Signatures of Selection in Natural and Cultured Abalone (*Haliotis midae*): A Population Genomics Study

by

Clint Rhode

Dissertation presented for the degree of Doctor of Philosophy in the Faculty of Science at Stellenbosch University

Supervisor: R. Roodt-Wilding, Ph.D.

Co-supervisor: A.E. Bester-van der Merwe, Ph.D.

Department of Genetics

March 2013

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: **12 February 2013**

Abstract

The South African abalone, *Haliotis midae,* commonly known as *perlemoen*, is an economically important gastropod mollusc. Historically, this species maintained a lucrative fisheries sector; however with increasingly lower landings there has now been a shift to aquaculture. Efforts to conserve natural populations and to improve abalone aquaculture production are thus running in parallel. Previous studies reported significant disparities in parental contributions in aquaculture populations that could explain the rapid divergence of commercial stocks from wild populations. Furthermore, subtle, but significant, population differentiation has also been reported for wild populations on the west-, south-, and east coast of the South African coastline. This study therefore aimed to investigate the evolutionary forces, in particularly selection, facilitating population divergence in wild and cultured *H. midae* populations using a population genomics approach. By using both microsatellite- and single nucleotide polymorphism (SNP) markers it was found that approximately 10% to 27% of the *H. midae* genome may be influenced by selection. When incorporating these loci into analyses of population differentiation (*e.g.* AMOVA, factorial correspondence analysis and estimates of genetic distance) there was a marked increase in genetic divergence between wild and cultured populations (especially when using microsatellite loci) and amongst populations from different geographic regions (particularly supported by the SNP loci). The differences in population clustering as highlighted by microsatellite- and SNP markers can most likely be attributed to the genomic distribution of the respective loci: The SNP markers were developed from EST sequences and therefore mostly represents protein structural variation; whereas the microsatellite markers, found to be putatively under selection, were mainly located in regulatory motifs. The results of this study therefore confirmed previous observations of divergence amongst wild- and cultured populations, but more importantly demonstrated that selection is an important factor driving this divergence. In wild populations selection probably facilitates adaptation to local environmental conditions, whilst amongst aquaculture

population adaptation to captivity, husbandry practices and artificial selection may be important determinants. There is evidence for population bottlenecks in wild- and cultured populations; nonetheless long-term effective population sizes seem to be large. Amongst the wild populations, however, short-term population sizes appear to be small most likely due to differential spawning rates amongst reproductively active animals leading to temporal fluctuation in genetic diversity. The results indicate that contact between wild and cultured abalone should be minimised to prevent any adverse effects due to outbreeding depression. With regards to conservation, an emphasis on maintaining adaptive diversity of the wild stocks might be warranted. Continued genetic monitoring is advisable for both wild and cultured abalone populations as to optimally manage the abalone resource for both conservation and commercial viability and sustainability.

Opsomming

Die Suid-Afrikaanse perlemoen, *Haliotis midae,* is 'n ekonomies belangrike buikpotige weekdier. Histories het hierdie spesie 'n winsgewende vissery gehandhaaf, maar met steeds dalende vangste is daar nou 'n verskuiwing na akwakultuur. Pogings om natuurlike populasies te bewaar en perlemoen te verbeter vir verhoogde akwakultuur produksie loop dus in parallel. Vorige studies het bevind dat beduidende verskille in ouerlike bydraes tot die nageslag, in akwakultuur populasies, kan verduidelik hoekom die populasies so vinnig divergeer van die wilde voorouers. Verder, is subtiele, maar betekenisvolle genetiese differensiasie tussen wilde populasies aan die wes-, suid-en ooskus van die land gevind. Hierdie studie is dus daarop gemik om ondersoek in te stel na die mate waartoe verskeie evolusionêre prosesse, in besonder seleksie, die populasie divergensie in beide wilde en gekweekte *H. midae* teweegbring deur gebruik te maak van ‘n populasie genomika benadering. Deur gebruik te maak van beide mikrosatelliet- en enkel nukleotied polimorfisme (ENP) merkers is dit bevind dat ongeveer 10% tot 27% van die *H. midae* genoom moontlik beïnvloed word deur seleksie. Met die gebruik van loki onder seleksie tydens die ontleding van populasie differensiasie (bv. AMOVA, faktoriaal korrespondensie analise en genetiese afstand ramings) was daar 'n merkbare toename in genetiese divergensie tussen wilde- en gekweekte populasies (veral wanneer mikrosatelliet loki gebruik is) en onder die populasies vanuit verskillende geografiese gebiede (veral ondersteun deur die ENP loki). Die verskille in die populasie groeperings soos uitgelig deur die mikrosatelliet- en ENP-merkers kan waarskynlik toegeskryf word aan die genomiese verspreiding van die onderskeie loki: Die ENP-merkers is ontwikkel vanaf uitgedrukte volgorde merker (UVM) volgordes en daarom verteenwoordig dit meestal proteïen strukturele veranderinge, terwyl mikrosatelliet merkers eerder in regulatoriese motiewe geleë is. Die resultate van hierdie studie steun dus vorige waarnemings, maar meer belangrik, het dit getoon dat seleksie ‘n betekenisvolle faktor in populasie divergensie in beide wilde en gekweekte populasies is. In wilde populasies fasiliteer seleksie waarskynlik die

aanpassing tot plaaslike omgewingstoestande terwyl seleksie onder die gekweekte populasies teweeggebring kan word as gevolg van aanpassing tot aanhouding, boerdery praktyke en kunsmatige seleksie. Daar is bewyse vir populasie bottelnekke in wilde- en gekweekte populasies; tog blyk langtermyn effektiewe populasiegroottes om redelik groot te wees. Onder die wilde populasies is egter gevind dat kort-termyn populasiegroottes klein kan wees, waarskynlik as gevolg van differensiële broeikoerse onder reproduktiewe diere. Dit het tot gevolg dat daar beduidende fluktuasies is in temporale genetiese diversiteit. Die resultate dui daarop dat kontak tussen wilde en gekweekte perlemoen tot 'n minimum beperk moet word om enige nadelige effekte weens uitteling depressie te voorkom. Verder, met betrekking tot bewaring, is ‘n klem op die handhawing van aangepaste genetiese diversitiet dalk geregverdig. Voortgesette genetiese monitering word aanbeveel vir beide wilde- en gekweekte perlemoen populasies ter wille van die optimale bestuur van die perlemoen hulpbron vir beide bewaring en kommersiële lewensvatbaarheid en volhoubaarheid.

Acknowledgements

I would like to extend gratitude to the following institutions for financial support (in alphabetical order): the International Foundation for Science; the National Research Foundation of South Africa and Stellenbosch University. The following persons and institutions are also thanked for aiding in the acquisition of biological specimens (in alphabetical order): Mr. S. Ashlin [Roman Bay Sea Farm (Pty) Ltd], Mr. R. Clark [Wild Coast Abalone (Pty) Ltd], Mr. B. Godfrey, Mr. G. Harkins, Mr. G. Isaacs, Mr. P. Pesch [Atlantic Sea Farm, Benguela Holdings (Pty) Ltd], Mr. N. Prins (Aquaculture Division, Stellenbosch University), Mr. M. Viljoen (Aquaculture Division, Stellenbosch University) and Mr. B. Webb (GBAP Marine Services cc). I am also grateful for the assistance of Dr. M. Gründlingh (South African Data Centre for Oceanography) and Ms. G. Linnow (South African Weather Service) in obtaining environmental data and Ms. A. van der Walt (Central Analytical Facility, Stellenbosch University) for assisting with bioinformatics analyses. Lastly, for their support and guidance through the course of my studies I would like to thank my supervisors, Prof. R. Roodt-Wilding and Dr. A.E. Bester-van der Merwe, the members of the Molecular Aquatic Research Group, friends and family.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

%	Percentage
(Pty) Ltd	Property Limited
<	Less than
>	Greater than
ΔSST-SumWin	Average Difference in Sea Surface Temperature between Winter and Summer
®	Registered Trademark
µl	Microlitre
µM	Micromole
2n	Diploid Chromosome Number
3’	Three prime
5’	Five prime
A	Adenine
A_e	Effective Number of Alleles
AFLP	Amplified Fragment Length Polymorphism
AMOVA	Analysis of Molecular Variance
A_n	Number of Alelles
A_R	Alellic Richness
AS	Cultured Population from Atlantic Sea Farm
BF	Bayes Factor
BLAST	Basic Local Alignment Search Tool
bp	Basepair
C	Cytosine
cDNA	complimentary DNA
CI	Confidence Intervals
cM	centiMorgan
CPEC	Cultured Population from the East Coast of South Africa
CPSC	Cultured Population from the South Coast of South Africa

CPWC	Cultured Population from the West Coast of South Africa
CR	Cape Recife Population
CTAB	cetyltrimethylammonium Bromide [$((C_{16}H_{33})N(CH_3)_3Br]$
D'	Linkage Disequilibrium Parameter
dH_20	Distilled Water
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
e.g.	*exempli gratia* (for example)
eQTL	Expressed Quantitative Trait Locus
EST	Expressed Sequence Tag
Evo-Devo	Evolutionary Developmental Biology
F_1	First Generation
F_2	Second Generation
FCA	Factorial Correspondence Analysis
F_{CT}	Derivative of Wright's Fixation Index adapted for hierarchical AMOVA (group of populations relative to the total population)
F_{is}	Wright's Fixation Index (individual relative to the sub-population, equal to the inbreeding coefficient - *f*)
F_{SC}	Derivative of Wright's Fixation Index adapted for hierarchical AMOVA (sub-population relative to the group of populations)
F_{st}	Wright's Fixation Index (subpopulation relative to the total population)
g	Grams
G	Guanine
GB	Gansbaai Population
gDNA	genomic Deoxyribonucleic Acid
GenBank Acc.#	GenBank Accession Number at www.ncbi.nih.nlm.gov
h.	hours
H_e	Expected Heterozygosity
H_o	Observed Heterozygosity
HWE	Hardy-Weinberg Equilibrium
I	Information Index

i.e.	*id est* (that is to say)
Inc.	Incorporated
LD	Linkage Disequilibrium
M	Molar (Moles per Litre)
MAF	Minor Allele Frequency
MAS	Marker Assisted Selection
mg/ml	Milligram per Millilitre
$MgCl_2$	Magnesium Chloride
min	Minutes
ml	Millilitre
mM	Millimole
mo.	months
MOPS	3-(N-morpholino) propanesulfonic acid
MYA	Million Years Ago
n	Haploid Chromosome Number
N/A	Not Applicable
N_e	Effective Population Size
ng	Nanograms
ng/ml	Nanogram per Millilitre
ng/µl	Nanogram per Microlitre
nr	Non-Redundent
ns	Not Significant
^oC	Degrees Celsius
P	Probability value (As a statistically significant limit)
PCR	Polymerase Chain Reaction
pmol	Picomol
PO	Posterior Probability
pp.	Pages
QTL	Quantitative Trait Locus
r	Relatedness
RNA	Ribonucleic Acid
RP	Riet Point Population

SD	Saldanha Bay Population
sec	Seconds
SNP	Single Nucleotide Polymorphism
SSR	Simple Sequence Repeat
SST	Sea Surface Temperature
SST-SumAve	Summer Average Sea Surface Temperature
SST-SumMax	Summer Maximum Sea Surface Temperature
SST-WinAve	Winter Average Sea Surface Temperature
SST-WinMin	Winter Minimum Sea Surface Temperature
STR	Short Tandem Repeat
T	Thymine
T_a	Annealing Temperature
Taq	*Thermus aquaticus* DNA Polymerase
™	Trademark
U	Units (enzyme)
UNEP-WCMC	United Nations Environmental Programme-World Conservation Monitoring Centre
US$	United States Dollar
UTR	Untranslated Region
v.	Version (as in computer software)
v/v	Volume per Volume
vs.	versus
w/v	Weight per Volume
WC	Cultured Population from Wild Coast Abalone
WPEC	Wild Population from the East Coast of South Africa
WPSC	Wild Population from the South Coast of South Africa
WPWC	Wild Population from the West Coast of South Africa
WS	Witsand Population
χ'^2	Standardised Chi-square as a measure of linkage disequilibrium
yrs.	Years
ZAR	South African Rand (monetary currency)

"There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved."

- Charles Darwin, 1859 -

(The Origin of Species by Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life)

CHAPTER 1

Introduction: Literature Review, Aims and Objectives

1.1. Population Genomics: A Framework for an Extended Evolutionary Synthesis?

1.1.1. The Modern Evolutionary Synthesis

The origin and diversity of life is the central theme of the biological sciences and probably the unifying feature that defines the discipline. As such, understanding and dissecting the processes that lead to and shape biological forms have sprouted a rich body of work in aim of explaining this quintessential biological phenomenon (Mitchell-Olds *et al*., 2007). Through these endeavours biology has arguably undergone two major "paradigm shifts" over the last two centuries. The first was Darwin's theory of evolution by means of natural selection, which advocated the principle of gradual modification, by heritable descent, through the actions of natural selection (Darwin, 1859). This theory elegantly explained the history and diversity of life on earth and perhaps more importantly provides a mechanism that can account for the correlation between biological forms and -functionality. The Darwinian theory transformed biology from a mostly teleological enterprise to a "modern" science under a mechanistic conceptual framework with testable hypotheses (Kutschera and Niklas, 2004; Pigliucci, 2007). The second is the modern evolutionary synthesis, which remains the principal theoretical construct governing the biological research agenda since its formulation in the early- to mid-1900's. The principle reason for this shift was the rediscovery of Mendel's work, at the turn of the 20th century that sparked controversy. Mendel's postulates of particulate inheritance (Mendel, 1866) seem to suggest that evolution could proceed in leaps, rather than gradually as proposed by the Darwinian theory (Mayr, 1993, 1996; Pigliucci, 2007; Rose and Oakley, 2001; Delisle, 2010). Through the seminal treatises of Fisher, Haldane, Wright and later on the more empirical works of Dobzhansky, Mayr and others it became possible to seamlessly merge the Mendelian and

Darwinian traditions leading to the development of a mathematical and theoretical framework under the banner of population- and quantitative genetics culminating in the modern evolutionary synthesis (*e.g.* Fisher, 1930; Haldane, 1932; Wright, 1932; Dobzhansky, 1937; Mayr, 1942; Simpson, 1944).

The primary tenets of the modern evolutionary synthesis are as follow: 1) All life originated from one or a few ancestral forms. 2) Populations of organisms are the units of evolution, where a population is defined as a group of individuals that can readily interbreed (Waples and Gaggiotti, 2006). 3) The evolutionary transitions of populations are gradual and can be explained by slight incremental genetic changes that are then predominantly "categorised/organised" by natural selection. 4) The occurrence of new genetic variation, and consequently phenotypic variation, are due only to chance events. The sources of this genetic variation are: gene flow between populations, recombination during sexual reproduction and ultimately mutation (Mayr, 1996; Kinoon, 2009a). 5) Speciation is thus the point in the gradual evolutionary processes where individuals from separated populations can no longer exchange genetic material, *i.e.* become reproductively isolated. Lastly, 6) Macro-evolutionary processes (the phylogenetic developments of higher taxa beyond species level) can be explained by the known genetic mechanisms of micro-evolutionary processes (such as gene flow, mutation, selection and random drift) over geological timescales (Kutschera and Niklas, 2004) (Figure 1.1).

Although the modern synthesis acknowledges the role of random drift in the evolution of organisms (particularly in small populations), at its core it is adaptationist in nature. As such, natural selection is the predominant mechanism explaining observable diversity (Kinoon, 2009a). In essence, natural selection is a two-step process as can be deduced from point 3 and 4 of the primary tenets: Phenotypic variation (indirectly genetic variation) must first arise. This is then followed by "sifting" the variation in terms of relative fitness, *i.e.* the actual selection. Broadly this "sifting" can be defined as the "non-random elimination" of individuals (or genotypes) that are not as suitably adapted to a particular environment relative to other individuals (genotypes) in that population. As such, the premise is that

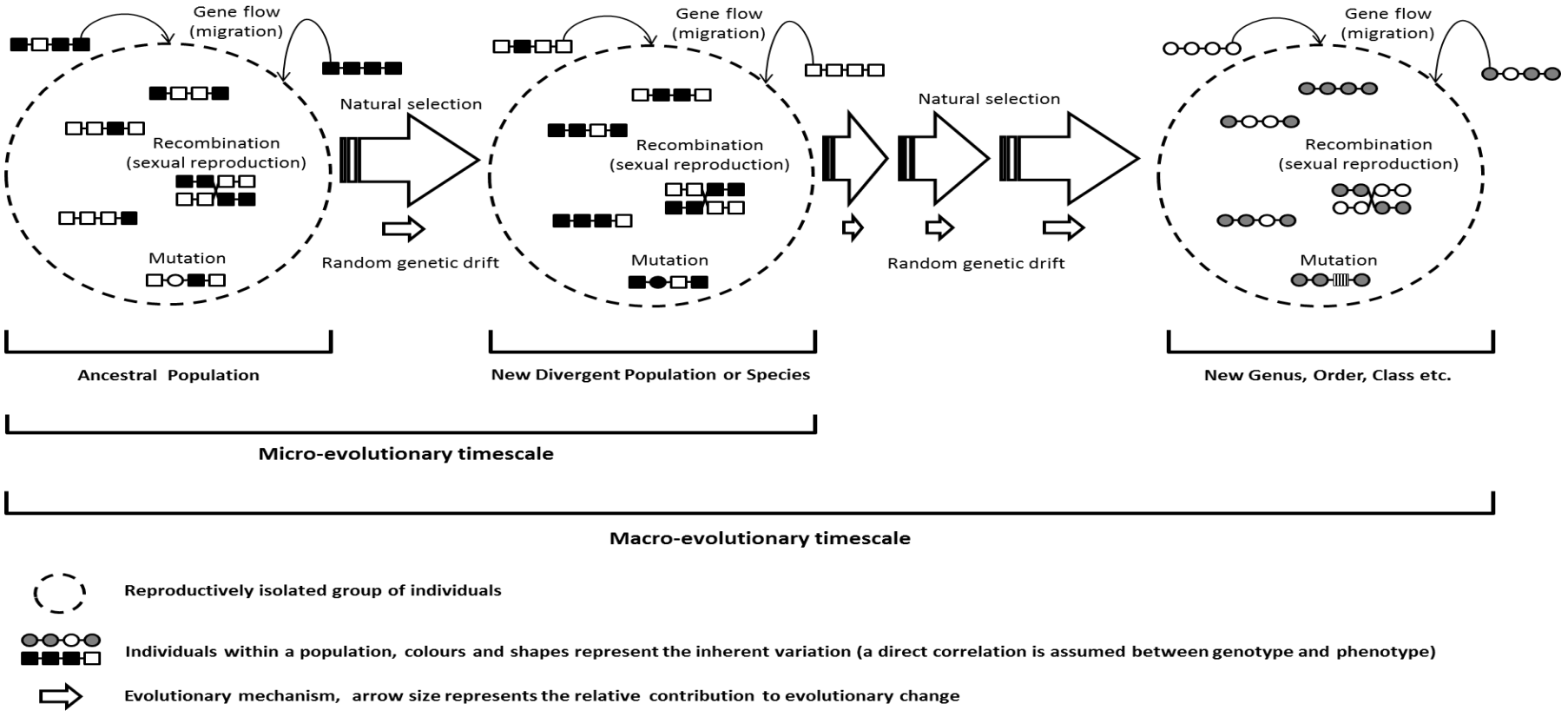

Figure 1.1: A schematic representation of the tenets of the modern evolutionary synthesis. The figure was adapted and modified from Kutschera and Niklas (2004).

once genetic variation occurs it becomes subject to predominantly natural selection (Stebbins and Ayala, 1981; Mayr, 1996; Kinoon, 2009a).

This simple paradigm works well if there is a strict linear relationship between genotype, phenotype and environment. However, over the last few decades the biosciences have seen a rapid development in advanced technology that has led to a deeper understanding of especially cellular and molecular processes of life. Foremost of these advances was the molecular revolution that consequently led to the genomics era. With the complete genome sequences of many organisms (*e.g.* Adams *et al*., 2000; Venter *et al*., 2001; Mouse genome sequencing consortium, 2002; Maglich et al., 2003), biology is firmly in the post-genomics age. The modern evolutionary synthesis, constructed before this revolution, could not anticipate the explosion of knowledge brought about by these developments and in particular the novel complexities of life now discovered; such as "selfish" genetic elements, genomic regions with no apparent function ("junk" DNA), phenotypic plasticity, epigenetic inheritance, epistatis and pleiotropic effects of genes (Pigliucci, 2007; Rose and Oakley, 2007; Koonin, 2009c). Furthermore, some argue that the modern evolutionary synthesis does not incorporate explicit formulations for developmental biology or ecology amongst its central theorems, although it is intuitively implied (*e.g*. Müller, 2007; Carroll, 2008; Schoener, 2011). For this reason it has been argued that the premises of the modern evolutionary synthesis are outdated and can no longer function as the fundamental paradigm for biological research in its current form and therefore a revision and reformulation as an extended- or post-modern evolutionary synthesis is needed (Pigliucci, 2007; Rose and Oakley, 2007; Koonin, 2009a, b).

Population genomics is a relatively new discipline that combines attributes of population genetics and functional genomics (Bonin, 2008). As such, it provides an unique interface between functional-/systems biology and evolutionary theory. A genomics perspective on population genetics may provide an ideal framework for an extended evolutionary synthesis in the post-genomics age.

1.1.2. Population Genomics as an Interface between Old Concepts and New Technologies

The rapid development of sequencing technology, so called, next-generation sequencing and high-throughput genotyping platforms has made it possible to obtain genome sequences and quantify genome-wide genetic variation faster and at lower cost than ever before (*e.g.* Mardis, 2008; Morozova and Marra, 2008; Ansorge, 2009; Harismendy *et al*., 2009; Stapley *et al*., 2010; Wheat, 2010). This wealth of genetic diversity data lend itself to analyses across various genomic regions both within and between populations or species. In its broadest sense population genomics can thus be defined as the simultaneous population genetic analysis of a large number of variable loci spanning across the genome in order to gain understanding of the various evolutionary processes, including mutation, random genetic drift, gene flow and selection (Black *et al*., 2001; Luikart *et al*., 2003; Beaumont and Balding, 2004; Stinchcombe and Hoekstra, 2008). This genome-wide analysis allows for the identification of aberrant (locus-specific) patterns of genic variation from "regular/generalised" genomic variation. Herein lies the major advantage of population genomics: demographic processes such as bottlenecks, population expansions, gene flow and random drift are expected to affect genetic variation throughout the genome in a similar manner. Patterns of genetic diversity caused by mutation, selection or recombination are expected to stand-out from genomic-background variation and would therefore be detectable within this context (Black *et al*., 2001; Luikart *et al*., 2003). For this reason, population genomics studies have readily been conducted in order to detect outlier-loci that may indicate genomic regions under selection (Table 1.1).

