 1995). The gill chamber is positioned next to the mouth under the shell pores and consists of a pair of large, bipectinate gills termed ctenidia. Ventilation occurs in a unidirectional manner where water is drawn in under the edge of the shell then flows over the gills and out the pores (Taylor & Ragg 2005). Furthermore, waste and gametes are carried out with the flow of water through the shell pores. The abalone is considered a primitive organism as it has no evident brain structure. It does however, have an anatomically and functionally complex circulatory system. The heart is situated on the left side of the body and is the source of blood flow through a network of fine vessels and sinuses. The surrounding tissues and muscles assist this intricate cardiac system (Baldwin *et al.* 2007).

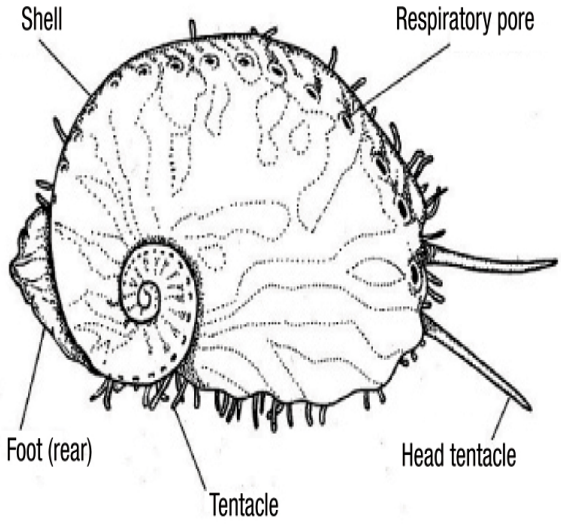


Figure 1.8a External anatomy of abalone (Fallu 1994).

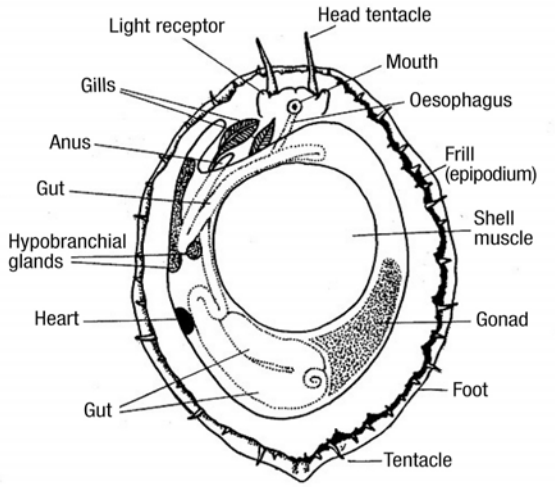


Figure 1.8b Internal anatomy of abalone (Fallu 1994).

The characteristics that differentiate Haliotidae include shell shape, shell to body ratio, and epipodial structures as well as the morphology of the respiratory pores (Lindberg 1992). Other visible features that differentiate the various species include size, colour of the shell (pink, green, black, etc.) and shape of the shell (flat, round, congruent, etc.; Geiger 2000). Abalone species are also unique in the environmental niches they occupy due to their food requirements and their ability to endure exposure to rocky territories (Nash 1991).

1.5.3 REPRODUCTION

Abalone reproduce *via* a form of external fertilisation referred to as broadcast spawning (Uki & Kikuchi 1984). The eggs and the sperm are released through the tremata with the respiratory

current. Gamete release among males and females of a population is a synchronous phenomenon, however the spawning events themselves are often rare and unpredictable (Counihan *et al*. 2001). Amongst wild haliotids, there is great variation in the length of the spawning season and the number of spawning events per season (McShane 1992). Research done over the last 40 years has indicated that natural spawning for most abalone species is unpredictable and is the result of endogenous (biological i.e. pheromones or other hormones) as well as exogenous (environmental i.e. water temperature or length of day) factors. Past theories have suggested that the presence of other gametes may induce spawning of the opposite sex, and that sex pheromones released at the same time as the gametes may also bring about the liberation of ova and sperm (Hardege *et al*. 1996; Zeek *et al*. 1998). Other theories have implied that abalone may delay the release of their gametes until more individuals surround them, and males may ejaculate more frequently when more abalone (male or female), are present (Uki & Kikuchi 1984). There is, however, evidence in studies done with other marine invertebrates that bring the two theories together. They show that fertilisation success is influenced to a large degree by population density because of the resultant concentration of gametes (Levitan 1991; Claerebuodt 1999). A study done by Counihan *et al*. (2001) suggests that a variety of species-specific mechanisms underlie the regulation of haliotid spawning.

1.5.3 LIFE CYCLE

After the gametes have been spawned, fertilisation takes place. As abalone are ectotherms they are dependent on seawater temperature for the success of fertilisation, timing of hatch-out and survival at early life stages (Fallu 1991). The eggs hatch as microscopic, free-living larvae, which drift with the ocean currents for about a week. This is referred to as the pelagic stage in which they remain for 5 – 10 days (McShane 1992). According to Hahn (1989), during this period 41 larval stages can be distinguished until the initiation of metamorphosis. There are however two main morphological phases: 1) the trocophore larval stage followed by 2) the veliger larval stage. Once the eggs are fertilised they are approximately 0.2 mm in diameter and begin a series of divisions until they reach the trocophore stage, which takes an average of about 20 hours, after which the trocophores develop into the free-swimming veliger larvae (Nash 1991). Within a week of hatching, the veligers develop a head and foot, and can detach themselves from the surfaces on which they land. Haliotid larvae survive on the nutrients provided by the egg yolk, and do not feed off their surroundings until settlement (Hahn 1989). On final settlement, circa 14 days after hatching, the veligers sink to the seabed and rock surfaces where they undergo their last stage of metamorphosis (Nash 1991). They shed their cilia, or swimming hairs, and begin to develop their adult shells. At this stage, they are termed “recruits” (Fallu 1991). Subsequently, they become photophobic and move into obscure habitats, when they become known as “juveniles” or “spat” (Nash 1991). According to Morse

(1992), metamorphosis and settlement of larvae are induced by a series of biotic and chemical substances such as microalgae and bacteria as well as abalone mucus. The chance that an individual larva will survive to adulthood is very low as mortality rates exceed 99% (McShane 1992). Once abalone grow into sexually mature adults, the life cycle can be repeated (Figure 1.9). However, sexual maturity is only reached within four to seven years, depending on the species.

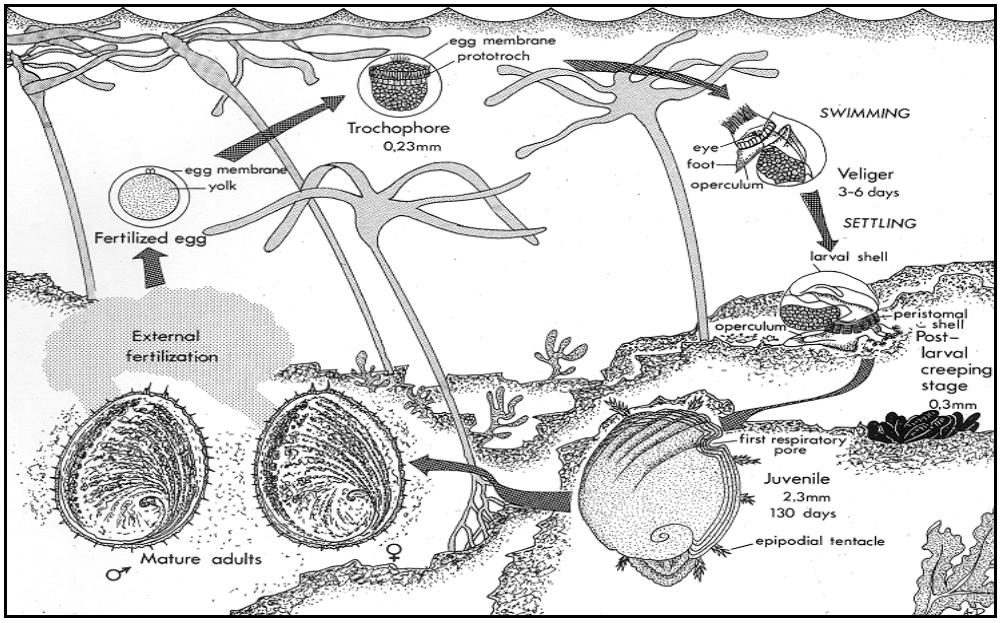


Figure 1.9 The life cycle of an abalone, from spawning to maturity (Tarr 1989).

1.5.4 FEEDING & HABITAT

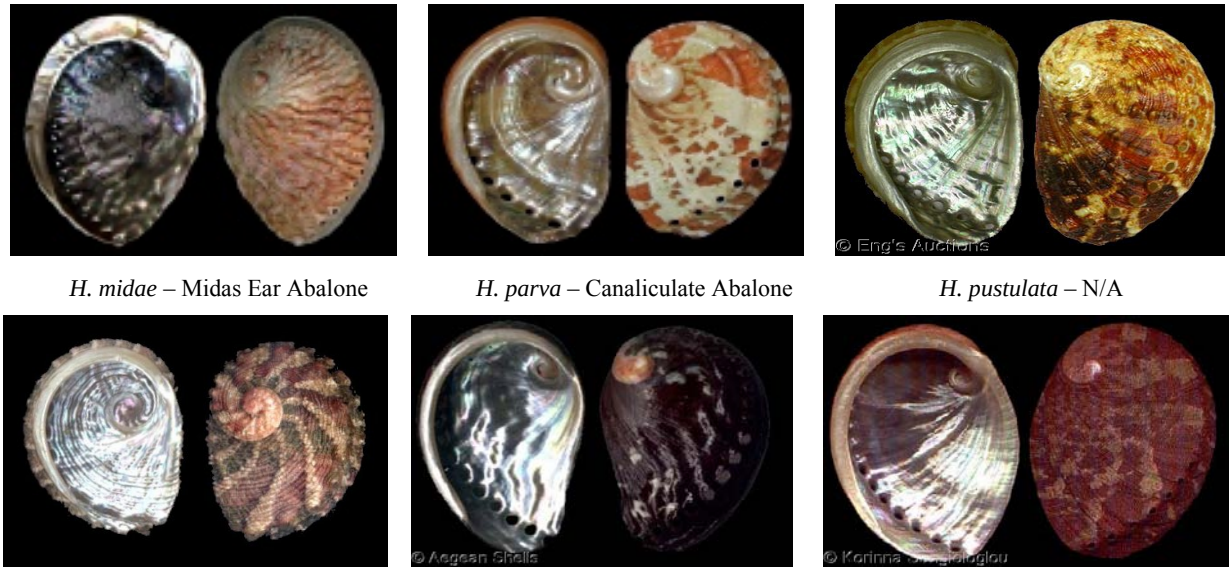
Abalone are herbivorous gastropods that are nourished entirely by algae. They are effective grazers but also selective about their choice of which algae – red, green and brown – they consume (Fallu 1991). Such preferences can sometimes be visualised by varying shell colour, and in some cases the colour of the foot. Abalone are sedentary, nocturnal feeders in nature, ranging over rocks in search of food from late afternoon to early morning. They tend to remain in one place once they have found a satisfactory feeding ground (Nash 1991). They are said to feed by trapping drifting pieces of food under their foot and ingesting small pieces at a time. Larvae depend on crustose corallines for settlement and feed off the diatoms or bacterial films (Day & Branch 2000), while juveniles prefer to feed on soft macro-algal species such as *Gracillaria* and *Ulva* (McShane 1992).

Worldwide, abalone are distributed in temperate and tropical coastal waters where there is a large supply of food resources (Nash 1991). They are found mostly stationary along reefs and rocky shores up to 30m deep and will only move when food becomes scarce after a long period (Geiger 2000). Shallow and turbulent waters are the preferred habitat and during the day, they can be found in rocky overhangs and dark crevices (Fallu 1991). Haliotids are not distributed globally but there

are four discrete regions of endemism namely South Africa, Australia, New Zealand and North America (Degnan *et al.* 2006).

1.6 SOUTH AFRICAN ABALONE

In southern Africa, there are five species of Haliotids endemic to the coastal waters, with a sixth species being found as far up as Mozambique. They include *H. midae, H. parva, H. pustulata, H. queketti, H. spadicea* and *H. speciosa*. These six abalone species can be distinguished from each other by their size as well as the shape and colour of their shells (Figure 1.10).



H. midae – Midas Ear Abalone | *H. parva* – Canaliculate Abalone | *H. pustulata* – N/A

H. queketti – Quekett's Abalone | *H. spadicea* – Blood-spotted Abalone | *H. speciosa* – Splendid Abalone

Figure 1.10 The shells of the six indigenous Southern African *Haliotis* species with generic names by which they are commonly known (www.gastropods.com/Taxon_pages/TN_Family_HALIOTIDAE.html).

1.6.1 *HALIOTIS MIDAE*

Haliotis midae, or "perlemoen" (from the Dutch paarlemoer, meaning mother-of-pearl) as it is locally known, is distributed between Cape Columbine on the West coast and Port St. John's on the East coast of South Africa, with the highest densities being found along the south-west coast (Figure 1.11; Godfrey 2003). Haliotids are found mostly in the rocky areas of strong wave action and the width of their shells is an adaptation to this unpredictable environment (Tarr 1989).

Haliotis midae is the largest of the South African species with the shell length reaching a maximum of up to 90mm and the width reaching 200mm in 30 years (Barkai & Griffiths 1986). Despite this, it is a slow-growing species only reaching sexual maturity around seven years. Interestingly, Wood and Buxton (1996) showed specimens of *H. midae* from the east coast to be faster growing than

those found on the west coast. This could be due to the varying degrees in temperature from the warmer east coast currents to the colder west coast currents, which suggests that water temperature does indeed play a role in the growth of abalone. The most intensive breeding periods for the animals occur during South African autumns and winters - March to September (Wood & Buxton 1996).

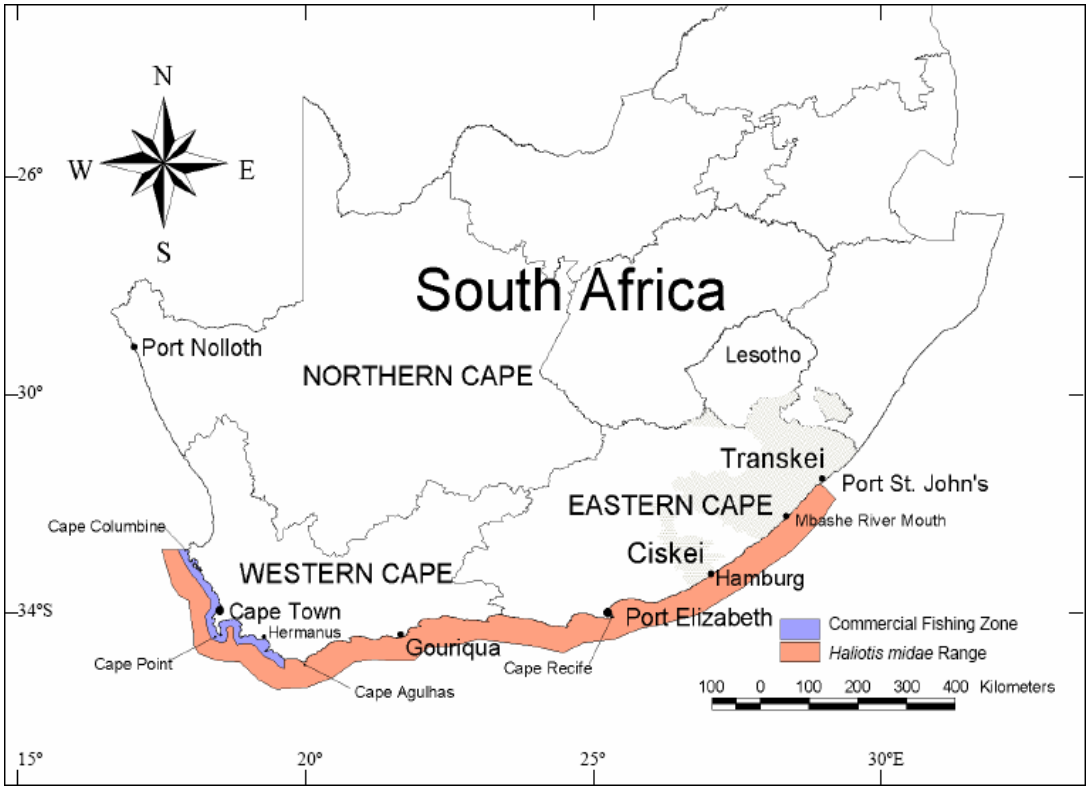


Figure 1.11 Distribution of *H. midae* and allocation of the highest density commercial fishing zones (Godfrey 2003).

Perlemoen have a reddish shell colour, with a pale cream or light brown muscular foot and yellow gills. When ripe, the female gonads are a green colour and those of the male are cream (Tarr 1989). The main sources of nutrition for adult *H. midae* vary with changes in the seasons. According to Barkai and Griffiths (1986), the kelps *Ecklonia maxima* and *Plocamium cornutum* form 80% of the abalone diet in summer, while in winter their diet includes various other macro-algae, particularly species of *Ulva* as there is a significant increase in the abundance of this particular type of algae during this period.

1.6.2 THREATS TO THE RESOURCE

The major reasons for the decline in the density of wild populations of *H. midae* include predation, loss of habitat and competition. All result in the mortality of juveniles. However, illegal harvesting and over fishing of natural stocks result in the most detrimental decline. Abalone are constantly being threatened by a number of predators in the wild throughout their various life stages. As eggs

and larvae, they are predated on by filter-feeding fish, while as juveniles, abalone often fall prey to the rock lobster, crabs, octopuses and fin fish (Tarr 1996). A major player in the decline of juveniles of *H. midae* is their natural predator, the Western Cape rock lobster *Jasus lalandii*. *Haliotis midae* also compete on a daily basis with the sea urchin *Parechinus angulosus* who utilizes abalone food sources as well as living space (Tarr 1996). On the contrary, sea urchins can also be beneficial to the survival of juveniles as they provide shelter for the concealment of young abalone, and therefore a form of protection against predators such as the rock lobster (Day & Branch 2000). *Jasus lalandii* play an indirect role in the loss of juvenile habitat as it feeds on the sea urchins, thereby exposing the juvenile abalone to potential dangers and forcing them to migrate (Tarr 1996).

In a more recent study, the effect of the removal of *P. angulosus* on the survival of abalone recruits and juveniles was investigated (Day & Branch 2002). The study concluded that, despite being in competition with sea urchins, recruits and juveniles will only survive in those areas in which urchins are found. In the absense of sea urchins, there is an increase in the accumulation of drift kelp which is an important food source for abalone. On the contrary, there is also a decline in the number of abalone recruits and juveniles. The experiments undertaken showed the crucial direct and indirect, as well as the positive and negative effects of *P. angulosus* on *H. midae* (Figure 1.12; Day & Branch 2002).

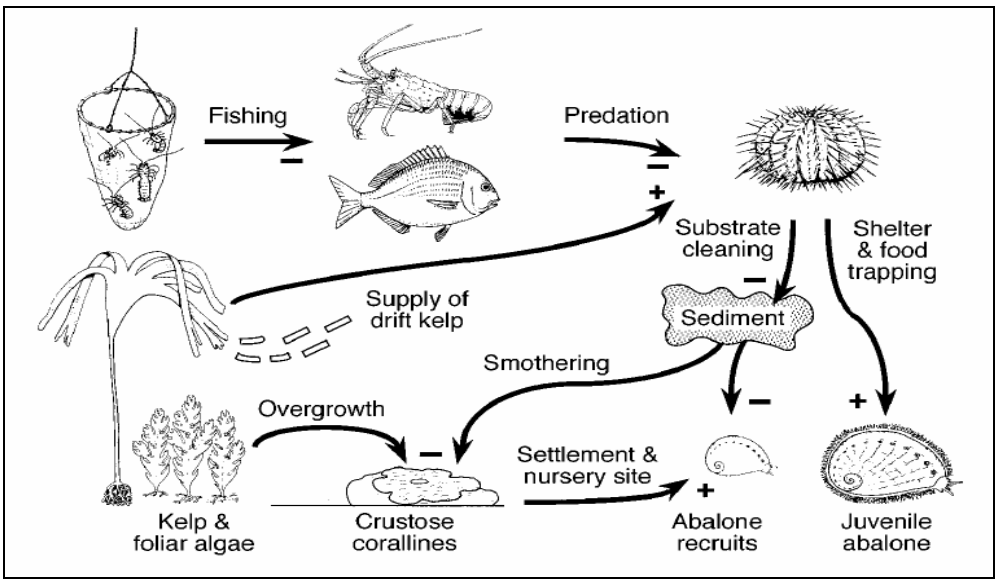


Figure 1.12 An illustration of the relationships, both positive and negative, encountered by *H. midae* in its natural environment (Day & Branch 2002).

i) Illegal Exploitation

Haliotis midae has become an extremely well known, lucrative marine resource over the past 20 years. The fact that it lives in exposed shallow waters makes it vulnerable to overexploitation by

commercial and recreational divers, as well as poachers (Godfrey 2003). Poaching has become the most important contributing factor in the decline of natural stocks of *H. midae*. It also seems to be having dire consequences in various facets of small fishing communities and the abalone industry that rely on the resource as means for survival. The overwhelming rise in illicit perlemoen trading came about in the early 1990's and by 2002, more abalone were being confiscated by law enforcement per year than was being harvested for commercial purposes, bringing it to the "brink of commercial extinction" (Steinberg 2005).

Corruption in the political sector has also contributed greatly to poaching. Law enforcers and other government authorities receive financial payment or other enticing incentives for "turning a blind eye" or assisting in illegal activities (Hauck & Sweijd 1999). Another fundamental obstacle to the survival of natural *H. midae* resources and commercial fisheries are organised crime syndicates; mostly Chinese businesses with connections to syndicates in the Far East that are involved in supplying the ever-growing black market. On the black market, fresh and frozen abalone meat fetches a very high price of up to R650/kg particularly in China where it is thought to delay senility and increase fertility (CITES 2007). When dried, the meat shrinks to one-tenth of its original size and the price per kilogram is increased ten-fold. Through crime syndicates, there is an influx of illicit abalone into the international market, which increases competition for legitimate industry (Steinberg 2005). Since the mid 1990's, illegal harvests amounted to more than 500 tons (Tarr 2003). More importantly, 50% of abalone being illegally fished were under the minimum legal size and therefore not yet able to reproduce, thus having dire consequences for the survival of the species.

ii) The "Formal Sector"

Poaching may be a major contributing factor to the demise of *Haliotis midae*, but legal fishing should not be ruled out as playing a significant role. In South Africa, the commercial fishery began in Gansbaai in 1949 when natives became more aware of the value that abalone meat had, particularly to the international market (Steinberg 2005). Overexploitation surfaced in the mid 1960's when harvesting of *H. midae* was a "free for all" occurrence and 2 800 tons of abalone were being brought in annually. During the 1970's annual catches were limited by quotas introduced by government, and the total allowable catch (TAC) per annum was drastically decreased to 700 tons (Tarr 2000). By the end of the 1990's the TAC had decreased by another 200 tons. The species is continually under threat, and therefore Marine and Coastal Management (MCM), a branch of the Department of Environmental Affairs and Tourism (DEAT), decreases the quota limit annually (Figure 1.13). Currently, the TAC stands at 125 tons per annum (CITES 2007). It is likely that

additional reductions may occur in years to come should the abundance of the species decrease further. This may consequently lead to the total closure of the abalone fishery.

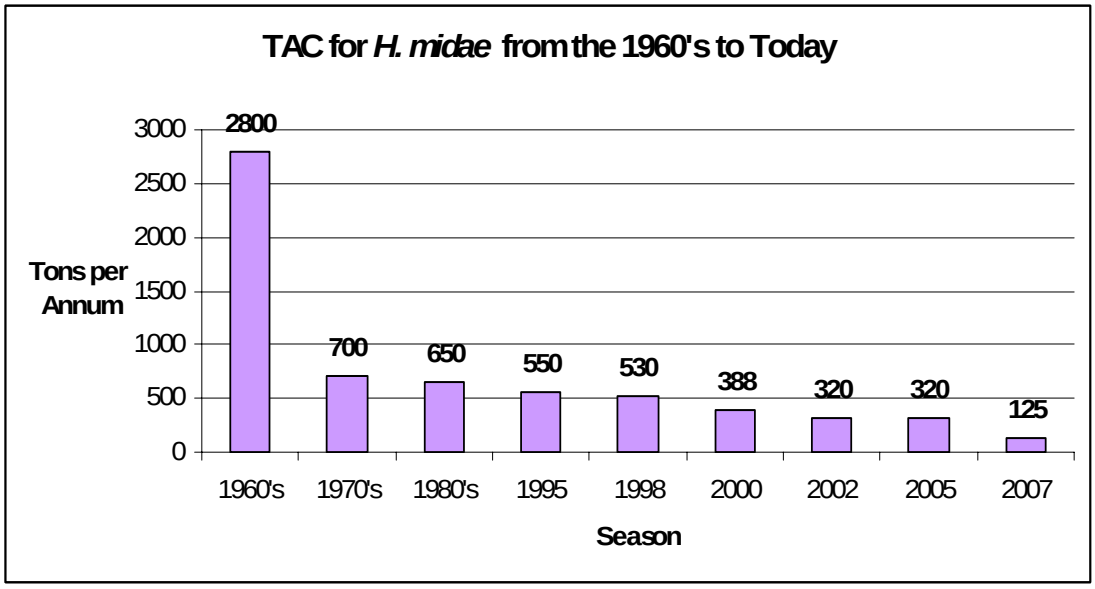


Figure 1.13 Indication of the decrease in TAC for the commercial harvesting of *Haliotis midae* over the last 50 years (Compiled from Tarr 2003 and CITES 2007).

Management of *H. midae* not only relies on the limitations of TAC but also on the minimum legal size (MLS) of harvested animals, bag limitations, closed seasons and recreational permits. For commercial and recreational divers the MLS of harvesting is 114 mm in shell length. This is the size reached by mature adults in eight to nine years. The MLS thus provides some degree of protection to undersized and immature animals (Tarr 2000). Recreational fishing has also been a contributing factor the decline of *H. midae* since its expansion in the early 1980's, hence the instatement of recreational abalone permits in 1983. In 1993, recreational catches peaked at 550 tons, equalling the TAC of commercial fisheries and consequently placing strain on the industry. In 2002, this resulted in the reduction of seasonal fishing lengths, restricting recreational fishing to weekends only and decreasing catches to just over 100 tons (Tarr 2003). Today, the removal of *H. midae* from its natural habitat is completely banned.

1.6.3 THE ABALONE MARKET

Abalone are primarily sold for profit and it is therefore important that the product be of the highest quality. According to Fallu (1991), it is fundamental to design an abalone farm "so that it suits the production of products designed for the market." Abalone is marketed in various forms including live, frozen, dried, cocktail and canned. Interestingly, the most consumed form happens to be the canned product. The major exporters of abalone are South Africa, Mexico, North America,

Australia and New Zealand. It is suggested that cultured abalone fetches a higher price in the Far East than fishery products (Oakes & Ponte 1996). Products are priced on market form as well as qualities such as colour, taste and texture. Fresh, live abalone can sell for approximately US$32/kg, while canned can sell for up to US$100/kg. The secretive processing methods for drying abalone meat can result in products often selling for anything from US$700 to US$2000 per kilogram (CITES 2007).

Live abalone is a favourite in Japan as it can be used for sushi, the staple Japanese diet, and is the best way to ensure the meat is fresh and of the best quality (Fallu 1991). Frozen abalone is the easiest to handle when exporting, and is almost as good as fresh meat products. Dried abalone weighs one-tenth the amount of its original flesh, therefore it is a product that stores well and is sold as a flavouring agent (Oakes & Ponte 1996). The biggest markets for cocktail sized abalone are those in Hong Kong and Japan. Canned abalone is the most common form of processing and the majority of the product is exported to China because it is the most preferred form for consumption (Oakes & Ponte 1996). Abalone shells are also a profitable and productive commodity because of their iridescent, mother-of-pearl layer. They can be used to make various pieces of jewellery, which can be sold for high prices.

1.7 ABALONE FARMING IN SOUTH AFRICA

Haliotis midae is the only one of the six indigenous species that is of commercial importance to the South African abalone industry. The other five species are exploited recreationally but are too rare and relatively small in comparison and are, therefore, of no use for commercial purposes (Cook 1998). In order to curb the rapid decline of the species in its natural surroundings, abalone farms were established in high-density coastal areas of the country. They provide a controlled environment in which to spawn and grow animals in artificial settings for commercial purposes. Even though abalone fisheries began in 1949, cultivation was only initiated in the early 1980's by funding from corporate fishing groups. Construction of farms commenced in the early 1990's and today there are 13 operational farms, most situated on the south-western Cape coast particularly in the Walker Bay region (Figure 1.14). The farms are all registered with the Abalone Farmers Association of Southern Africa (AFASA), and range from fully integrated farms with hatcheries, nurseries and grow-out facilities, to smaller farms designed to operate and specialise in one area (Cook 1998). With an annual production of around 750 tons, South Africa's abalone aquaculture industry has become an increasingly lucrative and internationally competitive business since its inception. Troell *et al*. (2006) stated that South Africa has become the largest producer of abalone

outside of the Asian market. The cultivation of abalone is a multi-million rand business and to date total investment has reached R190 million (Nick Loubser, AASA congress 2005). Abalone products for export include fresh, frozen, canned and dried forms. Highest prices are obtained for live, in the shell "medallion" sized animals between 80 to 100 mm (approximately 80g), sold to international restaurants (Sales & Britz 2001).

Figure 1.14 The location of commercial abalone farms in South Africa (Tarr 2000).

South African abalone farming was fortunate in that the basic technologies for cultivation were already developed by other abalone culturing countries such as Japan, Taiwan and the U.S.A (Britz 1991). It was, however, important to select and adapt appropriate techniques for the culture of *H. midae* to South African conditions. The key objectives of abalone farmers were to establish techniques for mass seed production, maximise growth rates and secure a sustainable source of feed for cultured animals (Britz 1991). Genade *et al*. (1988) were the first to demonstrate that the cultivation of *H. midae* could be conducted, and were instrumental in the successful spawning and rearing of the species in captivity (Cook 1998). The Californian model was most suitable to the South African conditions: this involved highly intensive, shore-based culture systems with seawater being continually pumped through concrete rearing tanks, through recirculation technology, and in which abalone are reared in high densities on a diet of kelp. Some farms have hatchery and grow-out facilities whilst others rely on purchasing juveniles from other hatcheries (Troell *et al*. 2006). The three main sources of abalone feed include harvested *Ecklonia maxima*, cultured algae (*Gracillaria*) and artificial feed (Britz 1991).

1.7.1 BROODSTOCK & SPAWNING

Broodstock animals can be obtained either from natural stocks or from animals grown on existing farms. The choice of parental stock for a successful hatchery is based on three important features

(Fallu 1991): 1) the *onset of sexual maturity*, which varies among species. Sexually mature animals are identifiable when one can distinguish between the sexes. This could take up to seven years in some species. 2) The *fecundity* of mature animals is associated with sexual maturity. Abalone gamete production is proportional to body size. Therefore, animals should be chosen when gonad size and fertility are at their peak. 3) *Growth rates* of abalone play an equal if not more important role. The faster the growth rate, the quicker the turnover, which results in better profits.

The control of spawning is a crucial measure for successful seed production in any hatchery. In their natural environment, abalone spawn spontaneously once or twice a year. However, in captivity gravid males and females can be induced to spawn all year round. *Haliotis midae* is a difficult animal to spawn when not in its natural habitat, therefore most farms in the South African industry condition their chosen broodstock in tanks for six to twelve months before inducing them to spawn (Figure 1.15a, b). Induced spawning is carried out with the chemical stimulant hydrogen peroxide (Danie Brink, Stellenbosch University, personal communication), a method first suggested by Morse (1984). The males and females are spawned simultaneously but in separate tanks. The ova and sperm are drained out of their parental tanks and placed together into one large tank containing filtered seawater, at temperatures between 18ºC and 22ºC, to allow natural fertilisation. Minimum levels of bacteria need to be maintained in order to prevent contamination and mortality of the eggs (Fallu 1991). Once the eggs have hatched, a 48-hour larval period of metamorphosis is maintained followed by the larvae being filtered into settlement tanks. Inside these tanks are a number of settlement plates inoculated with diatom biofilms and various micro-algal nutrients that help in stimulating the settlement of the larvae (Cook 1998). The spat remain in this nursery-like setting for up to six months; they are then transferred to larger tanks to go through a period of weaning.

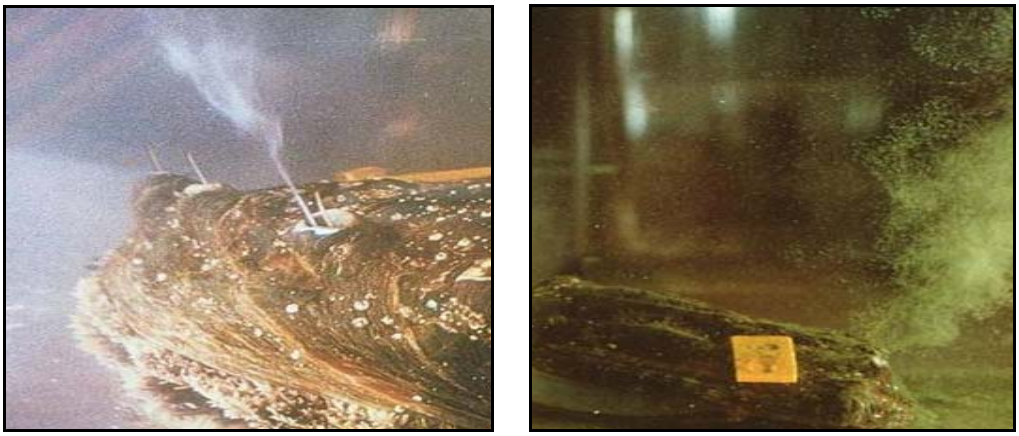


Figure 1.15 Broodstock animals spawning in captivity. Gametes are released via the respiratory pores in the shell. a) a male spawning and b) a female during spawning (Sorgeloos 1997).

1.7.2 LARVAL SETTLEMENT & SPAT REARING

Abalone larvae are positively buoyant, meaning that they float. At this point in most hatchery systems, larvae with "creeping ability" are transferred into settlement tanks (Jarayabhand & Paphavasit 1996). Within the settlement tanks there are a number of corrugated plastic or fibreglass sheets, referred to as plates, vertically suspended in rows (Nash 1991). The larvae remain afloat for a day or so, and then become neutrally buoyant (Fallu 1991). After a few weeks of floating, depending on the species and water temperature, larval settlement takes place and the larvae take up their residence on the settlement plates. *Haliotis midae* are then placed on a diet of artificial foods such as Abfeed™ (Marifeed Pty Ltd, South Africa) and seaweed-based pellets such as Midae Meal MM-1c (Eric-Piet Pty Ltd, Namibia), as well as natural seaweeds such as *Ecklonia*, *Gracillaria* and *Ulva* (Danie Brink, personal communication). Post-settlement larvae or spat, as they are termed, become dependant on the plates for survival. Water quality and plate density are important factors of settlement to ensure optimal growth and maximum survival of spat as they compete with continuously growing diatoms for space (Fallu 1991). The growth rate of the spat is very much dependent on the food they consume and the temperature of their surrounding water. To maintain hygienic conditions, uneaten food is regularly removed and water is often exchanged (Nash 1991). It is important and more economical that juveniles have a diet devoid of diatoms and micro-algae in order for them to be moved to the grow-out phase of the production cycle. The spat stage is the last stage before grow-out and is maintained for about four to six months.

1.7.3 GROW-OUT

Once the juveniles reach a suitable age (approximately one year) they are moved to rows of outdoor tanks referred, to as raceways (Figure 1.16). At this stage, they should be completely dependent on macro-algae and artificial foods for nutrition (Nash 1991). The next three to six years, depending on growth rates of individual species, are spent in the grow-out phase until the abalone reach their minimum market size (around 10 cm in shell length) and are ready for processing and export. At this point, young adult abalone can also be released into their innate environment in an attempt to restock natural populations (Landau 1992). This is referred to as ranching, a process that allows marketable sized abalone to be harvested from the wild and sold for profit after processing.

1.7.4 SOUTH AFRICAN ABALONE RANCHING

A subsidiary of the abalone aquaculture industry has been the advancement of abalone ranching. Hatchery produced larvae are carefully reared and reseeded into kelp beds outside their natural distribution. Once they have grown to a marketable size in their innate environments, they are used for commercial harvesting (Troell *et al.* 2006). Abalone seeding can also be applied to natural stock

enhancement where the main aim is to rehabilitate depleted and over-fished populations; this can potentially be used as a solution to the increasing pressures exerted by poachers and private divers. Interestingly this is not the main reason for ranching. The key factor for ranching and stock enhancement is to improve turnover rates among fisheries and therefore increase profit margins (Godfrey 2003). Over the last half of the 20th century, stock enhancement programmes have been initiated and implemented for a variety of fish and shellfish worldwide.

Figure 1.16 A layout of the raceways and manner in which juveniles are reared for 3-6 years until they reach export size at Roman Bay abalone farm in Gansbaai.

Sweijd *et al*. (1998) were the first to initiate the experimentation with the seeding of *Haliotis midae* in South Africa. The experiment took place on the north-west coast at Port Nolloth, about 400km north from the natural range of where *H. midae* is found. They chose that particular area because its environment resembles that of the ideal south-west coast which was once home to the now extinct *Haliotis saldanhae*, making it a suitable site for a reseeding programme. Two cohorts of abalone juveniles, four to six months old and grown in the hatchery of Port Nolloth Sea Farms, where transplanted into three different and carefully selected sites of the Stilbaai area around Port Nolloth. Another similar project was launched at Gouriqua situated along the south coast (De Waal *et al*. 2003). Although there were relatively high recovery rates, the ventures were not as successful as had been anticipated.

Although ranching of abalone, as solution to decreasing the rate of the extinction of the species seems like a feasible initiative, one needs to ask various important questions: How viable is it? With the influx of animals into the environment, will this not just give poachers greater motivation to continue their illegal trade? In addition, how genetically diverse will hatchery-reared animals be from their naturally grown counterparts? Industry will need to ensure success on all the levels, through various pilot studies, that will guarantee the best possible results.

1.8 AIMS & OBJECTIVES

An initiative has been put into place to assist the abalone farming industry in terms of improving productivity. When phenotypic selection becomes unreliable, genetic factors need to be brought in to aid in the enhancement of *Haliotis midae* to ensure optimal and more profitable production. An abalone initial five-year enhancement programme was launched at the beginning of 2006 and is funded and managed by three sectors. From the industry sector, five commercial abalone farms have randomly submitted a number of broodstock animals and their spawn, as well as funding for the project. From academia, Stellenbosch University is developing and supplying the genetic tools that will be used to perform analyses and is a source of funding for the project. From the government sector, the Innovation Fund has supplied half of the funding (IF project: TM1134FP).

The aims of this research are to:

1. Provide the molecular genetic tools that will identify hatchery-produced animals entered into a growth performance recording scheme (PRS) as being of good or poor quality in terms of their survival and growth rates.
2. Assist the involved hatcheries in using the results obtained for implementing the best possible solutions at an early stage and thereby enrich their stock with genotypes that are associated with "good performers."

The objectives of this project are to:

1. Isolate and characterise species-specific microsatellite markers using the most efficient technologies available.
2. Choose suitable markers, from those developed, to genotype broodstock and their offspring entered into the PRS.
3. Assign parentage to the fast and slow growing animals as guidelines to abalone farmers.

1.9 REFERENCES

Abuin M, Martinez P, Sanchez L (1996) Localisation of the repetitive telomeric sequence (TTAGGG)n in four salmonid species. *Genome* 39: 1035-1038

Agresti JJ, Seki S, Cnaani A, Poompuang S, Hallerman EM, Umiel N, Hulata G, Gall GAE, May B (2000) Breeding new strains of tilapia: development of an artificial centre of origin and linkage map based on AFLP an microsatellite loci. *Aquaculture* 185: 43-56

Araneda C, Neira R, Iturra P (2005) Identification of a dominant SCAR marker associated with colour traits in Coho salmon (*Oncorhynchus kisutch*). *Aquaculture* 247: 67-73

Avise JC (1994) Molecular tools. In: Molecular Markers, Natural History and Evolution. Chapman and Hall, New York, pp44-91

Baldwin J, Elias JP, Wells RMG, Donovan DA (2007) Energy metabolism in the tropical abalone, *Haliots asinine* Linné: Comparisons with the temperate abalone species. *Journal of Experimental Marine Biology and Ecology* 342: 213-225

Baranski M, Loughnan S, Austin CM, Robinson N (2006) A microsatellite linkage map of the blacklip abalone, *Haliotis rubra*. *Animal Genetics* 37: 563-570

Barkai R, Griffiths CL (1986) Diet of the South African abalone *Haliotis midae*. *South African Journal of Marine Science* 4: 37-44

Beaumont MA, Bruford MW (1999) Microsatellites in conservation genetics. In: Goldstein DB, Schlötterer C (eds), Microsatellites: Evolution and Applications. Oxford University Press, New York, pp165-182

Beaumont AR, Hoare K (2003) Biotechnology and Genetics in Fisheries and Aquaculture. Blackwell Science Ltd, Oxford, England.

Blancou J (2001) A history of the traceability of animals and animal products. *Revue Scientifique et Technique* 20: 420-425

Borrell YJ, Álvarez J, Vázquez E, Pato CF, Tapia CM, Sánchez JA, Blanco G (2004) Applying microsatellites to the management of farmed turbot stocks (*Scophthalmus maximus* L.) in hatcheries. *Aquaculture* 241: 133-150

Bowling AT (2001) Historical development and application of molecular genetic tests for horse identification and parentage control. *Livestock Production Science* 72: 111-116

Britz PJ (1991) Prospects for abalone aquaculture. In: Heath RGM (ed) *Aquaculture 90': Proceedings of a symposium by the Aquaculture Association of South Africa, Stellenbosch, South Africa 11-13 July 1990*. AASA, Pretoria, pp110-113

Canapa A, Cerioni PN, Barucca M, Olmo E, Caputo V (2002) A centromeric satellite DNA may be involved in heterochromatin compactness in gobiid fishes. *Chromosome Research* 10: 297-304

Chambers GK, MacAvoy ES (2000) Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology Part B* 126: 455-476

Chang DK, Metzgar D, Wills C, Boland CR (2001) Microsatellites in the eukaryotic DNA mismatch repair genes as modulators of evolutionary mutation rate. *Genome Research* 11: 1145-1146

Christiakov DA, Hellermans B, Volckaert FAM (2006) Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. *Aquaculture* 255: 1-29

Claerebuodt C (1999) Fertilisation success in spatially distributed populations of benthic free-spawners: a simulation model. *Ecology Model* 121: 2221-2233

Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES, 2007) *Fourteenth meeting of the Conference of the Parties*, *The Hague, Netherlands, 3-15 June 2007*

Cook P (1998) The current status of abalone farming in South Africa. *Journal of Shellfish Research* 17: 601-602

Counihan RT, McNamara DC, Souter DC, Jebreen EJ, Preston NP, Johnson CR, Degnan BM (2001) Pattern, synchrony and predictability of spawning of the tropical abalone *Haliotis asinina* from Heron Reef, Australia. *Marine Ecology Progress Series* 213: 193-202

Cunningham EP, Meghen CM (2001) Biological identification systems: genetic markers. *Revue Scientifique et Technique* 20: 491-499

Danzmann RG, Jackson TR, Ferguson MM (1999) Epistasis in allelic expression at upper temperature tolerance QTL in rainbow trout. *Aquaculture* 173: 45-58

Davis PJ, Hetzel JS (2000) Integrating molecular genetic technology with traditional approaches for genetic improvement in aquaculture species. *Aquaculture Research* 31: 3-10

Day E, Branch GM (2000) Relationships between recruits of abalone *Haliotis midae*, encrusting corallines and the sea urchin *Parechinus angulosus*. *South African Journal of Marine Science* 22: 137-144

Degnan SM, Imron, Geiger DL, Degnan BM (2006) Evolution in temperate and tropical seas: Disparate patterns in southern hemisphere abalone (Mollusca: Vetigastropoda: Haliotidae). *Molecular Phylogenetics and Evolution* 41: 249-256

De Waal SWP, Branch GM, Navarro R (2003) Interpreting evidence of dispersal by *Haliotis midae* juveniles seeded in the wild. *Aquaculture* 221: 299-310

Dieringer D, Schlötterer C (2003) Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequence of nine species. *Genome Research* 13: 2242-2251

Dodds KG, Tate MI, Sise JA (2005) Genetic evaluation using parentage information from genetic markers. *Journal of Animal Science* 83: 2271-2279

Dunham RA (2004) History of genetic biotechnology in aquaculture and fisheries. In: Aquaculture and Fisheries Biotechnology: Genetic Approaches. CABI Publishing, Oxfordshire, United Kindom, pp1-6

Dunham RA, Majumdar K, Hallerman E, Bartley D, Mair G, Hulata G, Liu Z, Pongthana N, Bakos J, Penman D, Gupta M, Rothlisberg P, Hoerstgen-Schwark G (2001) Review of the status of aquaculture genetics. In: Subasinghe RP, Bueno P, Phillips MJ, Hough C, McGladdery SE, Arthur JR (eds), *Technical Proceedings of the Conference on Aquaculture in the Third Millennium, Bangkok, Thailand, 20-25 February 2000*. NACA, Bangkok, and FAO, Rome, pp129-157

Edwards YJ, Elgar G, Clark MS, Bishop MJ (1998) The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish *Fugu rubripes*: perspectives in functional and comparative genome analysis. *Journal of Molecular Biology* 278: 843-854

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics* 16: 551-558

Epplen JT, Kyas A, Maueler W (1996) Genomic simple repetitive DNAs are targets for differential binding of nuclear proteins. *FEBS Letters* 389: 92-95

Fallu R (1991) Abalone Farming. Fishing News Books, Cornwall, England.

Féral JP (2002) How useful are the genetic markers in attempts to understand and manage marine biodiversity? *Journal of Experimental Marine Biology and Ecology* 268: 121-145

Ferguson MM, Danzmann RG (1998) Role of genetic markers in fisheries and aquaculture: useful tools or stamp collecting? *Canadian Journal of Fisheries and Aquatic Sciences* 55: 1553-1563

Fernández J, Toro MA (2006) A new method to estimate relatedness from molecular markers. *Molecular Ecology* 15: 1657-1667

Fessehaye Y, El-bialy Z, Rezk MA, Crooijmans R, Bovenhuis H, Komen H (2006) Mating systems and male reproductive success in Nile tilapia (*Oreochromis niloticus*) in breeding hapas: a microsatellite analysis. *Aquaculture* 256: 148-158

Food and Agricultural Organisation (FAO, 1995) Review of the state of world fishery resources: aquaculture. *Fisheries Circular* no. 886. Food and Agricultural Organisation of the United Nations, Rome

Foresti F, Oliveira C, Galetti Jr. PM, Almeida-Toledo LF (1993) Synaptonemal complex analysis in spermatocytes of tilapia, *Oreochromis niloticus* (Pisces, Cichlidae). *Genome* 36: 1124-1128

Fresco JR, Alberts BM (1960) The accommodation of noncomplementary bases in helical polyribonucleotides and deoxyribonucleic acids. *Proceedings of the National Academy of Science USA* 85: 311-321

Geiger DL (2000) Distribution and biogeography of the recent Haliotidae (Gastropoda: Vetigastropoda) world-wide. *Bollettino Malacologico* 35: 57-120

Genade AB, Hirst AL, Smith CJ (1988) Observations on the spawning, development and rearing of the South African abalone *Haliotis midae Linn*. *South African Journal of Marine Science* 6: 3-12

Godfrey BP (2003) The potential of abalone stock enhancement in the Eastern Cape province of South Africa. M.Sc. thesis, Rhodes University, Grahamstown.

Hahn KO (1989) CRC Handbook of culture of abalone and other marine gastropods. CRC Press, Florida, USA

Hancock JM (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Goldstein DB, Schlötterer C (eds), Microsatellites: Evolution and Applications. Oxford University Press, Oxford, England, pp1-9

Hardege JD, Bentley MG, Beckmann M, Muller C (1996) Sex pheromones in marine polychaetes: volatile organic substances (VOS) isolated from *Arenicola marina*. *Marine Ecology Progress Series* 139: 157-166

Håstein T, Hill BJ, Berthe F, Lightner DV (2001) Traceability of aquatic animals. *Revue Scientifique et Technique* 20: 564-583

Hauck M, Sweijd NA (1999) A case study of abalone poaching in South Africa and its impact on fisheries management. *ICES Journal of Marine Science* 56: 1024-1032

Hayes BJ, Gjuvsland A, Omholt S (2006) Power of QTL mapping experiments in commercial Atlantic salmon populations, exploiting linkage and linkage disequilibrium and effect of limited recombination in males. *Heredity* 97: 19-26

Henry NR (1995) Studies related to the artificial spawning and culture of the abalone, *Haliotis midae*, Linné 1978. M.Sc. thesis, University of Cape Town, Cape Town

Herbinger CM, Doyle RW, Pitman ER, Paquet D, Mesa KA, Morris DB, Wright JM, Cook D (1995) DNA fingerprint-based analysis of paternal and maternal effects on offspring growth and survival in commonly reared rainbow trout. *Aquaculture* 137: 245-256

Hoyheim B, Danzmann R, Guyomard R, Holm L, Powell R, Taggart J (1998) Constructing a genetic maps of salmonid fishes: Salmap. *Abstracts in Plant and Animal Genome* 6: 165

Hubert S, English LJ, Landau BJ, Guo X, Hedgecock D (2000) Microsatellite analysis of trisomic families in the Pacific oyster, *Crassostrea gigas* Thunberg. *Journal of Shellfish Research* 19: 615

Hubert S, Hedgecock D (2004) Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. *Genetics* 168: 351-362

Hulata G (2001) Genetic manipulations in aquaculture: a review of stock improvement by classical and modern technolohies. *Genetica* 111: 155-173

Jackson TR, Ferguson MM, Danzmann RG, Fishbook AG, Ihssen PE, O'Connell M, Crease TJ (1998) Identification of two QTL influencing upper temperature tolerance in three rainbow trout (*Oncorhynchus mykiss*) half-sib families. *Heredity* 80: 143-151

Jackson TR, Martin-Robichaud DJ, Reith ME (2003) Application of DNA markers to the management of Atlantic halibut (*Hippoglossus hippoglossus*) broodstock. *Aquaculture* 220: 245-259

Jarayabhand P, Paphavasit N (1996) A review of the culture of tropical abalone with special reference to Thailand. *Aquaculture* 140: 159-168

Jeffreys AJ, Neil DL, Neumann R (1998) Repeat instability at human minisatellites arising from meiotic recombination. *EMBO Journal* 17: 4147-4157

Jerry DR, Evans BS, Kenway M, Wilson K (2006) Development of a microsatellite DNA parentage marker suite for black tiger shrimp *Penaeus monodon*. *Aquaculture* 255: 542-547

Jerry DR, Preston NP, Crocos PJ, Keys S, Meadows JRS, Li Y (2004) Parentage determination of Kuruma shrimp *Penaeus* (*Marsupenaeus*) *japonicus* using microsatellite markers (Bate). *Aquaculture* 235: 237-247

Kai W, Kikuchi K, Fujita M, Suetake H, Fujiwara A, Yoshiura Y, Ototake M, Venkatesh B, Miyaki K, Suzuki Y (2005) A genetic linkage map for the tiger pufferfish, *Takifugu rubripes*. *Genetics* 171: 227-238

Kause A, Ritola O, Paananen T, Mäntysaari E, Eskelinen U (2002) Coupling body weight and its composition: a quantitative genetic analysis in rainbow trout. *Aquaculture* 211: 65-79

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738

Kimura M, Ohta T (1973) Stepwise mutational model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Science USA* 75: 2868-2872

Klug WS, Cummings MR (2000) Mitosis and meiosis. In: Concepts of Genetics, Sixth Edition. Prentice Hall, New Jersey, USA, pp17-44

Knaepkens G, Bervoets L, Verheyen E, Eens M (2004) Relationship between population size and genetic diversity in endangered populations of the European bullhead (*Cottus gobio*): implications for conservation. *Biological Conservation* 115: 403-410

Landau M (1992) Commonly cultures molluscs. In: Introduction to Aquaculture. John Wiley & Sons, Inc, Toronto, Canada, pp181-185

Levitan DR (1991) Influence of body size and population density on fertilisation success and reproductive output in a free-spawning invertebrate. *Biology Bulletin* 181: 261-268

Li Q, Kijima A (2006) Microsatellite analysis of gynogenetic families in the Pacific oyster, *Crassostrea gigas*. *Journal of Experimental Marine Biology and Ecology* 331: 1-8

Li Y, Wilson KJ, Byrne K, Whan V, Iglesis D, Lehnert SA, Swan J, Ballment B, Fayazi Z, Kenway M, Benzie J, Pongsomboon S, Tassanakajon A, Moore SS (2000) International collaboration on genetic mapping of the black tiger shrimp, *Penaeus monodon*: progress update. In: *Plant and Animal Genome VIII Abstracts*. Available at www.intl-pag.org/8/abstracts

Lindberg DR (1992) Evolution, distribution and systematics of Haliotidae. In: Shepherd SA, Tegner MJ, Guzman del Proo SA (eds) Abalone of the World. Fishing News Books, Oxford, England, pp3-18

Liu ZJ, Cordes JF (2004) DNA markers and their applications in aquaculture genetics. *Aquaculture* 238: 1-37

Liu X, Liu X, Zhang G (2007) Identification of quantitative trait loci for growth related traits in the Pacific abalone *Haliotis discus hannai* Ino. *Aquaculture Research* 38: 789-797

Lucas T, Macbeth M, Degnan SM, Knibb W, Degnan BM (2006) Heritability estimates in the tropical abalone *Haliotis asinina* using microsatellites to assign parentage. *Aquaculture* 259: 146-152

Lymbery AJ, Doupé RG, Jenkins G, Thorne T (2000) Genetics in the aquaculture industry. *Aquaculture Research* 31: 1-2

Mackay TFC (2001) The genetic architecture of quantitative traits. *Annual Reviews in Genetics* 35: 303-309

Martinez V (2006) Importance and implementation of molecular markers in selective breeding programmes for aquaculture species. In: *8th World Congress on Genetics Applied to Livestock Production, 13-18 August 2006, Belo Horizonte, Brazil*

McShane PE (1992) Early life history of abalone: a review. In: Shepherd SA, Tegner MJ, Guzman del Proo SA (eds) Abalone of the World. Fishing News Books, Oxford, England, pp120-138

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997) Genome mapping, molecular markers and markers-assisted selection in crop plants. *Molecular Breeding* 3: 87-103

Morse DE (1984) Biochemical and genetic engineering for improved production of abalone and other valuable molluscs. *Aquaculture* 39: 263-282

Morse DE (1992) Molecular mechanisms controlling metamorphosis and recruitment in abalone larvae. In: Shepherd SA, Tegner MJ, Guzman del Proo SA (eds) Abalone of the World. Fishing News Books, Oxford, England, pp107-119

Moxon RE, Wills C (1991) DNA microsatellites: agents of evolution? *Scientific American* 280(1): 94-99

Nash CE (1991) The production of abalone. In: World Animal Science C4: Production of Aquatic Animals. Elsevier, Amsterdam, The Netherlands, pp173-181

Neff BD (2001) Genetic paternity analysis and breeding success in bluegill sunfish (*Lepomis macrochirus*). *Journal of Heredity* 92: 111-119

Norris AT, Bradley DG, Cunningham EP (2000) Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* 182: 73-83

Oakes FR, Ponte RD (1996) The abalone market: Opportunities for cultured abalone. *Aquaculture* 140: 187-195

O'Brien SJ (1991) Molecular genome mapping: lessons and prospects. *Current Opinions in Genetic Development* 1: 105-111

Ozaki A, Sakamoto T, Khoo S, Nakamura K, Coimbra MR, Akutsu T, Okamoto N (2001) Quantitative trait loci (QTL) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Molecular Genetics and Genomics* 265: 23-31

Qin Y, Liu X, Zhang H, Zhang G, Guo X (2007) Identification and mapping of amplified fragment length polymorphism markers linked to shell colour in Bay scallop *Argopecten irradians* (Lamarck, 1819). *Marine Biotechnology* 9: 66-73

Robinson BD, Wheeler PA, Sundin K, Sikka P, Thorgaard GH (2001) Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *The American Genetics Association* 92: 16-22

Saavedra C, Bachère E (2006) Bivalve genomics. *Aquaculture* 256: 1-14

Sakamoto T, Danzmann RG, Gharbi K, Howard P, Ozaki A, Khoo SK, Woram RA, Okamoto N, Ferguson MM, Ihssen OE (1999) A microsatellite linkage map of rainbow trout (*Oncorhynchus mykiss*) characterized by large sex-specific differences in recombination rates. *Genetics* 155: 1331-1345

Sales J, Britz PJ (2001) Research on abalone (*Haliotis midae* L.) cultivation in South Africa. *Aquaculture Research* 32: 863-874

Salguiero P, Carvalho G, Collares-Pereira MJ, Coelho MM (2003) Microsatellite analysis of the genetic population structure of the endangered cyprinid *Anaecypris haspanica* in Portugal: implications for conservation. *Biological Conservation* 109: 47-56

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8: 552-560

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365-371

Seaton G, Haley CS, Knott SA, Kearsey M, Visscher PM (2002) QTL express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* 18: 339-340

Sekino M, Hara M (2007) Linkage maps for the Pacific abalone (genus *Haliotis*) based on microsatellite DNA markers. *Genetics* 175: 945-958

Selvamani MJP, Degnan SM, Degnan BM (2001) Microsatellite genotyping of individual abalone larvae: parentage assignment in aquaculture. *Marine Biotechnology* 3: 478-485

Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover. *Science* 191: 528-535

Soller M (1994) Marker-assisted selection: an overview. *Animal Biotechnology* 5 193-207

Somorjai IML, Danzmann RG, Ferguson MM (2003) Distribution of temperature tolerance quantitative trait loci in arctic charr (*Salvelinus alpinus*) and inferred homologies in rainbow trout (*Oncorhynchus mykiss*). *Genetics* 165: 1443-1456

Sorgeloos P (1997) Slide and overhead illustrations of molluscan embryo development. [CD-ROM]. Available at: http://www.dec.ctu.edu.vn/cdrom/cd3/general%20aspects/mollusc/les19/

Steinberg J (2005) The illicit abalone trade in South Africa. Available at: The Institute for Security Studies http://www.iss.co.za/pubs/papers/105/Paper105.htm

Strachan T, Read AP (2004) Genetic mapping of Mendelian characters. In: Human Molecular Genetics, 3rd Edition. Garland Science, London, pp397-413

Streelman JT, Kocher TD (2002) Microsatellite variation associated with prolactin expression and growth of salt-challenged tilapia. *Physiological Genomics* 9: 1-4

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution* 15: 199-203

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

Tarr RJQ (1989) Abalone. In: Payne AIL, Crawford RJM (eds) Oceans of Life off Southern Africa. Vlaeberg Publishers, Cape Town, South Africa, pp62-69

Tarr RJQ (2000) The South African abalone (*Haliotis midae*) fishery: a decade of challenges and change. In: Campbell A (ed) Workshop on Rebuilding Stock in British Columbia, in a special publication of *Canadian Journal of Fisheries and Aquatic Sciences* 130: 32-40

Tarr RJQ (2003) Perlemoen: The South African abalone. Available at http://www.environment.gov.za

Tarr RJQ, Williams PVG, Mackenzie AJ (1996) Abalone, sea urchins and rock lobster: A possible ecological shift that may affect traditional fisheries. *South African Journal of Marine Science* 17: 319-323

Taylor HH, Ragg NLC (2005) The role of body surfaces and ventilation in gas exchange of the abalone, *Haliotis iris*. *Journal of Comparative Physiology B* 175: 463-478

Troell M, Robertson-Andersson D, Anderson RJ, Bolton JJ, Maneveldt G, Halling C, Probyn T (2006) Abalone farming in South Africa: An overview with perspectives on kelp resources, abalone feed, potential for on-farm seaweed production and socio-economic importance. *Aquaculture* 257: 266-281

Uki N, Kikuchi S (1984) Regulation of maturation and spawning of an abalone, *Haliotis* (Gastropoda) by external environmental factors. *Aquaculture* 39: 247-261

Waldbieser GC, Bosworth BG, Nonneman DJ, Wolters WR (2001) A microsatellite-based genetic linkage map for channel catfish, *Ictalurus punctatus*. *Genetics* 158: 727-734

Wilkins NP (1981) The rationale and relevance of genetics in aquaculture: an overview. *Aquaculture* 22: 209-228

Wood AD, Buxton CD (1996) Aspects of the biology of the abalone *Haliotis midae* (Linné, 1758) on the east coast of South Africa. 1 Feeding Biology. *South African Journal of Marine Science* 17: 61-68

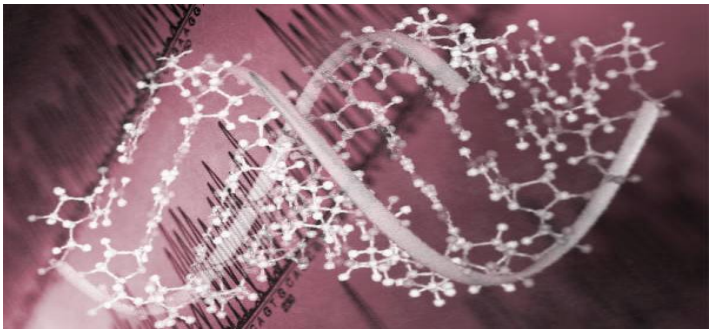
Young ND (1999) A cautiously optimistic vision for marker-assisted breeding. *Molecular Breeding* 5: 505-510

Yu Z, Guo X (2006) Identification and mapping of disease-resistance QTLs in the eastern oyster, *Crassostrea virginica* Gmelin. *Aquaculture* 254: 160-170

Zeek E, Harder T, Beckmann M (1998) Uric acid: the sperm release pheromone of the marine polychaete *Platyneris dumerilii*. *Journal of Chemical Ecology* 24: 13-22

Chapter 2

Microsatellite Isolation & Characterisation for the South African Abalone, *Haliotis midae*



The following chapter is presented in the form of a full-length manuscript in preparation for future publication.