Luikart *et al*. (2003) describes the population genomics approach as a four phase enterprise (Figure 1.2): Firstly, including/sampling as many individuals as possible over a large geographic range without making *a priori* assumptions on population structure. This avoids any bias that might be introduced by subjective sampling (Long and Langley, 1999; Pritchard *et al*., 2000; Wall and Pritchard, 2003). The second phase involves genotyping a sufficient number of marker-loci that provides good genome coverage. Typically, tens to hundreds of molecular markers should suffice, however this is highly dependent on the degree of linkage disequilibrium (LD) maintained within a population. Given an outbred, panmictic population LD might decay quite rapidly and therefore many more markers will be required (Goldstein and Weale, 2001). The number and distribution of markers allows for

the estimation of baseline levels of genetic diversity (*i.e.* putatively neutal variation) to which loci with unusual patterns of genetic diversity could be compared. Many molecular marker systems have been developed over the years (*e.g.* Vignal *et al*., 2002; Brumfield *et al*., 2003; Schlötterer, 2004; Seddon *et al*., 2005; Chistiakov *et al*., 2006).

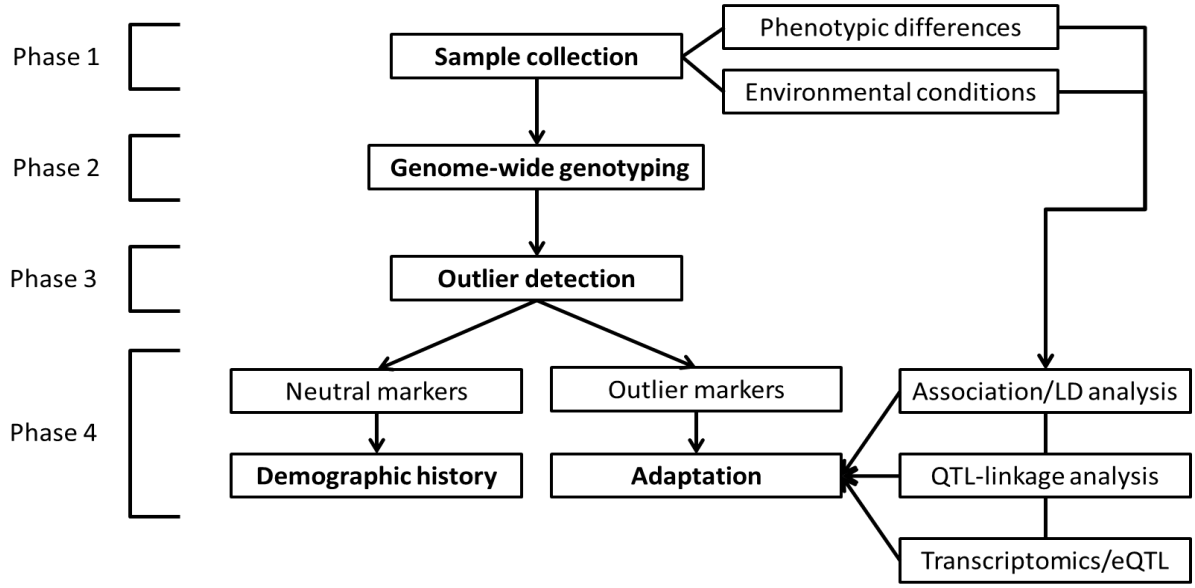

Figure 1.2: Flow diagram of the population genomics approach. Adapted and modified from Luikart *et al*. (2003).

Amplified fragment length polymorphisms (AFLPs) are typically used in non-model organisms, because it does not rely on *a priori* sequence information and provides an almost instantaneous genome-wide assessment (Table 1.1). Recently, however, microsatellite markers [also known as short tandem repeats (STRs) or short sequence repeats (SSRs)] and single nucleotide polymorphisms (SNPs) have gained popularity mostly for their co-dominant mode of inheritance, ease of genotyping by means of automated/semi-automated genotyping platforms and frequency within the genome. The use of EST-derived microsatellites and -SNPs are also particularly advocated because of the direct link to coding regions of the genome (Bonin, 2008). The third phase is testing for outlier loci. This is the pivotal step in the population genomics approach and various methods have been developed, including those that make use of within population diversity, population

divergence and extensive modelling of population history (Luikart *et al*., 2003; Nielsen, 2005; Storz, 2005; Biswas and Akey, 2006; Pavlidis *et al*., 2008 for reviews). The most popular to date, however, is the F_{st}-outlier procedures first developed by Beaumont and Nichols (1996) and subsequently re-implemented under both frequentist and Bayesian statistical frameworks (*e.g.* Antao *et al*., 2008; Foll and Gaggiotti, 2008; Excoffier *et al*., 2009). In essence, F_{st}-outlier tests exploit the genetic differentiation between populations to detect loci that have unusually high or low F_{st} values amongst the genome-wide loci sampled, due to heterogeneous genomic divergence (Nosil *et al*., 2009). It does this by first computing a null distribution for F_{st} values under the assumption of neutrality and then compares it to the empirical data. A locus is assumed under directional (positive) selection if it has an F_{st} value higher than expected and under balancing selection if it has an F_{st} value lower than expected under the precedence of neutrality. The fourth and final phase entails using the acquired data to better understand the evolutionary processes affecting the study populations.

Table 1.1: A summary of population genomics studies, using a variety of marker types and demonstrating relatively large numbers of candidate loci under selection thoughout the genome.

Organism	% of loci under selection	Marker type	Reference
Atlantic oyster (*Crassostrea virginica*)	1.9	AFLP	Murray and Hare, 2006
Hawk moth-pollinated violet (*Viola cazorlensis*)	2.6	AFLP	Herrera and Bazaga, 2008
Mosquito (*Aedes rusticus*)	9.0	AFLP	Paris *et al*., 2010
Mud minnow (*Fundulus heteroclitus*)	6.0	AFLP	William and Oleksiak, 2008
Lake whitefish (*Coregonus clupeaformis*)	3.2	AFLP	Campbell and Bernatchez, 2004
Common frog (*Rana temporaria*)	4.9	AFLP	Bonin *et al*., 2006
Ocellated lizard (*Lacerta lepida*)	4.1	AFLP	Nunes *et al*., 2011

Scandinavian wolf (*Canis lupus*)	6.0	Genomic microsatellite	Hagenblad *et al*., 2009
Cattle (*Bos taurus*)	9.5	Genomic microsatellite	Medugorac *et al*., 2009
Guppy (*Poecilia reticulata*)	5.5	Genomic SNP	Willing *et al*., 2010
Atlantic salmon (*Salmo sala*)	26	Genomic- and EST microsatellite	Vasemägi *et al*., 2005
Periwinkle snail (*Littorina saxatillis*)	7	EST SNP	Galindo *et al*., 2010
Zebrafish (*Danio rerio*)	8.5	EST SNP	Whiteley *et al*., 2011
Atlantic cod (*Gadus morhua*)	10.2	EST SNP	Nielsen *et al*., 2009a
White spruce (*Picea glauca*)	5.5	EST SNP	Namroud *et al*., 2008

From the onset using genome-wide neutral loci will increase both the precision and accuracy of population genetic parameter estimates, such as effective population size, population declines and -expansions, gene flow and population structure (Luikart *et al*., 2003). Wilding *et al*. (2001) investigated population divergence in *Littorina saxatilis* inhabiting various environmental niches. They constructed phylogenies based on outlier- and neutral loci as identified by a population genomics analysis. Phylogenies inferred using outlier loci grouped populations with similar shell morphologies, even though these populations inhabited distinct and separated geographic localities. On the contrary, phylogenies inferred using only neutral loci reflected the geographic orientation of populations and thus more accurately represented the demographic history of the populations. Groupings on the grounds of shell morphology, therefore reflect convergent evolution as a result of adaptation to similar environmental conditions. This illustrates the importance of removing outlier loci when estimating population demographic processes. As such, population genomics allows for a joint analysis of population demographic history and selection (Li *et al*., 2012).

Outlier loci may be detected, as such, for two reasons: firstly, due to statistical biases that arise by chance (*i.e.* type 1 errors; Akey *et al*., 2004; Beaumont, 2005) or due to a true biological effect (Bonin *et al*., 2006). In most cases this biological effect is ascribed to

selection (*e.g.* Table 1.1), however in practice it is often difficult to ascertain whether selection is truly the causative agent. A new mutation, a stochastic event (*i.e.* random drift) or even recombination could cause a genetic pattern similar to a signature of selection (Stinchcombe and Hoekstra, 2008). Various other genomic techniques have been developed to find genotype-phenotype correlations; most notable are the linkage-based QTL analysis and association/LD mapping studies (*e.g.* Goddard and Hayes, 2009; Massault *et al*., 2009; Hayes and Goddard, 2010; Piertney and Webster, 2010; Coluccio *et al*., 2011). The association of an outlier locus with a particular phenotype provides additional evidence for selection. Association/LD studies have become particularly popular since the development of genomic technologies with its reimplementation as genome-wide association studies using SNP-chips consisting of thousands of markers (Wang *et al*., 2005; McCarthy *et al*., 2008). Recently, association studies have also been developed to find correlations between genotype and environmental data (*e.g.* Joost *et al*., 2007; Coop *et al*., 2010). Furthermore, the association of particular genotypes are done without the necessity of breeding experiments or extensive pedigree information as is the case for conventional linkage analysis. Such dualistic approaches have been implemented in a number of organisms, including the ocellated lizard (*Lacerta lepida*; Nunes *et al*., 2011) and the Atlantic herring (*Clupea harengus*; Limborg *et al*., 2012) where population genomics data was combined with environmental associations. And in pearl millet (*Pennisetum glaucum*; Mariac *et al*., 2011) population genomics and phenotypic association tests were both used in order to find genotypes underlying adaptation.

Given the accompanying difficulties with the construction of linkage maps for organisms not easily kept and bred in captivity; there is an advantage in combining conventional linkage analysis with population genomics, especially for non-model organisms. Linkage mapping can provide an indication of the genomic position for loci of interest (Stinchcombe and Hoekstra, 2008). A striking example is the study by Rogers and Bernatchez (2005). Different ecotypes of whitefish (*Coregonus clupeaformis*) that demonstrated size dimorphism (dwarf and normal) were investigated. The authors showed that loci under divergent selection mapped to the same genomic region as known QTLs for growth rate. The addition of gene expression data may add an extra dimension to this framework. Integrating gene expression- with QTL analyses have become common with many “eQTL”

studies published over the last years using micro-arrays (West *et al*., 2007; Gilad *et al*., 2008). More recently, with the advent of massive parallel sequencing technologies comparisons of population-specific gene expression profiles (population transcriptomics) became possible (*e.g.* Giger *et al*., 2008; Normandeau *et al.,* 2009; Roelofs *et al*., 2009; Chelaifa *et al*., 2010; O'Niel *et al*., 2010). The simultaneous analysis of genomic and gene expression variation allows for the investigation of the extent of the influence of evolutionary/population genetic mechanisms on gene and genomic functionality (Khaitovich *et al*. 2004; Holloway *et al*., 2007). As such, a more comprehensive understanding of the genome as a dynamic system with gene pathways and regulatory networks could be attained. For example, loci under selection, across various chromosomes (or linkage groups), might be in strong linkage disequilibrium. This could be an indication of functional linkages that could be elucidated with gene expression profiling and the effects of this on phenotypes could be further tested by genotype-phenotype association studies.

1.1.3. Biological Anomalies and Complexity Explained Under a Population Genomics Framework

One of the primary findings of the molecular revolution was the abundance of molecular variation that seemingly contradicted the selectionist view of the modern synthesis. From this, arguably, the most important conceptual construct since the modern synthesis was formulated - the neutral theory (Kimura, 1968, 1983; King and Jukes, 1969) and more recently the nearly neutral theory (Ohta, 1992; Ohta and Gillespie, 1996). It states that the majority of molecular variation is selectively neutral (or nearly neutral) and thus will become fixed during the evolutionary process due to stochastic events, *i.e.* random genetic drift. As such, genomic "anomalies" including transposable elements, gene- and genome duplications and "junk" DNA could be explained by neutral processes (Kinoon, 2009c). Therefore, a population genomics view will intuitively assume a pluralistic mechanism for evolution where selection and neutral events are not mutually exclusive, but act in unison to explain a variety of biological observations. More interestingly population genomics may provide an opportunity to observe the interplay between neutrality and selection: As noted by Kimura (1991) and later Wagner (2005) molecular variation at any point in time may

come under selection due to changing environmental pressures (*e.g.* Hansen *et al*., 2012). Experimentally this could be tested by searching for signatures of selection among populations separated by time rather than space. This might be further supported by finds that non-coding DNA ("junk" DNA) may have functionally important roles in the genome, especially with regards to chromatin structure formations and maintenance (Glazko *et al*., 2003; Linnemann *et al*., 2009). Studies on *Drosophila* spp. suggest that up to 70% of non-coding DNA might be under selection (Andolfatto, 2005; Halligan and Keightley, 2006; Haddrill *et al*., 2008). Recent population genomics studies, also conclude that selection might be more abundant thoughout the genome than what was previously believed under a strict (nearly) neutral model. Many of these studies using genome-wide anonymous DNA markers (such as AFLPs) report that 1.9% to 9.5% of loci might be affected by selection across various species (Table 1.1). Provided that only approximately 1.5% of a genome represents actual protein coding DNA (Mouse Genome Sequencing Consortium, 2002) many of these loci must be located outside of coding regions.

Population genomics may elucidate the underlying biological complexities that arise from non-linear interaction of genes with one another (gene-by-gene) and the environment (gene-by-environment) that is seemingly not accounted for by the modern synthesis. As discussed earlier the relationship between loci under selection, transcriptional-, phenotypic- and environmental associations will highlight interdependent gene-networks. This could explain phenomena such as evolutionary capacitance (the release of cryptic variation under new environmental conditions/stressors) (Bergman and Siegal, 2003; Masel, 2005), phenotypic plasticity (multiple phenotypic forms of the same genotype under different environmental conditions) (Pigliucci *et al*., 2006; Valladares *et al*., 2006; Pigliucci, 2008; Lande, 2009) and epigenetics (alterations in gene function that cannot be explained by DNA sequence modification) (Bossdorf *et al*., 2008).

Gene expression analysis of Coregonine fish (*Coregonus* spp.), adapted to niche environments showed divergent gene expression patterns in different ecotypes of the same species, but comparable expression patterns in different species inhabiting similar niches (Derome and Bernatchez, 2006; Derome *et al*., 2006). In a similar study Cheviron *et al*. (2008) investigated transcriptomic profiles of rufous-collared sparrows (*Zonotrichia capensis*) at high and low altitudes and concluded that gene expression was highly plastic

and environmentally dependent: expression differentials seem to diminish when birds shared a common environment. From these observations it could be asked to what extent the genes identified are regulated by other genes that might be under selection? It is known that mechanisms underlying such plasticity are at least in part heritable (Schlichting and Smith, 2002; Li *et al*., 2006). A population genomics approach could provide some answers.

One of the foremost areas of research and a presumed pillar of an extended evolutionary synthesis, evolutionary-developmental biology or "evo-devo" has made particular headway in dissecting the regulatory networks that lead to the development of phenotypes (Müller, 2007; Carroll, 2008). These investigations are generally preceded by the evaluation of single or groups of candidate genes - candidate gene approach - most notably the homeotic (*Hox*) genes. The *Hox* genes code for a family of regulatory proteins (transcription factors) responsible for embryonic development that seem to be conserved in animal lineages from arthropods to mammals (Schierwater and DeSalle, 2001; Arthur, 2002; Gilbert, 2003). Although the "evo-devo" research programme has made substantial advances in the construction of genotype-phenotype maps, it is only now during the genomics era that the full extent of such regulatory pathways can be examined.

A population genomics framework may also facilitate more explicit formulation of ecological dynamics on organismal evolution. The modern synthesis is often criticised by marginalising ecological effects (Matthews *et al*., 2011; Shoener, 2011). The recent developments in landscape/seascape genetics, where geographic- and habitat-specific variables are incorporated into the analyses, have become particularly popular in order to identify specific trends in genetic variation among populations within and between species. (*e.g.* Manel *et al*., 2003; Selkoe *et al*., 2008; Storfer *et al*., 2010). In particular, there is renewed interest in understanding the genetic architecture of adaptation of organisms to certain environments (Table 1.1; Orr, 2005; Nadeau and Jiggins, 2010; Stapley *et al*., 2010). The "reverse ecology" approach of Li *et al*. (2008) may prove particularly useful; whereby adaptive genotypes are first identified by means of a population genomics scan. By performing additional association tests and functional assays, genotypes responsible for particular phenotypes and ecological functionalities could infered. Such analyses are not restricted to abiotic conditions, but the co-evolutionary mechanisms that arise from interaction between species can also be investigated. Egan *et al*. (2008) used a population

genomics scan to find genomic regions associated to host-specific adaptations in the leaf beetle (*Neochlamisus bebbianae*) and similar studies were done to investigate the co-evolution of lions (*Panthera leo*) and feline immunodeficience virus (FIV_{Ple}) (Antunes *et al*., 2008) and host-pathogen interactions between plants and fungi (Aguileta *et al*., 2010). It is comprehendible that such studies could be extrapolated to explain other community-ecological interactions.

Rose and Oakley's (2007) "new biology" is undoubtedly genomics centred. The modern evolutionary synthesis most likely fails to account for the diverse biological occurrences, because it is a reductionist construct that aims to explain all in terms of fluctuating gene frequencies; assuming linear relationships between genotype, phenotype and environment. As such, it cannot account for dynamic interactions. It must, however, be noted that for the most part these new observations are not necessarily incompatible with the tenets of modern synthesis. Therefore, an expansion of the modern synthesis would be more appropriate than an abandonment of its central tenets. Population genomics may not be able to explain all observations, but it does at least provide a partial framework for integrating genetic, phenotypic and ecological phenomena under a complex dynamic systems view of biology (Ge *et al*., 2003; Pigliucci, 2007, 2008; Badyaev, 2011; Weber, 2011).

1.2. A Review of Abalone Genetics and Genomics with reference to *Haliotis midae*: Perspectives on Biology, Aquaculture and Conservation

1.2.1. The Haliotids

Abalone (*Haliotidae*) are marine gastropod molluscs with approximately 56 extant species world-wide, distributed along the tropical and temperate waters off the coasts of all continents with the exception of Antarctica (Geiger, 2000). Among the gastropods abalone are distinct; characterised by a single, depressed shell that spirals clock-wise. On the left outer periphery of the shell a row of seven to twelve tremata or "respiratory pores" are generally observable; whilst the inner shell is layered with nacre. When the animal is in a relaxed state the well-defined, hypertrophied epipoduim is commonly seen under the shell (Geiger, 1999).

Section 1.2 is in preparation for submission as a review paper.

No single species has a global distribution, but four geographic regions of endemism persist: the North Pacific, Southern Africa, Australia and New Zealand (Lee and Vacquier, 1995; Geiger, 2000; Estes *et al.,* 2005). In South Africa five endemic species can be found [*Haliotis midae* (a), *H. spadicea* (b), *H. alfredensis* (c), *H. parva* (d) and *H. queketti* (e), Figure 1.3] of which *H. midae* is the most studied and the only economically valuable species. Commonly known as *perlemoen, H. midae* is the largest growing of all the endemics, with a relatively extensive distribution range across the temperate seaboard of South Africa, stretching from the Western Cape- to the Eastern Cape Province. *Haliotis midae* is an intertidal species prefering the rocky, kelp bed habitats up to 10m offshore (Tarr, 1989; Lindberg, 1992).

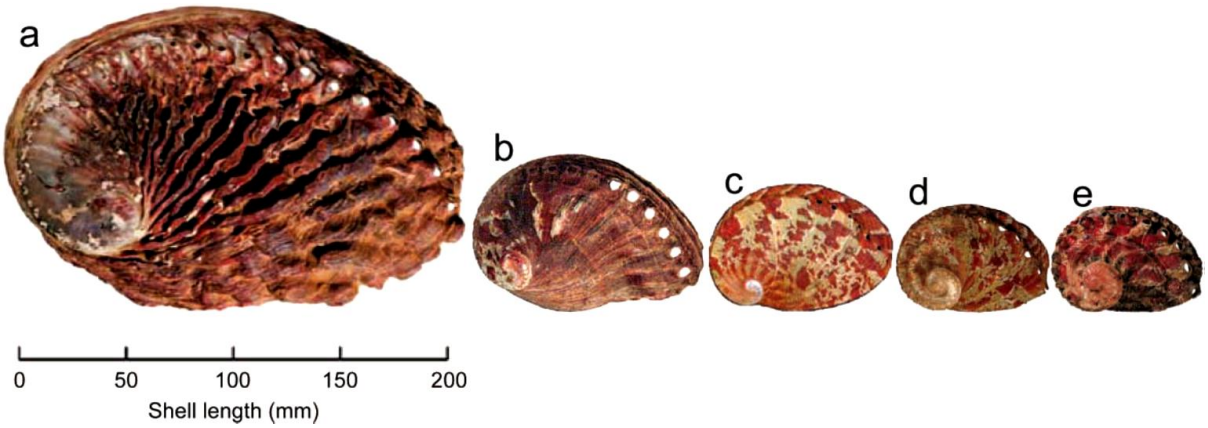

Figure 1.3: Five endemic abalone species of South Africa: *Haliotis midae* (a), *H. spadicea* (b), *H. alfredensis* (c), *H. parva* (d) and *H. queketti* (e). Figure taken from Bester-van der Merwe *et al*. (2012).

The adult abalone is for the most part a benthic, sessile animal that will rarely move once it has established a "home-site" (Tarr, 1995). Abalone are dioecious and broadcast spawning; as such, coordinated mass release of egg and sperm is generally seasonal (normally spring and/or autumn) and water temperature dependent (Tarr, 1989). Like many marine invertebrates, the abalone life cycle is complex. Larval development takes place through various phases, during which the larvae will undergo the gastropod indicative process of torsion. Abalone larvae are lecithotrophic and therefore the pelagic larval stages are short, approximately five to ten days. Larval settlement is poorly understood, but some

environmental factors such as substrate topology, water temperature and pheromonic actions of gamma-aminobutyric acid (GABA), secreted by the diatom filaments growing on the substrate, have been postulated to play a role. Once settled, the spat will continue to metamorphose into the adult specimen. In the wild, *H. midae* juveniles mature at about seven to ten years of age, during which time they progressively wean from grazing micro-algae to grazing macro-algae, such as sea weeds and kelp (Figure 1.4) (Tarr, 1989; Barkai and Griffiths, 1986; McShane, 1992; Day and Branch, 2000).