ABSTRACT

This study isolated and characterised 36 microsatellite loci in the South African abalone, *Haliotis midae*. Using the enrichment-based FIASCO method, markers were obtained from the screening of random amplified polymorphic DNA products and a size-selected genomic DNA library. The majority of microsatellites isolated were polymorphic and could be characterised structurally into perfect, imperfect or compound repeats. PCR primers designed to amplify 15 loci were used to screen 32 individuals from a natural *H. midae* population in Black Rock. All loci were polymorphic with an average of 14.3 alleles per locus. Significant deviations from HWE were observed at three loci due to an excess of homozygotes. Expected heterozygosity and PIC values were high, suggesting that loci could be useful for parentage, linkage and QTL mapping, MAS and population genetics.

2.1 INTRODUCTION

The meat of *Haliotis midae* is worth thousands of rand per kilogram, making it a highly valued and sought after marine resource and one that has been heavily exploited over the last decade. As a result of the combined effects of illegal poaching and commercial harvesting, "perlemoen" has become severely threatened. Subsequently, the species is on the rapid decline to extinction. Over the past 15 years or so, increased market demand and the decline in natural stocks, have resulted in abalone farms being established along the Western Cape coast. This has offered a possible solution to the vulnerability of the species. Farms launched by large fishing companies provide relief to the species through breeding and stocking programmes (Oakes & Pontes 1996). However, the establishment of these programmes draws a major concern as a potential decline in genetic diversity may occur due to captivity. Ultimately, the ability of abalone to adapt to farm surroundings will be reduced as a result of the specific associated environmental pressures (Evans *et al.* 2004a).

Molecular markers have proven to be applicable in the genetic management of a number of aquaculture species. One of these markers are microsatellites which form part of a group of molecular genetic markers referred to as simple tandem repeats (STR). Microsatellites are abundant and conveniently distributed throughout the non-coding regions of the genomes of all eukaryotic organisms, and are made up of units of highly variable repetitive DNA motifs with each array ranging from 2 to 6 base pairs in size (Avise 1994). In cases where the nucleotide sequences in the flanking regions of the microsatellite are known, specific primers can be designed to amplify the locus *via* PCR (Watkins 2007). Microsatellites are co-dominant DNA markers, inherited in a Mendelian fashion, with multiple alleles segregating at each locus making them highly polymorphic

markers. Polymorphisms are the result of differences in the number of repeat units at a particular locus which give rise to various alleles for said locus. These mutations are believed to be caused by polymerase slippage during DNA replication (Liu & Cordes 2004). Microsatellites have become increasingly popular to aquaculture geneticists over the last ten years and have proven to be exceptionally useful in a wide range of applications. Part of their appeal is the genetic basis of their variability, i.e. unique primers amplify a region of genomic DNA including a well defined repetitive structure that is responsible for the observed variation (Beaumont & Bruford 1999).

Microsatellite markers can be applied to investigate various phenomena on abalone farms such as parentage assignment, which aids in the selection of superior parental animals, and the genotyping of parental broodstock used on farms to maintain pedigree information for selective breeding and genetic mapping (Selvamani *et al*. 2001). In addition, the genetic structure of populations can be studied to determine levels of genetic variation within and among farmed populations as well as the wild populations from which they were sourced. Microsatellites can be used for the identification of QTL that can serve in MAS. A quantitative trait is one that has phenotypic variation as a result of the influence of polymorphic genes and environmental factors. The cause of this variation is due to the variety of alleles associated with the genetic locus of a specific trait (Abiola *et al.* 2003). As quantitative traits are controlled by several genes, each with a small additive effect, the identification of genes that control economically important traits is significant to abalone farmers. MAS uses genetic and phenotypic information for linking specific genetic markers with a particular phenotypic trait, such as enhanced growth. It refers to a selection process in which future broodstock for a particular phenotype or trait is chosen based on genotypes using molecular markers (Liu & Cordes 2004).

Species-specific microsatellite markers are considered to be more powerful genetic markers as they generate an accurate genetic composition of a particular species of interest. Specifically, these markers provide a more accurate and reliable means for data collection and analysis because of amplification and detection of all possible alleles for the particular locus within the specific species (Beaumont & Bruford 1999). They have been particularly useful in various population studies to investigate genetic diversity within and among populations, assisting with QTL mapping and MAS and exploring population structures of aquaculture species such as salmon (Slettan *et al*. 1997), tilapia (Lee & Kocher 1998), trout (Rexroad *et al*. 2002) and oyster (Hubert & Hedgecock 2004), as well as other *Haliotis* species (Evans *et al*. 2004b; Li *et al*. 2004; Sekino *et al*. 2005). These studies have assisted various farms in appropriate restocking and stock enhancement programmes (Bester *et al*. 2004).

Evans *et al*. (2004b) were the first to establish the genetic population structure for *Haliotis midae* in South Africa using microsatellite markers developed for the Australian abalone species, *H. rubra*. However, there is a great need for the characterisation of a large amount of DNA markers for *H. midae* specifically, as very little information regarding the genome is known. The abundance of microsatellites, polymorphic nature and sensitivity to PCR make them ideal markers for this particular study. In the current study the aim was to isolate unidentified species-specific microsatellite markers from *H. midae*. The technique chosen for the development of microsatellites for this project was the Fast Isolation by AFLP of Sequences Containing repeats, FIASCO, method (Zane *et al*. 2002). This methodology has been proven in various other studies to be efficient in achieving the required results.

2.2 MATERIALS AND METHODS

2.2.1 SOURCE MATERIAL

The material used for microsatellite development came from two wild *H. midae* animals obtained from A. van der Merwe (Department of Genetics, Stellenbosch University), who performed DNA isolation *via* a CTAB extraction method. DNA samples chosen from two different wild populations were used to construct the enrichment libraries. The material came specifically from Gansbaai (G) in the Walker Bay area and Robben Island (R) off the coast of Cape Town, both situated along the south-west coast of South Africa. Once isolation of microsatellite loci was completed, polymorphism testing was performed on a subset of individuals from Black Rock (B) located along the South African east coast to ensure consistency among the species.

2.2.2 DNA ENRICHMENT

i) Digestion & Ligation

Digestion and ligation of genomic DNA (gDNA), from G and R, were performed using the restriction enzyme *Mse* I and the *Mse* I AFLP adaptor (5'-TAC TCA GGA CTC AT-3'/5'-GAC GAT GAG TCC TGA G-3') in total reaction volumes of 25µl. The following reagents were used in accordance with the protocol: 10ng/µl gDNA, 1x One-Phor-All buffer (Amersham Biosciences), 5.0mM dithiothreitol (DTT), 50µg/ml bovine serum albumin (BSA; New England Biolabs®, Inc.), 1.0µM *Mse* I adaptor (Integrated DNA Technologies, Inc.), 200µM ATP (Amersham Biosciences), 2.5U *Mse* I (New England Biolabs®, Inc.) and 1.0U T4 DNA ligase (New England Biolabs®, Inc.). The reactions were incubated at 37ºC for 3 hrs.

ii) AFLP Amplification I

A 1 in 10 dilution of each digestion-ligation mixture was made and directly amplified in a 20µl reaction. Four previously designed AFLP-adaptor specific primers, *Mse* I-1, *Mse* I-2, *Mse* I-3 and *Mse* I-4 {*Mse* I-N (5'-GAT GAG TCC TGA GTA A**N**-3') Zane *et al*. 2002}, were mixed in a 1:1:1:1 ratio. Each reaction consisted of: 1x buffer, 1.5mM $MgCl_2$, 0.8mM dNTPs (0.2mM each dATP, dCTP, dGTP and dTTP), 0.6µM *Mse* I-N primer mix and 0.5U GoTaq® Flexi DNA polymerase (Promega). Five reactions per DNA sample were made according to the described conditions in order to determine the optimal number of cycles which would ensure the least amount of biased amplification. The PCR programmes were set to 30 sec at 94ºC, 1 min at 53ºC, and 1 min at 72ºC for 16, 18, 20, 22 and 24 cycles. PCR products were electrophoresed through a 2% (w/v) agarose gel {agarose powder, 1x TBE buffer (Tris-HCl, boric acid and EDTA) and 0.05µg/ml ethidium bromide (EtBr)} at 120 V for 1 hr, and visualised *via* ultraviolet (UV) light transillumination to determine whether amplification was successful. The image was captured on the Multigenius Bio Imaging System (SynGene).

iii) Hybridisation

Amplification under optimised PCR conditions was repeated, using the same thermal conditions as described above, and the products were used for hybridisation. Sample G DNA was hybridised with a biotinylated $(GACT)_6$ probe (Integrated DNA Technologies, Inc.) while sample R DNA was hybridised with a biotinylated $(CAA)_8$ probe (Integrated DNA Technologies, Inc.). DNA samples G and R were therefore used for the creation of tri- and tetranucleotide enriched libraries, respectively. Conditions for the 100µl hybridisation reactions were as follows: 25ng/µl AFLP amplification product, 0.8µM Bio-$(CAA)_8$ or Bio-$(GACT)_6$ probe, 4.2x SSC buffer (sodium chloride, sodium citrate and EDTA) and 0.07% (v/v) sodium-dodecyl sulphate (SDS). DNA was denatured at 95ºC for 3 min and probes were annealed at 25ºC for 25 min.

2.2.3 MICROSATELLITE ISOLATION

i) Selective Capturing

The gDNA hybridised to the biotinylated probes was selectively captured using streptavidin coated magnetic beads (Roche). Prior to capture, the beads were thoroughly washed in TEN_{100} buffer (10mM Tris-HCl, 1mM EDTA, 100mM NaCl; pH 7.5) and resuspended in 40µl of fresh TEN_{100} buffer. To minimise non-specific binding, 10µl of unrelated PCR product containing human gDNA was added to the beads. The hybridisation mixture was diluted in 300µl of TEN_{100}. Beads containing human DNA, and the hybridisation mixture, were then pooled and incubated at room temperature (22ºC) for 30 min on a slow vortex.

The bead-probe-DNA complexes were separated, by a magnetic field, from the hybridisation buffer which was then discarded. A series of three low-stringency washes was performed on the complexes to eliminate non-specific DNA. Washes were performed by adding 400µl of TEN_{1000} (10mM Tris-HCl, 1mM EDTA, 1M NaCl; pH 7.5) and incubating the mixture at room temperature for 5 min on a slow shaking vortex. Supernatant was discarded between washes. The last of the three washes (NS) was set aside for further use. Subsequently, a series of three high-stringency washes performed with 0.2x SSC and 1% SDS (v/v) in the same manner described above. The last of the three washes (S) was stored for further use.

Enriched-DNA was recovered from the bead-probe complex with two subsequent denaturation steps. The first step involved the suspension of the beads in 50µl TE buffer (10mM Tris-HCl, 1mM EDTA; pH 8.0) and the incubation thereof at 91ºC - 95ºC for exactly 5 min. The elution containing the target DNA (D_1) was immediately removed and stored for further use. The second step was done by treating the beads with 12µl NaOH (0.15M). The supernatant was then neutralised with 1µl acetic acid (0.1667M) and TE buffer (pH 8.0) was added to attain a volume of 50µl (D_2). The DNA from the four remaining elutions (NS, S, D_1 and D_2) was precipitated overnight with one volume isopropanol and 0.15M sodium acetate (CH_3COONa) and then centrifuged at 16 000rpm for 30 min. The supernatant was removed and the remaining pellets were dried at 55ºC for 15 min. The DNA was then resuspended in 50µl of SABAX water and stored to be used as templates for PCR.

ii) AFLP Amplification II

The PCR programme for the second round of AFLP amplification was set at the same conditions as the first round (see 2.2.2): denaturation at 94ºC for 30 sec, annealing at 53ºC for 1 min and elongation at 72ºC for 1 min (30 cycles). Amplification was performed with 2µl of each of the recovered washes, from both DNA samples, in a final volume of 10µl containing 1x PCR buffer, 1.5mM $MgCl_2$, 0.2mM of each dNTP (0.8mM of mix), 0.6µM of the *Mse* I-N primer mix and 0.5U GoTaq®. Enriched-DNA PCR products were electrophoresed through a 2% agarose gel at 120 V for 1 hr and 20 min, and visualised by UV-light transillumination. In addition, a 100bp Hyperladder® IV (Bioline) was included in electrophoresis to determine the sizes of the PCR fragments. Products from G and R representing the most unbiased amplification and the widest range of DNA fragments were used for TA-cloning.

iii) Cloning

Sterile LB medium {1% (w/v) Bacto-Tryptone, 0.5% (w/v) Bacto-Yeast, 0.5% (w/v) NaCl, 1.2% (w/v) Bacterial Agar, pH 7.5} containing 100µg/ml of the antibiotic Ampicillin (Roche Applied Science) and 80µg/ml X-Gal (for blue-white selection; ABgene) was made, poured evenly into Petri dishes and allowed to set in a laminar flow cabinet. A pDrive cloning kit with the pDrive vector (Figure 2.1a; Qiagen®) and a TOPO®TA cloning kit with the pCR®4-TOPO vector (Figure 2.1b; Invitrogen™ Life Technologies) was used to create microsatellite enriched libraries for samples G and R, respectively. Cloning reactions and transformations were performed according to manufacturers' instructions. The transformed cells from both transformation kits were grown on the freshly prepared selective LB medium plates overnight (18 - 20 hrs) at 37ºC. The two different cloning kits were used for the transformation reactions in order to determine the reliability and effectiveness of the vectors for microsatellite library construction.

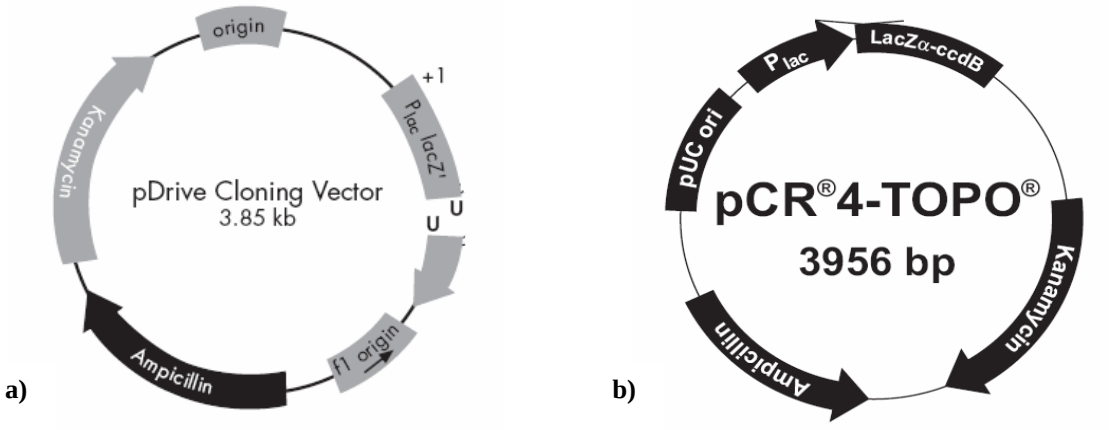


Figure 2.1 Maps of vectors used for cloning. a) the pDrive cloning vector and b) the TOPO®TA cloning vector.

iv) Screening of Colonies

White colonies grown overnight on the selective plates were screened using the M13 vector-specific primers (Forward 5'-GGT TTT CCC AGT CAC GAC-3'; Reverse 5'-GGA AAC AGC TAT GAC CAT G-3'). Colony PCR reactions were performed with picked colonies and contained: 1x PCR buffer, 2.0mM $MgCl_2$, 0.2mM of each dNTP (0.8mM of mix), 0.5µM of each M13 primer and 0.5U GoTaq®, in a final volume of 25µl. The PCR cycle was as follows: an initial denaturation step of 10 min at 94ºC followed by 25 cycles of 94ºC for 1 min, 55ºC for 1 min and 72ºC for 1 min, finalising in an elongation step of 72ºC for 10 min. The amplified products underwent electrophoresis through a 2% agarose gel for 1 hr and 20 min at 120 V, along with a 100bp size ladder to determine the sizes of the fragments of interest. Fidelity of the PCR was visualised by UV-light transillumination.

2.2.4 SEMI-AUTOMATED SEQUENCING

Recombinant clones, giving a positive signal for the insertion of DNA fragments between 400bp and 1kb to allow for efficient sequencing, were cleaned using the NucleoSpin® Extract kit (Macherey-Nagel) and sequenced in 10µl reactions using the BigDye v3 Terminator kit (Applied Biosystems). Sequencing was performed according to manufacturer's instructions in addition to a set DNA concentration of 20ng/µl and using the M13 forward primer. Sequence cycling was performed under the following conditions: 30 cycles of 96ºC for 10 sec, 50ºC for 10 sec and 60ºC for 4 min. There was no initial denaturation or final extension step in the particular programme. The sequences were electrophoresed on an ABI PRISM® 3100 DNA automated sequencer (Applied Biosystems).

2.2.5 PRIMER DESIGN

Sequences obtained from the positive clones were edited by removing vector and adaptor sequences using Chromas 1.45 (Conor McCarthy, North Carolina State University, USA), and subsequently analysed for potential microsatellite repeat sequences. The edited sequences were checked against the NCBI nucleotide database (BLASTN; http://www.ncbi.nlm.nih.gov/BLAST/Blast), as well as an internal laboratory database, to ensure a lack of significant homology with existing sequences of the same or other species as well as to guarantee the uniqueness of each sequence isolated. Sequences were compared with previously isolated abalone microsatellites (Bester *et al*. 2004) and with each other using BioEdit 5.0.9 (Tom Hall, Griffith University, Australia) to determine the uniqueness of the flanking regions as well as potential recombination between clones.

PRIMER 3 (Rozen & Skaletsky 2000) was used to design and analyse primers for the newly isolated microsatellite repeats. Factors taken into consideration during the designing process were dimer and loop formation, GC content, annealing temperatures (ranging from 55ºC to 65ºC), primer length (18 to 24 base pairs) and maximum product size. Primers were designed to flank the repeat sequences and not to contain repetitive units themselves. Primer sets were manufactured (Whitehead Scientific) and dissolved in TE buffer (10mM Tris-HCl, 1mM EDTA; pH 8.0) to a 100µM stock. Dilutions of the primers were made to prepare working stocks of 10µM.

i) PCR Optimisation of Primers

The initial step during optimisation was to perform a gradient PCR to confirm amplification of the markers at their specific annealing temperatures. A gradient PCR constitutes 12 reactions per primer pair, one for every annealing temperature within the given range from 50°C to 65°C. The annealing temperature best suited for a particular set of primers could thus be determined and the

gradient conditions with said temperature used to produce a successful PCR amplification reaction. The cycling conditions for analysing variation among subset individuals were performed with an initial denaturation at 94ºC for 5 min followed by 35 cycles of denaturation at 94ºC for 45 sec, annealing at the range from 50ºC to 65ºC for 45 sec and extension at 72ºC for 45 sec, and ending with a final elongation at 72ºC for 7 min. A standard PCR reaction mix of 10µl volume with 1µl gDNA contained: 1x PCR buffer, 2mM $MgCl_2$, 0.2mM of dNTPs, 0.2µM of both the forward and reverse primer, and 0.25U GoTaq®. DNA concentrations were adjusted accordingly during optimisation to ensure successful and efficient amplification.

The primer sets were also tested using TouchDown PCR (Rahman *et al*. 2000) with the same reaction composition as mentioned above. The PCR amplification of DNA using this protocol was initiated at a set denaturation step of 94°C for 5 min followed by 2 cycles of 94°C (denaturation), 65°C (annealing), and 72°C (extension) for 30 sec each. The programme continued with the stepwise lowering of the annealing temperature by 1°C from 65°C to 55°C after each consecutive 2-cycle step. When an annealing temperature of 55°C was reached the programme was maintained for 30 cycles of 94°C for 30 sec, 55°C for 30 sec and 72°C for 1 min, with a final elongation step of 72°C for 7 min. Products and a 100bp size marker were electrophoresed through a 2% agarose gel at 120 V for 1 hr and 20 min and visualised by UV-light transillumination.

ii) Polymorphism Testing

Optimal PCR reactions for each microsatellite locus were tested on a subset population of 8 different individuals from Black Rock (provided by A. van der Merwe, Stellenbosch University). The fidelity of these reactions along with a 100bp size marker were visualised by UV-light transillumination after 2% agarose gel electrophoresis. In addition, polyacrylamide (PAA) gel {12% (v/v) acrylamide, 1x TBE, 0.8% (v/v) APS, 0.16% (v/v) TEMED} electrophoresis was also performed at 150 V for 2 hrs because of the better resolution provided by this system (Hancock 1999). The Mighty Small (Hoefer) PAA gels were visualised by UV-light transillumination after 10-15 min of staining while gently shaking in an ethidium bromide solution (1x TBE; 0.05µg/ml EtBr).

2.2.6 PRIMER LABELLING

Primer sets amplifying polymorphic products, at the correct size with no or minimal non-specific binding, were chosen for labelling. Fluorescent labels were chosen on the basis of how many microsatellites could be amplified in a single genotyping reaction. Four fluorescent dyes were used, i.e. NED (black), FAM (blue), VIC (yellow) and PET (red). Microsatellites whose primers had the

same colour dyes were labelled specifically not to overlap the estimated size ranges of those situated adjacently. A total of 15 of the 36 microsatellite markers were chosen for genotype screening. Only one primer per microsatellite marker was labelled with an appropriate colour dye. Various factors were taken into consideration when choosing which primer (forward or reverse) to label. These included proximity to the repeat sequence, longer length, higher GC content at the 3'-end, and if at all possible no C or G at the first three base pairs of the 5'-end.

i) PCR Optimisation of Labelled Primers

Labelled primers were first tested using the same reaction concentrations and conditions as their unlabelled counterparts. As primer labelling can interfere with the annealing temperature of the primer and other PCR conditions, DNA and primer concentrations were, to some extent, adjusted to achieve the best possible results. Testing was done on the same subset of eight Black Rock individuals as the non-labelled primers, with gDNA concentrations ranging from as little as 0.5ng/µl to 10ng/µl.

2.2.7 GENOTYPING

A set of 32 individuals obtained from their natural environment (wild) at Black Rock, were genotyped in order to test the variability levels of the microsatellite loci whose primers had been labelled. Following PCR amplification under optimised conditions with DNA concentrations of approximately 5ng/µl, products were screened using the GeneScan™ 600 LIZ® (Applied Biosystems) automated sequencer and analysed using the GeneMapper® Software version 3.7 (Applied Biosystems). In some cases three or four loci could be amplified in a single PCR reaction. The 20µl reaction includes: 5ng/µl gDNA, 1.0x buffer, 2.5mM $MgCl_2$, 0.5mM dNTPs, 1.0 - 3.0µM of each primer (equal concentrations for both primers of the same locus), and 0.5U GoTaq®. The PCR cycle, (optimised by C. van Heerden, Stellenbosch University) was 4 min at 94ºC followed by 35 cycles of 1 min at 94ºC, 20 sec at 52ºC and 10 sec at 72ºC, with a final extension of 10 min at 72ºC. After analysis of the microsatellite polymorphisms, alleles were designated to each locus according to the PCR amplicon size relative to the fluorescent DNA fragment size viewed on the GeneMapper® chromatogram.

2.2.8 STATISTICAL ANALYSIS

The genotyping results obtained from the subset population of 32 individuals provided the basis for the statistical analyses performed. Various parameters were calculated using GENEPOP 3.4 software (an updated version of 1.2; Raymond & Rousset 1995). These included the number of alleles per locus, the observed (H_o) and expected (H_e) heterozygosities at each of the loci, and χ^2

tests for deviations from Hardy-Weinberg equilibrium (HWE). With this data, the allele frequencies of each allele for each locus could be calculated with the programme GENETIX 4.03 (Belkhir *et al.* 2000). To determine the null allele frequencies (*r*) of each microsatellite locus, an equation developed by Brookfield (1996) was used. This equation makes use of the expected and observed heterozygosities of the tested markers, and is represented as follows: $r = (H_e - H_o) / (1 + H_e)$.

2.3 RESULTS

2.3.1 DNA ENRICHMENT

i) AFLP Amplification I

PCR amplification under the various cycle conditions revealed products in the form of smears greater than 200bp, as was expected (Zane *et al.* 2002). The PCR reaction consisting of 16 cycles produced the best result, by providing the greatest range of DNA fragments with the least amount of biased amplification. This product is an indication that there is no specific DNA fragment, but rather a series of bands representing a range of fragments of various sizes flanked by *Mse* I adaptor sites (Figure 2.2). These fragments should, in theory, contain repetitive DNA. Sixteen PCR cycles were used for the hybridisation of the biotinylated probes.

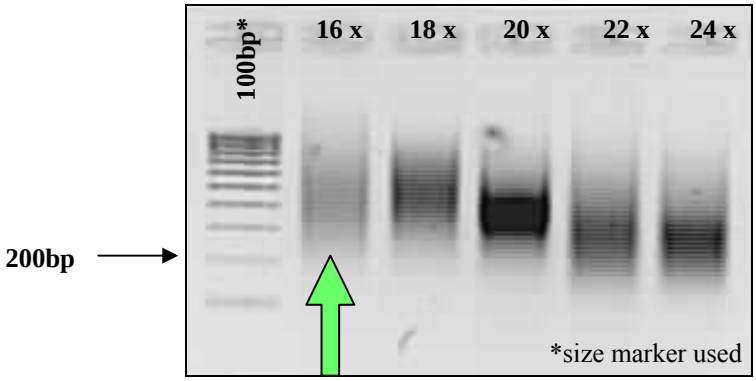


Figure 2.2 First AFLP amplification of sample R with increasing PCR cycle numbers. Products were separated by agarose gel electrophoresis.

2.3.2 MICROSATELLITE ISOLATION

i) AFLP Amplification II

The enriched DNA contained in the elutions from the four washes (NS, S, D_1, D_2) of each sample were expected to contain increasing proportions of repetitive fragments with *Mse* I-N priming sites at each end for specific amplification. The amplified products were once again visualised as smears

above 200bp, representing two highly enriched microsatellite libraries. Elution D_1 was selected to proceed with the cloning step (Figure 2.3).

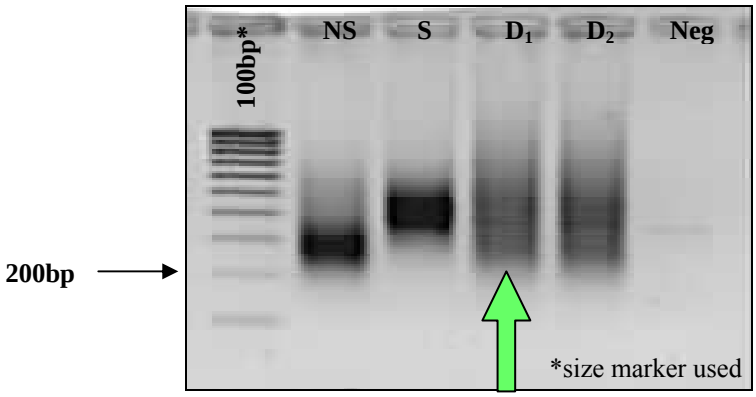


Figure 2.3 Second AFLP amplification of elution products for sample G represented as smears after separation by agarose gel electrophoresis.

ii) Cloning and Colony PCR

A total of 180 white colonies picked from the four plates containing sample G, were screened *via* colony PCR. The same cycle was used for the screening of inserts within 165 white colonies from the four plates containing sample R. Single bands visualised after gel electrophoresis gave an indication that the colonies possibly contained repetitive inserts (Figure 2.4). Products greater than 400bp were chosen for sequencing to ensure sufficient flanking regions during subsequent primer design.

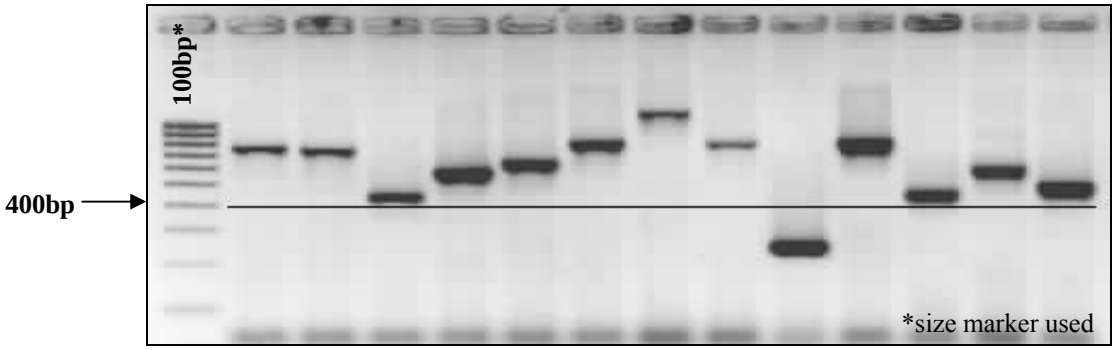


Figure 2.4 Representation of screened colonies after separation by agarose gel electrophoresis. Products larger than 400bp, indicating the presence of inserts of sufficient size, were chosen to be sequence for primer design.

2.3.3 SEQUENCING

A total of 230 colony PCR products were sequenced (110 from sample G and 120 from sample R). Of those inserts sequenced, an overall amount of 56 sequences contained repetitive DNA. The microsatellites found within the sequences varied in composition and were characterised by length and type according to Weber (1990), i.e. there were perfect and imperfect as well as compound sequences consisting of di-, tri-, tetra-, penta- and hexanucleotides (Figure 2.5).

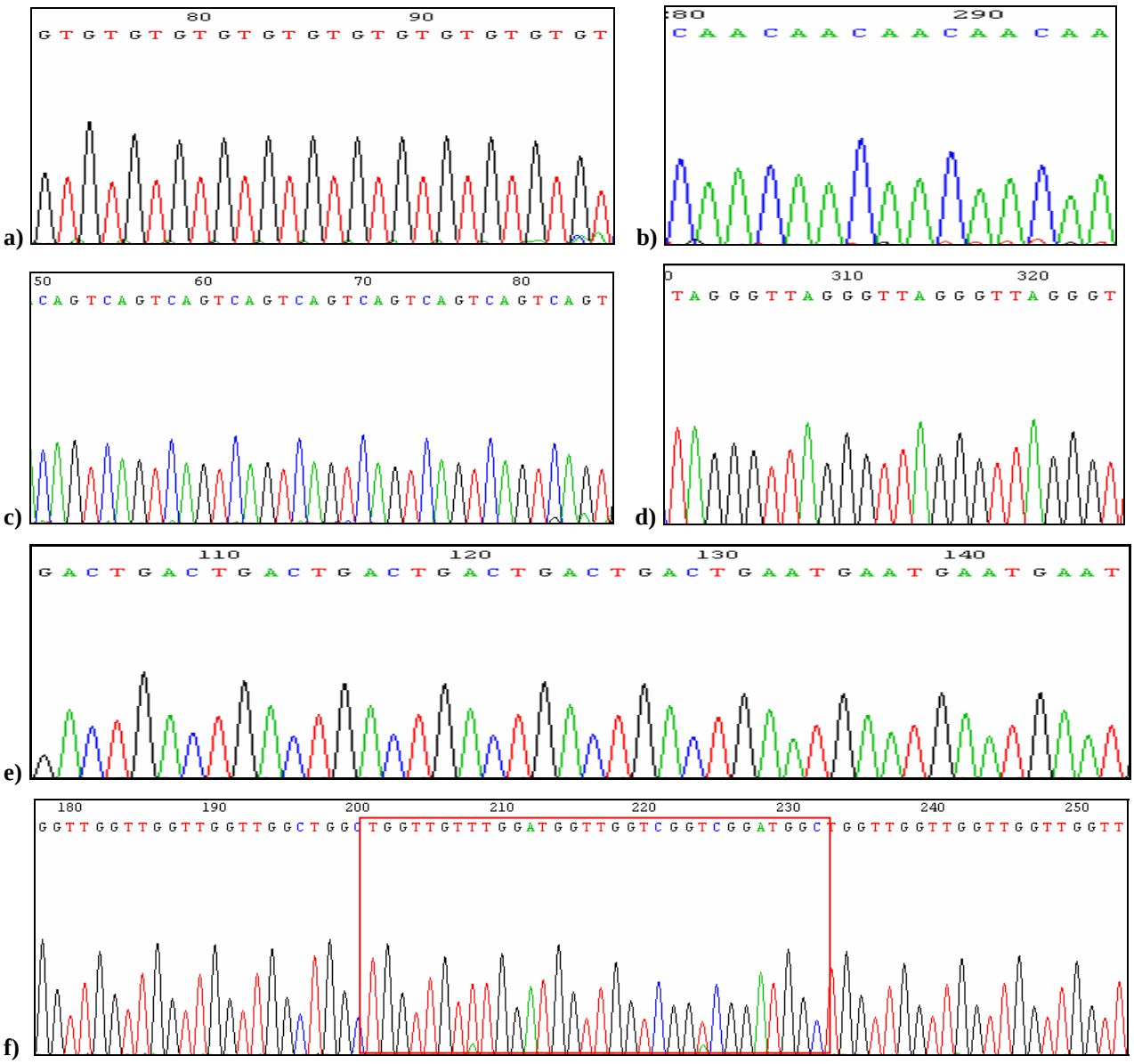


Figure 2.5 Chromatographs of some of the microsatellite sequences isolated. a) a perfect dinucleotide (2bp per unit); b) a perfect trinucleotide (3bp per unit); c) a perfect tetranucleotide (4bp per unit); d) a perfect pentanucleotide (6bp per unit); e) a perfect compound microsatellite with two types of tetranucleotide repeat units; f) an imperfect compound microsatellite with four types of tetranucleotide repeat units, the block indicates interruption between two repeat motifs.

2.3.4 PRIMER DESIGN & PCR OPTIMISATION

The resulting alignments from the BLASTN searches indicated the sequences flanking the repeats being mostly related to other marine species as well as other species of the genus *Haliotis*. Of the 56 microsatellites contained within the inserts, primers were designed for 54 microsatellite markers. Among the 54 markers identified, 7.4% were imperfect repeats, 18.5% were perfect compound repeats and 74.1% were perfect sequence repeats (Figure 2.6). The two markers for which primers were not designed had insufficient and inappropriate flanking sequences due to the locus being too close to the 5' end.

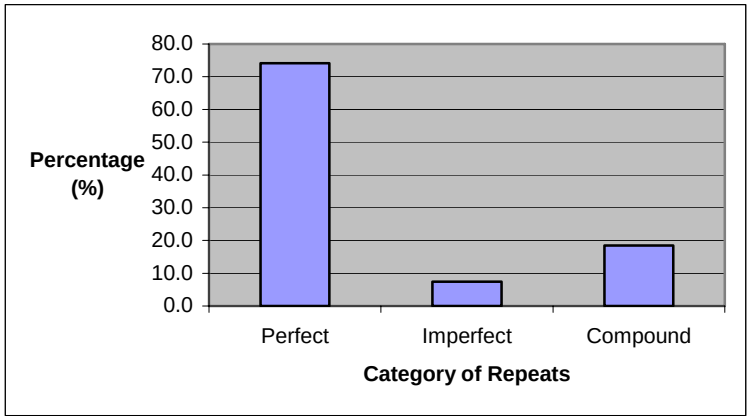


Figure 2.6 Percentages of perfect, imperfect and compound repeats in *H. midae* microsatellites.

A total of **36** microsatellite markers were successfully amplified and submitted to Genbank (Appendix A). Eighteen of the microsatellite markers demonstrated polymorphism, seven markers demonstrated bimorphism, ten were monomorphic, and one marker was representative of nonamplifying/null alleles, as shown by Figure 2.7. The TouchDown PCR method resulted in the successful amplification of 28 of the microsatellites.

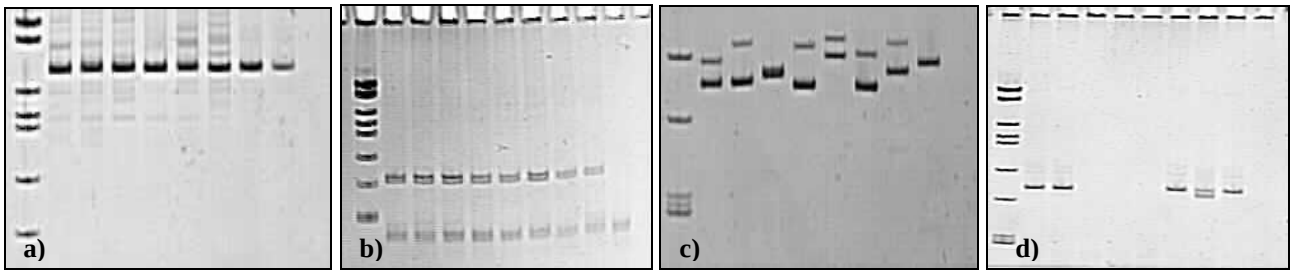


Figure 2.7 Polymorphism testing of four microsatellite loci via 12% polyacrylamide gel electrophoresis. a) A monomorphic marker on polyacrylamide gel. b) A bimorphic marker on polyacrylamide gel. c) A polymorphic locus as seen on polyacrylamide gel. d) A locus indicating null alleles on polyacrylamide gel.

2.3.5 GENOTYPING & ANALYSIS OF MICROSATELLITES

A total of 15 primers were selected for labelling based on their levels of polymorphims. Of the 15 pairs, 12 markers amplified successfully after labelling. Sequences of labelled primers were submitted to GenBank (Slabbert *et al.* in press). Loci HmNR48T, HmNR83D and HmNR268R proved difficult to score with the analysis software, even after extensive optimisation. Alleles were scored according to corresponding "peaks" observed with the GeneMapper® software. The following criteria were taken into consideration: For a heterozygote, the two peaks representing the alleles should not differ by more than 2/3 in height from each other particularly when the first peak corresponds with the shorter allele. The second allele is typically expected to have a smaller peak height as a result of allele dropout (Gagneux *et al.* 1997). Allele dropout is the effect of

amplification favouring the smaller allele, thereby resulting in a lower peak for the larger allele. Figure 2.8 shows the alleles of microsatellite loci HmNR180D and HmNR120T identified in two Black Rock individuals. The presence of minor peaks below the major peak observed at loci HmNR180D are an indication of "stuttering" (Johansson *et al.* 1992). This phenomenon is particularly common with dinucleotide repeat sequences. At locus HmNR120T, the peaks of shorter alleles are inclined to be larger than those of longer alleles; this is resultant of allele drop-out.

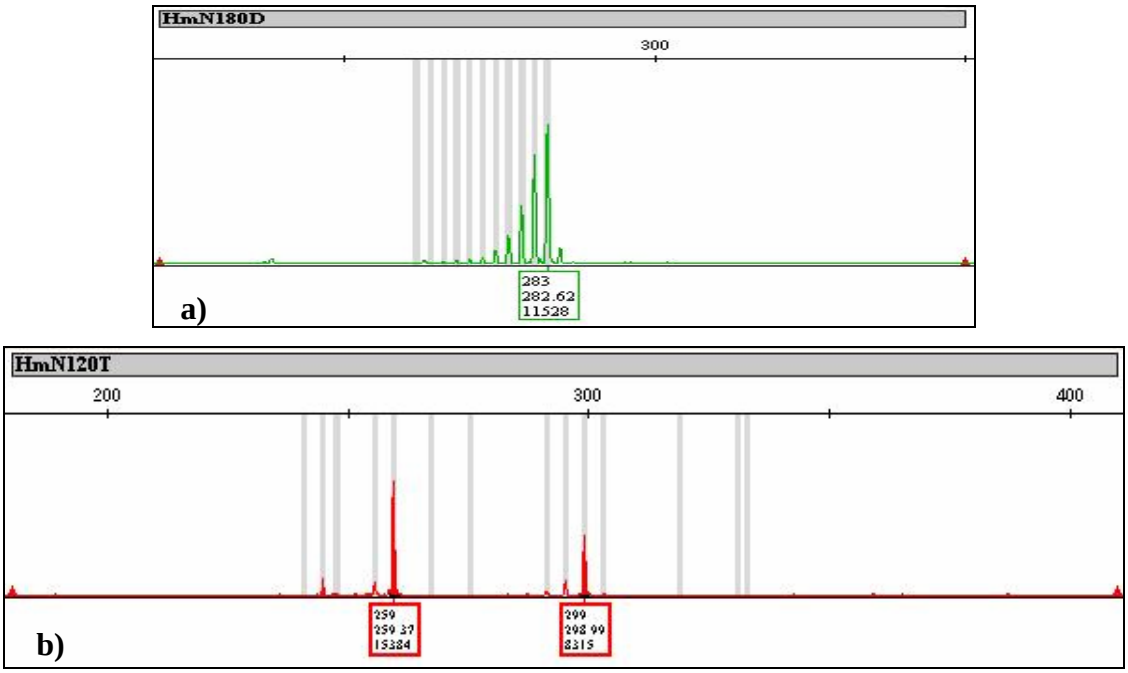


Figure 2.8 Electropherograms for two loci, HmNR180D and HmNR120T, representing anomalies associated with genotyping. a) the phenomenon referred to as stuttering: one can clearly see an increase in peak size and intensity with each increase of a repeat sequence. b) allele drop-out is distinguished by the smaller allele having a greater peak and intensity than the larger.

Repeat motifs, annealing temperature and relevant statistical information for the microsatellite loci are summarised in Table 2.1. Ten of the 12 markers were highly polymorphic where as the degree of variability was different at each locus. Locus HmNR120T had the highest number of alleles (24), with nine loci having alleles ranging from 10-21, and HmNR258T and HmNR289P had six and four alleles, respectively. Allele frequencies of the 12 microsatellites are summarised in Appendix B. At locus HmNR289, the frequency of the major allele was 0.859. At the other 11 loci, allele frequencies were overall low, with the most common allele at locus HmNR136D, being 0.438. The main cause of low frequencies in this set of markers is that the sample size was small, consisting of only 32 individuals. The expected heterozygosity ranged from 0.757 for HmNR54H

Table 2.1 Summary of primers chosen for labelling and microsatellite information.

Primer Name	Repeat Motif	T_a (ºC)	N_a	Range (bp)	N	H_e	H_o	r	*P-value*
HmNR20M	$(TCC)_5(TAC)_7$	60	11	187 – 289	27	0.849	0.852	-0.002	0.717
HmNR48T	$(CAGT)_{25}$	60	-	-	-	-	-	-	-
HmNR54H	$(TTAGGG)_4$	60 – 65	10	329 – 407	32	0.757	0.781	-0.014	0.255
HmNR83D	$(GTCA)_{12}(CTCA)_{15}$	50	-	-	-	-	-	-	-
HmNR106D	$(TG)_{15}$	60	16	329 – 389	31	0.879	0.807	0.039	0.082
HmNR120T	$(TGAG)_{23}$	62	24	235 – 347	29	0.943	0.897	0.024	0.216
HmNR136D	$(CA)_{11}$	60 – 65	20	211 – 309	32	0.797	0.750	0.026	0.156
HmNR180D	$(GT)_{24}$	65	12	269 – 297	31	0.906	0.516	0.205	*
HmNR185D	$(GT)_{13}$	65	11	132 – 160	31	0.879	0.613	0.142	*
HmNR191T	$(GAGT)_6$	62	16	241 – 497	31	0.856	0.807	0.027	0.109
HmNR224T	$(CATA)_{18}$	65	20	444 – 540	22	0.950	0.955	-0.002	0.805
HmNR258R	$(CAA)_{11}$	62	6	239 – 257	32	0.764	0.719	0.026	0.630
HmNR268R	$(CAA)_9$	60	-	-	-	-	-	-	-
HmNR281P	$(CTCAA)_{24}$	62	21	225 – 375	28	0.919	0.714	0.107	*
HmNR289P	$(GTTGT)_5$	65	4	301 – 316	32	0.257	0.250	0.006	0.479
Average			**14.3**			**0.8132**	**0.7216**		

T_a: optimal annealing temperatures; ***N_a:*** number of alleles; ***N:*** number of individuals in which successful amplification of the marker occurred; ***H_o:*** observed heterozygosity; ***H_e:*** expected heterozygosity; ***r:*** null allele frequency; **** ($P < 0.01$):*** significant deviation from HWE

to 0.950 for HmNR224T, with HmNR289P having a comparably low value of 0.257. Significant deviations from HWE for observed heterozygosities were observed at HmNR180D, HmNR185D and HmNR281P ($P < 0.01$). Departures from HWE occur when there is a homozygote excess.

2.4 DISCUSSION

Traditionally, the isolation of microsatellite markers from partial genomic libraries has been based on the screening of several hundreds of clones through colony hybridization with repeat-containing probes. For species with microsatellite-rich genomes this approach is relatively simple, alternatively, it may be tedious and inefficient for those species with low microsatellite frequencies (Zane *et al.* 2002). For successful isolation it is important to have information on the status of microsatellite DNA in the genome of the organisms being studied. However, this is not always possible as such data are scarce and incomplete for many taxa (Li & Kijima 2007). Recently, methods based on selective hybridisation have become a popular technique for the enrichment of microsatellites in aquaculture species (Li *et al.* 2002; Steven *et al.* 2005; An & Han 2006).

In this study, the FIASCO technique was used to isolate novel species-specific microsatellite markers from *Haliotis midae*. Bester *et al.* (2004) were the first to successfully isolate eleven polymorphic, species-specific microsatellite markers for *H. midae* using the FIASCO technique. This method proved to be an efficient protocol for obtaining informative, polymorphic loci. With a mean recovery of repetitive sequences of approximately 63%, the method was most satisfactory for the requirements of this particular project. During the isolation procedure, loci are selected from the upper end of the repeat length distribution in the genome, the region said to harbour the most polymorphic loci. This observation is the basis for species-specific markers because a similarly high level of polymorphism may not be found in related species (Zane *et al.* 2002).

The 54 microsatellite sequences isolated from 230 positive clones in the current study were categorised into perfect (74.1%), imperfect (7.4%) and compound (18.5%) repeats. The results obtained are consistent with those from previous studies in other aquaculture species; including the blue crab, *Callinectes sapidus* (Steve *et al.* 2005); the Pacific oyster, *Crassostrea gigas* (Li & Kijima 2004); and other species of abalone such as *H. discus hannai* (An & Han 2006; Li & Kijima 2007) and *H. rubra* (Baranski *et al.* 2006). Similar studies employing enrichment procedures have been equally successful with various other abalone species around the world, including *H. asinina* from Thailand (Klinbunga *et al.* 2003) and *H. discus hannai* from Japan (Sekino *et al.* 2005). Different methodologies based on comparable selective hybridisation protocols were used in these

studies and all showed to be successful in isolating microsatellites specific to the respective abalone species.

During the isolation procedure for the novel microsatellite markers, an interesting observation was made in the form of the hexanucleotide repeat HmNR54H {$(TTAGGG)_4$; Accession number: EF063103}. The specific microsatellite successfully amplified in all individuals tested. Research has shown that this type of sequence structure, $(TTAGGG)_n$, is a highly repetitive and putative telomeric nucleosome DNA sequence found to be structurally and functionally conserved among eukaryotic organisms (Meyne *et al.* 1989; Lejnine *et al.* 1995). The sequence has been detected in over 100 different eukaryotic species including humans, and data on its distribution in fish species have recently come to the fore, with its localisation being confirmed in species of sturgeon (Fontana *et al.* 1998), salmon (Perez *et al.* 1999), clam (Plohl *et al.* 2002), and more recently various species of abalone: *H. rufescens*, *H. fulgens* and *H. corrugata* (Figure 2.9; Gallardo-Escárate *et al.* 2005a; b).

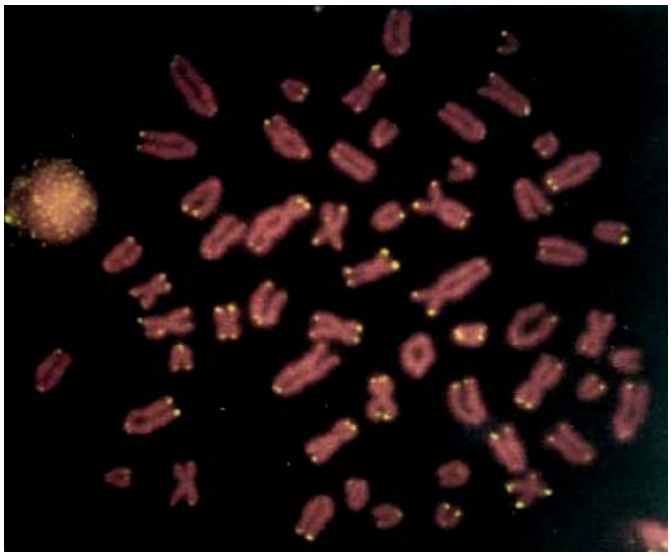


Figure 2.9 Fluorescent *in situ* hybridisation (FISH) of the telomeric sequence, $(TTAGGG)_n$, at interphase in the abalone *Haliotis fulgens*. There is a clear indication of the presence of the repeat at the telomere regions of the chromosomes (Gallardo-Escárate *et al.* 2005a).

This repeat motif is said to constitute a universal telomeric consensus sequence. Telomeres are functional nucleoprotein structures found at the physical ends of chromosomes which play a major role in chromosomal protection from DNA degradation and end-to-end fusions, as well as chromosomal rearrangement and loss. Therefore telomeric sequences such as $(TTAGGG)_n$ which are so highly conserved, are potentially responsible for the mechanics involved in these protective activities. The G- rich 5'-3' strand is enzymatically elongated by telomerase, a specialised reverse

transcriptase, to extend toward the 3' chromosomal termini (Perez *et al.* 1999). This telomeric sequence is therefore required to complete the telomerase-dependant replication at the tips of chromosomes and to preserve the stability of the chromosome arms, positioning and segregation (Plohl *et al.* 2002).

To determine the effectiveness of the isolated loci, the 12 fluorescently labelled primer pairs were genotyped in a subset of individuals. The stutter bands observed at several loci are a typically common artefact of PCR amplification of microsatellites. Often, alleles do not appear as a single product but rather as a series of peaks ascending in size and intensity. This event is due to slipped-strand mispairing during PCR and may cause problems when assigning alleles for genotyping (Weber 1990). Shorter repeats, such as di- and trinucleotides, tend to stutter more than longer repeats while an increase in repeat sequence arrays also increase stutter band intensity (Walsh *et al.* 1996).

The average number of repeats in this study was 15.1, which is similar to that found in the Pacific abalone *H. discus hannai* (13.4; Li & Kijima 2007). The level of polymorphism is proportional to the length of uninterrupted repeats, signifying the importance of the length of a microsatellite when considering its level of polymorphism (Rico *et al.* 1997). According to a model proposed by Kruglyak *et al.* (1998), differences in average microsatellite length between species, and therefore allelic diversity, can be explained by two parameters: the tendency for microsatellites to undergo replication slippage, which favours expansion, and the rate of point mutation, which results in the contraction of long array into two shorter ones. However, it was suggested by Lai and Sun (2003) that when slippage happens, the probability of contraction is higher than that of expansion, and mutation increases linearly with the number of repeat units.

The allelic variability found among the 12 amplifiable markers indicated that they could be useful in several applications, particularly when looking at genetic diversity within and among populations of a species. The mean number of alleles and average observed heterozygosity in this study were 14.3 (range 4-24) and 0.72 (range 0.25-0.95), respectively. These values are consistent with those of Bester *et al.* (2004; 12.4 and 0.82 respectively) and comparable to those in other species of *Haliotis* (Li *et al.* 2002; An & Han 2006). A low level of heterozygosity (0.25) at locus HmNR289P and deviations from HWE ($P < 0.01$) at loci HmNR180D, HmNR185D and HmNR281P were observed. Several contributing factors may be associated with significant departures from HWE, low heterozygosity levels and low allelic diversity. These include: 1) Presence of null alleles: Microsatellite null alleles are alleles that consistently fail to amplify during

PCR, and thus are not detected when individuals are genotyped (Pemberton *et al.* 1995). They are most commonly caused by point mutations at or near the 3' ends of one or both primer binding sites preventing primer annealing. Null alleles thus result in decreased heterozygosity in a sample by underestimating the total number of heterozygotes, thereby interfering with measurements of genetic diversity in a population. Other consequences are overestimation of allele frequencies, and decreases in estimates of relatedness and obstruction of individual and parentage identification (Dakin & Avise 2004). 2) Small sample size: Estimates of allelic variation are often biased as a result of sample sizes being inadequate and therefore affecting the number of and distances between alleles. The number of observed alleles in any population varies because of the number of individuals sampled, thus influencing the true allele frequency from one population to another (Foulley & Ollivier 2006). According to Ruzzante (1998), in order to obtain comparable measures of allelic richness, a sample size of 50-100 individuals is necessary. 3) Allele "dropout": Often, at a heterozygous locus, the amplification of one allele disturbs that of the other, particularly when there are large differences in the allele sizes (Walsh *et al.* 1992). Allele dropout is a consequence of the smaller allele being amplified earlier in the PCR process at the expense of the amplification of the larger allele (Gagneux *et al.* 1997). The effect is the artificial increase of homozygosity resulting in biased allele frequencies (Björklund 2005). 4) Presence of size homoplasy: Mistaking homoplasy for homology can lead to inaccurate estimation of genetic diversity within and among populations. Homoplasious alleles are identical in state (size) but not identical by descent (Ardren *et al.* 1999). Taylor *et al.* (1999) suggested that same-sized alleles will have the same sequence whether they are homologous or homoplasious, however compound microsatellite motifs are an exception. Compound loci have two different repeat units ensuring the advantage for detecting homoplasy as the mutations occurring in the different repeat regions can be easily recognised.

In conclusion, this phase of the study created a microsatellite library for the South African abalone *Haliotis midae* using an efficient technique. The statistical data obtained from those markers chosen for primer labelling were valuable in determining the usefulness of the markers for future applications. The high variability identified will ensure them as excellent tools for population and genetic diversity studies, parentage testing, and linkage analysis of genes related to economically important traits through MAS. Five of the twelve markers were chosen for the second phase of this study. They were applied to assign parentage to hatchery-reared *H. midae* offspring submitted to a growth performance recording scheme, which is part of a stock enhancement programme for the commercial abalone farms of the Western Cape.

2.5 REFERENCES

Abiola O, Angel JM, Avner P, Bachmanov AA, Belknap JK, Bennett B, Blankenhorn EP, *et al.* (2003) The nature and identification of quantitative trait loci: a community's view. *Nature Reviews Genetics* 4: 911-916

An HS, Han SJ (2006) Isolation and characterization of microsatellite DNA markers in the Pacific abalone, *Haliotis discus hannai*. *Molecular Ecology Notes* 6: 11-13

Ardren WR, Borer S, Thrower F, Joyce JE, Kapuscinski (1999) Inheritance of 12 microsatellite loci in *Oncorhynchus mykiss*. *Journal of Heredity* 90: 529-536

Avise JC (1994) Molecular tools. In: Molecular Markers, Natural History and Evolution. Chapman and Hall, New York, pp44-91

Baranski M, Rourke M, Loughnan S, Austin C, Robinson N (2006) Isolation and characterization if 125 microsatellite DNA markers in the blacklip abalone, *Haliotis rubra*. *Molecular Ecology Notes* 6: 1-7

Beaumont MA, Bruford MW (1999) Microsatellites in conservation genetics. In: Goldstein DB, Schlötterer C (eds) Microsatellites: Evolution and Applications. Oxford University Press, New York, pp165-182

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.