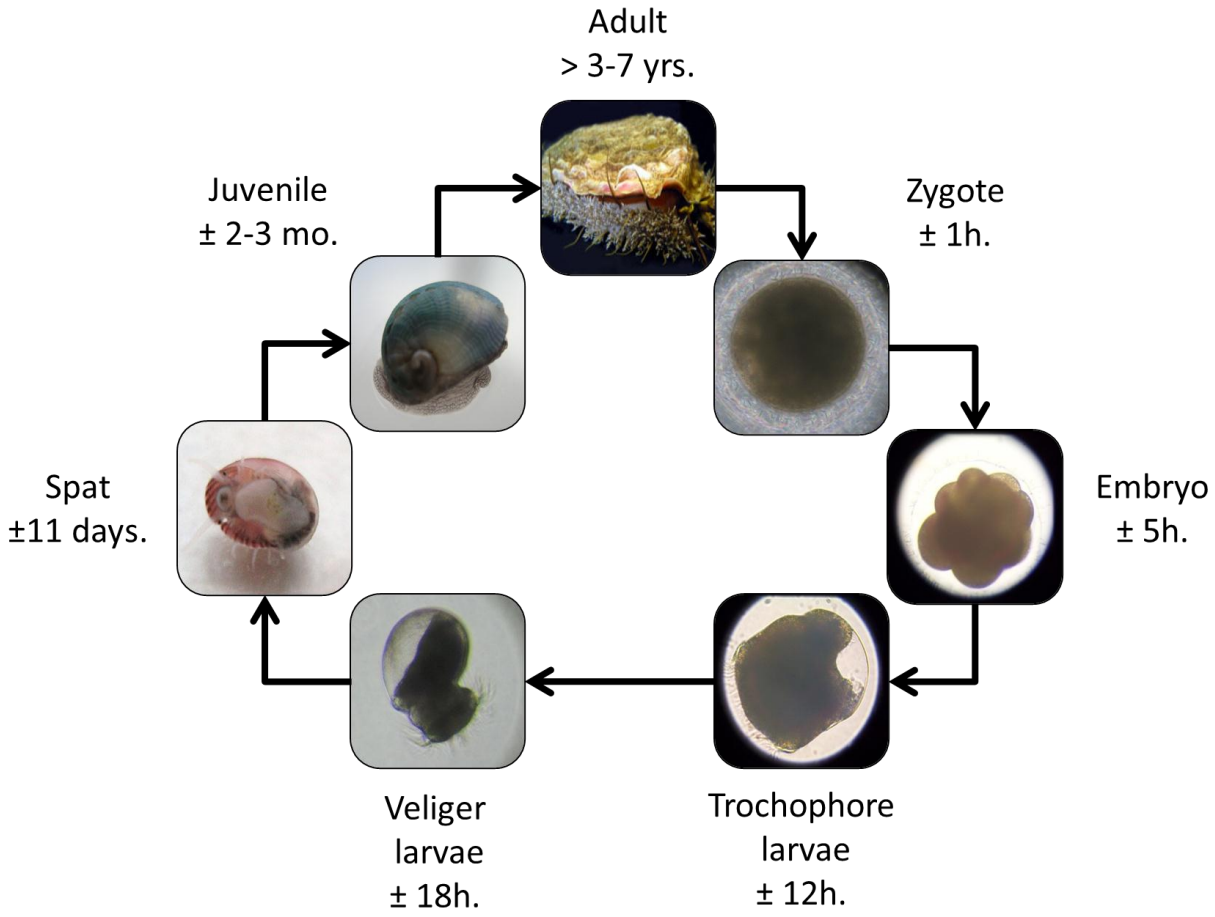

Figure 1.4: Abalone life cycle. Photos (by A. Roux) taken from the Molecular Aquatic Research Group's photo archives.

Globally, abalone is an economically important marine living resource, previously only as fisheries species and now also as an aquaculture species. South Africa's *perlemoen* is particularly revered for it meat quality and taste. With many of the worlds' fisheries in a dire state, an intensive research effort was initiated to resolve unknown biological questions in

aim of management and conservation of fisheries stocks and to find new innovative strategies to supply the high demand for abalone products. Much of this research was focused on genetics and using genetics as a tool for conservation and commercial production. These developments in abalone genetics and genomics are reviewed with specific reference to the South African abalone, *Haliotis midae*.

1.2.2. Origin of Abalone: From Fossils to Chromosomes and Genes

Molluscs were among the first animals to appear during the Cambrian explosion, approximately 570 million years ago (MYA), and are second only to the arthropods in terms of species diversity. The phylogeny of this phylum remains a matter of debate with the monophyletic or paraphyletic origins of the *Mollusca* still unresolved and several hypotheses being proposed (Bieler, 1992; Winnepenninckx *et al*., 1996, 1998; Passamaneck *et al*., 2004; Giribet *et al*., 2006; Wilson *et al*., 2010). The high degree of morphological diversity, convergent evolution of analogous structures and phenotypic plasticity has made phylogenetic deductions based on morphology difficult. Furthermore, the lack of strong phylogenetic signals at many genes has been attributed to the rapid radiation of molluscs (Rokas *et al*., 2005). However, a recent phylogenomic study places the *Gastropoda* and the *Scaphopoda* (tustk shells) as sister taxa with a monophyletic origin. Within the *Gastropoda* the *Vetigastopoda*, the order to which *Haliotis* belongs, forms a distinct clade (Smith *et al*., 2011). The early divergence of the *Vetigastopoda* from other gastropods has also been established (Winnepenninckx *et al*., 1998).

Inferring the origins of abalone purely from paleontological evidence is problematic: The fossil record is incomplete with only a few specimens in a limited number of geographic regions documented. Furthermore, of the proposed 35 fossil species many are based on the evaluation of only a single specimen; the loss of soft tissues during the mineralisation process and known shell morphological plasticity leading to ambiguities and uncertainties in the accuracy of this estimate (Geiger and Groves, 1999). Nonetheless, the earliest fossils date back to the late Cretaceous (Maastrichian) providing an estimate of the time of origin at approximately 70-80 (MYA) (Groves and Alderson, 2008).

However, a point of contention still exists on the place of origin: Abalone-like fossils from the Cretaceous were found in California (USA) as well as in Europe (the Netherlands and Sweden). There are various debates on whether these fossils represent true abalone, but most are in favour of the Californian fossils representing the ancestral archetype of these species (Vokes, 1935; Durham, 1979a, b; Sohl, 1987, 1992; Lindberg, 1992; Geiger and Groves, 1999; Groves and Alderson, 2008). The North American origin of abalone would support a westward radiation of abalone which is in accord with similar dispersal patterns of other marine molluscs (Squires, 1987).

Nonetheless, the high number of abalone species in the Indo-Pacific, suggest an alternative point of origin (Lindberg, 1992). Early cytogenetic studies also suggested a Mediterranean (the ancient Tethys Sea) origin of at least modern Haliotids, with a progressive increase in chromosome number from the Mediterranean (n = 14), through the Indo-Pacific (n = 16) to the North Pacific (n = 18) (Geiger and Groves, 1999) and Southern Pacific (including South African species; n = 18) (Li *et al.,* 1999; Van der Merwe and Roodt-Wilding, 2008). Chromosomal aneuploidy or polyploidy is often associated with speciation events and a progressive increase or decline in chromosome number from a radial centre frequently correlates with phylogenetic histories of such species or groups of species (Wang and Lan, 2000; Hipp *et al*., 2007). These observations lend support to the contrasting, alternative hypothesis of an eastward radiation of abalone. Under this hypothesis, the most likely extant species representing the ancestral form would be the European-Mediterranean abalone, *H. tuberculata*.

With the advent of molecular genetics this hypothesis gained credence, with molecular phylogenies based on various gene sequence data (including: lysin, hemocyanin, 16S rRNA, cytochrome oxidase) supporting the European origin and eastward radiation (Geiger, 2000; Coleman and Vacquier, 2002; Estes *et al*., 2005; Degnan *et al*., 2006; Streit *et al*., 2006) (Figure 1.5). Furthermore, the use of molecular genetic data established that there was a strong correlation between phylogeny and geographical distribution for Haliotids. Two clades predominate; a Northern Pacific clade (consisting of the North American and Japanese species) and an European-Australasian clade (consisting of the European, Australian, New Zealand and southern African species) (Estes *et al*., 2005; Degnan *et al*., 2006). Southern hemisphere species can furthermore be subdivided into two distinct

groupings, consisting of the southern African species and the Australasian species (Bester-Van der Merwe *et al*., 2012). This expanded radiation has been viewed as evidence for the division of *Haliotis* into two distinct genera as the extent of the genetic distance between the divergent clades is generally congruent with the recognition of such a partition (Brown and Murray, 1992; Lindberg, 1992). However, the matter remains under debate with Geiger and Poppe (2000) arguing in favour of a single genus with several proposed sub-genera.

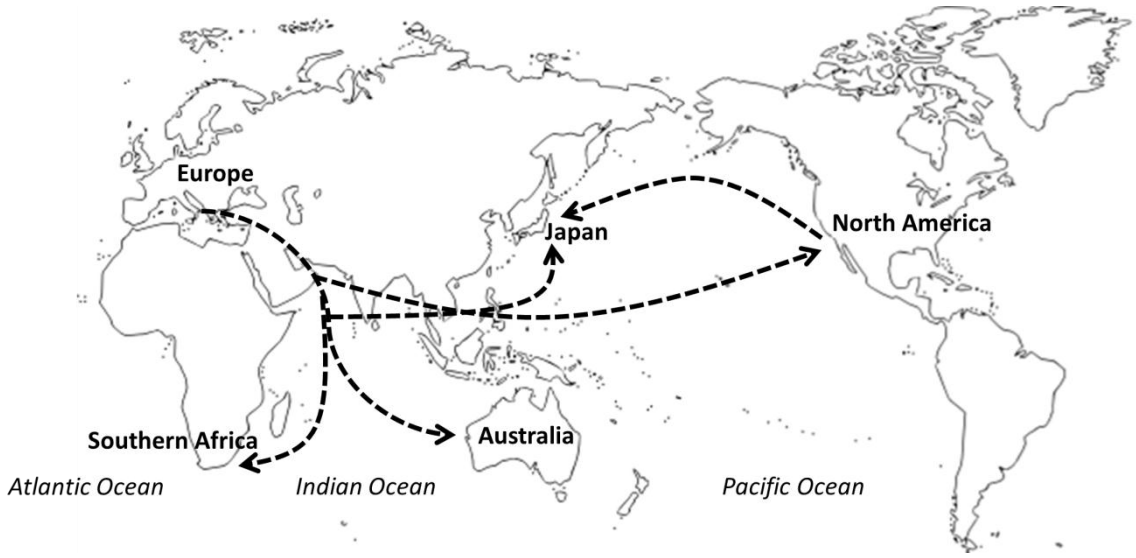

Figure 1.5: The eastward radiation of modern Haliotids from Europe to the major regions of endemism. This figure was adapted from the original by Streit *et al*. (2006).

Abalone fossils from South Africa date back to the Pleistocene (1.8 MYA), however it is suspected that this lineage probably predates these fossils; most likely dating back to the existence of Gondwana (Geiger and Grooves, 1999, Bester-van der Merwe *et al*., 2012). The South African species consists of a monophyletic group (Lee *et al*., 1995; Bester-Van der Merwe *et al*., 2012) with the Australian endemics, *H. rubra* and *H. laevigata* forming a sister grouping that is consistent with the southern geographic distribution (Estes *et al*., 2005; Degnan *et al*., 2006; Streit *et al*., 2006).

Species radiation of the South African abalone seems to be recent and may be due to founder dispersal and/or vicariance. Considering the short larval stages of abalone, the lack of evidence for colonisation *via* trans-oceanic dispersal and an ancient Gondwanan origin,

Bester-van der Merwe *et al*. (2012) argues in favour of bio-geographical vicariance for the origin of the South African endemics based on a combined NADH-dehydrogenase I and hemocyanin gene phylogeny. This may explain species radiation as a function of ecological adaptation to different environments. This is consistent with the three biogeographic provinces along the South African Coast: cool-temperate on the west-, warm-temperate on the south- and subtropical on the east coast (Emanuel *et al*., 1992). As such, the South African species is further divided into two clades mostly corresponding to niche requirements of the different species: The mostly intertidal and eurythermal *H. midae* and *H. spadicea* and species with more restricted habitats, *H. queketti*, *H. alfredensis* and *H. parva* (Figure 1.6). In the case where two species share a recent common ancestor and a range overlap (*e.g. H. midae* and *H. spadicea*), speciation could probably be attributed to rapidly evolving fertilisation genes, such as, sperm lysin and egg vitelline envelope receptors that create prezygotic barriers to reproduction. As broadcast spawning animals, the external fertilisation is highly dependent on chemotactic and recognition molecules on the surface of the gametes. If these are altered significantly, sperm and egg will no longer recognise "compatibility". As a population genetic event this could lead to the development of sympatric species (Lee *et al*., 1995; Kresge *et al*., 2001; Swanson *et al*., 2001; Galindo *et al*., 2003).

1.2.3. Contemporary and Historic Abalone Population Dynamics: The Case of Gene Flow

Within an evolutionary framework, population structure is dependent on the reproductive isolation of a group of individuals living within a set time and location. The isolation of groups of animals is directly correlated to the degree of gene flow amongst groups, which in turn is a function of species' dispersal capabilities (Waples and Gaggiotti, 2006). Marine species in general are characterised by their mobility. Even animals that are mostly static as adults often have planktonic larval stages that are easily dispersed by ocean

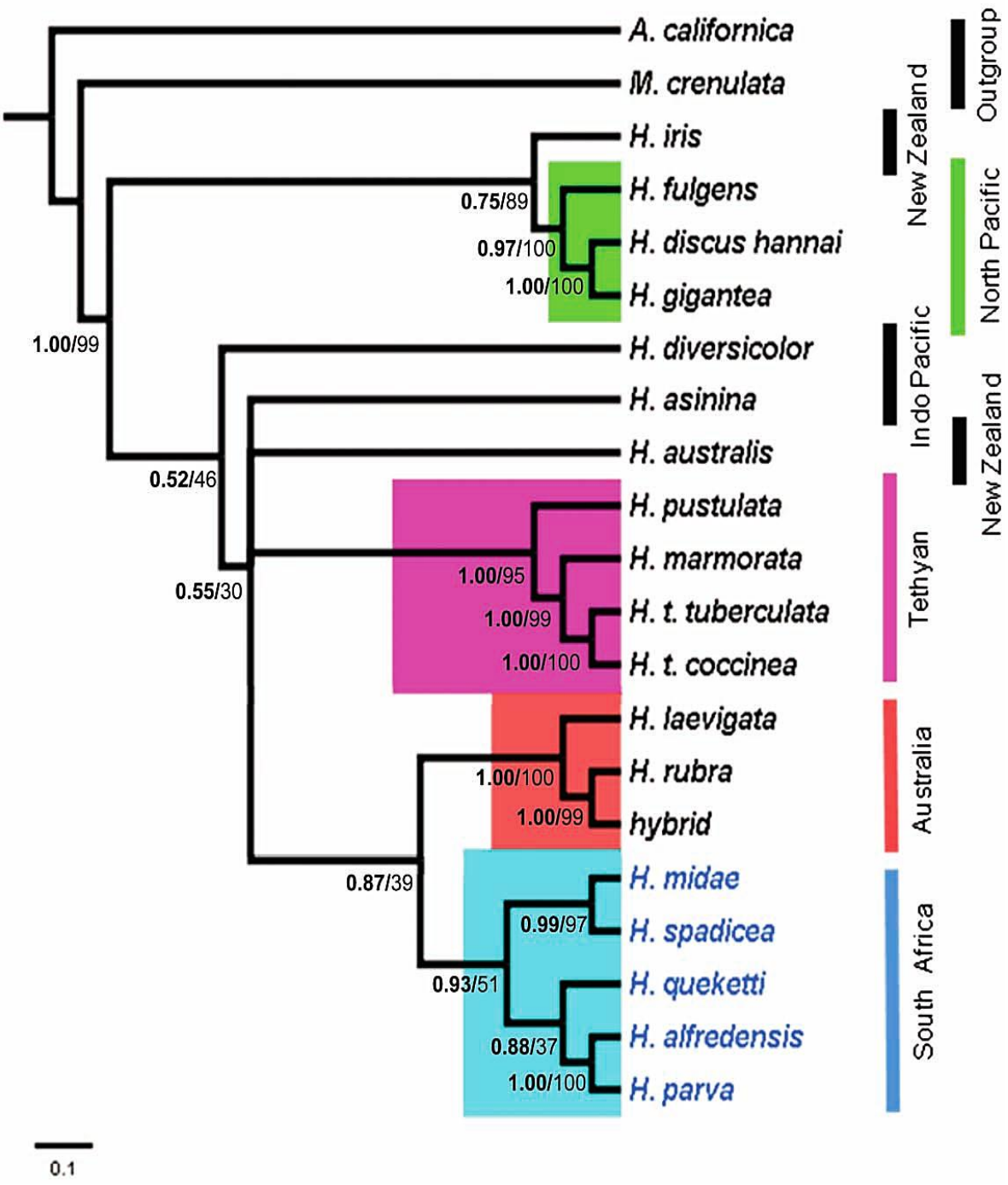

Figure 1.6: A Bayesian consensus phylogenetic tree showing the relationship between the South African- and other *Haliotis* species. Nodal values: Bayesian posterior probabilities (bold) and maximum likelihood bootstrap (plain text). Figure taken from Bester-van der Merwe *et al*. (2012).

currents. Furthermore, in marine environments physical barriers to gene flow seem to be limited and thus individual dispersal probability is assumed to be high. This led to the generally held belief that marine species consists of large panmictic populations (Hilbish, 1996). This view was initially supported by early studies that found genetic continuity among individuals sampled across a wide geographic distribution for many species (Gyllensten, 1985; Waples, 1987; Ward *et al*., 1994; Hilbish, 1996; Waples, 1998; Kyle and Boulding, 2000). However, recent investigations are challenging the prediction of universal panmixia of marine populations (Warner and Cowen, 2002; Gilg and Hilbish, 2003; Veliz *et al*., 2006). Furthermore, refined analyses using sophisticated statistical models and whole-genome data show significant fine-scale population structure (Beheregaray and Sunnucks, 2001;) and population divergence due to adaptation even within species with high gene flow amongst populations (Knutsen *et al*., 2003; Pampoulie *et al*., 2004; Hemmer-Hansen *et al*., 2007; Zhan *et al*., 2009; Nielsen *et al*., 2009a, b; André *et al*., 2011; Hoffman *et al*., 2011, 2012).

Haliotids too seem to contradict the general expectations of panmixia expected of marine populations. The relatively short larval phase limits the window of opportunity for dispersal in abalone (Prince *et al*., 1987; McShane, 1992). It is therefore expected that some population genetic structuring will be evident due to population "self-recruitment". This is a common phenomenon in sedentary marine organisms where the level of population structure is correlated to the dispersal capabilities of the larvae (Hellberg, 1996; Arndt and Smith, 1998; Kyle and Boulding, 2000; Gilg and Hilbish, 2003; Levin, 2006; Banks *et al*., 2007). As such, population structure has been reported for various abalone species, including *Haliotis rubra* (Brown, 1991; Huang *et al*., 2000; Temby *et al*., 2007; Miller *et al*., 2009), *H. coccoradiata* (Piggott *et al*., 2008), *H. cracherodii* (Hamm and Burton, 2000; Chambers *et al*., 2006; Gruenthal and Burton, 2008), *H. discuss* (Hara and Sekino, 2005; Sekino *et al*., 2005), *H. asinina* (Tang *et al.,* 2004), *H. corrugata* (Díaz-Viloria *et al*., 2009), *H. kamtschatkana* (Withler *et al*., 2003), and *H. rufescens* (Gruenthal *et al*., 2007).

Molecular genetic data suggests population heterogeneity of *H. midae* populations on the west and east coasts of South Africa giving rise to two major reproductive stocks with Cape Agulhas as the point of transition (Evans *et al.,* 2004a; Bester-van der Merwe *et al*., 2011). Evans *et al*. (2004a) found higher levels of genetic diversity as estimated by three

microsatellite loci and mitochondrial haplotypes on the west coast. They argued that the reduction of genetic diversity in the east coast population was a consequence of a population bottleneck. Subsequently, they proposed a eastward range expansion with a colonisation event from an ancestral population on the west coast, from where the new east coast population remained isolated from the parent population. The bottleneck was, therefore, produced by the founder effect, which the authors argued was particularly supported by the mitochondrial haplotype analysis.

Although founder effects are common (*e.g.* Brooker *et al*., 2000; Hundertmark and Van Daele, 2010; Keller *et al*., 2010; Tatarenkov *et al*., 2010) the west-east colonisation hypothesis of Evans *et al*. (2004a) contradicts the Mediterranean origin and eastward (south-eastward to southern Africa) radiation hypothesis for abalone (Geiger, 2000; Coleman and Vacquier, 2002; Estes *et al*., 2005; Degnan *et al*., 2006; Streit *et al*., 2006) and the Gondwanan biogeographic vicariance proposed for the origin of the South African species (Bester-van der Merwe *et al*., 2012). Furthermore, Bester-van der Merwe *et al*. (2011) could not find evidence for a reduction in the effective population size; instead long-term effective population size seemed to be stable across the species' range. The discontinuity of west- and east coast populations were however maintained, although population differentiation was subtle, with the retroflection of the Agulhas current (Dijkstra and de Ruijter, 2001) creating the major barrier to gene flow. A secondary barrier could possibly be caused by a thermal front in the Algoa Bay region, however this was not strongly supported. It could however explain a possible third distinct population on the south coast of the country (Bester-van der Merwe *et al*., 2011) which is in agreement with the three biogeographical provinces reflected in the population structure of many other offshore marine species in South Africa. This lends support to the idea that environmental and oceanographic features of the South African coast might be an important determinant for population structuring in marine organisms and that population divergence might further be correlated to adaptation (Ridgway *et al*., 1998; Teske *et al*., 2006, 2007, 2008; Zardi *et al*., 2007; Von der Heyden *et al*., 2008; Bester-van der Merwe *et al*., 2011).

Therefore, based on the evidence from eight microsatellite- and 12 SNP loci, Bester-van der Merwe *et al.* (2011) formulated an alternative hypothesis to the west-east colonisation hypothesis of Evans *et al*. (2004a). It is postulated that historically *H. midae* populations on

the west- and east coast become isolated within *refugia* during the last glacial maxima. Following glacial retreat, approximately 20 000 years ago, both populations experienced range expansion, also evident in other South African marine fauna (*e.g.* Tolley *et al*., 2005; Gopal *et al*., 2006; Matthee *et al*., 2006; Von der Heyden *et al*., 2007; Neethling *et al*., 2008). This range expansion ultimately culminated in the formation of a secondary contact zone on the south coast with contemporary population dynamics being maintained by the bio-physical characteristics of the biogeographical provinces (Figure 1.7).

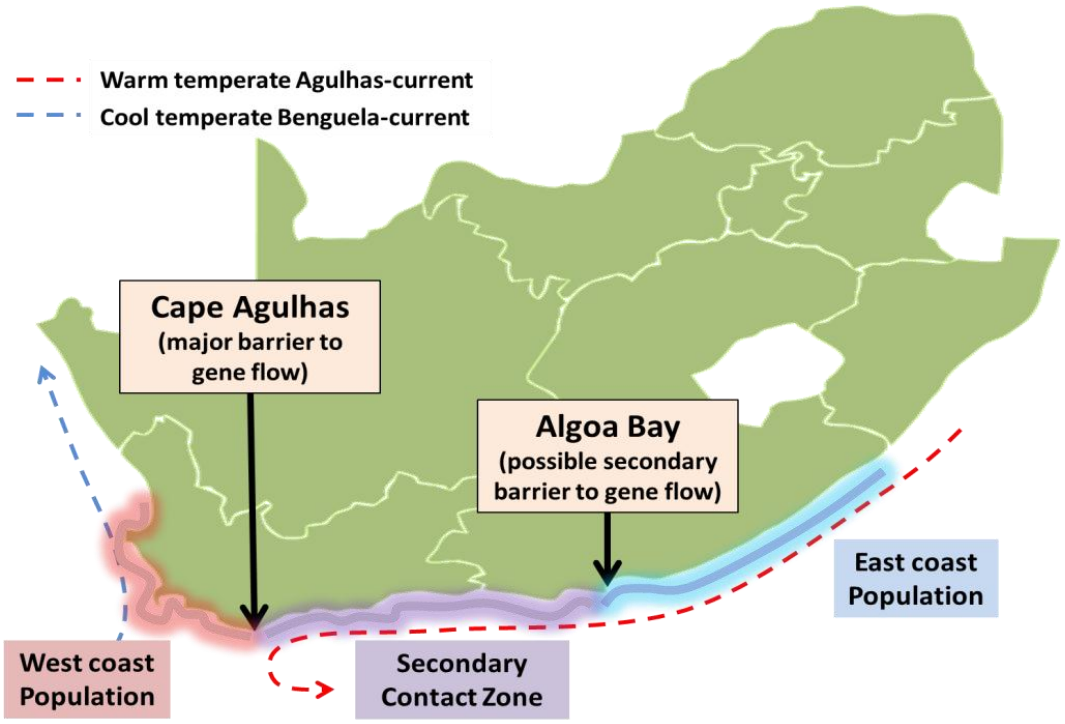

Figure 1.7: A graphical representation of the population structure and barriers to gene flow of *H. midae* around the South African coast.

Considering the aforementioned, the marginal decrease in genetic diversity indicated by the microsatellite loci, as reported by Evans *et al*. (2004a) is most likely an artefact of using only three markers and therefore probably also non-significant. The disparity as noted by the mitochondrial analysis could be attributed to various factors. Firstly, the non-recombining mitochondrial genome might be more susceptible to selective sweeps than nuclear loci (Ballard and Whitlock, 2004; Bazin *et al*., 2006; Peijnenburg *et al*., 2006;

Tarnowska *et al*., 2010). The mitochondrion plays a central role in cellular respiration, in particular electron transport and oxidative phosphorylation; these biochemical reactions are oxygen dependent. It is also known that warmer waters have a lower capacity to dissolve oxygen, whilst the oxygen requirement of ectothermic organisms generally increases under warmer temperatures due to increased metabolism; thus higher demand for oxygen lower supply. It is therefore comprehendible that such a selective sweep could have reduced mitochondrial diversity due to functional constraints on the respiratory mechanism to cope with the shortage of oxygen in the warmer waters of the east coast – this hypothesis however remains to be tested. Secondly, because the mitochondrial genome is only inherited maternally and thus always haploid, the effective population size for mitochondrial loci is only a quarter of the effective population size for nuclear loci. The effect of random drift on mitochondrial genetic diversity is therefore more pronounced (Buonaccorsi *et al*., 2001; Appleyard *et al*., 2002; Ballard and Whitlock, 2004; Shaw *et al*., 2004; Lukoschek *et al*., 2008; Larmuseau *et al*., 2010; Tarnowska *et al*., 2010). The observed aberrant pattern of genetic diversity could therefore be explained as a chance event. Lastly, biases in gender specific dispersal rates could explain this phenomenon; however there are limited examples of dispersal disparities between the sexes in marine organisms (*e.g.* Arnaud-Haond *et al*., 2003; Diaz-Almela *et al*., 2004; Consuegra and de Leaniz, 2007). And no evidence for this has been reported for abalone, therefore this is the most unlikely reason for the mitochondrial disparity observed in the study of Evans *et al*. (2004a).

1.2.4. Abalone Mass Production: An Industrial Revolution

Globally, abalone is a sought-after culinary delicacy especially in the Far East where it is also used in traditional medicine. The South African abalone generally fetches prices of 32 US$ to 34 US$ per kilogram, but this may reach as high as 1000 US$ depending on the market (UNEP-WCMC, 2010; Department of Agriculture, Forestry and Fisheries, 2012). There are approximately 14 economically important abalone species distributed across the major regions of endemism (Roodt-Wilding and Slabbert, 2006). Historically, the abalone fishery is probably one of the oldest fisheries in world; archaeological evidence suggest that stone-age-man have been harvesting abalone in South Africa as early as 125 000 years ago

(Tarr, 1989). The modern fishery in South Africa is reliant on a single species, *Haliotis midae*, and was initiated in 1949. The fishery reached a peak in 1965 when a record 2 800 metric tons of abalone were hauled from the ocean (Tarr, 1989, 1992). Since then the fishery went into decline with landings reaching an all-time low in 2007/8 at 75 metric tons. Declining landing raised concerns about the sustainability of harvesting practises and led to the suspension of commercial fishing activities in 2008 (Troell *et al*., 2006; Raemeakers *et al.,* 2011). The fishery was conditionally re-opened in 2010 to allow small-scale fisherman a means of income (Department of Agriculture, Forestry and Fisheries, 2010).

Declining and collapsing fisheries stocks is a global phenomenon for many species (McShane *et al*., 1994; Altstatt *et al*., 1996; Hobday *et al*., 2001; Hobday and Tegner, 2002). Increasing apprehension regarding the sustainability of fisheries and growing demand for fisheries products created the need to find alternative sources for these commodities; hence the dawn of aquaculture. Abalone aquaculture coincides, in most of the major producing countries, with the collapse of fisheries (Ebert, 1992; Garza and Bernal, 1992; Tong and Moss, 1992; Altstatt *et al*., 1996; Hobday *et al*., 2001). In China, research in abalone aquaculture dates back to the 1950's and therefore the country most likely has the oldest history of abalone culture (Nie, 1992; Fleming and Hone, 1996).

Other than declining wild stocks, initial interest in the culture of the local *perlemoen* was sparked by the demonstration that this species could be spawned and reared successfully in a captured environment during the 1980's (Genade, 1988). Initial success of abalone culture abroad and more favourable growth rates for *H. midae* in captivity spurred the establishment of the first aquaculture ventures in the 1990's (Troell *et al*., 2006). During 1996 the first exports were delivered (Cook, 1998); currently South Africa has 14 operating farms with a total output of 1015.44 metric tons, valued at 355 million ZAR (approximately 44 million US$). Abalone production accounts for just more than half of the total mariculture tonnage, but produces more than 90% of mariculture revenue (Department of Agriculture, Forestry and Fisheries, 2012).

In order to maintain market competitiveness, a substantial research effort was initiated to optimise and refine general husbandry and management of aquaculture stocks, with advances in animal handling, nutrition, reproduction and disease surveillance and

biosecurity (*e.g.* Britz, 1994; Britz *et al*., 1997; Mackay and Coyne, 2005; Vosloo and Vosloo, 2006; Simon and Booth, 2007; Roux *et al*., 2008; Simon *et al*., 2010; Mackay *et al*., 2011; Mouton and Gummow, 2011). To further enhance production output, a genetic improvement programme was initiated in 2006. The main aim of this programme was to exploit the inherent genetic variation to create a genetically superior domesticated abalone strain that would outperform the wild progenitor in terms of growth rate (Roodt-Wilding and Brink, 2011).

Phenotypic selection, where the breeding value of an individual is determined by phenotypic performance, is the backbone for conventional selective breeding. High performance individuals are selected to contribute to the next generation; as such the mean trait value is shifted in a desired direction over successive generations. Much progress has been made by traditional quantitative genetic approaches to genetically improve conventional livestock (*e.g.* cattle, sheep, swine and poultry). Progress has also been made in some aquaculture species, including: Atlantic salmon (*Salmo salar*, Gjedrem, 2000; Kjoglum *et al*., 2008), rainbow trout (*Onchorhynchus mykiss*, Pottinger *et al*., 1994), channel catfish (*Ictalurus punctatus*, Dunham and Brummet, 1999; Rezk *et al*., 2003), Nile tilapia (*Oreochromis niloticus*, Betsen *et al*., 1998; Eknath *et al*., 2007) and some molluscan species (Langdon *et al*., 2003; Toro *et al*., 2004; Zheng *et al*., 2006).

Irrespective of the advances made with conventional breeding methods; it remains a cumbersome process. This is especially applicable for phenotypes that are not readily measurable or only observed in a particular sex, traits with low heritability and species with long generation intervals. With the recent rapid development of molecular marker- and DNA sequence technologies it was decided to implement a dualistic strategy for the genetic improvement of *perlemoen*; using conventional- and molecular breeding methods in a complementary manner. This approach has been advocated for numerous aquaculture species (Davis and Hetzel, 2000; Hulata, 2001). As such, a base population consisting of more than a 1000 individuals was established and 426 full- and half-sib family groups were spawned and a performance recording scheme initiated (Roodt-Wilding and Brink, 2011).

Heritability estimates for growth-related traits (shell length and wet weight) for *H. midae* is estimated to be moderate, 0.21-0.25 at 60 months of age. It is interesting to note that

estimates increased with age, starting at 0.05-0.08 as the initial estimate taken at 6 months (Roodt-Wilding and Brink, 2011). This is however, not surprising as the effects of additive genetic variation may be thwart by unknown maternal and/or larval effects at early stages in development (Kube *et al*., 2007). These heritability estimates correlate well with estimates in other Haliotid species, including: *H. rufescens* (Jonasson *et al*., 1999), *H. rubra* (Li *et al*., 2005), *H. asinina* (Lucas *et al*., 2006), *H. discus hannai* (Deng *et al*., 2007), *H. laevigata* (Kube *et al*., 2007) and *H. diversicolor* (You *et al*., 2010a, b). It must however be noted that *H. midae* is the only species with estimates taken up to 60 months. Based on these heritability estimates and a selection intensity of 1%, response to selection is predicted to be 7-17% in accordance with gains achieved or predicted in other Haliotids (Kawahara *et al*., 1997; Lucas *et al.,* 2006; Kube *et al*., 2007; Robinson *et al*., 2010; You *et al*., 2010a; Roodt-Wilding and Brink, 2011).

Hayes *et al*. (2007a) demonstrated that by using marker assisted selection (MAS) a 13% advantage might be gained in comparison to using only conventional breeding strategies for the Australian abalone, *H. rubra*. The major advantage is that high performance individuals can be selected based on genotype at an early age, before the phenotype is expressed; thus the rate of genetic change in response to selection is accelerated due to a reduced generation interval. Molecular breeding is divided into four key phases: the development of molecular markers (DNA polymorphisms that segregates in a given population), linkage mapping, identification of QTL associated to particular phenotypes and the implementation of MAS (Poompuang and Hallerman, 1997; Collard *et al*., 2005).

Microsatellite- and SNP markers are the most widely applied molecular genetic markers in animal genetics at present. Both marker types are co-dominant, however microsatellites were long favoured due to their multi-allelic nature: thus having a high polymorphism information content, in comparison to biallelic SNPs. Nonetheless, SNPs are gaining popularity due to new sequence and genotyping technology advancements that has led to easy and quick discovery of many markers at reduced costs. The high frequency of SNPs throughout the genome and their lower genotyping error rate has also contributed to the use of this marker type (Beuzen *et al*., 2000; Brumfield *et al*., 2003; Lui and Cordes, 2004; Morin and McCarthy, 2007; Pérez-Enciso and Ferretti, 2010; Liu *et al*., 2011).

As in other abalone species (*e.g*. Selvamani, 2000, 2001; Hara and Sekino, 2005, 2007; Sekino *et al*., 2005; Baranski *et al*., 2006a), molecular marker development in *H. midae* was initially focused on developing microsatellites. To date more than 250 anonymous- and gene-linked microsatellites have been developed for *perlemoen* (Bester *et al.,* 2004; Slabbert *et al*., 2008, 2010, 2012; Rhode, 2010]. Rhode and Roodt-Wilding (2011) also concluded that microsatellite loci are not randomly distributed in the *H. midae* genome but rather particular motifs seem to associate to particular genomic regions, mostly genes and transposable elements. The development of SNP markers in aquaculture species has increased rapidly, *e.g.* finfish (He *et al*., 2003; Smith *et al*., 2005; Ryynanen and Primmer, 2006; Cenadelli *et al*., 2007; Hayes *et al*., 2007b; Wang *et al*., 2008, 2010; Hubert *et al*., 2009), shrimp (Gorbach *et al*., 2009) and molluscs (Elfstrom *et al.,* 2005; Quilang *et al*., 2007). The only other abalone species for which a concerted effort has been made to develop SNP markers is *H. discus hannai* (Qi *et al*., 2008, 2009, 2010; Zhang *et al*., 2010) and some recent attempts at comparative SNP analyses across various species (Kang *et al*., 2011). Small-scale SNP development studies were initiated in 2006/7 for *H. midae* (Bester *et al*., 2008; Rhode *et al*., 2008; Rhode, 2010). However, a large-scale endeavour commenced with the transcriptome sequencing project using the sequence-by-synthesis technology of Illumina® (Franchini *et al*., 2011). At present the South African abalone has more than 200 validated SNP markers (Bester *et al*., 2008; Rhode *et al*., 2008; Rhode, 2010; Blaauw, 2012; Du Plessis, 2012).

Using these molecular marker resources, the linkage map of *H. midae* has recently been completed (Vervalle *et al.,* in press). The integrated linkage map comprises of 186 microsatellite- and SNP markers and resolved into 18 linkage groups; corresponding to the karyotype haploid number (Van der Merwe and Roodt-Wilding, 2008; Franchini *et al.,* 2010). The estimated genome size based on genetic recombination (Vervalle *et al.,* in press) and flow cytometry (Franchini *et al.,* 2010) was also in agreement; estimating the *H. midae* genome at approximately 1400 cM (diploid genome size in physical basepairs: ±2.8 GB). Average marker spacing was 6.88 cM and thus provides a sufficient framework map for preliminary QTL analysis. Generally a marker spacing of 20 cM is deemed sufficient for such crude QTL detection (Massault *et al*., 2008). Currently linkage maps for only three other

abalone species are available: *H. diversicolor* (Shi *et al*., 2010; Zhan *et al*., 2012), *H. discus hannai* (Liu *et al*., 2006; Sekino and Hara, 2007) and *H. rubra* (Baranski *et al*., 2006b).

Five putative QTL for growth traits, explaining 15% to 33% of genetic variation, were detected in two families for *H. midae*. All loci were located on a single chromosome indicating that this chromosomal region probably harbours a major gene(s) regulating growth rate (Slabbert, 2010). Saturation of the linkage map and fine-mapping of QTL regions remains to be done. Baranski *et al*. (2008) identified nine QTL associated to growth traits in *H. rubra* explaining 16% to 47% of phenotypic variation. Various growth QTL were also found for *H. discus hanni* accounting for 8% to 18% of phenotypic variation (Liu *et al*., 2007). It is not uncommon to detect different QTL or deviations in the percentage of variation explained by a particular locus, as QTL are highly context dependent, *e.g.* differing environmental circumstances may activate different genes or populations with diverse genetic backgrounds may exhibit different patterns of LD and thus have different functional gene-networks. It therefore remains vital to validate QTL for particular populations under specific environmental conditions (Dekkers, 2004; Collard *et al*., 2005). The implementation of MAS or genomic selection (where whole genome data is used) in aquaculture species has been hindered by: (1) the current lack of high density linkage maps and (2) validated LD-QTL (QTL-marker loci that shows association to a particular trait throughout the population and not necessarily limited to a particular family). However, with the substantial benefits of molecular breeding it is envisioned to play an increasingly important role in future aquaculture genetic improvement programmes (Liu and Cordes, 2004; Wenne *et al*., 2007).

Molecular marker technology has not only served the aquaculture industry in the detection of genetic variants associated with economically important traits, but has become an important tool for management and record keeping of particular commercial populations/animals. Pedigree records are difficult to maintain for especially broadcast spawning animals, such as abalone, as individual spawnings are impractical under commercial settings. Thus, the use of molecular markers for parentage assignments and subsequent pedigree inference has become important (Evans *et al.,* 2000; Jerry *et al*., 2004; Dong *et al*., 2006; Lucas *et al*., 2006; Gheyas *et al*., 2009). This also aids the assessment of differential parental contributions that are common in highly fecund broadcast spawning molluscs. This could skew the estimate for heritability (Kube *et al.,* 2007), as well as the

genetic structure of the commercial population leading to exacerbated inbreeding if the majority of high performance individuals are selected from the same family group (Selvamani *et al*., 2001; Bentsen and Olesen, 2002; Park *et al*., 2006; Horreo *et al*., 2008; Lind *et al*., 2009; Van der Berg and Roodt-Wilding, 2010).

Molecular markers are also routinely used to assess genetic diversity of commercial stocks. This is particularly important when establishing a base population, where it is necessary to capture as much of the inherent genetic variation as possible. This will allow sustainable and long-term genetic gains under variable environments (Rauw *et al*., 1998; Elliott, 2000; Gamborg and Sandøe, 2005; Jensen and Andersson, 2005; Hayes *et al*., 2006; Flint and Woolliams, 2008; Cardellino and Boyazoglu, 2009). As such, it was found that the wild broodstock collected as the base population for the *H. midae* breeding programme had levels of genetic diversity comparable to the general wild population (Roodt-Wilding and Brink, 2011). It is also important to monitor genetic diversity in the subsequent culture-reared generations to evaluate the effects of a breeding programme on the genetic constitution of commercial stocks, *e.g.* effective population sizes, rate of inbreeding and relatedness and population differentiation (Brown *et al*., 2005 Hara and Sekino, 2007; Li *et al*., 2007; Dixon *et al*., 2008; Lind *et al*., 2009; De la Cruz *et al*., 2010; Praipue *et al*., 2010; An *et al*., 2011). In the South African abalone, it has been found that even after one generation of culture the genetic properties of the F_1-population could be altered so dramatically that it was distinct from the wild populations; it is postulated that this may be due to founder effects, random genetic drift and selection (Evans *et al*., 2004b; Slabbert *et al.,* 2009). Similar observations were made for other Haliotids (Hara and Sekino, 2007; Li *et al*., 2007; De la Cruz *et al*., 2010; Praipue *et al*., 2010; An *et al*., 2011).

Next-generation sequencing technology is currently revolutionising genetics and genomics (Mardis, 2008; Shendure and Ji, 2008; Varshney *et al*., 2009; Pérez-Enciso and Ferretti, 2010). The technology has already been applied to some aquaculture species for marker development and transcriptomic profiling of genes (Salem *et al*., 2010; Tymchuk *et al*., 2010; Hohenlohe *et al*., 2011; Liu *et al.,* 2011). As mentioned earlier it has also been used in *H. midae* (Franchini *et al*., 2011) and two other abalone species; *H. diversicolor* (Jiang *et al*., 2011) and *H. rufescens* (De Wit and Palumbi, 2012). What is evident from these studies is the high number of unique gene transcripts in abalone that has no known

homologous gene in model species, which are overrepresented in public gene databases such as NCBI. Van der Merwe *et al*. (2011) also conducted a differential gene expression experiment using the Illumina® technology to identify genes with differing expression profiles between fast- and slow growing South African abalone. Genes involved in growth-related physiological processes including insulin-related peptide receptors and insulin-like growth factor binding proteins and genes involved in stress tolerance such as heat shock proteins were identified. The identification of specific genetic variants in these genes and their association to the fast growth rate phenotype remains to be done; this will be necessary if gene expression data is to be readily incorporated into molecular breeding strategies.

From an evolutionary viewpoint, the industrialised domestication of abalone provides the opportunity for a unique "genetic experiment". When animals become subject to domestication three genetic processes, in particular, are involved: inbreeding, random genetic drift and selection (Mignon-Grasteau *et al*., 2005). The founder population, at the start of domestication, is generally much less than the original wild population. This consequently leads to a smaller effective population size, which in turn increases the rate of inbreeding and pronounces the effects of random genetic drift. The effects of selection is multi-dimensional: Firstly, there is relaxed natural selection on traits for survival in the wild; secondly there is increased natural selection for adaptation to the new captured environment and lastly, humans exert artificial selection for desirable traits, such as production characteristics. The rapid population differentiation of cultured abalone from their wild progenitor populations (Evans *et al*., 2004b; Slabbert *et al*., 2009) begs the question as to what extent each of the evolutionary forces are responsible for the development of the abalone domestic phenotypes? Furthermore, what is the interplay between natural selection and artificial selection? Recent studies in cattle and chickens reported signatures of selection in genomic regions of known QTL for production traits (*e.g.* Qanbari *et al*., 2010; Rubin *et al*., 2010). From studies on conventional livestock the importance of regulatory factors in gene expression and pleiotropic effects in the development of the "domestic syndrome" must also be noted (Schütz *et al*., 2002; Kerje *et al*., 2003; Dobney and Larson, 2006; Qanbari *et al*., 2010; Rubin *et al*., 2010).

1.2.5. Abalone Conservation: Preservation of Unique Genetic Resources

Due to its high market value, natural populations of *H. midae* has come under immense pressure due to overharvesting in the past and presently the illicit activities of poachers are severely hampering conservation efforts (Hauck and Sweijd, 1999; Tarr, 2000; Dichmont *et al*., 2000; Plagányi *et al*., 2001; Steinberg, 2005; Plagányi and Butterworth, 2010; Raemaekers *et al*., 2011). Management of the abalone resource was, until recently, mainly based on modelling ecological demography reliant on various parameters of reproductive biology and landing statistics, *e.g.* tonnage harvested per surface area, size distributions of harvested specimens, and other ecological factors such as predation (Dichmont *et al*., 2000; Tarr, 2000; Plagányi *et al*., 2001; Plagányi and Butterworth, 2010). With the promulgation of the Marine Living Resources Act (Republic of South Africa, 1998) a holistic ecosystems management approach was adopted for the management of marine *biota*. At present approximately 21% of the South Africa coast consists of marine protected areas, but only 9% are no-take zones (Von der Heyden, 2009).