Bester A, Slabbert R, D'Amato ME (2004) Isolation and characterisation of microsatellite markers in the South African abalone (*Haliotis midae*). *Molecular Ecology Notes* 4: 618-619

Björklund M (2005) A method for adjusting allele frequencies in the case of microsatellite allele drop-out. *Molecular Ecology Notes* 5: 676-679

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* 5: 453-455

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* 93: 504-509

Evans B, Bartlett J, Sweijd N, Cook P, Elliot NG (2004a) Loss of genetic variation at microsatellite loci in hatchery produced abalone in Australia (*Haliotis rubra*) and South African (*Haliotis midae*). *Aquaculture* 233: 109-127

Evans B, Sweijd NA, Bowie RCK, Cook PA, Elliot NG (2004b) Population genetic structure of the perlemoen *Haliotis midae* in South Africa: evidence of range expansion and founder events. *Marine Ecology Progress Series* 270: 163-172

Fontana F, Lanfredi M, Chicca M, Aiello V, Rossi R (1998) Localization of the repetitive telomeric sequence $(TTAGGG)_n$ in four sturgeon species. *Chromosome Research* 6: 303-306

Foulley JL, Ollivier J (2006) Estimating allelic richness and its diversity. *Livestock Science* 101: 150-158

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with non-invasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* 6: 861-868

Gallardo-Escárate C, Álvarez-Borrego J, Del Río-Portilla MA, Cross I, Merlo A, Rebordinos L (2005a) Fluorescence *in situ* hybridisation of rDNA, telomeric $(TTAGGG)_n$ and $(GATA)_n$ repeats in the red abalone *Haliotis rufescens* (Archaeogastropoda: Haliotidae). *Hereditas* 142: 73-79

Gallardo-Escárate C, Álvarez-Borrego J, Del Río-Portilla MA, Von Brand-Skopnik E, Cross I, Merlo A, Rebordinos L (2005b) Karyotype analysis and chromosomal localization by FISH of ribosomal DNA, telomeric $(TTAGGG)_n$ and $(GATA)_n$ repeats in *Haliotis fulgens* and *H. corrugata* (Archaeogastropoda: Haliotidae). *Journal of Shellfish Research* 24: 1153-1159

Hancock JM (1999) Microsatellites and other simple sequences: genomic content and mutational mechanisms. In: Goldstein DB, Schlötterer C (eds) Microsatellites: Evolution and Applications. Oxford University Press, New York, pp1-9

Hubert S, Hedgecock D (2004) Linkage maps of microsatellite DNA markers for the Pacific oyster *Crassostrea gigas*. *Genetics* 168: 351-362

Johansson M, Ellegren H, Anderson L (1992) Cloning and characterization of highly polymorphic porcine microsatellites. *Journal of Heredity* 83: 196-198

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

Kruglyak S, Durrett RT, Schug MD, Aquardo CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Science USA* 95: 10774-10778

Lai Y, Sun F (2003) Microsatellite mutations during the polymerase chain reaction: mean field approximations and their applications. *Journal of Theoretical Biology* 224: 127-137

Lee B, Kocher T (1998) Microsatellite mapping of the prolactin locus in the tilapia genome. *Animal Genetics* 29: 68-69

Lejnine S, Makarov VL, Langmore JP (1995) Conserved nucleoprotein structure at the ends of vertebrate and invertebrate chromosomes. *Proceedings of the National Academy of Science USA* 92: 2393-2397

Li Q, Kijima A (2004) Rapid isolation and characteristic analysis of microsatellite clones in the Pacific oyster. *Journal of Ocean University of China* 4: 364-370

Li Q, Kijima A (2007) Sequences characterization of microsatellite DNA sequences in Pacific abalone (*Haliotis discus hannai*). *Journal of Ocean University of China* 6: 47-52

Li Q, Park C, Kijima A (2002) Isolation and characterization of microsatellite loci on the Pacific abalone, *Haliotis discus hannai*. *Journal of Shellfish Research* 2: 811-815

Li Q, Park C, Endo T, Kijima A (2004) Loss of genetic variation at microsatellite loci in hatchery strains of Pacific abalone (*Haliotis discuss hannai*). *Aquaculture* 235: 207-222

Liu ZJ, Cordes JF (2004) DNA markers and their applications in aquaculture genetics. *Aquaculture* 238: 1-37

Meyne J, Ratliff RL, Moyzis RK (1989) Conservation of the human telomeric $(TTAGGG)_n$ among vertebrates. *Proceedings of the National Academy of Science USA* 89: 7049-7053

Oakes FR, Ponte RD (1996) The abalone market: Opportunities for cultured abalone. *Aquaculture* 140: 187-195

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* 4: 249-252

Perez J, Moran P, Garcia-Vasquez E (1999) Cloning and physical mapping of Atlantic salmon (*Salmo salar* L.) telomeric sequences. *Heredity* 82: 409-414

Plohl M, Prats E, Martínez-Lage A, González-Tizón A, Méndez J, Cornudella L (2002) Telomeric localisation of the vertebrate-type hexamer repeat, $(TTAGGG)_n$, in the wedgeshell clam *Donax trunculus* and other marine invertebrate genomes. *The Journal of Biological Chemistry* 277: 19839-19846

Rahman MH, Jaquish B, Khasa PD (2000) Optimisation of PCR protocol in microsatellite analysis with silver and SYBR® stains. *Plant Molecular Biology Reporter* 18: 541-542

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86: 248-249

Rexroad CE, Coleman RL, Hershberger WK, Killerfer J (2002) Thirty-eight polymorphic microsatellite markers for mapping in rainbow trout. *Journal of Animal Science* 80: 541-542

Rico C, Ibrahim KM, Rico I, Hewitt GM (1997) Stock composition in North Atlantic populations of whiting using microsatellite markers. *Journal of Fish Biology* 51: 462-475

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Ottawa, pp365-386

Ruzzante DE (1998) A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Sciences* 55: 1-14

Sekino M, Saido T, Fujita T, Kobayashi T, Takami H (2005) Microsatellite DNA markers of Ezo abalone (*Haliotis discus hannai*): a preliminary assessment of natural populations sampled from heavily stocked areas. *Aquaculture* 243: 33-47

Selvamani MJP, Degnan SM, Degnan BM (2001) Microsatellite genotyping of individual abalone larvae: parentage assignment in aquaculture. *Marine Biotechnology* 3: 478-485

Slabbert R, Ruivo NR, Van den Berg NC, Lizamore DK, Roodt-Wilding R (2007) Isolation and charaterisation of 63 microsatellite loci for the abalone, *Haliotis midae*. *Journal of the World Aquaculture Society*, in press.

Slettan A, Olsaker I, Lie O (1997) Segregation studies and linkage analysis of Atlantic salmon microsatellites using haploid genetics. *Heredity* 78: 620-627

Steven CR, Hill J, Masters B, Place AR (2005) Genetic markers in blue crab (*Callinectes sapidus*) I: isolation and characterization of microsatellite markers. *Journal of Experimental Biology and Ecology* 319: 3-14

Taylor JS, Sanny JSP, Breden F (1999) Microsatellite allele size homoplasy in the guppy (*Poecilia reticulata*). *Journal of Molecular Evolution* 48: 245-247

Walsh PS, Ehrlich HA, Higuchi R (1992) Preferential amplification of alleles: mechanisms and solutions. *PCR Methods and Applications* 1: 241-250

Walsh PS, Fildes NJ, Reynolds R (1996) Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Research* 24: 2807-2812

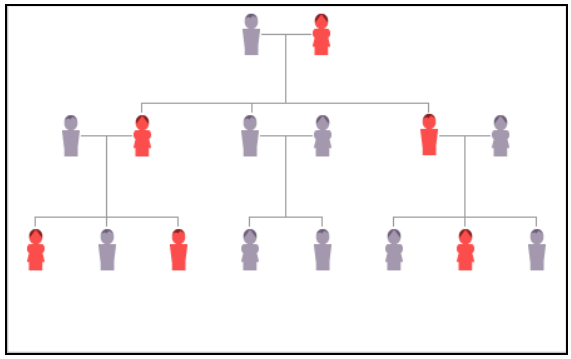
Watkins JC (2007) Microsatellite evolution: Markov transition functions for a suite of models. *Theoretical Population Biology* 71: 147-159

Weber JL (1990) Informativeness of human (dC-dA)n(dG-dT) polymorphisms. *Genomics* 7: 524-530

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

Chapter 3

Assignment of Parentage to Hatchery-Reared Abalone Submitted to a Growth Performance Recording Scheme



The following chapter is presented in the form of a full-length manuscript in preparation for future publication.

ABSTRACT

Microsatellite markers are rapidly becoming the molecular marker of choice for parentage assignment in aquaculture. Ten loci were compared in three broodstock groups of commercially-reared *Haliotis midae* and used in the parentage assignment of 200 juveniles per farm to unknown broodstock pairs. The CERVUS program was applied and data revealed that the markers proved to be useful in assigning approximately 60% of offspring to one breeding pair each of abalone farms I&J and Roman Bay. The application of microsatellite markers will broaden the scope of a breeding programme, enabling progeny with "good performing" genotypes to be tested for selection as future breeding animals.

3.1 INTRODUCTION

Farming of the abalone *Haliotis midae* began in South Africa during the early 1990's, when aquaculturists discovered that the species had immense commercial and economic potential. To remain competitive in this highly lucrative global market, the South African industry needs to increase overall performance, productivity and profitability of commercially-reared abalone. Intensive farming of abalone requires the application of specialised technologies to ensure competitiveness. There has been significant progress in the culturing of *H. midae*, with the successful establishment of management, feeding and reproduction expertise. Although the knowledge and skills of farm management is advanced, the application of genetic technologies to advance productivity has been relatively slow compared to that of other commercial abalone species around the world. For industry to meet the demand for good quality export products, the inclusion of a genetic improvement programme into farm management is necessary (Isberg *et al.* 2004). Genetic improvement is of great value to maximise efficiency in aquaculture and requires the selection of superior animals for breeding purposes with reference to traits of economic importance (Castro *et al.* 2006).

Increasing growth rate of commercial *H. midae* is extremely valuable to the sustainability of the industry and is becoming the top priority in hatcheries. The most efficient way to achieve genetic gain of growth-related traits is *via* selection within and between families. This requires knowledge of the heritability of those traits (Vandeputte *et al.* 2004). Breeders also need to know the influence non-genetic factors have on the expression of phenotype variation. Factors associated with environmental effects like water temperature and stocking density can considerably manipulate familial growth traits (Jerry *et al.* 2006). However, the lack of knowledge of reliable pedigree information and techniques for examining large pedigrees on commercial farms are the most

significant obstacles in the realisation of a first-rate breeding programme. Broodstock individuals are reared together in common tanks where they are allowed to spawn naturally, making it extremely difficult to assign progeny to the correct parents. Ideally, each family's offspring should be grown in a separate tank, enabling heritability to be accurately traced and comparisons between families to be made based on phenotypic observations. Experiments such as these, however, are expensive and require a great deal of suitable infrastructure. The best alternative to overcome these constraints would be to raise animals from different families in a single tank and use molecular genetic DNA technology, to determine the parentage of the pooled individuals. In this way, a common as well as commercially realistic environment is created where families can naturally compete against each other. This will reduce confounding environmental effects and therefore allow for better assessment of the genetic effects governing the targeted growth trait (Lucas *et al.* 2006).

Microsatellite DNA parentage analyses are increasingly becoming the methodologies of choice in aquaculture genetic improvement programmes. To preserve pedigree information in aquaculture rearing circumstances, the application of a few highly variable microsatellite markers can make it possible to assign individuals to specific parental pairs (Norris *et al.* 2000; Selvamani *et al.* 2001; Hara & Sekino 2003; Jerry *et al.* 2004; Webster & Reichart 2005; Castro *et al.* 2006; Lucas *et al.* 2006). Microsatellites are highly polymorphic and codominant, segregate under Mendelian conditions and are abundant throughout the genome, making them ideal genetic tools for DNA based identification systems. Those developed for the particular species of interest, i.e. species-specific microsatellites, make parentage assignment even more powerful. With enough variation at as little as five or six loci, multiple genotypes of individual broodstock animals will be unique to those parents and, therefore, to each fertilisation carried out (Beaumont & Hoare 2003). The major constraint to microsatellite based systems is the considerable economic cost of using this technology. However, by selecting the most powerful markers as opposed to using more markers that are polymorphic and combining these in a single step, the costs of utilising these markers can be reduced while efficiency is maximised. This is done *via* the establishment of a reliable single-step method, termed multiplexing, where several markers can be co-amplified in a single PCR.

3.2 ESTABLISHMENT OF A GROWTH PERFORMANCE RECORDING SCHEME

A proposal to improve the genetic quality of the breeding stocks of *Haliotis midae* at five commercial abalone farms was put forward as a collaborative effort towards the genetic

improvement of farmed stocks in South Africa. Very little information is available concerning the genetic quality of commercial stocks of *H. midae*, and a lack of co-operation is hampering the flow of exchange of information within the industry. The five farms in question are situated in the Western Cape, the heart of the abalone fishery in southern Africa. These include Abagold®, Aquafarm Development, HIK Abalone, I&J and Roman Bay. The establishment of a growth performance recording scheme (PRS) is to aid these farms in producing the best quality products on the market. It will make provisions for: the evaluation and comparison of growth performance as an indicator of genetic qualities, the identification of genetically superior genotypes amongst current broodstock groups, and the organisation of an efficient protocol for the assessment of future breeding populations. In this particular study, offspring hatched at the five participating farms were reared together, in the same tanks, at the five locations under each farm's management practices. This ensured the reduction of environmental influences, giving each group an equal opportunity to be reared under the same conditions. Animals were tagged in order to trace and compare the performance of growth traits (weight and shell length) of the F_1 progeny at each of the farms submitted to the PRS.

Microsatellite markers developed in chapter 2, along with available microsatellites, will be optimised for a multiplex marker suite and used to assign parentage to PRS offspring. Superior performing offspring can then be identified and returned to their original hatcheries where they can be used as future broodstock to produce quality offspring of superior genotypes, while the broodstock who produced these superior progeny can also be preferentially used in the hatcheries. This study will also aid in determining whether genotype by environment interactions influence the familial growth performance status by investigating if superior families at one location are presenting the same superior qualities at the other four locations. If this is the case then the future of selective breeding at these farms will yield successful results.

3.3 MATERIALS AND METHODS

3.3.1 EXPERIMENTAL DESIGN

A predetermined amount of juveniles (i.e. approximately 3000 per farm) from five participating commercial abalone hatcheries were entered into the PRS. The broodstock selected for the PRS were animals used by the farms in their everyday commercial production lines. All dams and sires were spawned simultaneously, in July 2005, to ensure uniformed age of all juvenile animals. In January 2006, each group of F_1 offspring were tagged, then split into five equal subgroups of 600 individuals and distributed to the various locations. At each location, the offspring were once again

divided equally into three replicates, with each subgroup containing 200 animals from each farm (Figure 3.1).

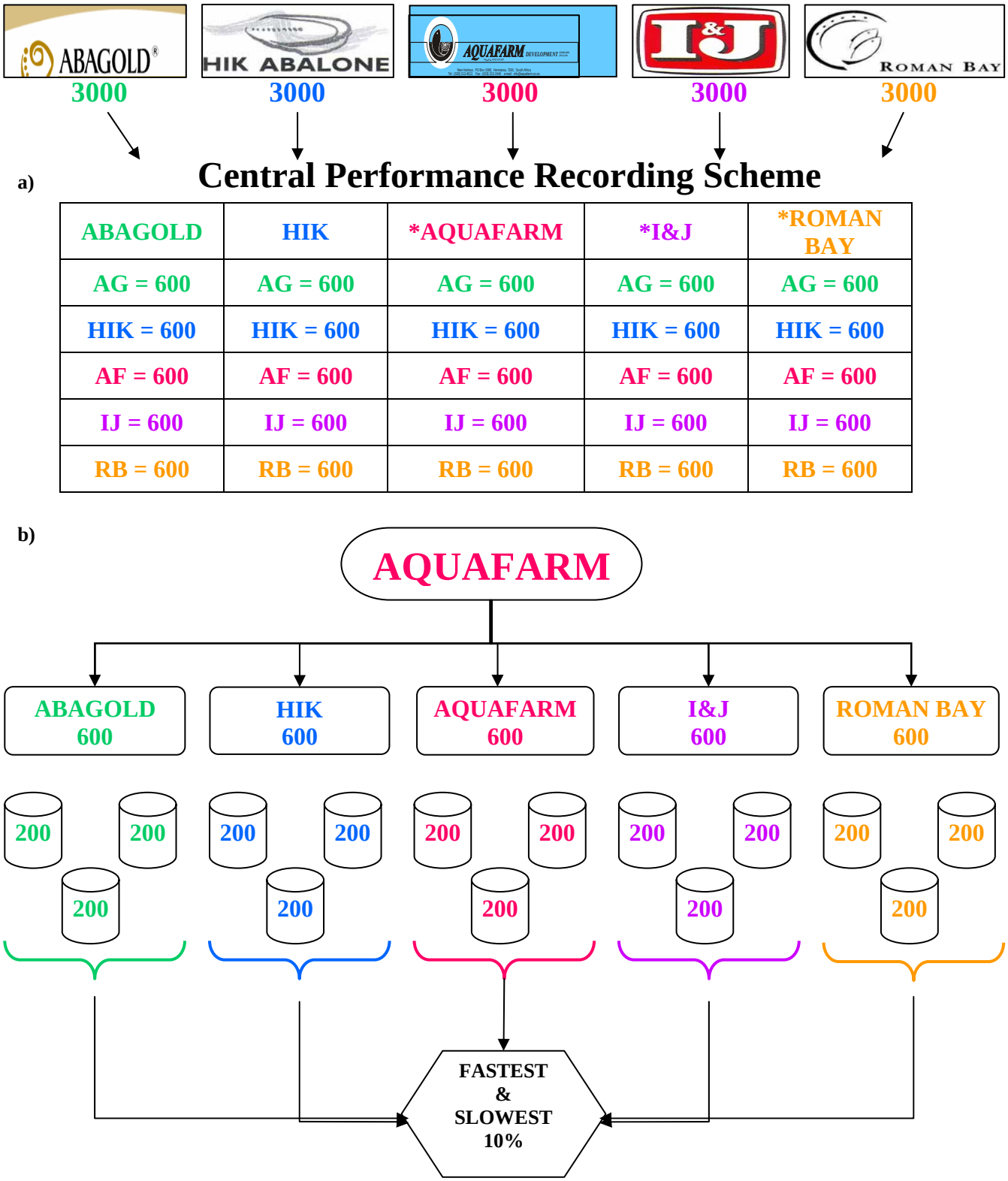


Figure 3.1 A schematic representation of the structure of the PRS. a) Represents all the offspring from all 5 participating farms and how the animals are divided at each location. * indicates locations worked on in this study. b) An indication of how the offspring animals are distributed into three replicates at each of the farm locations.

Animals were then left under the regulation of management at each particular location. Management and feeding practices of each farm were strictly conformed to in order to eliminate as many environmental influences on the juveniles as possible.

3.3.2 SOURCE MATERIAL

Participating farms are situated in the Walker Bay region of the Western Cape, specifically in Hermanus and Gansbaai. Three groups of broodstock, as well as their respective F_1 offspring, from hatcheries Aquafarm, I&J and Roman Bay, were used as the source material for this particular project. Sampling and analyses for the two remaining farms were performed by N. van den Berg (Department of Genetics, Stellenbosch University). A total of 200 offspring (100 of the fastest growing and 100 of the slowest growing), were sampled at each of the five participating locations (i.e. 20 largest animals and 20 smallest animals sampled at each of the five locations).

3.3.3 ABALONE SAMPLING

Sampling of tagged broodstock and offspring was performed on the respective farms. With the aid of farm workers, non-destructive sampling of broodstock was achieved. Animals were temporarily removed from their tanks and three epipodial tentacles were cut from each animal. Epipodia were stored in 99% ethanol (EtOH) until DNA isolation could be performed. At Aquafarm however, not all broodstock animals (approximately half) could be sampled as farm staff could not trace the animals, therefore resulting in the loss of a large fraction of this farm's broodstock data. Selected offspring of appropriate wet weight and shell length were sampled during July 2006 when the animals were approximately 1 year old. Offspring at this stage of life are rather small, therefore several epipodia were removed to obtain sufficient DNA for isolation and further analysis.

3.3.4 PREPARATION OF BROODSTOCK DNA

Total gDNA from epipodial clips was extracted using the CTAB DNA extraction method (Saghai-Maroof *et al*. 1984). This technique, which was performed over three days, proved to be efficient for the requirements of this project. The tentacles were immersed in 600µl of digestion buffer {2% (v/v) CTAB, 1.4M NaCl, 0.2% (v/v) β-mercapto-ethanol, 20mM EDTA, 100mM Tris-HCl; pH 8.0} with the addition of proteinase K (10mg/ml; Roche), and incubated overnight in a water bath at 60ºC. Eppendorf tubes containing the buffer-tissue sample mixture were shaken by hand to ensure the breakdown of tissue within the buffer. A half volume of phenol (300µl) along with a half volume of chloroform:isoamyl alcohol (24:1; 300µl), was added and the tubes were mixed by inversion to form a milky emulsion then centrifuged at 12 000 rpm for 10 min at 4ºC. The aqueous phase (600µl) was transferred to a clean eppendorf tube. The previous two steps were repeated

until the interface between the aqueous and organic phases was clean. A 2/3 volume of isopropanol (400µl) was added to the clean aqueous phase and stored at -20ºC overnight. Precipitated DNA was centrifuged at 16 000 rpm for 20 min at 4ºC. The isopropanol was then removed, the DNA pellet washed with 70% (v/v) cold EtOH (200µl) and centrifuged under the previous conditions. The ethanol was removed and the pellet allowed to dry at 55ºC for 30 min. The dried pellet of gDNA was finally dissolved in 50µl of SABAX water and stored at -20ºC for further use.

3.3.5 PREPARATION OF OFFSPRING DNA

Isolation of DNA from offspring was performed in a semi-automated manner due to the minute amount of sample tissue available. Epipodial tentacles were placed in a solution of 300µl C1 lysis buffer (Macherey-Nagel) and 2µl proteinase K (10mg/ml), and incubated overnight in a water bath at 60ºC. The lysed samples were subsequently placed in the TissueLyser apparatus (Qiagen) with metallic beads for 3 min for further breakdown of tissue, and placed back into the water bath for a further 30 min. Samples were transferred to the Genesis RSP 200 Instrument (Tecan) for DNA extraction.

The purity and concentrations of both the broodstock and offspring DNA were spectrophotometrically measured using the NanoDrop® ND-100 Spectrophotometer (Nanodrop Technologies).

3.3.6 GENOTYPING

i) Choice of Microsatellite Loci

The choice of markers for parentage assignment was based on the various statistical factors obtained from the subset population of 32 wild individuals in chapter 2, i.e. relatively high allele numbers and frequencies, heterozygosities and null allele frequencies with $r < 0.1$ (see chapter 2). Microsatellites constituted by compound repeats were also avoided as they may result in size homoplasy. This phenomenon occurs as a result of internal mutations where different copies of a particular locus allele are identical in size but not in origin (Estoup *et al.* 2002). Markers were also chosen based on their ability to be amplified and analysed together in one multiplex microsatellite genotyping reaction and one genotyping panel, respectively. Nonambiguous microsatellite alleles that could be consistently scored were the key deciding factors for the markers chosen for this phase of the project.

The number of markers with dinucleotide repeat motifs also had to be limited, particularly when using them for parentage, as they are infamous for exhibiting a phenomenon referred to as

"stuttering" (see section 2.3.5). This can more than often obscure the classification of genetic profiles at those particular microsatellite loci when used in parentage analysis (Liu & Cordes 2004). Although stuttering is not as frequent in tri- and tetranucleotide repeats, it has been observed in these types of motifs, however, it is not as prevalent (Jerry *et al.* 2006). Therefore markers with larger repeat motifs would be more robust systems and most advantageous for the assignment of parentage.

The most variable microsatellite loci from the pool developed in chapter 2 (Pool 1), along with others from a pool of markers (Pool 2) developed by N. van den Berg (Department of Genetics, Stellenbosch University), were used to generate microsatellite DNA profiles. Five markers from each pool, ten in total, were chosen for parentage assignment based on the criteria previously mentioned (Table 3.1).

ii) PCR Protocols

Four markers developed from pool 1; HmNR54H, HmNR120T, HmNR136D and HmNR281P were amplified together in one standard multiplex PCR reaction using the following conditions in a final volume of 20µl: approximately 5ng/µl of gDNA, 1.0x PCR buffer, 2.5mM $MgCl_2$, 0.5mM dNTP mix, 0.1µM of each primer for loci HmNR54H and HmNR136D, and 0.2µM of each primer for loci HmNR120T and HmNR281P, and 0.5U GoTaq® Flexi DNA polymerase (Promega). Standard cycling conditions involved an initial denaturation at 94ºC for 4 min, 35 cycles of 94ºC for 1 min, 52ºC for 10 sec and 72ºC for 20 sec, followed by a final elongation at 72ºC for 10 min. The fifth marker from pool 1, HmNR224T, had to be amplified individually as it required a lengthy final extension step. The PCR conditions were as follows: 5 min initial denaturation at 94ºC, 30 cycles of 1 min denaturation at 94ºC, 30 sec annealing at 65ºC and 30 sec elongation at 72ºC. The final extension of 45 min at 72ºC was used to ensure the complete addition of dNTPs, particularly adenine, to the PCR product (Smith *et al.* 1995). Amplification of HmNR224T was performed in 10µl volumes containing approximately 5ng/µl gDNA, 1.0x PCR buffer, 2.0mM $MgCl_2$, 0.2mM dNTP mix, 0.2µM of both the forward and reverse primers and 0.25U GoTaq®. This PCR product was then added to the final product of the multiplex PCR in a 1:1 ratio, and all 5 markers were genotyped in one reaction (one genotyping panel, Appendix C).

Four loci from pool 2, namely HmNST7T, HmNS28D, HmNS38T and HmNS58R, were also amplified using multiplex PCR conditions. These four however, had to be separated into two reactions each containing two different markers on account of overlapping that occurred between markers HmNST7T and HmNS58R. The first of the two PCR reactions contained markers

Table 3.1 Microsatellite loci chosen for the assignment of parentage.

Locus Name	Repeat Sequence	Primer Sequences (5'-3')	Label	PCR Conditions	N_a	H_e	R	Allele Range (bp)	Accession Number
Pool 1									
HmNR54H	$(TTAGGG)_4$	F-5' CAT TCT ACA TTC GAC ATT CG 3' R-5' TAA CAC TAA GTC CCT CAC CC 3'	VIC	Multiplex	10	0.7570	-0.0305	329-407	EF063103
HmNR120T	$(TGAG)_{23}$	F-5' ACC TGC TCT TTA GCT CAG ATG G 3' R-5' TTG AGC ATG AGT CGT TGA GC 3'	PET	Multiplex	24	0.9440	0.0199	235-347	EF121745
HmNR136D	$(CA)_{11}$	F-5' GAG TAA TAT GGG CAC CTC G 3' R-5' GTT TGG AAT GTC TGA TTG GA 3'	VIC	Multiplex	20	0.7970	0.0283	211-309	DQ825710
HmNR224T	$(CATA)_{18}$	F-5' TGT CCA TAG CAG CCC CTT AC 3' R-5' ACA TCT TGT TGC CGT TGT TG 3'	FAM	65ºC	20	0.9500	-0.0143	444-540	EF512269
HmNR281P	$(CTCAA)_{24}$	F-5' AAC CTT CAG TAA CCC ATG C 3' R-5' TGA ATA GGC ACC ATA AAG GG 3'	FAM	Multiplex	21	0.9190	0.1138	225-375	EF512274
Pool 2									
HmNS28D	$(CA)_{16}$	F-5' CAG TCA ATT TTC ATC GCA TT 3' R-5' AGG TCG TTT TTC TCC TTC AG 3'	FAM	Multiplex	19	0.9630	-0.0274	123-185	EF033332
HmNS38T	$(TCAC)_{10}$	F-5' CTG AGA CCC AAA GTT TTC TTT A 3' R-5' ATC TAT GTT CAG GGT GTC AGT G 3'	PET	Multiplex	11	0.8360	0.0472	402-474	EF 367113
HmNS58R	$(GTT)_8$	F-5' TGC CAC TCA AAT GTT CCT TA 3' R-5' CTA TTT CAG GTG TCC CCA GT 3'	NED	Multiplex	10	0.8780	0.0265	233-272	EF 367119
HmNST7T	$(CACT)_{26}$	F-5' CAC ATG GGT ACA ATG TGT GAA G 3' R-5' GGT AGC ACT GTT TCT CAC GA 3'	VIC	Multiplex	21	0.9580	-0.0127	228-328	EF 455618
HmNS56D	$(CA)_{20}$	F-5' TTC GGC AAG TGA ATG TCT AG 3' R-5' CCG AGT TTG GAA TGT CTG AT 3'	NED	60ºC	16	0.8550	-0.0434	211-253	EF 455619

N_a: number of alleles;
H_e: expected heterozygosity;
r: Brookfield's null allele frequency

HmNST7T and HmNS38T, both with primer concentrations of 0.2µM in the standard reaction. The annealing temperature of the standard multiplex PCR protocol was increased to 58ºC as this proved to have better scoring results.

The second PCR containing HmNS28D and HmND58R was amplified under standard conditions with appropriate modifications to primer concentrations, i.e. 0.2µM for primers of locus HmNS28D, and 0.1µM for primers of the HmNS58R locus. The final locus in this pool of markers, HmNS56D, was amplified individually under the same reaction conditions as HmNR224T, with a primer concentration of 0.1µM using the following cycling protocol: an initial extension at 94ºC for 5 min, 30 cycles of 94ºC for 30 sec, 60ºC for 1 min and 72ºC for 1 min, ending off with a final elongation at 72ºC for 10 min. The product from this reaction was then added, in a 1:1 ratio, to the product from the reaction containing HmNST7T and HmNS38T. Therefore, for the 5 markers of pool 2, there were three PCR reactions and two genotyping reactions (two genotyping panels, Appendix C).

iii) Scoring Microsatellites

The most convenient technique today for visualising and scoring microsatellite markers is to fluorescently label the end of one primer of a pair and then measure the size of PCR amplicons on an automated sequencer (Webster & Reichart 2005). By varying the dye colours at each locus, multiple loci can be run in one multiplex reaction and the products for each locus can be simply characterised.

On completion of PCR amplification with fluorescently labelled primers, 2µl of each product was screened via the GeneScan™ 600 LIZ® sequencer. GeneScan™ assigns a size standard to each sample which is then used in GeneMapper® software as the parameter for assigning allele sizes to the PCR products. The size standard is important as it determines how efficiently alleles are compared between individual samples at a particular locus. Individual alleles at each locus were scored using the GeneMapper® programme and labelled according to their sizes, in base pairs, represented by fluorescent peaks on the programme chromatogram profiles. Typically, the highest peaks or those that produce the strongest fluorescent signals, represent the most likely alleles particularly if they are within the expanse of the microsatellite range seen during polymorphism testing (see chapter 2). Allelic patterns should be consistent between broodstock and F_1 individuals. All parents and their respective offspring were assayed at all 10 loci. Those F_1 individuals typed at less than 4 loci were eliminated from further analyses.

3.3.7 PARENTAGE ASSIGNMENT

Upon completion of microsatellite genotyping of all individuals, parentage assignments for each of the three hatcheries were performed using the statistical likelihood-based approached implemented in the programme CERVUS 2.0 (Marshall *et al.* 1998). This method of analysis is applied to obtain the identity of the most likely parent, with no available genetic data, with a predetermined level of confidence. For each offspring tested, the programme calculates the highest log-likelihood ratio, expressed as LOD scores (in this case the natural logarithm of the combined likelihood ratio), for each candidate parent and compares it to the LOD scores of alternative candidates. According to Marshall *et al.* (1998), the LOD scores of the programme follow certain parameters designed to ensure correct assignment. These advocate that a LOD score equal to zero suggests that all candidate parents of a particular gender are equally as likely to be a parent as any randomly selected male/female; a positive LOD score implies that the alleged parent is more likely to be the true parent than a randomly selected male/female, while a negative LOD score indicates that the candidate parent do not share a common set of alleles with the offspring. The programme also calculates allele frequencies at each locus for each population, along with the probabilities of exclusion, expected heterozygosity, PIC, null allele frequency and Hardy-Weinberg using a Yates correction when chi-square values (df = 1). CERVUS makes provisions for an error rate which takes into account possible imperfections in the data. All these factors add to the successful assignment of parents to their offspring.