Von der Heyden (2009) argued that the current marine protected areas are not optimally designed as to reflect the population genetic structure of marine organisms around the South African coast and made recommendations on how to incorporate genetic data into reserve design within the context of ecosystems management. The key is to consider the genetic connectivity between different populations, *i.e.* larval dispersal probability and the directionality of currents and to protect genetically distinct populations to ensure conservation of all biodiversity (Palumbi, 2003 for a review). What makes genetic assessment such a powerful tool is that it allows for an understanding of a population's historic demography that could be more important in the prediction of a species' long-term population dynamics (Moritz, 2002).

Bester-van der Merwe *et al*. (2011) recommended that the barriers to gene flow around the South African coast, Cape Agulhas and to a lesser extent Algoa Bay, should be taken into consideration when delineating management units for *perlemoen*. What was clear is that population groups west and east of the Cape Agulhas are distinct reproductive stocks. However, the transition/admixture zone on the south coast between Cape Agulhas and Algoa Bay may warrant the recognition of a third management unit. Admixture zones may

be particularly important in the maintenance of adaptive diversity on either side of the transition, but may still allow some gene flow between the two extreme ends; thereby sustaining the evolutionary potential of a species across ecological niches (Riginos and Cunningham, 2005; Counterman *et al*., 2010). Various degrees of gene flow between west- and east coast populations of *H. midae* is known to occur; it would therefore be sensible to not conserve management units in isolation, but rather also preserve the connectivity between populations (Bester-van der Merwe *et al*., 2011).

A thorough understanding of population structure is also important when considering stock enhancement or ranching activities. The main concern is that if the genetic constitution of reseeded animals is not similar to that of the local population where release is intended, it could lead the erosion of the evolutionary/adaptive potential of the natural population. This will occur by lowering the fitness of hybrid individuals due to outbreeding depression (Naylor *et al*., 2005; Jones *et al*., 2006; Roodt-Wilding, 2007; Hara *et al*., 2008; Camara and Adopalas, 2009; Zhang H *et al*., 2010). With studies reporting significant differentiation between cultured and wild populations, concerns on the utility of culture reared animals for reseeding initiatives are justified, not only in South Africa (Evans *et al* 2004b; Slabbert *et al*., 2009) but also world-wide (Hara and Sekino, 2007; Li *et al*., 2007; De la Cruz *et al*., 2010; Praipue *et al*., 2010; An *et al*., 2011). Numerous studies have been conducted on abalone ranching globally and in South Africa with variable results (Gaffney *et al*., 1996; Sweijd *et al*., 1998; De Waal *et al*., 2003; Gutierrez-Gonzalez and Perez-Enriquez 2005; Dixon *et al*., 2006; Hamasaki and Kitada, 2008). With the advent of genomic scans for selection it has become possible to accurately account for adaptive diversity and to explicitly formulate this in conservation strategies that in the past relied mostly on a small number of presumably neutral markers. As such, management and evolutionary significant units can be more accurately determined and individuals for stock enhancement may be adaptively matched to populations where release is intended (Medugorac *et al*., 2009; Allendorf *et al*., 2010; Ouborg *et al*., 2010; Tymchuk *et al*., 2010).

1.3. Aims and Objectives

Haliotids, as an economically important genus has been subject to much research over the past years. Genetics and more recently genomics has played an unequivocal role in elucidating fundamental questions, such as, on the origin of abalone and general biological phenomena. As a tool for the commercial exploitation of abalone, genetics and genomics have become irreplaceable. Even though genetic resources such as marker maps and genome sequence data are rudimentary in comparison to conventional livestock and even other more industrialised aquaculture species such as salmon, much progress has been made in the development of such resources especially in the South African abalone. Some questions however, remain unanswered. For example, adaptation to environmental conditions has been hypothesised to be an important driving force for the development of new species of abalone from the initial point of radiation. But to what extent can the differential contributions of random genetic drift and selection account for population divergence in the in natural- and captive populations and which loci are responsible for the development of complex phenotypes? New experimental and statistical approaches in genomics may aid in answering such questions.

The aim of this study was thus to elucidate the evolutionary processes that contribute to the development of divergent ecotypes in wild abalone and to ascertain the effects of domestication on the genetic constitution of cultured abalone. As such, the occurrence of signatures of selection was investigated under a population genomics framework, where genome-wide patterns of genetic diversity was assessed. A standard population genetics analysis was first conducted to determine the extent of population differentiation within and between wild- and cultured populations using genomic and EST microsatellite markers. This microsatellite analysis was then expanded, using a population genomics approach (150 markers) to validate the initial analysis and identify functionally important loci that may be under selection and could explain population divergence as a function of adaptation. Lastly, a newly developed SNP assay was used to conduct a temporal investigation to assess the fluctuations in genome-wide genetic diversity across space and time of the South African abalone, *Haliotis midae.* The obtained results will then be interpreted in terms of general biological phenomena and applications in abalone conservation and aquaculture in South Africa.

References

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, Lewis SE, Richards S, Ashburner M, Henderson SN, Sutton GG, Wortman JR, Yandell MD, Zhang Q, Chen X, *et al*. (2000) The genome sequence of *Drosophila melanogaster*. Science 287: 2185-2195.

Aguileta G, Lengelle J, Marthey S, Chiapello H, Rodolphe F, Gendrault A, Yockteng R, Vercken E, Devier B, Fontaine MC, Winker P, Dossat C, Cruaud C, Couloux A, Giraud T (2010) Finding candidate genes under positive selection in non-model species: examples of genes involved in host specialization in pathogens. Mol Ecol 19: 292-306.

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA, Kruglyak L (2004) Population history and natural selection shape patterns of variation in 132 genes. PLoS Biol 2: 1591-1599.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11: 697-710.

Altstsatt JM, Ambrose RF, Engle JM, Haaker PL, Lafferty KD, Raimondi PT (1996) Recent declines in black abalone *Haliotis cracherodii* on the mainland coastal of central California. Mar Ecol Prog Ser 142: 185-192.

An HA, Wook J, Kim HC, Myeong J (2011) Genetic characterization of five hatchery populations of Pacific abalone (*Haliotis discus hannai*) using microsatellite markers. Int J Mol Sci 12: 4836-4849.

Andolfatto P (2005) Adaptive evolution of non-coding DNA in *Drosophila*. Nature 437: 1149-1152.

André C, Larsson LC, Laikre L, Bekkevold D, Brigham J, Carvalho GR, Dahlgren TG, Hutchinson WF, Mariani S, Mudde K, Ruzzante DE, Ryman N (2011) Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity 106: 270-280.

Ansorge W (2009) Next-generation DNA sequencing techniques. New Biotechnology 25: 195-203.

Antao T, Lopes A, Lopes RL, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a F_{st}-outliner method. BMC Bioinformatics 9: 323.

Antunes A, Troyer JL, Roelke ME, Pecon-Slattery J, Packer C, Winterbach C, Winterbach H, Hemson G, Frank L, Stander P, Siefert L, Driciru M, Funston PJ, Alexander KA, Prager KC, Mills G, Wildt D, Bush M, O'Brien SJ, Johnson WE (2008) The evolutionary dynamics of the lion *Panthera leo* revealed by host and viral population genomics. PLoS Genet 4: e1000251.

Appleyard SA, Ward RD, Williams R (2002). Population structure of the Patagonian toothfish around Heard, McDonald and Macquarie Islands. Antarct Sci 14: 364-373.

Arnaud-Haond S, Monteforte M, Blanc F, Bonhomme F (2003) Evidence for male-biased effective sex ratio and recent step-by-step colonization in the bivalve *Pinctada mazatlanica*. J Evol Biol 16:790-796.

Arndt A, Smith J (1998) Genetic diversity and population structure in two species of sea cucumber: differing patterns according to mode of development. Mol Ecol 7: 1053-1064.

Arthur W (2002) The emerging conceptual framework of evolutionary developmental biology. Nature 415: 757-764.

Badyaev AV (2011) Origin of fittest: link between emergent variation and evolutionary change as a critical question in evolutionary biology. Proc R Soc B 278: 1921-1929.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13:729-744

Banks SC, Piggott MP, Williamson JE, Bové U, Holbrook NJ, Beheregaray LB (2007) Oceanic variability and coastal topography shape local genetic structure in a long dispersing sea urchin. Ecology 88:3055-3064.

Baranski M, Rourke M, Loughnan S, Austin CM, Robinson N (2006a) Isolation and characterization of 125 microsatellite DNA markers on blacklip abalone, *Haliotis rubra*. Mol Ecol Notes 6: 740-746.

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

Baranski M, Rourke M, Loughnan S, Hayes B, Austin C, Robinson N (2008) Detection of QTL for growth rate in the blacklip abalone (*Haliotis rubra* Leach) using selective DNA pooling. Anim Genet 39: 606-614.

Barkai R, Griffiths CL (1986) Diet of the South African abalone *Haliotis midae*. S Afr J Mar Sci 4: 37-44.

Bazin E, Glémin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. Science 312: 570-572.

Beaumont MA (2005) Adaptation and speciation: What can F_{st} tell us? Trends Ecol 20: 435-440.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B 263: 1619-1626.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol 13: 969-980.

Beheregaray LB, Sunnucks P (2001) Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis.* Mol Ecol 10: 2849-2866.

Bentsen HB, Olesen I (2002) Designing aquaculture mass selection programs to avoid high inbreeding rates. Aquaculture 204: 349-359.

Bentsen HB, Eknath AE, Palada-de Vera MS, Danting JC, Bolivar HL, Reyes RA, Dionisio EE, Longalong FM, Circa AV, Tayamen MM, Gjerde B (1998) Genetic improvement of farmed tilapias: growth performance in a complete diallel cross experiment with eight strains of *Oreochromis niloticus*. Aquaculture 160: 145-173.

Bergman A, Siegal ML (2003) Evolutionary capacitance as a general feature of complex gene networks. Nature 424: 549-552.

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

Bester AE, Roodt-Wilding R, Whitaker HA (2008) Discovery and evaluation of single nucleotide polymorphisms (SNPs) for *Haliotis midae*: a targeted EST approach. Anim Genet 39: 321-324.

Bester-van der Merwe AE, Roodt-Wilding R, Volckaert FAM, D'Amato ME (2011) Historical isolation and hydrodynamically constrained gene flow in declining populations of the South-African abalone, *Haliotis midae.* Conserv Genet 12: 543-555.

Bester-van der Merwe AE, D'Amato ME, Swart BL, Roodt-Wilding R (2012) Molecular phylogeny of South African abalone, its origin and evolution as revealed by two genes. Mar Biol Res 8: 727-736.

Beuzen ND, Stear MJ, Chang KC (2000) Molecular markers and their use in animal breeding. Vet J 160: 42-52.

Bieler R (1992) Gastropod phylogeny and systematics. Annu Rev Ecol Syst 23: 311-338.

Biswas S, Akay JM (2006) Genomic insights into positive selection. Trends Genet 22: 437-446.

Blaauw S (2012) SNP screening and validation in *Haliotis midae*. Unpublished M.Sc. thesis, Stellenbosch University, South Africa.

Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: Genome-wide sampling of insect populations. Annul Rev Entomol 46: 441-469.

Bonin A (2008) Population genomics; A new generation of genome scans to bridge the gap with functional genomics. Mol Ecol 17: 3583-3584.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). Mol Biol Evol 23: 773-783.

Bossdorf O, Richards CL, Pigliucci M (2008) Epigenetics for ecologists. Ecol Lett 11: 106-115.

Britz PJ (1994) The development of an artificial feed for abalone farming. S Afr J Sci 90: 6-7.

Britz PJ, Hecht T, Mangold S (1997) Effect of temperature on growth, feed consumption and nutritional indices of *Haliotis midae* fed a formulated diet. Aquaculture 152: 191-203.

Brooker AL, Benzie JAH, Blair D, Versini J-J (2000) Population structure of the giant tiger prawn *Penaeus monodon* in the Austrailian waters determined using microsatellite markers. Mar Biol 136: 149-157.

Brown LD (1991) Genetic variation and population structure in the Blacklip abalone, *Haliotis rubra*. Aust J Mar Freshwater Res 42: 77-90.

Brown LD, Murray ND (1992) Genetic relationships within the genus *Haliotis*. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books, Wiley-Blackwell, New Jersey. pp. 19-23.

Brown RC, Woolliams JA, McAndrew BJ (2005) Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247: 219-225.

Brumfield RT, Beerli P, Nickers DA, Edwards SV (2003) The utility of single nucleotide polymorphism in inferences of population history. Trends Ecol Evol 18: 249-256.

Buonaccorsi VP, McDowell JR, Graves JE (2001). Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). Mol Ecol 10: 1179-1196.

Camara MD, Vadopalas B (2009) Genetic aspects of restoring olympia oysters and other native bivalves: balancing the need for action, good intentions, and the risks of making things worse. J Shellfish Res 28: 121-145.

Campbell D, Bernatchez L (2004) Genetic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Mol Biol Evol 21: 945-956.

Cardellino RA, Boyazoglu J (2009) Research opportunities in the field of animal genetic resources. Livest Sci 120: 166-173.

Carroll SB (2008) Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. Cell 134: 25-36.

Cenadelli S, Maran V, Bongioni G, Fusetti L, Parma P, Aleandri R (2007) Identification of nuclear SNPs in gilthead seabream. J Fish Biol 70: 399-405.

Chambers MD, Van Blaricom GR, Hauser L, Utter F, Friedman CS (2006) Genetic structure of black abalone (*Haliotis cracherodii*) populations in the California islands and central California coast: Impacts of larval dispersal and decimation from withering syndrome. J Exp Mar Biol Ecol 331: 173-185.

Chelaifa H, Hahé F, Ainouche M (2010) Transcriptome divergence between the hexaploid slat-march sister species *Spartina maritime* and *Spartina alterniflora* (Poaceae). Mol Ecol 19: 2050-2063.

Cheviron ZA, Whitehead A, Brumfield RT (2008) Transcriptomic variation and plasticity in rufous-collared sparrows (*Zonotrichia capensis*) along an altitudinal gradient. Mol Ecol 17: 4556-4569.

Chistiakov DA, Hellemans 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.

Coleman AW, Vacquier VD (2002) Exploring the phylogenetic utility of ITS sequences for animals: A test case for abalone (*Haliotis*). J Mol Evol 54: 246-257.

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142: 169-196.

Coluccio MP, Sanchez SE, Kasulin L, Yanovsky MJ, Botto JF (2011) Genetic mapping of natural variation in a shade avoidance response: ELF3 is the candidate gene for a QTL in hypocotyl growth regulation. J Exp Bot 62: 167-176.

Consuegra S, de Leaniz CG (2007). Fluctuating sex ratios, but no sex-biased dispersal, in a promiscuous fish. Evol Ecol 21: 229-245.

Cook PA (1998) The current status of abalone farming in South Africa. J Shellfish Res 17: 601-602.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. Genetics 185: 1411-1423.

Counterman BA, Araujo-Perez F, Hines HM, Baxter SW, Morrison CM, Lindstrom DP, Papa R, Ferguson L, Joron M, French-Constant RH, Smith CP, Nielsen DM, Chen R, Jiggins CD, Reed RD, Halder G, Mallet J, McMillan WO (2010) Genomic hotspots for adaptation: The population genetics of Mullerian mimicry in *Heliconius erato*. PLOS Genet 6: e1000796.

Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London.

Davis GP, Hetzel DJS (2000) Integrating molecular genetic technology with traditional approaches for genetic improvement in aquaculture species. Aquac Res 31: 3-10.

Day E, Branch GM (2000) Relationships between recruits of abalone *Haliotis midae*, encrusting corallines and the sea urchin *Parechinus angulosus*. S Afr J Mar Sci 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*). Mol Phylogenet Evol 41: 249-256.

Dekkers JCM (2004) Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. J Anim Sci 82 (E Suppl): E313-E328.

De la Cruz F-L, Del Rio Portilla M, Gallardo-Escarate C (2010) Genetic variability of cultured populations of red abalone in Chile: An approach based on heterologous microsatellites. J Shellfish Res 29: 709-715.

Delisle RG (2010) What was really synthesised during the evolutionaty synthesis? A historiographic proposal. Studies in the History and Philosophy of Biological and Biomedical Sciences 42: 50-59.

Deng YW, Xiao L, Zhang GF, Zhao HG (2007) Genetic parameters estimated for growth traits at early stage of Pacific abalone, *Haliotis discus hannai* Ino. Acta Oceanol Sin 26: 90-95.

Department of Agriculture, Forestry and Fisheries (2010) The conditional re-opening of commercial abalone fishery in the 2010 season - Progress report to cabinet. Available from www.pmg.org.za/files/docs/110208reopening_0.ppt. [Accessed 31 October 2012]

Department of Agriculture, Forestry and Fisheries (2012) Aquaculture Annual Report 2011, South Africa.

Derome N, Bernatchez L (2006) The transcriptomics of ecological convergence between 2 limnetic Coregonine fishes (*Salmonidae*). Mol Biol Evol 23: 2370-2378.

Derome N, Duchesne P, Bernatchez L (2006) Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchill) ecotypes.Mol Ecol 15: 1239-1249.

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

De Wit P, Palumbi SR (2012) Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol Ecol doi: 10.1111/mec.12081

Diaz-Almela E, Boudry P, Launey S, Bonhomme F, Lapegue S (2004) Reduced female gene flow in the European flat oyster *Ostrea edulis*. J Hered 95: 510-516.

Díaz-Viloria N, Cruz P, Guzmán-Del Próo SA, Perez-Enriquez R (2009) Genetic connectivity among pink abalone *Haliotis corrugata* populations. J Shellfish Res 28:599-608.

Dichmont CM, Butterworth DS, Cochrane KL (2000) Towards adaptive approaches to management of the South African abalone *Haliotis midae* fishery. S Afr J Mar 22: 33-42.

Dijkstra HA, de Ruijter WPM (2001) On the physics of the Agulhas current: steady retroflection regimes. J Phys Oceanogr 31: 2971-2985.

Dixon CD, Day RW, Huchette SMH, Shepherd SA (2006) Successful seeding of hatchery-produced juvenile Greenlip abalone to restore wild stocks. Fish Res 78: 179-185.

Dixon TJ, Coman GJ, Arnold SJ, Sellars MJ, Lyons RE, Dierens L, Preston NP, Li Y (2008) Shifts in genetic diversity during domestication on Black Tiger shrimp, *Penaeus monodon*, monitored using two multiplexed microsatellite systems. Aquaculture 283: 1-6.

Dobney K, Larson G (2006) Genetics and animal domestication: new windows on an elusive process. J Zool 269: 261-271.

Dobzhansky T (1937) Genetics and the origin of species. Columbia University Press, New York.

Dong SR, Kong J, Zhang TS, Meng XH, Wang RC (2006) Parentage determination of Chinese shrimp (*Fenneropenaeus chinensis*) based on microsatellite DNA markers. Aquaculture 258: 283-288.

Dunham RA, Brummett RE (1999) Response of two generations of selection to increased body weight in channel catfish, *Ictalurus punctatus* compared to hybridization with blue catfish, *I. furcatus*, males. J Appl Aquac 9: 37-45.

Du Plessis (2012) Medium-throughput SNP genotyping and linkage mapping in *Haliotis midae.* Unpublished MSc thesis, Stellenbosch University, South Africa.

Durham JW (1979a) A fossil *Haliotis* from the Galápagos Island. The Veliger 21: 369-372.

Durham JW (1979b) California Cretaceous *Haliotis.* The Veliger 21: 373-374.

Ebert EE (1992) Abalone aquaculture: a North American regional review. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books. Wiley-Blackwell, New Jersey. pp. 570-582.

Egan SP, Nosil P, Funk DJ (2008) Selection and genomic differentiation during ecological speciation: Isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Evolution 62: 1162-1181.

Eknath AE, Bentsen HB, Ponzoni RW, Rye M, Nguyen NH, Thodesen J, Gjerde B (2007) Genetic improvement of farmed tilapias: Composition and genetic parameters of a synthetic base population of *Oreochromis niloticus* for selective breeding. Aquaculture 273: 1-14.

Elfstrom CM, Gaffney PM, Smith CT, Seeb JE (2005) Characterization of 12 nucleotide polymorphisms in weathervane scallop. Mol Ecol Notes 5: 406-409.

Elliott NG (2000) Genetic improvement programmes in abalone: what is the future. Aquac Res 31: 51-59.

Emanuel BP, Bustamante RH, Branch GM, Eekhout S, Odendaal FJ (1992) A zoogeographic and functional approach to the selection of marine reserves on the West coast of South Africa. S Afri J Mar Sci 12: 341-354.

Estes JA, Lindberg DR, Wray C (2005) Evolution of large body size in abalone (*Haliotis*): Patterns and implications. Paleobiology 31: 591-606.

Evans BS, White W, Elliott NG (2000) The use of microsatellite markers for parentage analysis in Australian blacklip and hybrid abalone. J Shellfish Res 19: 511.

Evans BS, Sweijd NA, Bowie RCK, Cook PA, Elliott NG (2004a) Population genetics structure of the perlemoen *Haliotis midae* in South Africa: evidence of range expansion and founder events. Mar Ecol Prog Ser 270: 163-172.

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

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103: 285-298.

Fisher RA (1930) The genetical theory of natural selection. Clarendon Press, Oxford.

Fleming AE, Hone PW (1996) Abalone aquaculture. Aquaculture 140: 1-4.

Flint APF, Woolliams JA (2008) Precision animal breeding. Philos T Roy Soc B 363: 573-590.

Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180: 977-993.

Franchini P, Slabbert R, Van der Merwe M, Roux A, Roodt-Wilding R (2010) Karyotype and genome size estimation in *Haliotis midae*: Estimators to assist future studies on the evolutionary history of the Haliotidae. J Shellfish Res 29: 945-950.

Franchini P, Van der Merwe M, Roodt-Wilding R (2011) Transcriptome characterization of the South African abalone *Haliotis midae* using sequencing-by-synthesis. BMC Research Notes 4: 59.

Gaffney PM, Rubin VP, Hedgecock D, Powers DA, Morris G, Hereford L (1996) Genetic effects of artificial propagation: signals from wild and hatchery populations of red abalone in California. Aquaculture 143: 257-266.

Galindo BE, Vacquier VD, Swanson WJ. 2003. Positive selection in the egg receptor for abalone sperm lysin. Proc Natl Acad Sci USA 100: 4639-4643.

Galindo K, Grahame JW, Bultin RK (2010) An EST-based genome scan using 454 sequencing in the marine snail *Littorina saxatilis*. J Evol Biol 23: 2004-2016.

Gamborg C, Sandøe P (2005) Sustainability in farm animal breeding: a review. Livest Prod Sci 92: 221-231.

Garza AES, Bernal RS (1992) Development and present status of abalone culture in Mexico. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books, Wiley-Blackwell, New Jersey. pp. 538-546.

Ge H, Walhout AJM, Vidal M (2003) Integrating 'omic' information: a bridge between genomics and systems biology. Trends Genet 19: 551-560.

Geiger DL (1999) A total evidence cladistic analysis of the Haliotidae (Gastrooda, Vetigastropoda). Unpublished PhD thesis, University of Southern California, USA.

Geiger DL (2000) Distribution and biogeography of the Haliotidae (Gastropoda: Vetigastropoda) world-wide. Boll Malacol 35: 57-120.

Geiger DL, Groves LT (1999) Review of fossil abalone (Gastropoda: Vetigastropoda: Haliotidae) with comparison to recent species. J Paleont 73: 872-885.

Geiger DL, Poppe GT (2000) A conchological iconography. The family Haliotidae. Conchbooks, Germany.

Genade AB, Hirst AL, Smit CJ (1988) Observations on the spawning, development and rearing of the South African abalone *Haliotis midae* Linn. S Afr J Mar Sci 6: 3-12.

Gheyas AA, Woolliams JA, Taggart JB, Sattar MA, Das TK, McAndrew BJ, Penman DJ (2009) Heritability estimation of silver carp (*Hypophthalmichthys molitrix*) harvest traits using microsatellite based parentage assignment. Aquaculture 294: 187-193.

Giger T, Excoffier L, Amstutz U, Day PJR, Champigneulle A, Hansen MM, Kelso J, Largiadèr CR (2008) Population transcriptomics of life-history variation in the genus *Salmo*. Mol Ecol 17: 3095-3108.

Gilad Y, Rifkin SA, Pritchard JK (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. Trends Genet 24: 408-415.

Gilbert SF (2003) Opening Darwin's black box: teaching evolution through developmental genetics. Nat Rev Genet 4: 735-741.

Gilg MR, Hilbish TJ (2003) Dispersal patterns of mussel larvae throughout a hybrid zone in Southwest England. Evolution 57: 1061-1077.

Giribet G, Okusu A, Lindgren AR, Huff SW, Schrödl M, Nishiguchi MK (2006) Evidence for a clade composed of molluscs with serially repeated structures: monoplacophorans are related to chitons. Proc Natl Acad Sci USA 103: 7723-7728.

Gjedrem T (2000) Genetic improvement of cold water fish species. Aquac Res 31: 25-33.

Glazko GV, Koonin EV, Rogozin IB, Shabalina SA (2003) A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions. Trends Genet 19: 119-124.

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10: 381-391.

Goldstein DB, Weale ME (2001) Population genomics: Linkage disequilibrium hold the key. Curr Biol 11: R576-R579

Gopal K, Tolley KA, Groeneveld JC, Matthee CA (2006) Mitochondrial DNA variation in spiny lobster *Palinurus delagoae* suggests genetically structured populations in the south-western Indian Ocean. Mar Ecol Prog Ser 319: 191-198.

Gorbach DM, Hu ZL, Du ZQ, Rothschild MF (2009) SNP discovery in *Litopenaeus vannamei* with a new computational pipeline. Anim Genet 40: 106-109.

Groves LT, Alderson JM (2008) Earliest record of the genus *Haliotis* (Mollusca: Gastropoda) from the Late Cretaceous (Campanian) of Los Angeles County, California. The Veliger 50: 24-26.

Gruenthal KM, Burton RS (2008) Genetic structure of natural populations of the California black abalone (*Haliotis cracherodii* Leach, 1814), a candidate for endangered species status. J Exp Mar Biol Ecol 355: 47-58.

Gruenthal KM, Acheson LK, Burton RS (2007) Genetic structure of natural populations of California red abalone (*Haliotis rufescens*) using multiple genetic markers. Mar Biol 152: 1237-1248.

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.

Gyllensten I (1985) The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous, and freshwater species. J Fish Biol 26: 691-699.

Haddrill PR, Bachtrog D, Andolfatto P (2008) Positive and negative selection on noncoding DNA in *Drosophila simulans*. Mol Biol Evol 25: 1825-1834.

Hagenblad J, Olsson M, Parker HG, Ostranders EA (2009) Population genomics of the inbred Scandinavian wolf. Mol Ecol 18: 1341-1351.

Haldane JBS (1932) The causes of evolution. Longmans, Green and Co., London.

Halligan DL, Keightley PD (2006) Ubiquitous selective constraints in the *Drosophila* genome revealed by a genome-wide interspecies comparison. Genome Res 16: 875-884.

Hamasaki K, Kitada S (2008) The enhancement of abalone stocks: lessons from Japanese case studies. Fish Fish 9: 243-260.

Hamm DE, Burton RS (2000) Population genetics of black abalone, *Haliotis cracherodii*, along the central California coast. J Exp Mar Biol Ecol 254: 235-247.

Hansen MM, Olivieri I, Waller DM, Nielsen EE, The GeM Working Group (2012) Monitoring adaptive genetic responses to environmental change. Mol Ecol 21: 1311-1329.

Hara M, Sekino M (2005) Genetic difference between Ezo-awabi *Haliotis discus hannai* and Kuro-awabi *H. discus discus* populations: Microsatellite-based population analysis in Japanese abalone. Fisheries Science 71: 754-766.

Hara M, Sekino M (2007) Genetic differences between hatchery stocks and natural populations in Pacific abalone (*Haliotis discus*) estimated using microsatellite DNA markers. Mar Biotechnol 9: 74-81.

Hara M, Onoue S, Taniguchi N (2008) Assessing the impact of releasing exogenous hatchery-reared juveniles of Pacific Abalone, *Haliotis discus*. Rev Fish Sci 16: 278-284.

Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA (2009) Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biol 10: R32.

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

Hayes B, Goddard M (2010) Genome-wide association and genomic selection in animal breeding. Genome 53: 876-883.

Hayes B, He J, Moen T, Bennewits J (2006) Use of molecular markers to maximise diversity of founder populations for aquaculture breeding programs. Aquaculture 255: 573-578.

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

Hayes B, Laerdahl JK, Lien S, Moen T, Berg P, Hindar K, Davidson WS, Koop BF, Adzhubei A, Hoyheim B (2007b) An extensive resource of single nucleotide polymorphism markers associated with Atlantic salmon (*Salmo salar*) expressed sequences. Aquaculture 265: 82-90.

He C, Chen L, Simmons M, Li P, Kim S, Liu ZJ (2003) Putative SNP discovery in inter-specific hybrids of catfish by comparative EST analysis. Anim Genet 34: 445-448.

Hellberg ME (1996) Dependance of gene flow on geographic distance in two solitary corals with different larval dispersal capabilities. Evolution 50: 1167-1175.

Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). Heredity 99: 592-600.

Herrera CN, Bazaga P (2008) Population-genomc approach reveals adaptive floral divergence in discrete populations of hawk moth-pollinated violet. Mol Ecol 17: 5378-5390.

Hilbish TJ (1996) Population genetics of marine species: the interaction of natural selection and historically differentiated populations. J Exp Mar Biol Ecol 200: 67-83.

Hipp AL, Rothrock PE, Reznicek AA, Berry PE (2007) Chromosome number changes associated with speciation in sedges: a phylogenetic study in *Carex* section *Ovales* (Cyperaceae) using AFLP data. Aliso 23: 193-203.

Hobday AJ, Tegner MJ (2002) The warm and the cold: Influence of temperature and fishing on the local population dynamics of the red abalone. CalCOFI Rep 43: 74-96.

Hobday AJ, Tegner MJ, Haaker PL (2001) Over-exploitation of a broadcast spawning marine invertebrate: Decline of the white abalone. Rev Fish Biol Fisher 10: 493-514.

Hoffman JI, Peck LS, Linse K, Clarke A (2011) Strong population genetic structure in a broadcast-spawning Antarctic marine invertebrate. J Hered 102: 55-66.

Hoffman JI, Clarke A, Clark MS, Fretwell P, Peck LS (2012) Unexpected fine-scale population structure in a broadcast-spawning Antarctic marine mollusc. PLoS ONE 7: e32415.

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Mol Ecol Res 11 (Suppl. 1): 117-122.

Holloway AK, Lawniczak MKN, Mezey JG, Begun DJ, Jones CD (2007) Adaptive gene expression divergence inferred from population genomics. PLoS Genet 3: e187.

Horreo JL, Machado-Schiaffino G, Griffiths A, Bright D, Stevens J, Garcia-Vazquez E (2008) Identification of differential broodstock contribution affecting genetic variability in hatchery stocks of Atlantic salmon (*Salmo salar*). Aquaculture 280: 89-93.

Huang BX, Peakall R, Hanna PJ (2000) Analysis of genetic structure of blacklip abalone (*Haliotis rubra*) populations using RAPD, minisatellite and microsatellite markers. Mar Biol 136: 207-216.

Hubert S, Bussey JT, Higgins B, Curtis BA, Bowman S (2009) Development of single nucleotide polymorphism markers for Atlantic cod (*Gadus morhua*) using expressed sequences. Aquaculture 296: 7-14.

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

Hundertmark KJ, Van Daele LJ (2010) Founder effect and bottleneck signatures in an introduced, insular population of elk. Conserv Genet 11: 139-147.

Jensen P, Andersson L (2005) Genomics meets Ethology; A new route to understanding domestication, behaviour, and sustainability in animal breeding. Ambio 34: 320-324.

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

Jiang J-Z, Zhang W, Guo Z-X, Cai C-C, Su Y-L, Wang R-X, Wang J-Y (2011) Functional annotation of an expressed sequence tag library from *Haliotis diversicolor* and analysis of its plant-like sequences. Marine Genomics 4: 189-196.

Jonasson J, Stefansson SE, Gudnason A, Steinarsson A (1999) Genetic variation for survival and shell length of cultured red abalone (*Haliotis rufescens*) in Iceland. J Shellfish Res 18: 621-625.

Jones JW, Hallerman EM, Neves RJ (2006) Genetic management guidelines for captive propagation of freshwater mussels (*Unionoidea*). J Shellfish Res 25: 527-535.

Joost S, Bonin A, Bruford MW, Després L, Conord C, Erhardt G, Taberlet P (2007) A Spatial Analysis Method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. Mol Ecol 16: 3955-3969.

Kang J-H, Appleyard SA, Elliott NG, Jee Y-J, Lee BJ, Kang SW, Baek MK, Han YS, Choi T-J, Lee YS (2011) Development of genetic markers in abalone through construction of a SNP database. Anim Genet 42: 309-315.

Kawahara I, Noro T, Omori M, Hasekura O, Kijima A (1997) Genetic progress for growth in different selected populations of abalone *Haliotis discus hannai*, at different hatcheries. Fish Genetics and Breeding Science 25: 81-90.

Keller SR, Olsen MS, Silim S, Schroeder W, Tiffin P (2010) Genomic diversity, population structure, and migration following rapid range expansion in the Balsam Poplar, *Populus balsamifera.* Mol Ecol 19: 1212-1226.

Kerje S, Carlborg O, Jacobsson L, Schütz K, Hartmann C, Jensen P, Andersson L (2003) The twofold difference in adult size between the red junglefowl and White leghorn chickens is largely explained by a limited number of QTLs. Anim Genet 34: 264-274.

Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Pääbo S (2004) A neutral model of transcriptome evolution. PLoS Biol 2: 682-689.

Kimura M (1968) Evolutionary rate at the molecular level. Nature 217: 624-626.

Kimura M (1983) The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.

Kimura M (1991) Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics. Proc Natl Acad Sci USA 88: 5969-5973.

King JL, Jukes TH (1969) Non-Darwinian evolution. Science 164: 788-798.

Kjoglum S, Henryon M, Aasmundstad T, Korsgaard I (2008) Selective breeding can increase resistance of Atlantic salmon to furunculosis, infectious salmon anaemia and infectious pancreatic necrosis. Aquac Res 39: 498-505.

Knutsen H, Jorde PE, André C, Stenseth N (2003) Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. Mol Ecol 12: 385-394.

Koonin EV (2009a) The origin at 150: is a new evolutionary synthesis in sight? Trends Genet 25: 473-475.

Koonin EV (2009b) Towards a postmodern synthesis of evolutionary biology. Cell Cycle 8: 799-800.

Koonin EV (2009c) Darwinian evolution in the light of genomics. Nucleic Acids Res 37: 1011-1034.

Kresge N, Vacquier VD, Stout CD (2001) Abalone lysin: the dissolving and evolving sperm protein. BioEssays 23: 95-103.

Kube PD, Appleyard S, Elliott NG (2007) Selective breeding greenlip abalone (*Haliotis leavigata*): preliminary results and issues. J Shellfish Res 26: 821-824.

Kutschera U, Niklas KJ (2004) The modern theory of biological evolution: an expanded synthesis. Naturwissenschaften 91: 255-276.

Kyle CJ, Boulding EG (2000) Comparative population genetic structure of marine gastropods (*Littorina* spp.) with and without palegic larval dispersal. Mar Biol 137: 835-845.

Lande R (2009) Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. Evol Biol 22: 1435-1446.

Langdon C, Evans F, Jacobson D, Blouin M (2003) Yields of cultured Pacific oysters *Crassostrea gigas* Thunberg improved after one generation of selection. Aquaculture 220: 227-244.

Larmuseau MHD, Raeymaekers JAM, Hellemans B, Van Houdt JKJ, Volckaert FAM (2010) Mito-nuclear discordance in the degree of population differentiation in a marine goby. Heredity 105: 532-542.

Lee YH, Vacquier VD (1995) Evolution and systematics in *Haliotidae* (Mollusca, Gastropoda): Inference from DNA sequences of sperm lysin. Mar Biol 124: 267-278.

Lee YH, Ota T, Vacquier VD (1995) Positive selection is a general phenomenon in the evolution of abalone sperm. Mol Biol Evol 12: 231-238.

Levin L (2006) Recent progress in understanding larval dispersal: new directions and digressions. Integ Comp Biol 46: 282-297.

Li J, Li H, Jakobsson M, Li S, Sjödin P (2012) Joint analysis of demography and selection in population genetics: Where do we stand and where could we go. Mol Ecol 21: 28-44.

Li Q, Shu J, Yu R, Tian C (2007) Genetic variability of cultured populations of the Pacific abalone (*Haliotis discus hannai* Ino) in China based on microsatellites. Aquac Res 38: 981-990.

Li X, Havenhand J, Hone P (1999) Karyotype of the greenlip abalone *Haliotis laevigata*. In: Fleming AE (ed) Proceedings of the 6th Annual Abalone Aquaculture Workshop. April 1999. Sydney, Australia. Abalone Aquaculture Subprogram, Fisheries Research and Development Corporation, Canberra, Australia.

Li X, Ponsoni R, Brien B, Nguyen NH, Butterworth A (2005) Selective breeding of farmed abalone in Australia: preliminary genetic analysis of the data from blacklip abalone families established in Victoria in the summer of 2000/2001. In: Fleming AE (ed.) Proceedings of the 12th annual abalone aquaculture workshop. August 2005. McLaren Vale, Australia. Abalone Aquaculture Subprogram, Fisheries Research and Development Corporation, Canberra, Australia.

Li Y, Álvarez OA, Gutteling EW, Tijsterman M, Fu J, Riksen JAG, Hazendonk E, Prins P, Plasterk RHA, Jansen RC, Breitling R, Kammenga JE (2006) Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. PLoS Genet 2: e222.

Li YF, Costello JC, Holloway AK, Hahn MW (2008) "Reverse ecology" and the power of population genomics. Evolution 62: 2984-2994.

Limborg MT, Helyar SJ, De Bruyn M, Taylor MI, Nielsen EE, Ogden R, Carvalho GR, FPT Consortium, Bekkevold D (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). Mol Ecol 21: 3686-703.

Lind CE, Evans BS, Knauer J, Taylor JJU, Jerry DR (2009) Decreased genetic diversity and a reduced effective population size in cultured silver-lipped pearl oysters (*Pinctada maxima*). Aquaculture 286: 12-19.

Lindberg DR (1992) Evolution, distribution and systematics of Haliotidae. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books, Wiley-Blackwell, New Jersey. pp. 3-18.

Linnemann AK, Platts AE, Krawetz SA (2009) Differential nuclear scaffold/matrix attachment marks expressed genes. Hum Mol Genet 18: 645-654.

Liu S, Zhou Z, Lu J, Sun F, Wang S, Liu H, Jiang Y, Kucuktas H, Kaltenboeck L, Peatman E, Liu Z (2011) Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. BMC Genomics 12: 53.

Liu X, Liu X, Gou X, Gau Q, Zhao H, Zhang G (2006) A preliminary genetic linkage map of the Pacific abalone *Haliotis discus hannia* Ino. Mar Biotechnol 8: 386-397.

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

Liu ZJ, Cordes JF (2004) DNA marker technologies and their application in aquaculture genetics. Aquaculture 238: 1-37.

Long AD, Langley CH (1999) The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. Genome Res 9: 720-731.

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.

Luikart G, Englang PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4: 981-994.

Lukoschek V, Waycott M, Keogh JS (2008). Relative information content of polymorphic microsatellites and mitochondrial DNA for inferring dispersal and population genetic structure in the olive sea snake, *Aipysurus laevis*. Mol Ecol 17: 3062-3077.

Macey BM, Coyne VE (2005) Improved growth rate and disease resistance in farmed *Haliotis midae* through probiotic treatment. Aquaculture 245: 249-261.

Macey BM, Christison KW, Mouton A (2011) *Halioticida noduliformans* isolated from cultured abalone (*Haliotis midae*) in South Africa. Aquaculture 315: 187-195.

Maglich JM, Caravella JA, Lambert MH, Willson TM, Moore JT, Ramamurthy L (2003) The first completed genome sequence from a teleost fish (*Fugu rubripes*) adds significant diversity to the nuclear receptor superfamily. Nucleic Acids Res 31: 4051-4058.

Manel S, Schwartz M, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18: 189-197.

Mardis ER (2008) The impact of next generation sequencing technology on genetics. Trends Genet 24: 133-141.

Mariac C, Jehin L, Saidou A-A, Thuillet AC, Couderc M, Sire P, Jugde´ H, Adam H, Bezancon G, Pham J-L, Vigouroux Y (2011) Genetic basis of pearl millet adaptation along an environmental gradient investigated by a combination of genome scan and association mapping. Mol Ecol 20: 80-91.

Masel J (2005) Evolutionary capacitance may be favored by natural selection. Genetics 170: 1359-1371.

Massault C, Bovenhuis H, Haley CS, De Koning DJ (2008) QTL mapping designs for aquaculture. Aquaculture 285: 23-29.

Massault C, Hellemans B, Louro B, Batargias C, Van Houdt JKJ, Canario A, Volckaert FAM, Bovenhuis H, Haley C, De Koning DJ (2009) QTL for body weight, morphometric traits and stress response in European sea bass *Dicentrarchus labrax*. Anim Genet 41: 337-345.

Matthee CA, Fourie F, Oosthuizen WH, Meyër MA, Tolley KA (2006) Mitochondrial DNA sequence data of the Cape fur seal (*Arctocephalus pusillus pusillus*) suggests that population numbers may be affected by climatic shifts. Mar Biol 148: 899-905.

Matthews B, Narwani A, Hausch S, Nonaka E, Peter H, Yamamichi M, Sullam KE, Bird KC, Thomas MK, Hanley TC, Turner CB (2011) Toward an integration of evolutionary biology and ecosystem science. Ecol Lett 14: 690-701.

Mayr E (1942) Systematics and the Origin of Species. Columbia University Press, New York.

Mayr E (1993) What was the evolutionary synthesis? Trends Evol Ecol 8: 31-34.

Mayr E (1996) The modern evolutionary theory. J Mammal 77: 1-7.

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little L, Ioannidis JPA, Hirschhorn JN (2008) Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 9: 356-396.

McShane PE (1992) Early life history of abalone: a review. In: Shepherd SA, Tegner Mia J, Guzman Del Proo SA (eds) Abalone of the World. Biology, Fisheries and Culture. Fishing News Books, Wiley-Blackwell, New Jersey, pp. 120-140.

McShane PE, Mercer SF, Naylor JR (1994) Spatial variation and commercial fishing of New Zealand abalone (*Haliotis iris* and *Haliotis australis*). New Zeal J Mar Fresh 28: 345-355.

Medugorac I, Medugorac A, Russ I, Veit-Kensch CE, Taberlet P, Luntz B, Mix HM, Föster M (2009) Genetic diversity of European cattle breeds with high effective population size. Mol Ecol 19: 3394-3410.

Mendel G (1866) Versuche über Pflanzen-Hybriden. Verh Naturforsch Ver Brünn 4: 3-47 (English translation (1901): Experiments on plant hybridization. J R Hortic Soc 26: 1-32).

Mignon-Grasteau S, Boissy A, Bouix J, Faure J-M, Fisher AD, Hinch GN, Jensen P, Le Neindre P, Morméde P, Prunet P, Vandeputte M, Beaumont C (2005) Genetics oa adaptation and domestication in livestock. Livest Prod Sci 93: 3-14.

Miller KJ, Maynard BT, Mundy CN (2009) Genetic diversity and gene flow in collapsed and healthy abalone fisheries. Mol Ecol 18: 200-211.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat Rev Genet 8: 845-856.

Morin PA, McCarthy M (2007) Highly accurate SNP genotyping from historical and low-quality samples. Mol Ecol Notes 7: 937-946.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biol 51: 238-254.

Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92: 255-264.

Mouse genome sequencing consortium (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520-562.

Mouton A, Gummow B (2011) The occurrence of gut associated parasites in the South African abalone, *Haliotis midae*, in Western Cape aquaculture facilities. Aquaculture 313: 1-6.

Müller GB (2007) Evo-devo: extending the evolutionary synthesis. Nat Rev Genet 8: 943-949.

Murray MC, Hare MP (2006) A genomic scan for divergent selection in a secondary contact zone between Atlantic and Gulf of Mexico oysters *Crassostrea virginica*. Mol Ecol 15: 4229-4242.

Nadeau NJ, Jiggins CD (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. Trends Genet 26: 484-492.

Namroud M-C, Beaulieu J, Juge N, LaRoche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in the white spruce. Mol Ecol 17: 3599-3613.

Naylor R, Hindar K, Fleming IA, Goldburg R, Williams S, Volpe J, Whoriskey F, Eagle J, Kelso D, Mangel M (2005) Fugitive salmon: Assessing the risks of escaped fish from net-pen aquaculture. BioScience 55: 427-437.

Neethling M, Matthee CA, Bowie RCK, von der Heyden S (2008) Evidence for panmixia despite barriers to gene flow in the southern African endemic, *Caffrogobius caffer* (Teleostei: Gobiidae). BMC Evol Biol 8: 325.