In this particular study, both parents are unknown; making all broodstock individuals putative candidates. Firstly, the most likely female candidate for each offspring was chosen on the basis of positive LOD scores. Once identified, the female was run against all male broodstock individuals and vice versa. The confidence level at which a candidate parent was accepted in an assignment was set at 95%, with a typing error rate of 1.0% and 10 000 replication cycles. The genotypes of offspring that were assigned with less than 95% confidence were manually compared against candidate parents and mismatches were evaluated. Offspring that could not be confidently assigned to the correct parent group were eliminated from further analysis.

A second programme was used in an attempt to assign parentage in the case of the Aquafarm, because of the missing broodstock data. Family Assignment Program (FAP) version 3.2 (Taggart 2007) is a custom made software programme that uses a predictive analysis method which assumes that all individuals are the progeny of the possible parental combinations for which all genotype data is available. By using exclusion principles it predicts the resolving power of specific parental

genotypes for unambiguous discrimination, and assigns a family of origin to progeny *via* the genotypic data.

3.4 RESULTS

3.4.1 SOURCE MATERIAL & ABALONE SAMPLING

Broodstock samples from two of the three farms were successfully obtained, however not all the PRS broodstock from Aquafarm could be traced back for sampling, resulting in incomplete broodstock data which hampered the assignment of this farm's offspring to their true parents. At each of the five farm locations, the first replicate containing 200 offspring animals was selected as the source for the project material. A total of 40 individuals {top (mean = 44.3 g & 6.5 mm) and bottom (mean = 27.5 g &1.6 mm) 10%} per farm from each of the five locations were weighed, shell lengths measured and sampled. Animals were sorted to ensure the largest and smallest could be easily distinguished (Figure 3.2).

Figure 3.2 A representation of the replicates situated on all the farm locations and how the animals were sorted into faster growing and the slower growing individuals to be sampled for the growth performance recording scheme.

3.4.2 DNA EXTRACTIONS

The protocols used for the isolation of DNA from broodstock and offspring proved to be consistent in producing reliable template DNA for microsatellite amplification. When testing the concentrations, the majority of the broodstock DNA were very high (up to 1000ng/µl) and therefore had to be diluted into appropriate working stocks of 5-10ng/µl. The DNA isolated via the automated method from the small amount of tissue of the offspring individuals, also produced sufficient concentrations for microsatellite analyses.

3.4.3 STATISTICAL ANALYSES

Broodstock and offspring were assigned microsatellite profiles according to the ten loci chosen. The markers chosen for the suite to assign parentage were highly polymorphic in the pooled

broodstock and F_1 data for each farm. Results of marker diversity were obtained *via* CERVUS and are available in tables 3.2, 3.3 and 3.4.

i) Farm: Aquafarm

Microsatellite profiles for all ten loci were obtained for 95 broodstock animals. A summary of loci diversity can be observed in Table 3.2. The total number of alleles per locus observed in the broodstock ranged from 9 (HmNR54H) to 33 (HmNR281P), with a mean value of 22.9. The mean expected heterozygosity and PIC were 0.847 (range 0.529 – 0.955) and 0.839 (range 0.529 – 0.948), respectively (Table 3.2). HmNR136D showed a significant departure from Hardy-Weinberg equilibrium after applying exact tests ($P < 0.001$). Probabilities of exclusion when no parent information was available (Excl 1) ranged from 0.183 to 0.818 and from 0.370 to 0.900 when information for one parent was available (Excl 2). However when exclusionary powers were combined for all ten loci, values above 0.999 for both Excl 1 and Excl 2 were obtained. Null allele frequencies estimated by CERVUS ranged between approximately 0% (HmNR54H) and 25% (HmNR136D) based on heterozygote deficiency in the broodstock. Since the probability of detecting null alleles depends on the frequency in the broodstock, a critical value of 0.05 should be taken into consideration as values greater than this threshold may compromise parentage determination (Marshall *et al*. 1998). As seen in Table 3.2, 5 markers had null frequencies greater than 0.05. In particular, HmNR136D had a frequency greater than 0.2.

Table 3.2 Estimates of genetic diversity for the 10 loci analysed in Aquafarm broodstock. For each marker the number of alleles (N_a), number of heterozygotes (Het) and homozygotes (Hom), heterozygosity observed (H_o) and expected (H_e), polymorphic information content (PIC), probabilities of exclusion (Excl 1: no parent known; Excl 2: one parent known), and null allele frequencies were calculated by CERVUS 2.0.

Locus	N_a	Het	Hom	H_o	H_e	PIC	Excl 1	Excl 2	Null Freq
HmNR54H	9	69	26	0.726	0.736	0.691	0.331	0.507	-0.0005
HmNR120T	32	87	8	0.916	0.948	0.940	0.793	0.884	+0.0144
HmNR136D*	27	47	48	0.495	0.837	0.824	0.543	0.707	**+0.2554**§
HmNR224T	24	77	16	0.828	0.950	0.942	0.798	0.888	**+0.0658**§
HmNR281P	33	74	21	0.779	0.955	0.948	0.818	0.900	**+0.0997**§
HmNST7T	24	87	8	0.916	0.943	0.935	0.778	0.875	+0.0115
HmNS28D	32	84	11	0.884	0.938	0.930	0.765	0.866	+0.0272
HmNS38T	14	61	34	0.642	0.764	0.786	0.436	0.617	**+0.0985**§
HmNS56D	16	49	46	0.516	0.529	0.529	0.183	0.370	+0.0397
HmNS58R	18	63	32	0.663	0.869	0.869	0.614	0.761	**+0.1409**§
Average	**22.9**			**0.737**	**0.847**	**0.839**			
Total							**0.999978**	**1.000000**	

***** Not in HWE ($p<0.01$); § null freq > 0.05

ii) Farm: I&J

A total of 36 broodstock animals were assigned genotype profiles for ten loci. Genetic diversity estimates of the ten loci is summarised in Table 3.3. The mean number of alleles was 18.1, ranging

from 8 alleles for HmNR54H, to 26 alleles for HmNR120T. The mean expected heterozygosity and PIC values were 0.872 and 0.848, respectively. Values for expected heterozygosity and PIC ranged from 0.757 (HmNR54H) to 0.953 (HmNR244T), and from 0.708 (HmNR54H) to 0.937 (HmNR224T), respectively. All loci conformed to HWE after applying exact p-tests ($P > 0.01$), indicating segregation in a Mendelian fashion. When no parent information was available (Excl 1), exclusion probabilities per locus ranged from 0.276 for HmNS38T to 0.785 for HmNR224T with a total combined exclusionary power above 0.999. When the genotype of one parent was available (Excl 2), probabilities of exclusion ranged from 0.462 to 0.879 with a combined power greater than 0.999. Null allele frequencies ranged from < 0 to 0.167 (HmNR281P).

Table 3.3 Estimates of genetic diversity for the 10 loci analysed in I&J broodstock. For each marker the number of alleles (N_a), number of heterozygotes (Het) and homozygotes (Hom), heterozygosity observed (H_o) and expected (H_e), polymorphic information content (PIC), probabilities of exclusion (Excl 1: no parent known; Excl 2: one parent known), and null allele frequencies were calculated by CERVUS 2.0.

Locus	N_a	Het	Hom	H_o	H_e	PIC	Excl 1	Excl 2	Null Freq
HmNR54H	8	23	13	0.639	0.757	0.708	0.351	0.528	**+0.0808**§
HmNR120T	26	35	1	0.972	0.950	0.934	0.776	0.087	-0.0192
HmNR136D	17	32	4	0.889	0.872	0.847	0.573	0.729	-0.0173
HmNR224T	24	33	3	0.917	0.953	0.937	0.785	0.879	+0.0122
HmNR281P	22	24	12	0.667	0.948	0.931	0.766	0.867	**+0.1673**§
HmNST7T	20	36	0	1.000	0.937	0.919	0.734	0.846	-0.0403
HmNS28D	19	33	3	0.917	0.929	0.910	0.712	0.832	+0.0004
HmNS38T	11	21	15	0.583	0.671	0.634	0.276	0.462	**+0.0813**§
HmNS56D	20	29	7	0.806	0.780	0.760	0.440	0.624	-0.0486
HmNS58R	14	34	2	0.944	0.919	0.899	0.680	0.810	-0.0224
Average	**18.1**			**0.833**	**0.872**	**0.848**			
Total							**0.999969**	**1.000000**	

(§ null freq > 0.05)

iii) Farm: Roman Bay

Microsatellite profiles for the ten analysed loci were obtained for 33 broodstock animals (Table 3.4). The spectrum of alleles ranged from a minimum number of 7 (HmNR54H) to a maximum number of 27 (HmNR281P), with a mean value of 18.9. Mean expected heterozygosity and PIC were 0.886 and 0.861, ranging from 0.740 and 0.685 for HmNR54H to 0.958 and 0.941 for HmNR281P, respectively. All loci conformed to HW expectations after applying exact p-tests ($P > 0.01$). Probabilities of exclusion for the 10 loci typed ranged from 0.321 to 0.796 when no parent genotypes were available (Excl 1), and from 0.496 to 0.886 when parent information was available for one individual (Excl 2) at loci HmNR54H and HmNR281P, respectively. Combined exclusionary powers for the 10 loci were above 0.999 for both Excl 1 and Excl 2. Frequencies of null alleles in the broodstock ranged from < 0% to 30% for all loci. Locus HmNS38T demonstrated a high number of observed homozygotes, a value larger than that of observed heterozygotes, thus

validating the high null allele frequency. This marker may be responsible for false exclusion due to single mismatches or inconclusive assignments.

Table 3.4 Estimates of genetic diversity for the 10 loci analysed in Roman Bay broodstock. For each marker the number of alleles (N_a), number of heterozygotes (Het) and homozygotes (Hom), heterozygosity observed (H_o) and expected (H_e), polymorphic information content (PIC), probabilities of exclusion (Excl 1: no parent known; Excl 2: one parent known), and null allele frequencies were calculated by CERVUS 2.0.

Locus	N_a	Het	Hom	H_o	H_e	PIC	Excl 1	Excl 2	Null Freq
HmNR54H	7	22	10	0.688	0.740	0.685	0.321	0.496	+0.0138
HmNR120T	25	29	3	0.906	0.948	0.929	0.765	0.866	+0.0120
HmNR136D	20	27	5	0.844	0.889	0.865	0.613	0.760	+0.0120
HmNR224T	21	28	4	0.875	0.911	0.890	0.669	0.801	+0.0054
HmNR281P	27	22	10	0.688	0.958	0.941	0.796	0.886	**+0.1600**§
HmNST7T	20	29	3	0.906	0.949	0.930	0.765	0.867	+0.0152
HmNS28D	24	32	0	1.000	0.956	0.938	0.786	0.880	-0.0308
HmNS38T	12	15	17	0.469	0.855	0.827	0.535	0.700	**+0.2999**§
HmNS56D	18	23	9	0.719	0.754	0.731	0.397	0.586	-0.0107
HmNS58R	15	27	5	0.844	0.899	0.875	0.629	0.773	+0.0241
Average	**18.9**			**0.794**	**0.886**	**0.861**			
Total							**0.999978**	**1.000000**	

(§ null freq > 0.05)

3.4.4 PARENTAGE ASSIGNMENT

Assignments of progeny to their respective parental pairs were done according to the LOD score results obtained from CERVUS. According to Slate *et al.* (2000), a LOD score of 3.0 or more is used to confirm parentage while a LOD score less than 3.0 is regarded as inconclusive. As only a percentage of F_1 individuals could be successfully assigned to only one parent, factors which may influence parentage results should be taken into consideration. These include mismatches at one or more loci, genotyping errors, incomplete data sets and the presence of null alleles.

i) Farm: Aquafarm

Parentage assignment for Aquafarm was unsuccessful when both programmes were used due to the lack of broodstock data. CERVUS assigned progeny to potential parents, however all LOD scores were below the required value of 3.0 for unambiguous assignment, leaving the results inconclusive for the F_1 individuals. When FAP was applied, only ten individuals of the total of 198 could be assigned unambiguously to a broodstock pair. This is an indication that the large majority of unassigned individuals are most likely the progeny of those broodstock animals that were not sampled.

ii) Farm: I&J

Microsatellite profiles, based on the ten chosen microsatellite DNA markers, were obtained for a total of 198 progeny, 98 fast growers and 100 slow growers. The two missing fast growers showed

PCR amplification at only three of the ten loci and were therefore excluded from further analysis. Unambiguous assignment to a single broodstock couple was confidently achieved for 62.6% of the progeny (124 juveniles), while 20.7% (41 individuals) were confidently assigned to just one parent (35 to a single female and 6 to a single male; LOD > 3.0). The final 16.7% (33 individuals) remained inconclusive with LOD scores of less than 3.0, and were therefore eliminated from further analyses. Mismatches at three or more loci are possibly the cause of the data being inconclusive. It results in one offspring being assigned to more than two sets of parents who may, by chance, all have the same genotypes at the loci being typed. Male and female contributions to both fast and slow growing animals were determined (Figure 3.3) as well as the parental pairs (Appendix D) that contributed to both growth groups in order to determine a trend, where possible, between the faster and slower growing F_1 progeny.

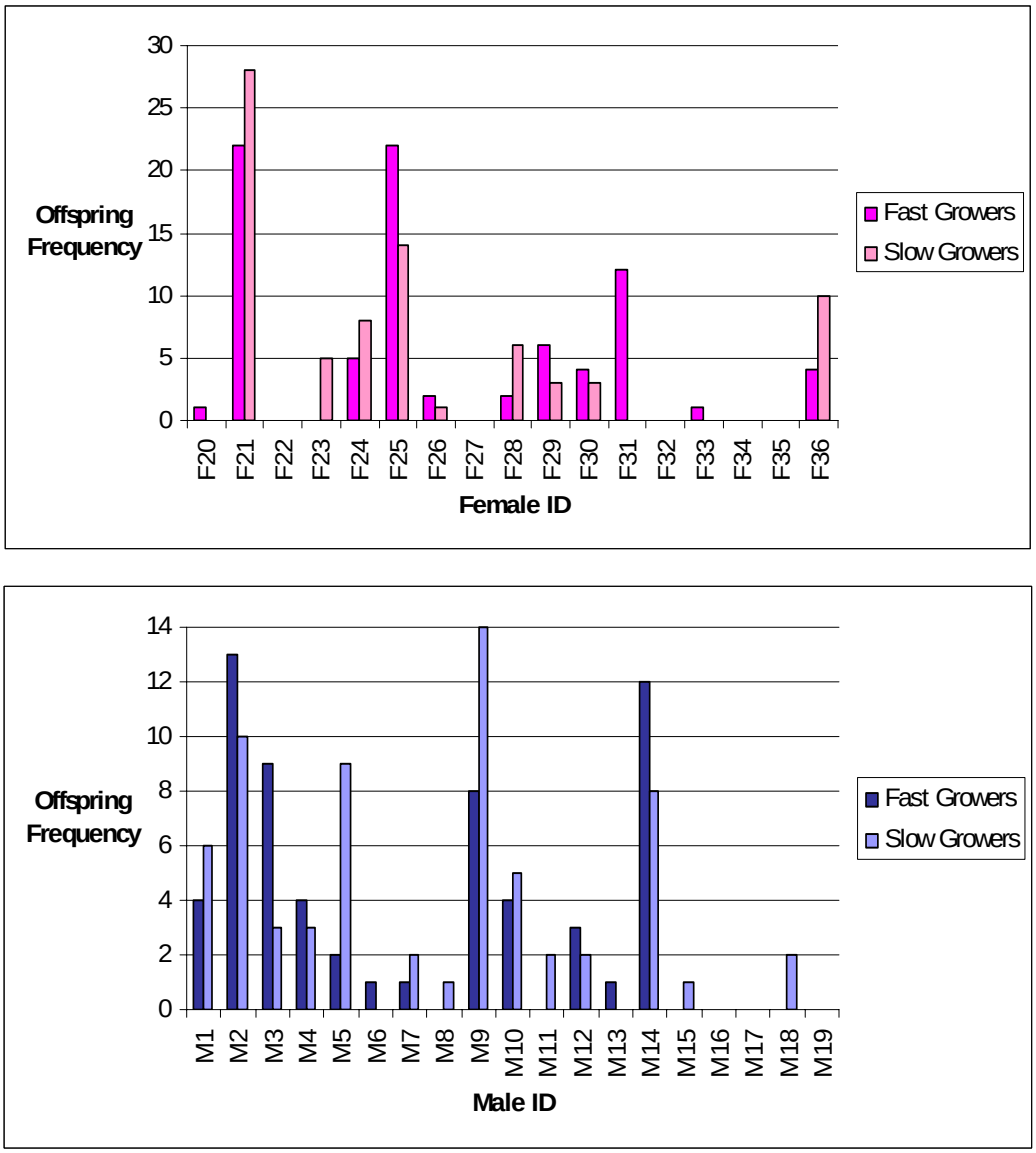


Figure 3.3 Parental contributions from I&J to fast and slow growing F_1 animals submitted to the growth PRS. Indications are clear as to which are the dominant contributing female and male broodstock.

The resulting female pedigree structure, based on 159 offspring, indicates that there are two dominant female contributors to both the slow and the fast growers: individuals F21 and F25, making up a total of 54.1% (86 offspring) of the female contributions to the progeny assigned. Female brooder F31 contributed more to fast growing offspring (12 individuals), while F36 spawned more slow growing offspring (10 individuals), however, on the commercial level of stock production the number of individuals spawned by the two is not very significant. Five of the female individuals contributed to none of the conclusively assigned offspring. The contribution of male broodstock is to some extent more homogeneous, with the number of individuals contributing to progeny being greater than that of the females. Characteristics of male pedigree organisation (130 offspring) designate individuals M2, M9 and M14 as being dominant male contributors to both fast and slow growing offspring, comprising a total of 50.0% of F_1 progeny assigned (65 individuals), while M3 and M5 contributed equally (9 offspring) to fast growers and slow growers, respectively. A total of three males contributed to none of the unambiguously assigned offspring individuals.

The main characteristic of the total pedigree structure of juveniles assigned to one couple (Appendix D) is that it consists of two females (F21 and F25) dominating spawning of fast and slow growers with eight and ten males, respectively, which contributed unequally. For the fast growing animals, the maximum amount of contributions by a given broodstock pair was six individuals. This was only seen by two couples, namely F21/M14 and F25/M2. For the slow growing animals one couple, F21/M9, contributed to six individuals of the total amount. Parental pairs F21/M2 and F21/M3 contributed equally to slow and fast growing offspring, respectively, spawning four individuals in total. Other broodstock pair contributions for both groups of offspring were between one and three individuals. The parental contributions shown may not indicate any specific trend; however individual broodstock members were able to be singled out as dominant contributors in comparison to the spawning performance of others. This is an expected result since the expected spawning behaviour of commercially reared *Haliotis midae* would be to have at least 2-3 females coupled with multiple males.

iii) Farm: Roman Bay

A total of 199 progeny animals, 100 fast and 99 slow growing were screened and subjected to parentage assignment. Parentage assignment of progeny (199 juveniles) using the ten loci revealed that 122 juveniles (61.3%) could be unambiguously assigned to a single pair of parents, whereas 39 individuals (19.6%) were confidently assigned to a single parent, 33 to a single female and 6 to a single male (LOD > 3.0). Inconclusive results (LOD < 3.0) obtained for 26 individuals (13.1%) were the result of multiple male and female assignments. In addition, the remaining 12 juveniles

(6.0%) were left unassigned due to no match at more than two loci, which is most likely attributed to the absence of PCR products for such loci. Individuals with no assignment as well as those with inconclusive assignments were left out of further analyses. Contributions of male and female broodstock to 161 juveniles are shown in Figure 3.4. Parental pair contributions to individuals assigned to a single pair can be seen in Appendix E.

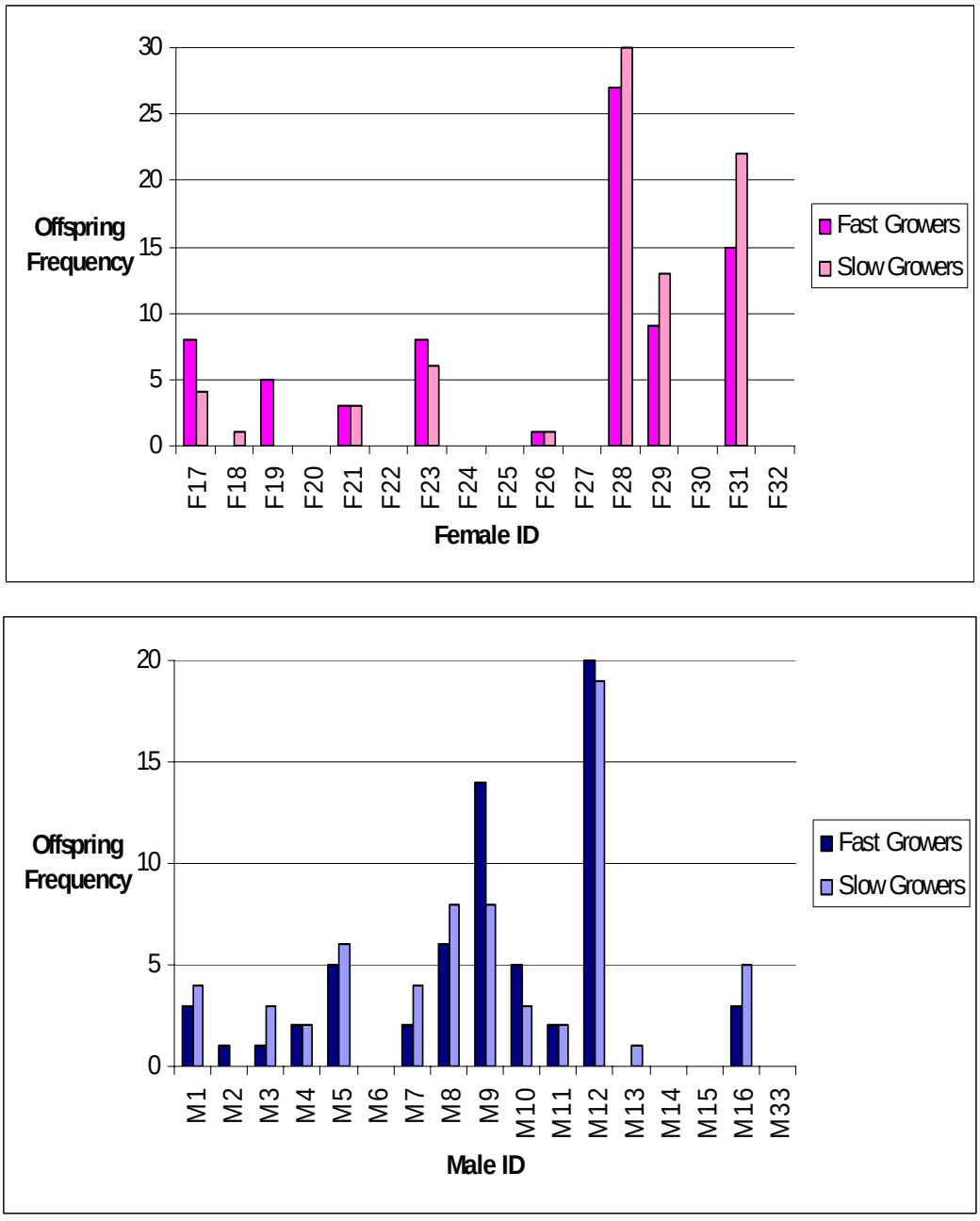


Figure 3.4 Parental contributions from Roman Bay to fast and slow growing F_1 animals submitted to the growth PRS. Indications are clear as to which are the major contributing female and male broodstock.

Inspection of the female distribution graph, including 155 offspring, demonstrates a clear dominance, to both fast and slow growers, by three individuals namely: F28, F29 and F31.

Together, the three females were responsible for maternal contribution to 116 offspring (74.8%). Females F17 and F23 also made contributions to both sets of progeny, but on a much smaller scale of 12 and 14 individuals, respectively (16.8%). A total of seven females made no contributions toward conclusively assigned offspring. Distribution analysis of male broodstock individuals reveals somewhat of a more uniform contribution pattern. However, two individuals do stand out as major contributors to fast and slow growing offspring (128 individuals). Male brooders M9 and M12 dominated the contributions to resulting progeny, being the male parent to 61 assigned individuals (47.7%). Individuals M5 and M8 made less of a contribution by being assigned as the male parent to 11 and 12 offspring (18.0%), respectively. A total of four males contributed to none of the unambiguously assigned offspring individuals.

A distinctive parental pair contribution can be seen in the pedigree distribution graph of all progeny assigned to a single couple (Appendix E). Brooders F28 and M12 dominated spawning by producing 31 offspring, consisting of 14 fast growers and 17 slow growers. Minor contributing couples to fast growing juveniles were F17/M9 with seven and F28/M5 with five individuals, while F29/M10 and F31/M12 each contributed to four individuals. Slow growing progeny confirmed minor contributions by broodstock pairs F31/M12, F28/M5 and F28/M16 to ten, six and four individuals, respectively. Remaining couples for fast and slow growers contributed between one and three individuals.

3.5 DISCUSSION

In a mass spawning system, such as that of *Haliotis midae*, the actual number of animals spawning and individual contributions are most often unknown. Unknown contributions of broodstock may have consequences for the replacement of broodstock in that the relative performances of individual animals cannot be established. For this reason, knowledge of pedigree structures and parental contributions are of particular importance in selection programmes such as the one being established for the five participating abalone farms. Microsatellite-based traceability methods have demonstrated to be very useful for the accurate attainment of pedigree information in many abalone species (Selvamani *et al.* 2001; Li *et al.* 2003; Gutierrez-Gonzalez & Perez-Enriquez 2005; Lucas *et al.* 2006; Sekino & Hara 2007). The current study used a microsatellite suite of ten loci to assign parentage to hatchery-reared abalone in an effort to determine broodstock contributions in a preliminary evaluation of growth performance. The three farms for which parentage assignment was performed will be discussed in the following sections.

FARM: AQUAFARM

Despite the high number of microsatellite used with the marker suite used, the lack of relevant broodstock data resulted in unsuccessful parentage assignment. This emphasises the importance of having completed data sets and pedigree information of the broodstock in order to guarantee accurate parentage assignment. Incomplete data sets are a problem for many researchers doing various studies as reliable pedigree information plays an important role in numerous aspects of genetic enhancement programmes, particularly in identification of QTLs, MAS and the management of broodstock in maintaining genetically diverse commercial stocks through the minimisation of inbreeding. To obtain information on the pedigree of this farm, it is advisable that all broodstock animals in the hatchery be screened in order to assign parents to juveniles of the PRS.

FARM: I&J

The ten loci analysed showed high genetic variation in the broodstock, with a mean expected heterozygosity and a mean allele number of 0.872 and 18.1, respectively. Given the high combined exclusion power of the set of markers, correct assignment could be achieved. However, single parent-offspring mismatches did occur, even with the modest error rates applied by CERVUS. Although the majority of null allele frequencies were low, three of the ten loci showed frequencies above the recommended threshold of 0.05. A total of 124 of 198 (62.6%) F_1 animals were successfully and unambiguously assigned to a single broodstock pair. The failure to assign the remaining 33 individuals (16.7%) conclusively might be explained by poor data quality (genotyping errors) or high levels of relatedness among broodstock (relation between brooders was unknown). The pedigree information obtained allowed for tracing of possible genotypic patterns of faster and slower growing offspring to determine whether specific broodstock couples were contributing more to fast growers and if others were contributing more to slow growers. The results obtained suggested no distinct pattern in parental contributions from broodstock, however there was a trend with regards to overall contribution of individuals to the combination of fast and slow growing animals. Broodstock contributions to both fast and slow growing progeny showed two dominant female brooders contributing to more than 50% of offspring analysed. The majority of male brooder contributions to offspring were the result of three individuals having contributed to 50% of progeny.

FARM: ROMAN BAY

From the 199 progeny typed for ten loci, 122 (61.3%) were unambiguously assigned to a male and female parent. Inconclusive results were obtained for 26 (13.1%) who were assigned to multiple

pairs of parents, and the remaining 12 (6.0%) were not assigned to any parent. Pedigree distribution data was comparable to that of I&J for fast and slow growing offspring. Female contributions were dominated by three brooders contributing to more than 70% of offspring, while two male brooders made up approximately 50% of the contributions. Although exclusion probabilities were above 0.999, and mean allele number and expected heterozygosity were high (18.9 and 0.886, respectively), assignment power of the ten markers was not as high as expected. Overall, null allele frequencies were low. However, there were two loci with null allele frequencies far above 0.05, particularly HmNS38T which had a frequency of approximately 0.3 due to the high number of homozygotes typed.

Using the parentage assignment data for I&J and Roman Bay, male and female broodstock individuals who had parented approximately 60% of both fast and slow growing offspring were determined. However, these data revealed a less than considerable variation in the contribution of each broodstock to the new generation: only a few males and females from relatively small groups of broodstock (≈ 34) produced most of the offspring. The ability to assign parentage is dependent upon the number of loci available and their ability to exclude non-parents; this is greatly related to the allelic diversity of the loci. Although the number of parents exerts a diminishing effect on the number of loci needed, the effect is reduced with an increase in the number of alleles (Selvamani *et al*. 2001). Given that the average number of alleles for *H. midae* microsatellite loci in this study was 18.5, the number of loci used should have been more than sufficient to assign 100% of the progeny. Nevertheless, roughly 40% of offspring remained assigned to just a single parent, had inconclusive assignments or were left unassigned, suggesting potential errors that may have affected parentage assignment. These sources of error include genotyping/scoring errors, the presence of null alleles and other less common factors such as mutations and relatedness.

GENOTYPING ERRORS

Microsatellite genotyping errors are a major source of problems with regards to parentage assignment. They occur when the genotype determined for an individual, after molecular analysis, does not correspond with the authentic microsatellite profile of that individual. Errors can occur in the genotypes of offspring, mother or father, or a combination of the three. They can be generated at every step of the genotyping process (sampling, DNA isolation, data scoring and analysis) by a variety of factors including human errors, technical anomalies and chance (Bonin *et al.* 2004). The microsatellite markers used in this study were highly polymorphic which could be the cause of many of the potential errors associated with genotyping because of the larger number of alleles available for scoring. Increasing the number of markers may be a solution to the problem; however,

genotyping errors become more prominent as the number of loci used increases. Conversely, using fewer but more informative loci as a reduction method may also prove problematic as error rates increase for more polymorphic loci (Hoffman & Amos 2005).