Nei ZQ (1992) A review of abalone culture in China. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books, Wiley-Blackwell, New Jersey. pp. 592-602.

Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer C, Taranger G-L, Ogden R, Carvalho GR (2009a) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC Evol Biol 9: 276.

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009b) Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18: 3128-3150.

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39: 197-218.

Normandeau E, Hutchings JA, Fraser DJ, Bernatchez L (2009) Population-specific gene expression responses to hybridization between farm and wild Atlantic salmon. Evol Appl 2: 489-503.

Nosil P, Funk DJ, Ortiz-Marrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18: 375-402.

Nunes VL, Beaumont MA, Butlin RK, Paulo OS (2011) Multiple approaches to detect outliers in a genome scan for selection in the ocellated lizards (*Lacerta lepida*) along an environmental gradient. Mol Ecol 20: 193-205.

Ohta T (1992) The nearly neutral theory of molecular evolution. Annu Rev Ecol Syst 23: 263-286.

Ohta T, Gillespie JH (1996) Development of neutral and nearly neutral theories. Theor Popul Biol 49: 128-142.

O'Niel ST, Dzurisin JDK, Carmichael RD, Lobo NF, Emrich SJ, Hellmann JJ (2010) Population-level transcriptome sequencing of nonmodel organisms *Erynnis propertius* and *Papilio zelicaon*. BMC Genomics 11: 310.

Orr HA (2005) The genetic theory of adaptation: A brief history. Nat Rev Genet 6: 119-127.

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. Trend Genet 26: 117-187.

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. Ecol Appl 13: 146-158.

Pampoulie C, Gysels E, Maes GE, Hellemans B, Leentjes V, Jones AG, Volckaert FAM (2004) Evidence for fine-scale genetic structure and estuarine colonisation in a potential high gene flow marine goby (*Pomatoschistus minutus*). Heredity 92: 434-445.

Paris M, Boyer S, Bonin A, Collado A, David J-P, Despres L (2010) Genome scan in the mosquito *Aedes rusticus*: population structure and detection of positive selection after insecticide treatment. Mol Ecol 19: 325-337.

Park C, Li Q, Kobayashi T, Kijima A (2006) Inbreeding depression traits in Pacific abalone *Haliotis discus hannai* by factorial mating experiments. Fish Sci 72: 774-780.

Passamaneck YJ, Schander C, Halanych KM (2004) Investigation of molluscan phylogeny using large-subunit and small-subunit nuclear rRNA sequences. Mol Phylogenet Evol 32: 25-38.

Pavlidis P, Hutter S, Stephan W (2008) A population genetic approach to map recent positive selection in model species. Mol Ecol 17: 3585-3598.

Peijnenburg KTCA, Fauvelot C, Breeuwer AJ, Menken SBJ (2006). Spatial and temporal genetic structure of the planktonic *Sagitta setosa* (Chaetognatha) in European seas as revealed by mitochondrial and nuclear DNA markers. Mol Ecol 15: 3319-3338.

Pérez-Enciso M, Ferretti L (2010) Massive parallel sequencing in animal genetics: wherefroms and wheretos. Anim Genet 41: 561-569.

Piertney SB, Webster MI (2010) Characterising functionally important and ecologically meaningful genetic diversity using a candidate gene approach. Genetica 138: 419-432.

Piggott MP, Banks SC, Tung P, Beheregaray L (2008) Genetic evidence for different scales of connectivity in a marine mollusc. Mar Ecol Prog Ser 365: 127-136.

Pigliucci M (2007) Do we need an extended evolutionary synthesis? Evolution 61: 2743-2749.

Pigliucci M (2008) Is evolvability evolvable? Nat Rev Genet 9: 75-82.

Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. J Exp Biol 209: 2362-2367.

Plagányi EE, Butterworth DS (2010) A spatial- and age-structured assessment model to estimate the impact of illegal fishing and ecosystem change on the South African abalone *Haliotis midae* resource. Afr J Mar Sci 32: 207-236.

Plagányi EE, Butterworth DS, Brandão A (2001) Toward assessing the South African abalone *Haliotis midae* stock using an age-structured production model. J Shellfish Res 20: 813-827.

Poompuang S, Hallerman EM (1997) Toward detection of quantitative trait loci and marker-assisted selection in fish. Rev Fish Sci 5: 253-277.

Pottinger TG, Moran TA, Morgan JAW (1994) Primary and secondary indices of stress in the progeny of rainbow trout (*Oncorhynchus mykiss*) selected for high and low responsiveness to stress. J Fish Biol 44: 149-163.

Praipue P, Klinbunga S, Jarayabhand P, (2010) Genetic diversity of wild and domesticated stocks of Thai abalone, *Haliotis asinina* (Haliotidae), analyzed by single-strand conformational polymorphism of AFLP-derived markers. Genet Mol Res 9: 1136-1152.

Prince JD, Sellers TL, Ford WB, Talbot SR (1987) Experimental evidence for limited dispersal of haliotid larvae (genus *Haliotis*, Mollusca: Gastropoda). J Exp Mar Biol Ecol 106: 243-263.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Qanbari S, Pimentel ECG, Tetens J, Thaller G, Lichtner P, Sharifi AR, Simianer H (2010) A genome-wide scan for signatures of selection in Holstein cattle. Anim Genet 41: 377-389.

Qi H, Liu X, Wu F, Zhang G (2008) Characterization of 12 single nucleotide polymorphisms (SNPs) in Pacific abalone, *Haliotis discus hannai*. Mol Ecol Res 8: 974-976.

Qi H, Liu X, Zhang G, Wu F (2009) Mining expressed sequences for single nucleotide polymorphisms in Pacific abalone *Haliotis discus hannai*. Aquac Res 40: 1661-1667.

Qi H, Liu X, Wu F, Zhang G (2010) Development of gene-targeted SNP markers for genomic mapping in Pacific abalone *Haliotis discus hannai* Ino. Mol Biol Reports 37: 3779-3784.

Quilang J, Wang S, Li P, Abernathy J, Peatman E, Wang Y, Wang L, Shi Y, Wallace R, Guo X, Liu Z (2007) Generation and analysis of ESTs from the eastern oyster *Crassostrea virginica* Gmelin and identification of microsatellite and SNP markers. BMC Genomics 8: 157-167.

Raemaekers S, Hauck M, Bürgener M, Mackenzie A, Maharaj G, Plagányi EE, Britz PJ (2011) Review of the causes of the rise of the illegal South African abalone fishery and consequent closure of the rights-based fishery. Ocean Coast Manage 54: 433-445.

Rauw WM, Kanis E, Noordhuizen-Stassen EN, Grommers FJ (1998) Undesirable side effects of selection for high production efficiency in farm animals: a review. Livest Prod Sci 56: 15-33.

Republic of South Africa (1998) Marine Living Resources Act no. 18 of 1998.

Rezk MA, Smitherman RO, Williams JC, Nichols A, Kucuktas H, Dunham RA (2003) Response to three generations of selection for increased body weight in channel catfish, *Ictalurus punctatus*, grown in earthen ponds. Aquaculture 228: 69-79.

Rhode C (2010) Development of gene-linked molecular markers in South African abalone (*Haliotis midae*) using an *in silico* mining approach. Unpublished MSc thesis, Stellenbosch University, South Africa.

Rhode C, Slabbert R, Roodt-Wilding R (2008) Microsatellite flanking regions: A SNP mine in South African abalone (*Haliotis midae*). Anim Genet 39: 329.

Ridgway TM, Stewart BA, Branch GM, Hodgson AM (1998) Morphological and genetic differentiation of *Patella granularis* (Gastropoda: Patellidae) along the coast of South Africa. J Zool 245: 317-333.

Riginos C, Cunningham CW (2005) Local adaptation and species segregation in two mussel (*Mytilus edulis*×*Mytilus trossulus*) hybrid zones. Mol Ecol 14: 381-400.

Robinson N, Li X, Hayes B (2010) Testing options for the commercialization of abalone selective breeding using bioeconomic simulation modeling. Aquac Res 41: e268-e288.

Roelofs D, Janssens TKS, Timmermans MJTN, Nota B, Mariën J, Bochdanovits Z, Ylstra B, Van Staalen NM (2009) Adaptive differences in gene expression associated with heavy metal tolerance in the soil arthropod *Orchesella cincta.* Mol Ecol 18: 3227-3239.

Rogers SM, Bernatchez L (2005) Integrating QTL mapping and gnome scans towards the characrerization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeafotmis*). Mol Ecol 14: 351-361.

Rokas A, Krüger D, Carroll SB (2005) Animal evolution and the molecular signature of radiations compressed in time. Science 310: 1933-1938.

Roodt-Wilding R (2007) Abalone ranching: a review on genetic considerations. Aquac Res 38: 1229-1241.

Roodt-Wilding R, Slabbert R (2006) Molecular markers to assist the South African abalone industry. S Afr J Sci 102: 99-102.

Roodt-Wilding R, Brink B (2011) Selection and sea snails: the South African story. Phil Trans Genet 1: 1-40.

Rose MR, Oakley TH (2007) The new biology: beyond the modern synthesis. Biol Direct 2: 30.

Roux A, Sandenbergh L, Roodt-Wilding R (2008) Preliminary investigation to determine the cytotoxicity of various cryoprotectants on southern African abalone (*Haliotis midae*) embryos. Cryobiology 57: 308-311.

Rubin C-J, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Halldöök F, Besnier F, Carlborg O, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. Nature 464: 587-591.

Ryynanen HJ, Primmer CR (2006) Single nucleotide polymorphism (SNP) discovery in duplicated genomes: intron-primed exon-crossing (IPEC) as a strategy for avoiding amplification of duplicated loci in Atlantic salmon (*Salmo salar*) and other salmonid fishes. BMC Genomics 7: 192.

Salem M, Rexroad CE, Wang J, Thorgaard GH, Yao J (2010) Characterization of the rainbow trout transcriptome using Sanger and 454-pyrosequencing approaches. BMC Genomics 11: 564.

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

Schierwater B, DeSalle R (2001) Current problems with the zootype and the early evolution of Hox genes. J Exp Zool 291: 169-174.

Schlichting CD, Smith H (2002) Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. Evol Ecol 16: 189-211.

Schlötterer C (2004) The evolution of molecular markers – just a matter of fashion? Nat Rev Genet 5: 63-68.

Schoener TW (2011) The newest synthesis: Understanding the interplay of evolutionary and ecological dynamics. Science 331: 426-429.

Schütz K, Kerje S, Carlborg O, Jacobsson L, Andersson L, Jensen P (2002) QTL analysis of a red junglefowl White Leghorn intercross reveals trade-off in resource allocation between behaviour and production traits. Behav Genet 32: 423-433.

Seddon JM, Parker HG, Ostrander EA, Ellegren H (2005) SNPs in ecology and conservation studies: a test in the Scandinavian wolf population. Mol Ecol 14: 503-511.

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

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.

Selkoe KA, Henzler CM, Gaines SD (2008) Seascape genetics and the spatial ecology of marine populations. Fish Fish 9: 363-377.

Selvamani MJP, Degnan SM, Paetkau D, Degnan BM (2000) Highly polymorphic microsatellite loci in the Heron Reef population of the tropical abalone *Haliotis asinina*. Mol Ecol 9: 1184-1186.

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

Shaw PW, Arkhipkin AI, Al-Khairulla H (2004). Genetic structuring of Patagonian toothfish populations in the Southwest Atlantic Ocean: the effect of the Antarctic Polar Front and deep-water troughs as barriers to genetic exchange. Mol Ecol 13: 3293-3303.

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26: 1135-1145.

Shi Y, Guo X, Gu Z, Wang A, Wang Y (2010) Preliminary genetic linkage map of the abalone *Haliotis diversicolor* Reeve. Chin J Oceanol Limn 28: 549-557.

Simon CA, Booth A (2007) Population structure and growth of polydorid polychaetes that infest the cultured abalone, *Haliotis midae*. Afr J Mar Sci 29: 499-509.

Simon CA, Bentley MG, Caldwell GS (2010) 2,4-decadienal: exploring a novel approach for the control of polychaete pests on cultured abalone. Aquaculture 310: 52-60.

Simpson GG (1944) Tempo and mode in evolution. Columbia University Press, New York.

Slabbert R (2010) Identification of growth related quantitative trait loci within the abalone *Haliotis midae,* using comparative microsatellite bulked segregant analysis. Unpublished PhD dissertation. Stellenbosch University, South Africa.

Slabbert R, Bester AE, D'Amato ME (2009) Analysis of genetic diversity and parentage of the abalone *Haliotis midae* Linnaeus using microsatellite markers. J Shellfish Res 28: 1-7.

Slabbert R, Ruivo NR, Van den Berg NC, Lizamore DL, Roodt-Wilding R (2008) Isolation and characterisation of 63 microsatellite loci for the abalone, *Haliotis midae*. J World Aquacult Soc 39: 429-435.

Slabbert R, Hepple J, Venter A, Nel S, Swart L, Van den Berg NC, Roodt-Wilding R (2010) Isolation and segregation of 44 microsatellite loci in the South African abalone, *Haliotis midae* L. Anim Genet 41: 332-333.

Slabbert R, Hepple J, Rhode C, Bester-van der Merwe AE, Roodt-Wilding R (2012). New microsatellite markers for the abalone *Haliotis midae* developed by 454 pyrosequencing and *in silico* analyses. Genet Mol Res 11: 2769-2779.

Smith CR, Elfstrom CM, Seeb L, Seeb JE (2005) Use of sequence data from rainbow trout and Atlantic salmon for SNP detection in Pacific salmon. Mol Ecol 14: 4193-4203.

Smith SA, Wilson NG, Goetz FE, Feehery C, Andrade SCS, Rouse GW, Giribet G, Dunn CW (2011) Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480: 364-367.

Sohl NF (1987) Cretaceous gastropods: contrast between Tethys and the temprate provinces. J Paleont 61: 1085-1111.

Sohl NF (1992) Upper Cretaceous gastropods (Fissurellidae, Haliotidae, Scissurellidae) from Puerto Rico and Jamaica. J Paleont 66: 414-434.

Squires JL (1987) Eocene molluscan paleontology of the Whitaker Peak area, Los Angeles and Ventura Counties, California. Los Angeles County Museum of History Contributions in Science, 338.

Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate J (2010) Adaptation genomics: the next generation. Trends Ecol Evol 25: 705-712.

Stebbins GL, Ayala FJ (1981) Is a new evolutionary synthesis necessary? Science 213: 967-971.

Steinberg J (2005) The illicit abalone trade in South Africa. Institute for Security Studies paper 105.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: Finding genes underlying ecological important traits. Heredity 100: 158-170.

Storfer A, Murphy MA, Spear SF, Holderegger R (2010) Landscape genetics: where are we now. Mol Ecol 19: 3496-3514.

Storz JF (2005) Using genome scans of DNA polymorphisms to infer adaptive population divergence. Mol Ecol 14: 671-688.

Streit K, Geiger DL, Lieb B (2006) Molecular phylogeny and the geographic origin of Haliotidae traced by haemocyanin sequences. J Mollus Stud 72: 105-110.

Swanson WJ, Aquadro CF, Vacquier VD (2001) Polymorphism in abalone fertilization proteins is consistent with the neutral evolution of the egg's receptor for lysin (VERL) and positive Darwinian selection of sperm lysin. Mol Biol Evol 18: 376-383.

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

Tang S, Tassanakajon A, Klinbunga S, Jarayabhand P, Menasveta P (2004) Population structure of tropical abalone (*Haliotis asinina*) in coastal waters of Thailand determined using microsatellite markers. Mar Biotechnol 6: 604-611.

Tarnowska K, Chenuil A, Nikula R, Féral J-P, Wolowicz M (2010) Complex genetic population structure of the bivalve *Cerastoderma glaucum* in a highly fragmented lagoon habitat. Mar Ecol Prog Ser 406: 173-184.

Tarr RJQ (1989) Abalone. In: Payne AIL, Crawford RJM (eds) Oceans of Life off Southern Africa. Vlaeberg Publishers, Cape Town. pp. 62-68.

Tarr RJQ (1992) The abalone fishery of South Africa. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books, Wiley-Blackwell, New Jersy. pp. 438-447.

Tarr RJQ (1995) Growth and movement of the South African abalone *Haliotis midae*: a Reassessment. Mar Freshwater Res 46: 583-590.

Tarr RJQ (2000) The South African abalone (*Haliotis midae*) fishery: a decade of challenges and change. Can Spec Publ Fish Aquat Sci 130: 32-40.

Tatarenkov A, Healey CIM, Avise JC (2010) Microgeographic population structure of green swordail fish: genetic differentiation despite abundant migration. Mol Ecol 19: 257-268.

Temby N, Miller K, Mundy C (2007) Evidence of genetic subdivision among populations of blacklip abalone (*Haliotis rubra* Leach) in Tasmania. Mar Fresh Res 58: 733-742.

Teske PR, McQuaid CD, Froneman CD, Barker NP (2006) Impacts of marine biogeographic boundaries on phylogeographic patterns of three South African estuarine crustaceans. Mar Ecol Prog Ser 314: 283-293.

Teske PR, Froneman CD, Barker NP, McQuaid CD (2007) Phylogeographic structure of the caridean shrimp *Palaemon peringueyi* in South Africa: further evidence for intraspecific genetic units associated with marine biogeographic provinces. Afr J Mar Sci 29: 253-258.

Teske PR, Papadopoulos I, Newman BK, Dworschak P, McQuaid C, Barker NG (2008) Oceanic dispersal barriers, adaptation and larval retention: an interdisciplinary assessment of potential factors maintaining a phylogeographic break between sister lineages of an African prawn. BMC Evol Biol 8: 341.

Tolley KA, Groeneveld JC, Gopal K, Matthee CA (2005) Mitochondrial DNA panmixia in spiny lobster *Palinurus gilchristi* suggests a population expansion. Mar Ecol Prog Ser 97: 225-231.

Tong LJ, Moss GA (1992) The New Zealand culture system for abalone. In: Shepard SA, Tegner MJ, Guzmán del Próo SA (eds) Abalone of the World: Biology, Fishery and Culture. Fishing News Books, Wiley-Blackwell, New Jersey. pp. 583-591.

Toro JE, Alcapan AC, Ojeda JA, Vergara AM (2004) Selection response for growth rate (shell height and live weight) in the Chilean blue mussel (*Mytilus chilensis* Hupe 1854). J Shellfish Res 23: 753-757.

Troell M, Robertson-Anderson 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.

Tymchuk WV, O'Reilly P, Bittman J, MacDonald D, Schulte P (2010) Conservation genomics of Atlantic salmon: variation in gene expression between and within regions of the Bay of Fundy. Mol Ecol 19: 1842-1859.

UNEP-WCMC (2010) Review of *Haliotis midae*. United Nations Environment Programme - World Conservation Monitoring Centre, Cambridge, United Kingdom.

Valladares F, Sanchez-Gomez V, Zavala MA (2006) Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications J Ecol 94: 1103-1116.

Van den Berg NC, Roodt-Wilding R (2010) Parentage assignment in *Haliotis midae* L.: A precursor to future genetic enhancement programs for the South African abalone. Aquac Res 41: 1387-1395.

Van der Merwe, M, Roodt-Wilding R (2008) Chromosome number of the South African abalone *Haliotis midae*. Afr J Mar Sci 30: 195-198.

Van der Merwe M, Franchini P, Roodt-Wilding R (2011) Differential growth-related gene expression in abalone (*Haliotis midae*). Mar Biotechnol 13: 1125-1139.

Varshney RK, Nayak SN, May GD, Jackson SA (2009) Next-generation sequencing technologies and their implications for crop genetics and breeding. Trends Biotechnol 27: 522-530.

Vasemägi A, Nilsson J, Primmer CR (2005) Express Sequence Tag-linked microsatellite as a source for gene-associated polymorphisms for detecting signatures of selection in Atlantic Salmon (*Salmo salar* L.). Mol Biol Evol 22: 1067-1076.

Veliz D, Duchesne P, Bourget E, Bernatchez L (2006) Stable genetic polymorphism in heterogeneous environments: balance between asymmetrical dispersal and selection in the acorn barnacle. J Evol Biol 19: 589-599.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng ZH, Chen L, Skupski M, *et al*. (2001) The sequence of the human genome. Science 291: 1304-1351.

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other molecular markers and their use in animal genetics. Genet Sel Evol 34: 275-305.

Vokes HE (1935) A new species of *Haliotis* from the Pliocene of southern California. J Paleont 9: 251-252.

Von der Heyden S (2009) Why do we need to integrate population genetics into South African marine protected area planning? Afr J Mar Sci 31: 263-269.

Von der Heyden S, Lipinski MR, Matthee CA (2007) Mitochondrial DNA analyses of the Cape hakes reveal an expanding population for *Merluccius capensis* and population structuring for mature fish in *Merluccius paradoxus*. Mol Phylogenet Evol 42: 517-527.

Von der Heyden S, Prochazka K, Bowie RCK (2008) Significant population structure and asymmetric gene flow patterns amidst expanding populations of *Clinus cottoides* (Perciformes, Clinidae): application of molecular data to marine conservation planning in South Africa. Mol Ecol 17: 4812-4826.

Vosloo A, Vosloo D (2006) Routes of water loss in South African abalone (*Haliotis midae*) during aerial exposure. Aquaculture 261: 670-677.

Wagner A (2005) Robustness, evolvability, and neutrality. FEBS Lett 579: 1772-1778.

Wall JD, Pritchard JK (2003) Haplotype blocks and linkage disequilibrium in the human genome. Nat Rev Genet 4: 587-597.

Wang S, Sha Z, Sonstegard TS, Liu H, Xu P, Somridhivej B, Peatman E, Kucuktas H, Liu Z (2008) Quality assessment parameters for EST-derived SNPs from catfish. BMC Genomics 9: 450.

Wang S, Peatman E, Abernathy J, Waldbieser G, Lindquist E, Richardson P, Lucas S, Wang M, Li P, Thimmapuram J, Liu L, Vullaganti D, Kucuktas H, Murdock C, Small BC, Wilson M, Liu H, Jiang Y, Lee Y, Chen F, *et al.* (2010) Assembly of 500,000 inter-specific catfish expressed sequence tags and large scale gene-associated marker development for whole genome association studies. Genome Biol 11: R8.

Wang W, Lan H (2000) Rapid and parallel chromosomal number reductions in muntjac deer inferred from mitochondrial DNA phylogeny. Mol Biol Evol 17: 1326-1333.

Wang WYS, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: theoretical and practical concerns. Nat Rev Genet 6: 109-118.

Waples RS (1987) A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution 41: 385-400.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered 89: 439-450.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15: 1419-1439.

Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels m marine. freshwater, and anadromous fishes. J Fish Biol 44: 213-232.

Warner RR, Cowen RK (2002) Local retention of production in marine populations: evidence, mechanisms, and consequences. Bull Mar Sci 70: S245-S249.