According to various studies, the most serious problem associated with microsatellite genotyping errors occur when PCR amplification is unreliable, resulting in the random failure of alleles to amplify (Gagneux *et al.* 1997; Miller *et al.* 2002; Hoffman & Amos 2005; Castro *et al.* 2006). Referred to as 'allelic dropout' (Walsh *et al.* 1992), this anomaly leads to heterozygote individuals being ambiguously typed as homozygotes. There are several mechanisms that could potentially lead to the preferential PCR amplification of the smaller of two alleles (Walsh *et al.* 1992). First, preferential amplification can result from significant GC% differences between alleles if the conditions of the reaction allow the denaturation of one allele but not the other (differential denaturation). Second, when PCR products from different alleles differ in length, preferential amplification of the shorter allelic product can occur when *Taq* DNA polymerase is limiting. Third, when the initial number of genomes sampled is very small, haphazard fluctuation in the number of copies of each allele can result in what appears to be preferential amplification (Taberlet *et al.* 1996). Finally, less efficient priming of DNA synthesis of one allele versus another can occur because of mismatches between the primer and the specific allelic template, resulting in preferential amplification of the smaller allele.

The number of microsatellite genotyping errors increases with sample size, however, they more often than not go undetected because they are by and large inconspicuous (Sobel *et al.* 2002). Genotyping errors affect both allele frequency estimates and precise discrimination of different genotypes which, consequently, create an artificial excess of homozygotes (Taberlet *et al.* 1996), false departure from HWE (Xu *et al.* 2002), and can distort individual identification and parentage assignment (Miller *et al.* 2002). With the growing number of studies showing that even modest error rates can seriously disturb estimates of kinship and parentage assignment, the importance of microsatellite genotyping errors has become increasingly recognised. However, few have looked at the origin of genotyping errors (Hoffman & Amos 2005). Each time a homozygote is observed, there is the possibility that it is actually a true heterozygote where consecutive dropout errors have occurred. The effect of an artificial increase in homozygosity then results in biased allele frequencies and ambiguous assignment of parentage. Marshall *et al.* (1998) determined that assuming genetic data to be error free, will lead to overestimation of the confidence in parentage assignment when data includes errors and result in the false exclusion of true parents.

NULL ALLELES

A classical source of incompatibilities in assigning parentage with microsatellites is the presence of null alleles (Pemberton *et al.* 1995). Microsatellite null alleles are consistently nonamplifying alleles that give the impression of a homozygous state in a heterozygous individual. They result typically from mutations in one or both flanking sequences of a locus, thus lacking a functional PCR primer binding site (Jones *et al.* 1998). Early research has shown that mutations disrupting primer binding sites range from single nucleotide point mutations (Ishibashi *et al.* 1996; Lehmann *et al.* 1996), to deletions of up to twelve base pairs (Ede & Crawford 1995).

Identification of null alleles is an important consideration for parentage assignment as they cause false exclusions when heterozygotes are incorrectly scored as homozygotes. Fortunately, null alleles can usually be detected when there is a significant departure from HWE through heterozygote deficiencies (Jones & Ardren 2003), or from expected Mendelian inheritance patterns (Castro *et al.* 2004). In programmes, such as CERVUS, null allele frequencies are estimated *via* a statistical approach on the basis of heterozygote deficiencies. Dakin and Avise (2004) reported that when the null allele frequency has a *p* value greater than 0.20, mean exclusion probabilities can be much higher than true values. For this reason, the marker with such a frequency should not be used for the parentage analysis, hence promoting the use of several microsatellite loci for which the combined exclusion probability is calculated as a function of the product of single locus exclusion probabilities. Marshall *et al.* (1998) suggested that when null allele frequencies are above 5% ($p > 0.05$), markers should be excluded as they are considered to compromise pedigree deductions. On the contrary, null alleles at low frequencies are more difficult to detect.

Null alleles are also said to cause complications in parentage analyses *via* their potential to falsely exclude true parents when offspring need to be assigned. An apparent homozygous offspring could actually be the progeny of an adult displaying a different homozygous genotype if in fact both are heterozygotes for a null allele that the offspring inherited from that parent (Dakin & Avise 2004). For example: progeny of a sire heterozygous for a null allele that receive the null allele will appear to be homozygous for their maternal allele. If the maternal allele is different from the sire's observed allele then no common allele will be observed between the sire and his progeny, and paternity will be erroneously rejected. Such a problem can occur even when null allele frequencies are low. Even greater rates of false exclusion are encountered when null alleles are present at multiple loci (Liewlaksaneeyanawin *et al.* 2002).

MUTATIONS & RELATEDNESS

Variability of microsatellites depends on their mutation. Mutation rates are said to range from 10^{-2} to 10^{-4} per locus per generation (Ellegren 2000; Schlötterer 2000). Microsatellite mutations may generate mismatches between genuine relatives at considerable frequencies, subsequently resulting in the false exclusion of one or both true parents. Although mutations are alterations of single alleles rather than pairs of alleles, treating them as an error is preferable to using them as a basis of exclusion (Marshall *et al.* 1998). Data sets in which offspring are collected separately from parents are most vulnerable to undetected mutations as transmission of alleles remains unnoticed. Another less common and potential challenge of parentage assignment is relatedness among broodstock individuals. Individuals obtained from the wild to become founders in commercial hatcheries may have been sourced *via* non-random sampling, thus bring in similar genotypes. When there is a relation between candidate parents, the power of parentage identification is decreased. It might also explain the fact that several offspring are assigned to one parent and multiple parents of the opposite sex, or multiple parents of both sexes. In such cases, more markers are needed to correctly assign offspring (Chatziplis *et al.* in press). Relatedness is, however, difficult to determine particularly in a study such as the current one. The broodstock used by the hatcheries are the founder populations of the farms. The animals were obtained from the wild and therefore there is no pedigree information to determine their relationships, however potential relationships can be determined using molecular techniques.

DIFFERENTIAL BROODSTOCK CONTRIBUTIONS

The results obtained in the current study from successful parentage assignment, provided an accurate account of pedigree information on which broodstock animals were contributing parents to F_1 offspring. Although juveniles were divided into two groups based on growth performance, there were no indications that certain couples contributed to progeny of a specific growth group (fast or slow). In fact, those parental pairs contributing did so similarly to both fast and slow growing offspring. It is therefore assumed that in this study, no observable genotypic links can be associated with fast or slow growers, but certain genotypes were responsible for producing the majority of F_1. A recent study undertaken by Hara and Sekino (2007) also demonstrated unequal contribution of parents to the next generation, despite normalising the number of gametes from males and females. These findings may be indicative of poor fertilisation success or sperm and egg quality, genetic incompatibility of gametes or differences in survival among families. The unequal contribution of parents is, however, not surprising as it is likely to occur in commercial operations such as those for *H. midae* where multiple parents are induced to spawn simultaneously. Nevertheless, it would be

advisable to gain a deeper understanding of these and other factors that may cause differential contributions among families reared together.

The indication of certain brood animals not producing offspring may be the result of unsuccessful fertilisation within the hatcheries. Without a doubt, fertilisation among broodstock is extremely important and absolutely necessary to the successful sustainability of any commercial hatchery. A study carried out on the commercial abalone *H. laevigata* maintained that unsuccessful fertilisation was a key limiting factor on recruitment (Babcock & Keesing 1999). The success of fertilisation is dependent upon the synchronised spawning of male and female brooders (see section 1.5.3). Once eggs and sperm have been spawned, fertilisation occurs randomly in the surrounding water. But it is important to ensure that all brooders are fecund and are in fact able to spawn enough gametes of good a high quality. Good quality gametes are important to subsequent development and reliant upon their biochemical content. It was found in *H. rubra* that the ratio between yolk diameter and egg diameter in female gametes was central in determining egg quality. A bigger egg size did not necessarily result in better fertilisation performance, but rather the yolk size: egg size ratio (Litaay & Da Silva 2001). Furthermore, it was recently confirmed that a larger cytoplasm diameter offered more sites for sperm to enter thus increasing success of fertilisation (Huchette *et al.* 2004). Sperm concentration play as much of an important role as the quality of eggs. In the cultured abalone *H. tuberculata* it was found that increasing sperm concentrations to 10^6 sperm/ml increased the success of fertilisation (Baker & Tyler 2001). However, values above these resulted in the destruction of ova or abnormal development of larvae due to polyspermy, when several spermatozoa enter an egg instead of just a single sperm. The study also suggested that the length of time sperm cells remain in contact with the ovum influences fertilisation success as the most significant changes occur in the first few minutes after adding sperm to the egg.

Inbreeding, or mating between individuals who share genes from a common ancestor, is a central cause of genetic incompatibility among gametes. Genetic incompatibility refers to the negative interactions between genetic elements inherited from the parents that cause inviability in zygotes, thus resulting in no offspring development due to a loss in genetic fitness (Puurtinen *et al*. 2005). In genetically closed populations such as that of an abalone hatchery, the potential for inbreeding is high. Despite the high fecundity and variable fertility of abalone, as well as the relatively large numbers of broodstock in hatcheries, just a few individuals may contribute to the next generation of seed. This may result in an increase in the rate of inbreeding and therefore differential survival of families or interfamily competition (Elliot 2000).

Another possible factor for survival difference among families (i.e. certain parental pairs producing offspring and other seemingly producing none) is intraspecific competition among juveniles in cultured surroundings. Intraspecific competition is an important source of growth and is often associated with organisms cultured in high population densities where food and space are limited (Beal 2006). There have been few studies of intraspecific competition, and those published have focused on the effects it has on growth. However, it is possible to apply the concept to this study. A study performed on the cultured mussel *Mytilus galloprovincialis*, showed differences in familial genotypes having significant effects on direct and associated competition for growth. The differences between genotypes could be due to additive and dominant genetic effects and would thus provide an estimate of genotypic variance for competition (Brichette *et al.* 2001). This suggests that intraspecific competition can be observed at a genetic level among individuals as well as among families, as was the case in the current study.

In conclusion, although 100% parentage assignment was not achieved, some positive results were achieved. Dominant brooders were identified for I&J and Roman Bay farm, as well as individuals who made no contributions to assigned offspring. Non-contributing broodstock could be removed from hatcheries to make room for new brood animals, thereby potentially sustaining a high level of genetic diversity. Furthermore, the microsatellite markers used for parentage assignment should be selected with extreme care. Not only should the criteria used in this study be considered (allelic diversity, PIC, *etc*.), but more emphasis should be placed on the occurrence of null alleles. In future, those with high null allele frequencies for a particular broodstock should be avoided from being used in assignments of those specific individuals. The estimation of genetic parameters (heritability and genetic correlation) in growth traits of juveniles and adults is also essential in any future breeding programme to enhance growth rate. Heritability estimates of growth in particular are necessary for the purpose of predicting the genetic gain by selection, thus ensuring successful selective breeding programmes.

3.6 REFERENCES

Babcock R, Keesing J (1999) Fertilization biology of the abalone *Haliotis laevigata*: laboratory and field studies. *Canadian Journal of Fisheries and Aquatic Science* 59: 1668-1678

Baker MC, Tyler PA (2001) Fertilization success in the commercial gastropod *Haliotis tuberculata*. *Marine Ecology Progress Series* 211: 205-213

Beal BF (2006) Relative importance of predation and intraspecific competition in regulating growth and survival of juveniles of the soft-shell clam, *Mya arenaria* L., at several spatial scales. *Journal of Experimental Marine Biology and Ecology* 336: 1-17

Beaumont AR, Hoare K (2003) Genetic considerations in the hatchery. In: Biotechnology and Genetics in Fisheries and Aquaculture. Blackwell Science Ltd, Oxford, England, pp73-90

Bonin A, Bellemain E, Bronken-Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* 13: 3261-3273

Brichette I, Reyero MI, García C (2001) A genetic analysis of intraspecific competition for growth in mussel cultures. *Aquculture* 192: 155-169

Castro J, Bouza C, Presa P, Pino-Querido A, Riaza A, Ferreiro I, Sánchez L, Martínez P (2004) Potential sources of error in parentage assignment of turbot (*Scophthalmus maximus*) using microsatellite loci. *Aquaculture* 242: 119-135

Castro J, Pino A, Hermida M, Bouza C, Riaza A, Ferreiro I, Sánchez L, Martínez P (2006) A microsatellite marker tool for parentage analysis in Senegal sole (*Solea senegalensis*): Genotyping errors, null alleles and conformance to theoretical assumptions. *Aquaculture* 261: 1194-1203

Chatziplis D, Batargias C, Tsigenopoulos CS, Magoulas A, Kollias S, Kotoulas G, Volckaert F, Hayley C (2007) Mapping quantitative trait loci in European sea bass (*Dicentrarchus labrax*): the BASSMAP pilot study. *Aquaculture,* in press.

Dakin EE, Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* 93: 504-509

Ede AJ, Crawford AM (1995) Mutations in the sequence flanking the microsatellite at the *KAP8* locus prevent the amplification of some alleles. *Animal Genetics* 26: 43-44

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics* 16: 551-558

Elliot NG (2000) Genetic improvement programmes in abalone: what is the future? *Aquaculture Research* 31: 51-59

Estoup A, Jarne P, Cornuet J (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* 11: 1591-1604

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with non-invasive genotyping based on nuclear DNA amplified from shed hairs. *Molecular Ecology* 6: 861-868

Gutierrez-Gonzalez JL, Perez-Enriquez R (2005) A genetic evaluation of stock enhancement of blue abalone *Haliotis fulgens* in Baja California, Mexico. *Aquaculture* 247: 233-242

Hara M, Sekino M (2003) Efficient detection of parentage in a cultures Japanese flounder *Paralichthys olivaceus* using microsatellite DNA markers. *Aquaculture* 217: 107-114

Hara M, Sekino M (2007) Parentage testing for hatchery-produced abalone *Haliotis discus hannai* based on microsatellite markers: preliminary evaluation of early growth of selected strains in mixed family farming. *Fisheries Science* 73: 831-836

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* 14: 599-612

Huchette SMH, Soulard JP, Koh CS, Day RW (2004) Maternal variability in the blacklip abalone, *Haliotis rubra* Leach (Mollusca: Gastropoda): effect of egg size on fertilisation success. *Aquaculture* 231: 181-196

Isberg SR, Chen Y, Barker SG, Moran C (2004) Analysis of microsatellites and parentage testing on saltwater crocodiles. *Journal of Heredity* 95: 445-449

Ishibashi Y, Saitoh T, Abe S, Yoshida MC (1996) Null microsatellite alleles due to nucleotide sequence variation in the grey-sided vole *Clethrionomys rufocanus*. *Molecular Ecology* 5: 589-590

Jerry DR, Preston NP, Crocos PJ, Keys S, Meadows JRS, Li Y (2004) Parentage determination of Kuruma shrimp *Penaeus* (*Marsupenaeus*) *japonicus* using microsatellite markers (Bate). *Aquaculture* 235: 237-247

Jerry DR, Preston NP, Crocos PJ, Keys S, Meadows JRS, Li Y (2006) Application of DNA parentage analysis for determining relative growth rates of *Penaeus japonicus* families reared in commercial ponds. *Aquaculture* 254: 171-181

Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations. *Molecular Ecology* 12: 2511-2523

Jones AG, Stockwell CA, Walker D, Avise JC (1998) The molecular basis of a microsatellite null allele from the White Sands pupfish. *Journal of Heredity* 89: 339-342

Lehmann T, Hawley WA, Collins FH (1996) An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics* 144: 1155-1163

Li Q, Park C, Kobayashi T, Kijima A (2003) Inheritance of microsatellite DNA markers in the pacific abalone *Haliotis discus hannai*. *Marine Biotechnology* 5: 331-338

Liewlaksaneeyanawin C, Ritland CE, El-Kassaby YA (2002) Inheritance of null alleles for microsatellites in the white pine weevil (*Pissodes strobe* [Peck] [Coleoptera: Curculionidae]). *Journal of Heredity* 93: 67-70

Litaay M, Da Silva SS (2001) Reproductive performance indices based on physical characteristics of female blacklip abalone *Haliotis rubra* L. *Journal of Shellfish Research* 20: 673-677

Liu ZJ, Cordes JF (2004) DNA markers and their applications in aquaculture genetics. *Aquaculture* 238: 1-37

Lucas T, Macbeth M, Degnan SM, Knibb W, Degnan BM (2006) Heritability estimates for growth in the tropical abalone *Haliotis asinina* microsatellites to assign parentage. *Aquaculture* 259: 146-152

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology* 7: 639-655

Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics* 160: 357-366

Norris AT, Bradley DG, Cunningham EP (2000) Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* 182: 73-83

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* 4: 249-252

Puurtinen M, Ketola T, Kotiaho JS (2005) Genetic compatibility and sexual selection. *Trends in Ecology and Evolution* 20: 157-158

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 Science USA* 81: 8014-8018

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365-371

Selvamani MJP, Degnan SM, Degnan BM (2001) Microsatellite genotyping of individual abalone larvae: Parentage assignment in aquaculture. *Marine Biotechnology* 3: 478-485

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

Smith JR, Carpten JD, Brownstein MJ, Ghosh S, Magnuson V, Gilbert DA, Trent JM, Collins FS (1995) Approach to genotyping errors caused by nontemplated nucleotide addition by *Taq* DNA polymerase. *Genome Research* 5: 312-317

Sobel E, Papp JC, Lange K (2002) Detection and integration of genotyping errors in statistical genetics. *American Journal of Human Genetics* 70: 496-508

Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* 24: 3189-3194

Taggart JB (2007) FAP: an exclusion-based parental assignment program with enhanced predictive functions. *Molecular Ecology Notes* 7: 412-415

Vandeputte M, Kocour M, Mauger S, Dupont-Nivet M, De Guerry D, Rodina M, Gela D, Vallod D, Chevassus B, Linhart O (2004) Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). *Aquaculture* 235: 223-236

Walsh PS, Ehrlich HA, Higuchi R (1992) Preferential amplification of alleles: mechanisms and solutions. *PCR Methods and Applications* 1: 241-250

Webster MS, Reichart L (2005) Use of microsatellites for parentage and kinship analysis in animals. *Methods in Enzymology* 395: 222-238

Xu J, Turner A, Little J, Bleecker ER, Meyers DA (2002) Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Human Genetics* 111: 573-574

Chapter 4

Conclusions, Implications & Future Prospects

4.1 CONCLUSIONS

The FIASCO method was used for the characterisation of novel species-specific microsatellite markers for *H. midae*. Two vectors, the pDrive cloning vector and the pCR®4-TOPO® vector, were used to perform the cloning steps of this method. The TOPO vector proved to be the more successful of the two vectors, with 66% recovery of repetitive sequences from screened colonies compared with 60% recovery obtained *via* the pDrive vector. This is indicative that the methodology used worked efficiently in both systems in achieving the required results. This technique proved to be prompt and simple for the current project, eliminating unnecessary steps and decreasing costs, and requiring basic molecular skills with the use of limited laboratory equipment. However, the technique did have limitations. On a long term scale, the isolation of molecular markers is not only laborious but also relatively expensive because of the need to screen many polymorphic loci and, therefore, the requirement of fluorescently labelled primers specific to those loci. In comparison to the large amount of colonies screened, relatively few highly informative markers were yielded. Furthermore, the PCR optimisation of designed primers for the newly isolated microsatellite loci proved to be rather time consuming and complicated in some cases.

A total of 54 microsatellites were isolated from 56 positive clones from the partial genomic libraries, while 36 sequences were used to design primers for marker amplification. Perfect, imperfect as well as compound repeat sequences were observed. Tetranucleotides were the most common type of microsatellites identified, constituting a total of 50% of repetitive sequences isolated, with 31% representing simple repeat sequences and 19% expressed in the form of compound repeats. The most frequently occurring tetranucleotide motif was the $(CAGT)_n$ microsatellite. The dinucleotides isolated comprised 25% of the total amount of developed markers, with the $(GT)_n$ pattern representing the most common motif. The remaining 25% is represented by a variety of repeat motifs including tri-, penta- and hexanucleotides in perfect and compound states.

Based on high levels of polymorphism, codominant features and the ability to be successfully amplified and scored, 15 of the 36 microsatellites were selected for labelling. These 15 loci were tested on 32 wild abalone individuals to obtain relevant allele information and statistical fidelity. Allele numbers and frequencies, observed and expected heterozygosities, PIC and null allele frequencies were determined in order to establish microsatellite diversity and determine which markers would be the most appropriate and effective for the assignment of parentage.

The efficiency of parentage assignment was less than expected, with 62.6% for I&J and 61.3% for Roman Bay being unambiguously assigned. Contributions from broodstock pairs showed certain individuals as dominant brooders of fast and slow growing offspring, however there were no specific couples contributing to each growth subset. Few individuals revealed no contributions to assigned individuals and could therefore possibly be considered for elimination from the farm hatcheries. Parentage assignment for Aquafarm was not at all possible using CERVUS as approximately half of the total number of the PRS broodstock could not be sampled due to the incorrect tagging of those animals, which resulted in them being untraceable. Only ten individuals (5%) could be traced back to their parents of origin when FAP was applied.

The ten microsatellite markers chosen for parentage assignment showed sufficient genetic variability in the *Haliotis midae* broodstock at each of the farms. Their high combined exclusionary powers and high number of alleles were assurances of their collective power. Ideally, it would have been preferable to attain a higher assignment rate, but it should be kept in mind that as with any system, microsatellites are not free from shortcoming. There are certain limitations that may impede the reliability of results, including: the presence of null alleles (non-amplifying alleles), mutations, allele stuttering and genotyping errors, which introduce error and can negatively impact on parentage assignment.

The objectives of this study, as stated on page 36, were thus achieved:

1. Species-specific microsatellite markers were successfully isolated and characterised, and submitted to GenBank.
2. Suitable markers were selected, on the basis of their information content, for genotyping broodstock and respective offspring of the growth PRS.
3. Parentage was successfully assigned to offspring of two farms submitted to the growth PRS.

The objective set forward by the performance recording scheme was to track family performance in communally stocked baskets by obtaining microsatellite information about contributing broodstock animals to the overall growth performance of the export production lines of three abalone farms in the Hermanus area of the Western Cape. This was a preliminary study in a sub-component of a greater project; however, the information obtained may be used to the benefit of the farms in order to assist them at a later stage in the long term management of their commercial stocks.

4.2 IMPLICATIONS

Currently, with the aid of a growth performance recording scheme, the growth rate of commercially farmed *Haliotis midae* is being traced *via* phenotypic (i.e. weight and shell length) and genotypic (i.e. microsatellite profiles) characteristics on five abalone farms. Access to the genotypic data of both contributing broodstock and their juvenile offspring will greatly increase the speed at which the genetic improvement of this valuable commodity can be realised. The use of microsatellites to provide genetic profiles for parental and offspring individuals has proven to be a most satisfactory starting point. The contribution of the genotype to variance for the growth rate trait is the first step in growth rate analyses and in its eventual improvement by genetic means. With the knowledge of the microsatellite profiles of all participating individuals, parentage assignments were undertaken in an effort to trace back performance or a lack thereof to certain broodstock animals and establish a protocol by which farm management can allocate parameters to manage, improve and sustain abalone aquaculture.

In South Africa, the abalone species *Haliotis midae* is fast declining due to illegal and commercial farming practices. Over the last decade, exploitation of the species has placed the species in a threatening and vulnerable position. Management of the fishery will eventually include the translocation of farmed adult broodstock to depleted areas, as well as developed seeding programmes. This long term management approach must rely on a complete understanding of the species basic biology, especially the genetic structure of breeding populations used. The success achieved in the production and management of *Haliotis midae* under artificial conditions has resulted in an increased capability to produce and maintain stocks.

For aquaculture geneticists and industry alike, the ultimate aim for any improvement programme is the production of animals with increased profitability to their owner. Profitability comes in the form of enhanced metrical traits which include mature size and weight, growth rate, tolerance to environmental stress, or a combination of all. Costly breeding experiments likely to yield little genetic improvement can be avoided in the future or replaced by programmes with an improved chance of success, if some preliminary genotypic data are available. From a genetic viewpoint, the 'optimal' combination of alleles for survival and growth in culture are unlikely to be permanent in wild populations, and provides the foundation for considerable improvement of valuable traits upon domestication. Selective breeding therefore endeavours to optimise enhancement in commercially desirable traits, while avoiding the negative effects of inbreeding and connect selection in response to undesirable qualities (Lucas *et al.* 2006).

The development of a successful selection programme requires monitoring of genetic and environmental factors, but also taking farm operations and economic implications into consideration. Reducing environmental influences is a principal concern as they can compromise genetic improvement and reduce the efficiency of selection methods. Family based breeding programmes have several advantages over mass selection based programmes as there is better control over inbreeding, potential to improve a wider variety of traits and allow for accurate estimates of the effects of QTL alleles segregating within each family (Hayes *et al*. 2007). Baranski *et al.* (2006) identified QTL affecting growth rate in *Haliotis rubra* using four full sib families of more than 1000 individuals each. The use of microsatellite markers linked to the QTL can provide accurate estimation of breeding values for animals before phenotypic information is available.

Microsatellite DNA markers have a range of uses in selective breeding and broodstock management. Their greatest advantage to genetic enhancement is their ability to screen and assess populations by examining all the genotypes of a specific population. The capability to be analysed by PCR rapidly in larger numbers of samples ensures the strength and capacity microsatellites have to produce data within the time-scales of the abalone farming system. Parentage assignment of commercial stocks is one of the principal applications of microsatellites to any breeding programme. They can identify the parents of superior-performing progeny in collective rearing environments and therefore assist in the selection process. Once genotype information on all candidates is collected, family selection becomes possible and using data on family relationships becomes part of the selection standard. "Walk-back" selection is an example of such a method, whereby fish are selected on size and then on family relationships following microsatellite analyses (Doyle & Herbinger 1995).

When using microsatellites, it is recognised that various errors can occur during the different steps of the genotyping process. Errors may arise due to technical causes, like allele dropout and the presence of null alleles, as well as by human error in data acquirement. Error rates of approximately 2% are common in microsatellite studies, with genotyping errors and the presence of null alleles being the most significant sources of false exclusion and incompatibilities in parentage assignment (Bonin *et al.* 2004; Hoffman & Amos 2005; Castro *et al.* 2006). Nevertheless, even modest error rates can significantly compromise estimations of parentage assignment, especially when true parents are excluded on the basis of a single mismatch.

According to Vandeputte *et al.* (2004), to measure the effectiveness of selective breeding for growth-related traits (weight and length) in aquaculture the heritability of those traits has to be

known, however when it comes to commercial *Haliotis midae* in this study, genomic knowledge is practically non-existent. Microsatellites function as genetic tools to trace heritability patterns among parental and offspring individuals. This was a pilot study to determine if the heritability of growth traits from currently used commercial broodstock to F1 individuals in a growth performance recording scheme could be traced. The data obtained, via microsatellite analysis and parentage assignment, from two of the farms i.e. I&J and Roman Bay, showed no significant trends with regard to certain broodstock contributing to fast and slow growing offspring. Data obtained from Aquafarm could not be used at all as a result of incomplete broodstock information. Unequal parental contributions to the F1 may be accounted for by several causes such as sperm quality, genetic incompatibility of gametes, or differences in survival among families. However, unequal contributions of parents to the next generation are expected to occur in the common seeding operations for abalone where multiple parents are induced to spawn synchronously.

As with most broodstocks in aquaculture, those used at the three locations of interest were caught from wild populations, making them the founders at these locations. It is, therefore, important in maintaining the genetic diversity of future generations. However, without sufficient knowledge of pedigree information, maintenance of genetic differentiation is a difficult task. Accordingly, the importance of parentage assignment is stressed in order to trace progeny to their family of origin and obtain reliable pedigree data for appropriate stock management and most importantly, for successful genetic improvement programmes. The current study was a pilot project in the genetic enhancement of growth related traits *via* broodstock management of *Haliotis midae*. It incorporated a within-family selection method, "walk-back" selection, which exploits the high fecundity of aquatic organisms to achieve intense selection while minimising inbreeding through tracing offspring with the means of DNA markers back to their parental origins (Doyle & Herbinger 1995). Superior individuals can then be selected from families that are grown together without interfering in standard commercial operations.

In the last decade there has been increased interest in applying selective breeding methods for growth rate in aquaculture species. However, for improvement programmes to be successful, it is important that breeders know how non-genetic sources can impact on the expression of growth-related traits. Environmental effects such as stock density and water temperature (physical), as well as feed and anaesthetics (chemical), can significantly influence familial growth traits. Significant genotype-phenotypic effects on growth performance have been observed in Atlantic salmon (Herbinger *et al*. 1999) rainbow trout (Fishback *et al.* 2002), and more recently, shrimp (Jerry *et al.* 2006). These studies suggest that environmental effects could considerably impact breeding

programmes. This is particularly important when progeny are reared under broad physical and chemical environments.

4.3 FUTURE PROSPECTS

Considering various aspects of the FIASCO technique, there may be a few chosen areas which can be improved upon. 1) To increase the variety of microsatellite repeat sequences a multiple repeat-containing probe could be used during hybridisation as suggested by Steven *et al.* (2005). The probe used in this specific study was a mixture of equal parts of eight different repeat motifs. The screen resulted in a wider variety of repeats in comparison to the use of a single type of probe. 2) In an effort to increase the enrichment efficiency, a simple step described by Diniz *et al.* (2007) could be applied in future: The basis of the methodology revolves around a double-enrichment process to recover more repetitive sequences. The technique is much the same as the FIASCO method but makes use of an additional enrichment after the first hybridisation and selective capture, using the same probes. Diniz *et al.* (2007) reported a 100% success rate was achieved using this technique. A high frequency of positive clones was found in the recombinant colonies and the sequences of the libraries produced by the double-enrichment all contained repeat motifs. 3) To avoid over amplification before cloning takes place, care should be taken in the form of the number of PCR cycles performed. In the current study, 30 cycles were used to amplify captured DNA, which may have, according to Hu *et al.* (2007), resulted in the under-representation of a variety of repeat containing fragments and the formation of duplicates within the enriched DNA. They recommend that only 20 cycles be used for a more successful enrichment. 4) To overcome sequencing difficulties posed by the presence of repetitive sequences, dGTP chemistry could be applied to the M13 vector-specific primer during the sequencing reaction (Wen 2001). Short tandem repeats, such as microsatellites, have the potential to cause premature termination during PCR. Premature termination results in the formation of secondary structures, which give rise to base ambiguities and signal dropout during sequencing. Applying dGTP chemistry will allow primers to sequence through repeats and prevent secondary structure formation. Nevertheless, for the most efficient results one should invest time and resources in developing microsatellites that amplify consistently and avoid the temptation to use the first polymorphic loci identified.

With little knowledge of the genome of *Haliotis midae*, it is important to continue developing microsatellite markers, along with other DNA markers, such as AFLPs and SNPs, which together, could assist in gaining further information on the species. Currently, all the microsatellites developed thus far are being applied in the identification of QTL related to growth rate. In addition,

a population and phylogeography study is being conducted. AFLPs have been characterised and, together with the markers developed in the current study, are being implemented in the development of a preliminary linkage map for *Haliotis midae*.

DNA markers provide basic genetic information about broodstock chosen for selection programmes. Obtaining microsatellite-based data at the beginning of a broodstock domestication programme will be an important tool for acquiring knowledge on pedigree structure. Using more variable microsatellite markers accompanied by many unique/rare alleles will further improve the efficiency of parentage assignment, particularly for hatchery-produced strains where a number of broodstock parents are used (Hara & Sekino 2007). Conversely, Jackson *et al.* (2003) suggested using fewer loci with higher PIC values. However, success is dependent on the species of interest.

With the continuous gain in knowledge of the genetic structure of cultured *Haliotis midae*, improvement of this commercially important product can successfully be achieved in the very near future. DNA marker technology and pedigree information play key roles in ensuring that the relevant information is obtained, and will assist as an important tool of genetic enhancement programmes.

4.4 REFERENCES

Baranski M, Loughnan S, Austin C, Robinson N (2006) Linkage maps of microsatellite DNA markers for the blacklip abalone *Haliotis rubra*. *Animal Genetics* 37: 563-570

Bonin A, Bellemain E, Bronken-Eidesen P, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* 13: 3261-3273

Castro J, Pino A, Hermida M, Bouza C, Riaza A, Ferreiro I, Sánchez L, Martínez P (2006) A microsatellite tool for parentage analysis in Senegal sole (*Solea senegalensis*): Genotyping errors, null alleles and conformance to theoretical assumptions. *Aquaculture* 261: 1194-1203

Diniz FM, Iyengar A, Lima PS, Maclean N (2007) Application of a double-enrichment procedure for microsatellite isolation and the use of tailed primers for high throughput genotyping. *Genetics and Molecular Biology* 30: 380-384

Doyle RW, Herbinger CM (1995) Broodstock improvement strategies based on DNA fingerprinting: examples and cost-benefit analyses. *Aquaculture* 137: 271-284

Fishback AG, Danzmann RG, Ferguson MM, Gibson JP (2002) Estimates of genetic parameters and genotype by environment interactions for growth traits in rainbow trout (*Oncorhynchus mykiss*) as inferred using molecular pedigrees. *Aquaculture* 206: 137-150

Hara M, Sekino M (2007) Parentage testing for hatchery-produced abalone *Haliotis discus hannai* based on microsatellite markers: preliminary evaluation of early growth of selected strains in mixed family farming. *Fisheries Science* 73: 831-836

Hayes B, Baranski M, Goddard ME, Robinson N (2007) Optimisation of marker assisted selection for abalone breeding programs. *Aquaculture* 265: 61-69

Herbinger CM, O'Reilly PT, Doyle RW. Wright JM, O'Flynn F (1999) Early growth performance of Atlantic salmon full-sib families reared in single family tanks versus mixed family tanks. *Aquaculture* 173: 105-116

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology* 14: 599-612

Hu J, Zhan A, Lu W, Hu W, Bao Z (2007) An efficient procedure for isolating microsatellite DNAs from sea cucumber (*Apostichopus japonicus*). *Journal of Ocean University of China* 6: 12-15

Jackson TR, Martin-Robichaud DJ, Reith ME (2003) Application of DNA markers to the management of Atlantic halibut (*Hippoglossus hippoglossus*) broodstock. *Aquaculture* 220: 245-259

Jerry DR, Preston NP, Crocos PJ, Keys S, Meadows JRS, Li Y (2006) Application of DNA parentage analyses for determining relative growth rates of *Penaeus japonicus* families reared in commercial ponds. *Aquaculture* 254: 171-181

Lucas T, Macbeth M, Degnan SM, Knibb W, Degnan BM (2006) Heritability estimates for growth in the tropical abalone *Haliotis asinina* using microsatellites to assign parentage. *Aquaculture* 259: 146-152

Steven CR, Hill J, Masters B, Place AR (2005) Genetic markers in blue crab (*Callinectes sapidus*) I: isolation and characterization of microsatellite markers. *Journal of Experimental Biology and Ecology* 319: 3-14

Vandeputte M, Kocour M, Mauger S, Dupont-Nivet M, De Guerry D, Rodina M, Gela D, Vallod D, Chevassus B, Linhart O (2004) Heritability estimates for growth-related traits using microsatellite parentage assignment in juvenile common carp (*Cyprinus carpio* L.). *Aquaculture* 235: 223-236

Wen L (2001) Two-step cycle sequencing improves base ambiguities and signal dropouts in DNA sequencing reactions using energy-transfer-base fluorescent dye terminators. *Molecular Biotechnology* 17: 135-142

Appendices

Appendix A: Summary of microsatellites isolated with relevant information

(* represents the markers whose primers were selected for labelling)

Locus Name	Repeat Sequence	Core Sequence Structure	Accession number	Primer Sequences (5'-3')
HmNR6D	TG_{17}	Perfect	EF063096	[§]F - TAA TCA CAT CGG GTC TGT CT [†]R - TTG TCT GTC ATC GCA CGT
HmNR8D	CA_{13}	Perfect	DQ825708	F - ACC CAT GAC CAA CTG AGC R - TGT TAG ACA GCC TAG AAT CCA
***HmNR20M**	$(TCC)_5(TAC)_7$	Perfect compound	EF063097	F - CTA CAA CAA ACG CCG ATG R - TGC AGT AAT AGG GGT ACC AG
HmNR31T	$(TCAC)_7$	Perfect	EF063098	F - GAA CAT GCT ATG ACA AAG TGA C R - TTC AGT GAG TAT TTG CTC AGT C
HmNR32T	$(CAGT)_{32}$	Perfect	EF063099	F - TGA CAG ACT GAG CAA ATA CTC A R - ATG AAT GAC TTG GAC ACC G
HmNR40T	$(ACTG)_{23}$	Perfect	EF063100	F - AAA AGT CAT ATC TCG GGT TAT G R - AAC AAT ACT CAG TCA GTC AGT CAG
***HmNR48T**	$(CAGT)_{25}$	Perfect	EF063101	F - AAA CTA AGT CAG TCA GTC AGT CAG R - TCA CGT AAT ATC CTG ATG GC
HmNR50M	$(GACT)_7(GAAT)_4$	Perfect compound	EF063102	F - TGA AAT ATG ATG ACC GAT GG R - CAG CAC CCC ACA TTG TCT
***HmNR54H**	$(TTAGGG)_4$	Perfect	EF063103	F - CAT TCT ACA TTC GAC ATT CG R - TAA CAC TAA GTC CCT CAC CC
HmNR56T	$(CAGT)_{23}$	Perfect	EF063104	F - AAC ATC TAG GAC AGC CAG C R - TTT GAC TGA CTG ACT GAC TGA C
HmNR71M	$(TCAG)_7(TCAC)_{16}$	Perfect compound	EF063105	F - ACA TTT CCG GTG TCA TCC R - TTT CCA GTT ACA TTT CCC TG

***HmNR83M**	$(GTCA)_{12}(CTCA)_{15}$	Perfect compound	EF1211742	F - CCA CAT GGG TAC ATT GTG TG R - ATC TGT TTC GTC CTT AGT GCT C
HmNR98T	$(GTCA)_{85}$	Perfect	EF1211743	F - CCA GGC AGT CCG AGT TAG TG R - TTG TTG ACC ATA CGA GAT GAG C
***HmNR106D**	TG_{15}	Perfect	DQ825709	F - TCC TTG GCC AGA ATA ACC R - TAT ATG GTC TGC ATC GCT G
HmNR107T	$(CTCA)_{8}$	Perfect	EF1211744	F - GCA TCG GCT TCA TTT CAG R - TCC TTT CGA CTG TGT CTG C
***HmNR120T**	$(TGAG)_{23}$	Perfect	EF1211745	F - ACC TGC TCT TTA GCT CAG ATG G R - TTG AGC ATG AGT CGT TGA GC
***HmNR136D**	CA_{11}	Perfect	DQ825710	F - GAG TAA TAT GGG CAC CTC G R - GTT TGG AAT GTC TGA TTG GA
HmNR138T	$(CAGT)_{32}$	Perfect	EF121746	F - CAC CAT AGT TCA CCA TAC AGT CAG R - TTA GGA ACA TTG CTG AGA CCT G
HmNR173M	$(GT)_{14}(CAGT)_{28}$	Perfect compound	EF121747	F - AAA TAC GTG TGT GTG TGT GT R - TTC ATG TGT AGG CTA AAG G
***HmNR180D**	$(GT)_{24}$	Perfect	EF121748	F - ACA AGG AGG CGT GAA ATC TGC R - GCA TTG TTA CCC CCT ACA AAG ACC
HmNR182D	$(AC)_{16}$	Perfect	EF121749	F - ATC TCA CCG CCC TTT TCT CT R - TGA TGG ATA GGT GTT GGA TGA G
***HmNR185D**	$(GT)_{13}$	Perfect	EF121750	F - TAG AGT TCA TGT GTG TAC GTG TGC R - TAC CTG TAA CGC GCT TGC T
HmNR189D	$(GT)_{39}$	Perfect	EF121751	F - TAA CTA TTT CCA AAT CCG CAG R - AAG GGA CTA GTC CTG CAG GT
***HmNR191T**	$(GAGT)_{6}$	Perfect	EF121752	F - TTA GTT TTA CGC CGC ACT C R - CCA CAT GGG TAC AAA GTC C
HmNR193D	$(GT)_{13}$	Perfect	EF121753	F - ATC AAG GTA GTT GTA GCC AG R - ATT CGA AGT CAG TTG TCA TG

HmNR200M	$(GAGAGT)_5(GAGT)_{14}$	Perfect compound	EF121754	F - AAC CCA GTT GAG TGA GTG AGT G R - TAT ACA CCC CTT GGA ATC AGC
HmNR202M	$(CAA)_4(CAG)_4$	Perfect compound	EF512266	F - AAC AAC ACC CGT AGC ATC R - AGT TAT CTC CCT TCT GTC AG
HmNR218R	$(CAA)_5$	Perfect	EF512267	F - CTC ATT CGT CAG CAT TTA TCA C R - CTC ACT CAC TCA CTT GTT TGT TG
HmNR220M	$(CA)_{11}(CGCA)_5$	Perfect compound	EF512268	F - TTA TTT TGG TCT GCC GTT CG R - TGT TGT TGT TCT TTG CGT GTG
***HmNR224T**	$(CATA)_{18}$	Perfect	EF512269	F - TGT CCA TAG CAG CCC CTT AC R - ACA TCT TGT TGC CGT TGT TG
HmNR225M	$(GGTT)_4(GGCT)_2\ldots(GGTT)_5$	Imperfect compound	EF512270	F - ATG AGA AAA CCC AGT CGT GAG G R - AGC AAG CCA ACC AAA CAA CC
HmNR228R	$(GTT)_7$	Perfect	EF512271	F - ACT CAC TCA CTC ACT CCA ACG R - GTT TTA CGC CGC TTT TAG C
***HmNR258R**	$(CAA)_{11}$	Perfect	EF512272	F - GCA TCG CCT GAT TTG ATT C R - CAG AAG GGT GGG TTG TAG TAT G
***HmNR268R**	$(CAA)_9$	Perfect	EF512273	F - CAA TCC AGT AAT CAA CAG C R - GAT GTA TGT CGG GTT TTG
***HmNR281P**	$(CTCAA)_{24}$	Perfect	EF512274	F - AAC CTT CAG TAA CCC ATG C R - TGA ATA GGC ACC ATA AAG GG
***HmNR289P**	$(GTTGT)_5$	Perfect	EF512275	F - GCA AGA CAG ACA TCC AAG AC R - TAC AAA TCC CGA CAC AAG AG

§ F represents the forward primer sequence

† R represents the reverse primer sequence

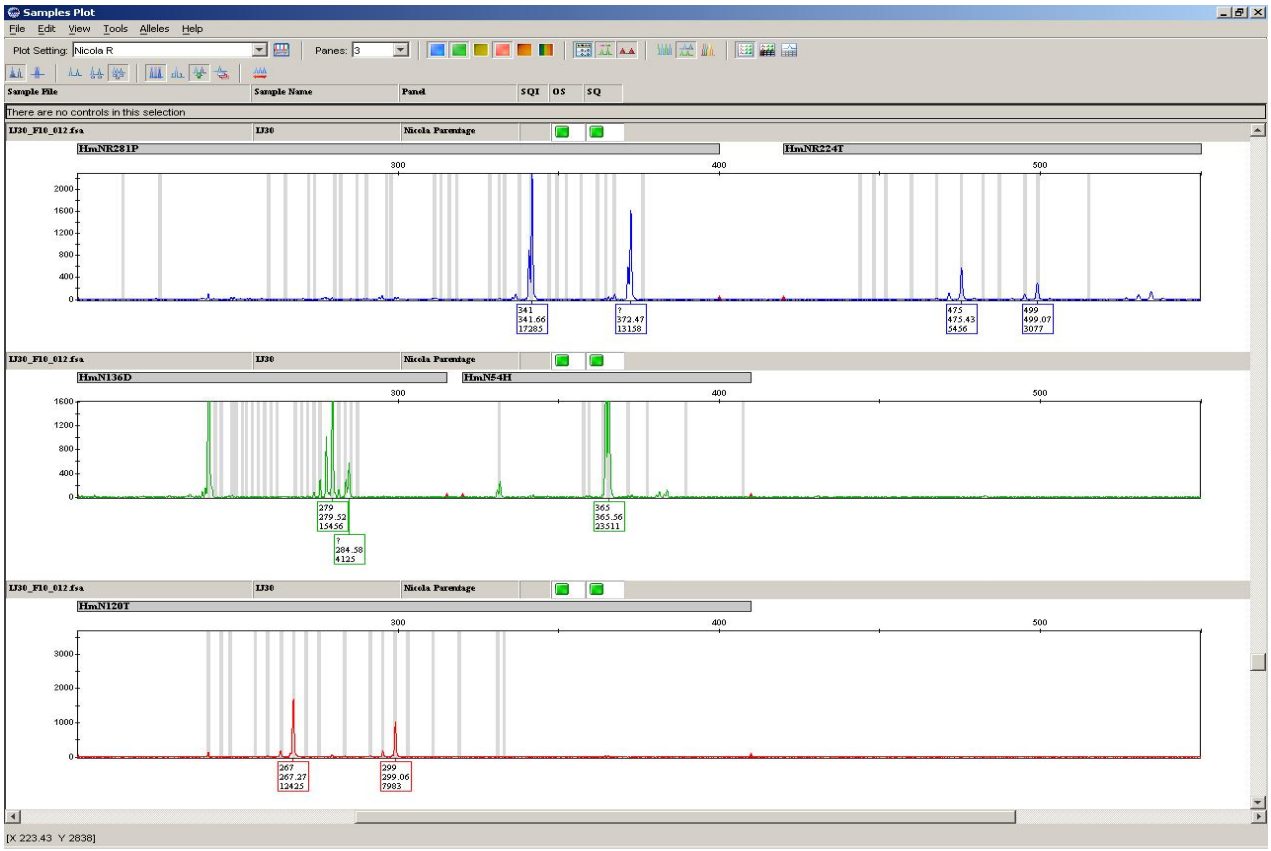
Appendix B: Allele frequency distributions for markers with fluorescently labelled primers

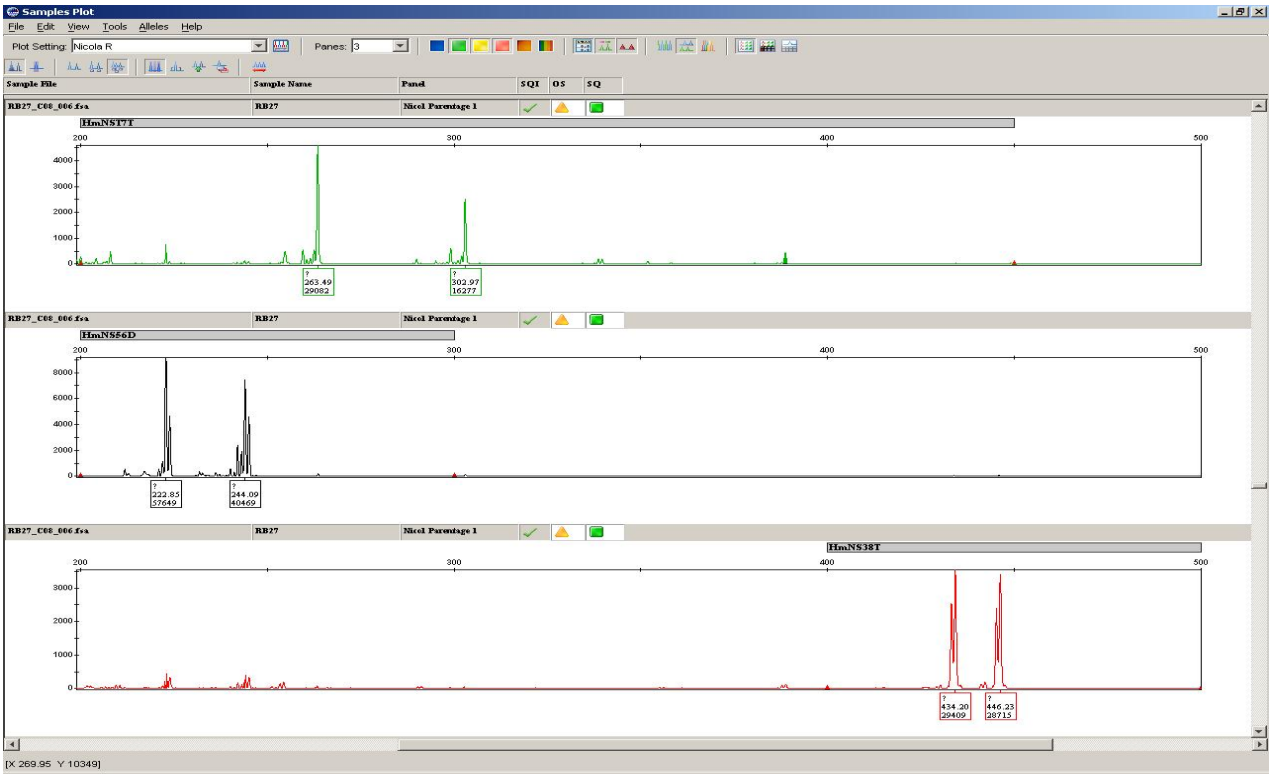
(N = number of individuals tested; coloured backgrounds represent colour of labels)

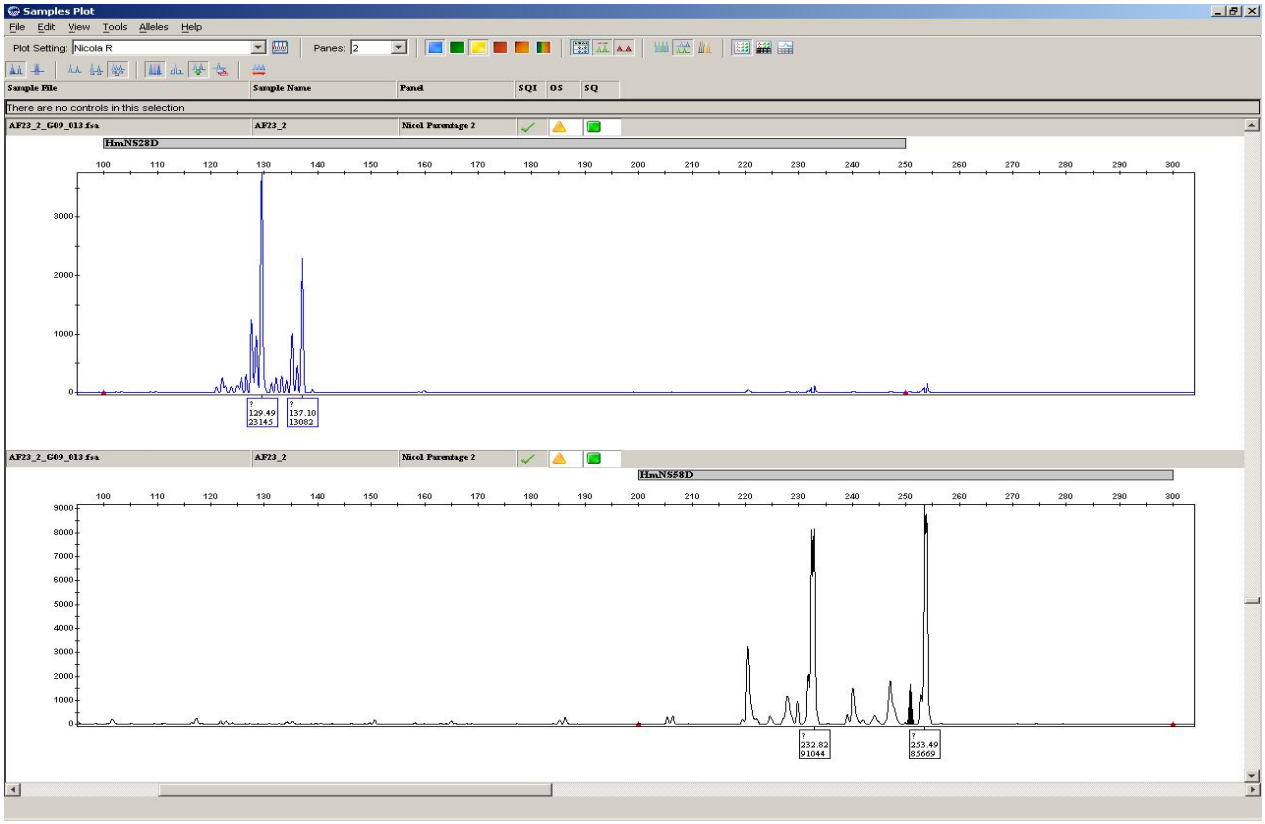
HmNR20M		HmNR54H		HmNR106D		HmNR120T		HmNR136D		HmNR185D		HmNR191T	
(N)	27	(N)	32	(N)	31	(N)	29	(N)	32	(N)	31	(N)	31
1	0.0556	1	0.0156	1	0.0161	1	0.0172	1	0.0156	1	0.1774	1	0.2258
2	0.1296	2	0.0156	2	0.0161	2	0.1552	2	0.0156	2	0.1452	2	0.0161
3	0.0370	3	0.3906	3	0.0161	3	0.0172	3	0.4375	3	0.1935	3	0.0161
4	0.0185	4	0.2188	4	0.2419	4	0.0345	4	0.0156	4	0.1613	4	0.1290
5	0.0741	5	0.2188	5	0.0323	5	0.0172	5	0.0781	5	0.0645	5	0.0645
6	0.2593	6	0.0625	6	0.0161	6	0.0345	6	0.0156	6	0.0645	6	0.2742
7	0.2407	7	0.0156	7	0.0161	7	0.1379	7	0.0469	7	0.0161	7	0.0323
8	0.0370	8	0.0313	8	0.0161	8	0.0345	8	0.0781	8	0.0806	8	0.0645
9	0.1111	9	0.0156	9	0.0161	9	0.0172	9	0.0313	9	0.0484	9	0.0161
10	0.0185	10	0.0156	10	0.1935	10	0.0345	10	0.0313	10	0.0323	10	0.0161
11	0.0185			11	0.0484	11	0.0345	11	0.0469	11	0.0161	11	0.0161
				12	0.1129	12	0.0345	12	0.0156			12	0.0161
				13	0.0645	13	0.0690	13	0.0156			13	0.0323
				14	0.0645	14	0.0517	14	0.0313			14	0.0161
				15	0.0161	15	0.0345	15	0.0313			15	0.0484
				16	0.1129	16	0.0172	16	0.0156			16	0.0161
						17	0.0345	17	0.0313				
						18	0.0345	18	0.0156				
						19	0.0172	19	0.0156				
						20	0.0345	20	0.0156				
						21	0.0862						
						22	0.0172						
						23	0.0172						
						24	0.0172						

HmNR180D		HmNR224T		HmNR258T		HmNR281P		HmNR289P	
(N)	31	(N)	22	(N)	32	(N)	28	(N)	32
1	0.0161	1	0.0227	1	0.1719	1	0.0179	1	0.0625
2	0.0645	2	0.0227	2	0.3906	2	0.0357	2	0.8594
3	0.0645	3	0.0227	3	0.2031	3	0.2321	3	0.0625
4	0.1290	4	0.1136	4	0.1406	4	0.0179	4	0.0156
5	0.1290	5	0.0455	5	0.0625	5	0.0536		
6	0.1129	6	0.0227	6	0.0313	6	0.0536		
7	0.0645	7	0.0455			7	0.0357		
8	0.1613	8	0.0227			8	0.0179		
9	0.0645	9	0.0682			9	0.0179		
10	0.1290	10	0.1136			10	0.0179		
11	0.0484	11	0.0455			11	0.0357		
12	0.0161	12	0.0227			12	0.0714		
		13	0.0227			13	0.0179		
		14	0.1136			14	0.0357		
		15	0.0682			15	0.0179		
		16	0.0455			16	0.0536		
		17	0.0455			17	0.1250		
		18	0.0227			18	0.0714		
		19	0.0227			19	0.0357		
		20	0.0909			20	0.0179		
						21	0.0179		

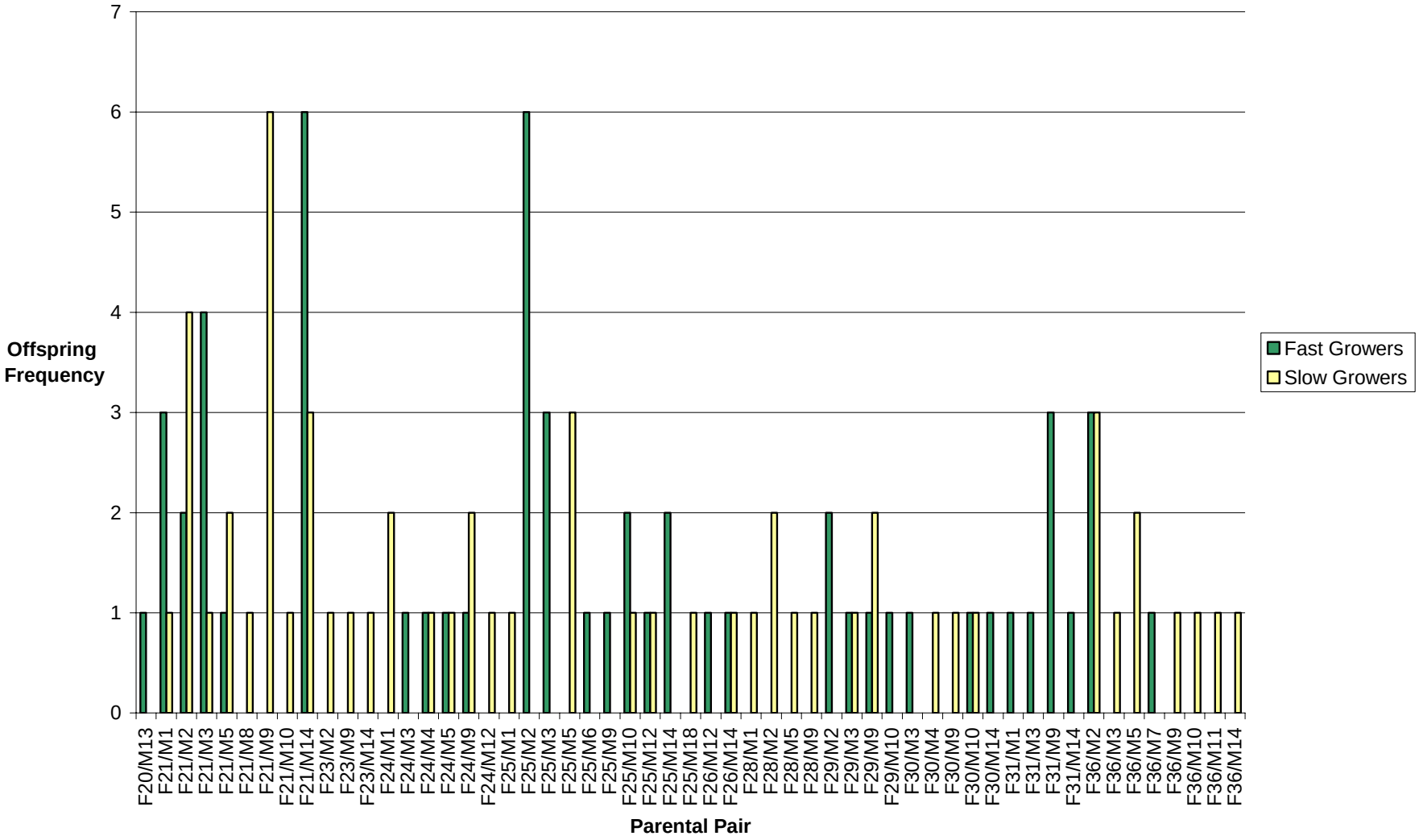
Appendix C: The three panels used for allocating genotypes of 10 microsatellite markers to animals of the PRS







Appendix D: Parental pair contributions to offspring (I&J)



Appendix E: Parental pair contributions to offspring (Roman Bay)