Weber BH (2011) Extending and expanding the Darwinian synthesis: the role of complex systems dynamics. Studies in the History and Philosophy of Biological and Biomedical Science 42: 75-81.

Wenne R, Boudry P, Hemmer-Hansen J, Lubienieki KP, Was A, Kause A (2007) What role for genomics in fisheries management and aquaculture? Aquat Living Resour 20: 241-255.

West MAL, Kim K, Kliebenstein DJ, Van Leeuwen H, Michelmore RW, Doerge RW, St. Clair DA (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. Genetics 175: 1441-1450.

Wheat C (2010) Rapidly developing functional genomics in ecological model systems *via* 454 transcriptome sequencing. Genetica 138: 433-451.

Whiteley AR, Bhat A, Martins EP, Mayden RL, Arunachalam M, Uusi-Heikkilä S, Ahmad ATA, Shrestha J, Clark M, Stemple D, Bernatchez L (2011) Population genomics of wild and laboratory zebrafish (*Danio rerio*). Mol Ecol 20: 4259-4276.

Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. J Evol Biol 14: 611-619.

Williams LM, Olecksiak MF (2008) Signatures of selection in natural populations adapted to chronic pollution. BMC Evol Biol 8: 282.

Willing E-M, Bentzen P, Van Oosterhout C, Hoffmann M, Cable J, Breden F, Weigel D, Dreyer C (2010) Genome-wide single nucleotide polymorphisms reveal populations history and adaptive divergence in the wild guppies. Mol Ecol 19: 968-984.

Wilson N, Rouse G, Giribet GW (2010) Assessing the molluscan hypothesis *Serialia* (Monoplacophora + Polyplacophora) using novel molecular data. Mol Phylogenet Evol 54: 187-193.

Winnepenninckx B, Backeljau T, De Wachter R (1996) Investigation of molecular phylogeny on the basis of 18S rRNA sequences. Mol Biol Evol 13: 1306-1317.

Winnepenninckx B, Seiner G, Backeljau T, De Wachter R (1998) Details of gastropod phylogeny inferred from 18S rRNA sequences. Mol Phylogenet Evol 9: 55-63.

Withler RE, Campbell A, Li SR, Brouwer D, Supernault KJ, Miller KM (2003) Implications of high levels of genetic diversity and weak population structure for the rebuilding of northern abalone in British Columbia, Canada. J Shellfish Res 22: 839-847.

Wright S (1932) Evolution in Mendelian populations. Genetics 16: 97-159.

You W, Ke C-H, Luo X, Wang D-X (2010a) Divergent selection for shell length in two stocks of small abalone *Haliotis diversicolor*. Aquac Res 41: 921-929.

You W, Ke C-H, Luo X, Wang D-X (2010b). Heritability of growth traits for small abalone *Haliotis diversicolor* estimated from sib matings. J Shellfish Res 29: 705-708.

Zardi GI, McQuaid CD, Teske PR, Barker NP (2007) Unexpected genetic structure of indigenous (*Perna perna*) and invasive (*Mytilus galloprovincialis*) mussel populations in South Africa. Mar Ecol Prog Ser 337: 135-144.

Zhan A, Hu J, Hu X, Zhou Z, Hui M, Wang S, Peng W, Wang M, Bao Z (2009) Fine-scale population genetic structure of Zhikong scallop (*Chlamys farreri*): Do local marine currents drive geographical differentiation. Mar Biotechnol 11: 223-235.

Zhan X, Fan F, You W, Yu J, Ke C-H (2012) Construction of an integrated map of *Haliotis diversicolor* using microsatellite markers. Mar Biotechnol 14: 79-86.

Zhang H, Scarpa J, Hare MP (2010) Differential fertilization success between two populations of Eastern oyster, *Crassostrea virginica*. Biol Bull 219: 142-150.

Zhang Z, Zhan A, Liu X (2010) A panel of genic single nucleotide polymorphism (SNP) markers for the Pacific abalone, *Haliotis discus hannai*. Conserv Genet Res 2 (Suppl 1): 133-135.

Zheng H, Zhang G, Liu X, Guo X (2006) Sustained response to selection in an introduced population of the hermaphroditic bay scallop *Argopecten irradians irradians* Lamarck (1819). Aquaculture 255: 579-585.

CHAPTER 2

A Population Genetic Analysis of Abalone Domestication Events in South Africa

Abstract

Abalone culture is South Africa's largest aquaculture sector in terms of revenue. Nonetheless, the industry is in its formative years with production rarely going beyond the first generation. Little is known on how aquaculture affects the genetic constitution of abalone populations now kept under artificial environments. With the implementing of advanced breeding strategies, for selective breeding for production traits and conservation efforts, it is important to first elucidate factors that affect patterns of genetic diversity in the F_1 generation cultured populations. The present study found no significant decrease in genetic diversity between wild and cultured populations as based on heterozygosity and allelic content of genomic- and EST-microsatellite loci. However, estimates for pairwise genotypic differentiation, F_{st}, AMOVA and factorial correspondence analysis suggest the genetic heterogeneity of cultured populations and their significant differentiation from the wild progenitor populations. As expected, the cultured population showed reduced effective population sizes, but relatedness remained low. It is postulated that both neutral and selective evolutionary forces are responsible for the observed patterns of genetic variability within and amongst populations.

Keywords: Abalone; Aquaculture; Conservation; Domestication; Genetic Diversity; *Haliotis midae*

The results of this chapter was published: Rhode et al. *(2012)* Aquaculture *356-357: 235-242.*
The results of this chapter was presented: The International Abalone Symposium. *May 2012, Hobart, Tasmania, Australia (oral presentation);* The International Symposium on Genetics in Aquaculture XI. *Jun. 2012, Auburn University, Alabama, USA. (oral presentation)* & The South African Genetics and Bioinformatics Society Conference. *Sept. 2012, Stellenbosch University, South Africa (poster presentation).*
- See Appendix D -

2.1. Introduction

South Africa has five endemic abalone species of which *Haliotis midae*, locally known as *perlemoen* is the largest growing and most abundant. Its large size, widespread distribution and high market value make this abalone a prime target for commercial exploitation. Growing concerns regarding the sustainability of the abalone fishery lead the South African government to impose increasingly stricter regulations culminating in the closure of the fishery in 2008 for fear of a collapse of natural populations (Hauck and Sweijd, 1999; Day and Branch, 2000; Raemaekers *et al.*, 2011). The fishery has now reopened to allow subsistence fishers to draw income from the natural resource (Department of Agriculture, Forestry and Fisheries, 2010). With the rapid decline of the fishery and a steady increase in the global demand for abalone products, emphasis was placed on aquaculture as an alternative means to expand the industry in a sustainable manner. Currently, South Africa has 14 abalone aquaculture facilities operating at various levels of production with a total output of 1015.44 metric tons, valued at 355 million ZAR (approximately 47.33 million US$) (2010 estimates, Department of Agriculture, Forestry and Fisheries, 2012).

These aquaculture facilities mostly operate on an open system basis where wild animals are collected and kept at the facility as broodstock for the production of seed animals that will enter the market as product. Matings are done, for the most part, at random with mass-spawning induced (*via* chemical or physical means) under semi-natural conditions. As the industry develops, producers will increasingly retain seed animals that demonstrate favourable production characteristics as potential broodstock in selective breeding programmes – effectively closing the aquaculture reproductive cycle. The emphasis on breeding for production traits has left traditional livestock industries reliant on substantial veterinary interventions to maintain animal health (Rauw *et al.*, 1998). Furthermore, concerns about limited genetic resources for future adaptability and genetic improvement, especially in the light of climate change, have been raised in recent years (Notter, 1999; Gamborg and Sandoe, 2005; Medugorac *et al.*, 2009; Ajmone-Marsan *et al.*, 2010; Groeneveld *et al.*, 2010; Hoffman, 2010). Reports of effective population sizes for highly commercialised breeds of cattle and sheep are as little as 50 and consequently these populations suffer from inbreeding (Taberlet *et al.*, 2008). The aquaculture industry is in a

favourable position as it can learn from the mistakes of traditional animal breeding, by carefully monitoring the progress made by domestication and selective breeding from the beginning. There is also increasing interest in using aquaculture reared animals for restocking of wild populations with the specific intent of ranching, i.e. recollecting seeded animals after a period of maturation in the wild. This has raised questions about the impact of such activities on the conservation of natural populations (Roodt-Wilding, 2007; Bester-van der Merwe *et al.*, 2011).

The establishment of a comparatively small founder population from an entire wild population could result in a population bottleneck as it limits the number of individuals that could effectively contribute to the next generation and may exacerbate the effects of random genetic drift. These founder effects are intensified by the reproductive strategy of abalone: Abalone are highly fecund, broadcast spawners and parental contributions are often unequal (Lind *et al.*, 2009; Slabbert *et al.*, 2009). Thus when retaining F_1 animals for broodstock replacement, the probability of inbreeding increases. If not monitored this could lead to an excessive loss of genetic diversity and a decrease in fitness of the overall commercial stock. Declines in genetic diversity have been reported for numerous aquaculture species including abalone (also the South African abalone) (Alarcón *et al.*, 2004; Evans *et al.*, 2004; Hara and Sekino, 2007; Li *et al.*, 2007; Lind *et al.*, 2009; De la Cruz *et al.*, 2010). A reduction in genetic diversity due to artificial selection is an expected consequence of any breeding programme; however it should occur in a controlled manner and a sufficient level should be maintained that will ensure sustainable breeding and continued genetic gains in the long run. On the contrary if ranching is considered, the genetic constitution of the cultured stock should be equivalent to that of the wild population where release is intended. It is therefore important to evaluate the genetic properties of F_1 cultured populations.

The use of molecular or DNA marker estimates of genetic diversity has become common practise and increasingly, emphasis is placed on the use of gene-associated molecular markers (Serapion *et al.*, 2004; Kucuktas *et al.*, 2009; Ma *et al.*, 2011). Gene-associated markers provide an evaluation of genetic diversity at coding regions and may be of greater value than anonymous markers as it provides information on regions of the genome that are directly responsible for phenotypic variation. Previously, gene-associated SNP markers were

developed for *H. midae* using a traditional EST sequencing protocol (Bester *et al.*, 2008). Recently, a substantial EST resource, generated *via* the next generation sequencing platform of Illumina®, was created for *Haliotis midae* (Franchini *et al.*, 2011) and used to develop additional gene-associated molecular markers (Hepple, 2010; Blaauw, 2012; Jansen, 2012). This study aimed to evaluate the genetic properties of three F_1- generation culture populations of *H. midae*, using EST- and genomic-derived microsatellite markers, on three different aquaculture facilities and to compare estimates with the wild progenitor populations.

2.2. Materials and Methods

2.2.1. Study Populations and Specimens

All necessary permits to collect and transport abalone for research purposes were obtained from the Department of Agriculture, Forestry and Fisheries (Republic of South Africa). Ninety six F_1 cultured animals were collected from three aquaculture facilities, one from the west- (CPWC), south- (CPSC) and east- (CPEC) coast of South Africa (32 animals per facility). These animals were randomly selected, across spawning cohorts, in order to attain a representative sample of the total F_1 population on each respective facility. Animals were aged between three and four years and had gone through the entire production system, including several grading procedures according to each facility's specifications.

Muscle and gill tissues were collected from each individual and placed in 70% ethanol and stored at -20°C until DNA extraction could be performed *via* the standard CTAB method (Saghai-Maroof *et al.*, 1984). For comparison, 96 wild animals previously collected from the west- (Saldanha Bay, WPWC), south- (Witsand, WPSC) and east- (Riet Point, WPEC) coast of South Africa (32 animals each) were also used (populations are described in Bester-van der Merwe *et al.*, 2011). These populations represent the ancestral progenitor populations for each of the respective cultured populations corresponding to the geographic region.

2.2.2. Population Genetic Analysis of Study Populations

To evaluate genetic diversity amongst the study populations, eight tetranucleotide genomic-microsatellites previously developed *via* the FIASCO protocol (Bester *et al.*, 2004; Slabbert *et al.*, 2008) and eight tetranucleotide EST-microsatellites (Hepple, 2010; Jansen, 2012) were used. All PCR reactions were conducted in a final volume of 10µl and according to the specifications of the authors. Allele scoring was done using GeneMapper® v.4 (Applied Biosystems). Tetranucleotide repeats are mutationally more stable and generally allows for easy and more reliable allele scoring. In addition, to minimise biases and in order to make direct comparisons between genomic and EST-microsatellites, similar repeat motifs were used across the molecular marker classes. Micro-checker v.2.2.3 (Van Oosterhout *et al.*, 2004) was used to test for possible genotyping errors, and the presence of null alleles (null allele estimates as per the method of Brookfield, 1996).

Hardy-Weinberg equilibrium (exact probability test, 500 batches, 10 000 iterations), expected and observed heterozygosity and locus-specific F_{is} was calculated using Genepop v.4.0 (Rousset, 2008). The number of alleles and allelic richness was computed using FStat v.2.9.3.2 (Goudet, 1995) and a Kruskal-Wallis test was performed to evaluate significant differences in number of alleles, allele richness and observed and expected heterozygosity amongst populations and molecular marker classes. Furthermore, the probability of linkage disequilibrium between all pairs of loci was calculated *via* an exact test using Genepop. Neutrality was tested using the Slatkin exact test (10 000 permutations) based on the Ewens sampling theory (Slatkin, 1994) in Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010) as well as the F_{st}-outlier procedure as implemented in Lositan v.1.44 (10 000 permutations assuming the infinite alleles model) (Antao *et al.*, 2008).

To evaluate population differentiation, pairwise F_{st} between populations (with Bonferonni correction at the 5% nominal level), was calculated in FStat and an exact test for pairwise genotypic differentiation was done in Genepop. Both methods were used in order to distinguish whether population differentiation was a consequence of unique allelic combinations, as it has been argued that the exact test may provide a more powerful estimate in such cases (Goudet *et al.*, 1996; Balloux and Lugon-Moulin, 2002). A locus by locus molecular analysis of variance (AMOVA, 10 000 permutations) was also computed in

Arlequin; populations were grouped as either cultured or wild and to visualise population distinctness, a factorial correspondence analysis plot was drawn in Genetix v.4.05.2 (Belkhir *et al.*, 2004).

Effective population sizes were calculated using the heterozygous excess test and the moment-based temporal test (using progenitor wild populations as generation zero) in NeEstimator v.1.3 (Peel *et al.*, 2004); as well as a linkage disequilibrium test (minimum allele frequency, 0.02) in LDNe v.1.0 (Waples, 2006). To further investigate the occurrence of recent bottlenecks, the Wilcoxon signed rank test (Luikart *et al.*, 1998), assuming the infinite alleles model, in Bottleneck v.1.2.02 (Piry *et al.*, 1999) was used. Finally, mean relatedness was calculated for each population using the method of Queller and Goodnight (1989) in Kinship v.2.0 (Konovalov *et al.*, 2004).

2.3. Results

2.3.1. Genetic Diversity within and between Wild and Cultured Populations

Micro-checker indicated no significant genotyping errors while evidence of null alleles at few loci for particular populations was present. There were no significant differences ($P > 0.05$) in number of alleles, allelic richness, observed- and expected heterozygosity across all populations, cultured and wild (Kruskal-Wallis test results: A_n: $P = 0.992$; R_s: $P = 0.641$; H_o: $P = 0.907$; H_e: $P = 0.427$), but significant differences ($P < 0.05$) in estimates were detected between genomic-microsatellites and EST-microsatellites except for observed heterozygosity (H_o) [A_n: $P = 8.450e^{-5}$; R_s: $P = 1.784e^{-5}$; H_o: $P = 0.093$; H_e: $P = 2.957e^{-5}$]. Over all populations, the number of observed alleles were more than the number of alleles in individual populations (*e.g.* locus *HmRS27T*: amongst populations A_n ranged between 23 and 32, total number of observed alleles across all populations was 52). Average F_{is}-values ranged from -0.001 to 0.236 and most loci conformed to Hardy-Weinberg expectations within populations. However, over all populations 11 of the 16 loci showed deviations from HWE (Appendix A: Table S2.1, S2.2, S2.3). Lositan and Slatkin's exact test showed evidence of non-neutral behaviour at particular loci, with candidates for both balancing and

directional selection (Appendix A: Table S2.1, S2.2; Figure 2.1). Only five of the 120 pairs of loci demonstrated significant ($P < 0.05$) linkage disequilibrium: *HmAD102T – HmLCS1T* ($P = 0.0101$); *HmAD102T – HmNS6T* ($P = 0.0$); *HmidILL-128551T – HmidILL-006622T* ($P = 0.0061$); *Hm128551T – HmidILL-071359P* ($P = 0.0$); *HmLCS67T – HmidILL-084787T* ($P = 0.0032$).

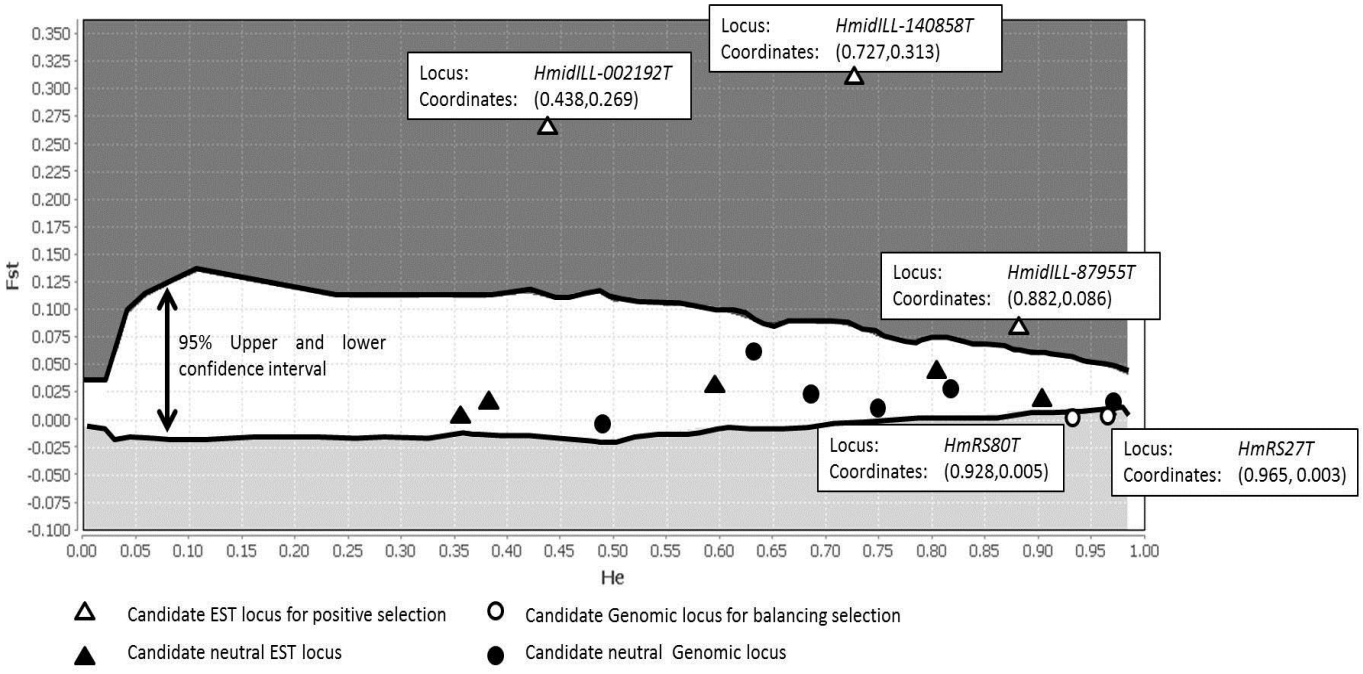

Figure 2.1: Lositan results indicating outlier loci as candidate loci under positive and balancing selection.

2.3.2. Population Differentiation, Effective Population Size and Relatedness

Pairwise genotypic differentiation as estimated by the exact test for genomic- and EST-markers suggests highly significant population differentiation between cultured and wild populations as well as significant differentiation amongst cultured populations from different facilities, with many of the *P*-values reaching the 0.01 statistically significant level (Table 2.1). Pairwise F_{st} estimates, calculated using genomic markers, ranged from 0.000-0.046, with values showing less genetic differentiation amongst wild populations and between cultured populations and wild progenitors. However, EST-marker estimates of pairwise F_{st} were more congruent with the results found for the exact test for genotypic differentiation (range: 0.014-0.200; Table 2.2) showing significant differentiation between most populations. Population differentiation was also supported by the AMOVA results,

with significant differentiation amongst groups (F_{CT} = 0.045, *P* = 0.001), within groups (F_{SC} = 0.026, *P* = 0.000) and over all groups and populations (F_{ST} = 0.070, *P* = 0.000) (Table 2.3). Factorial correspondence analysis clearly showed two clusters, one consisting of the wild populations and one containing the cultured populations (Figure 2.2).

Table 2.1: Exact test *P*-values for pairwise genotypic differentiation as implemented in Genepop v.4, using genomic-microsatellites (shaded area) and EST-microsatellites (unshaded area).

	CPWC	CPSC	WPEC	WPWC	CPEC	WPSC
CPWC	-	0.004**	0.000**	0.000**	0.000**	0.000**
CPSC	0.022*	-	0.000**	0.000**	0.000**	0.000**
WPEC	0.017*	0.002**	-	0.000**	0.000**	0.000**
WPWC	0.030*	0.064	0.937	-	0.000**	0.000**
CPEC	0.000**	0.000**	0.005**	0.003**	-	0.000**
WPSC	0.001**	0.037*	0.036*	0.197	0.000**	-

**statistical significance at the 5% nominal level; ** statistical significance at the 1% nominal level.*

Table 2.2: Pairwise F_{st}-values for populations as calculated in Fstat v.2.9.3.2. using genomic-microsatellites (shaded area) and EST-microsatellites (unshaded area).

	CPWC	CPSC	WPEC	WPWC	CPEC	WPSC
CPWC	-	0.014	0.083*	0.093*	0.069*	0.114*
CPSC	0.003	-	0.097*	0.104*	0.086*	0.140*
WPEC	0.006	0.012	-	0.018*	0.176*	0.019
WPWC	0.012	0.015	0.000	-	0.143*	0.022*
CPEC	0.024*	0.028*	0.014	0.024	-	0.200*
WPSC	0.016	0.011	0.017	0.005	0.046*	-

**statistical significance at the 5% nominal level.*

Effective population sizes, as estimated by the temporal method, were generally low; as little as 4.1 for the CPSC data (based on EST-markers only). The point estimate calculated by the Heterozygosity excess test postulated a large (infinite) effective population for all populations with the exception of the WPEC population. The LD estimates of effective population size conformed to what was expected: generally high effective population sizes in the wild populations with a reduction in effective population sizes amongst cultured populations. This population bottleneck was further supported by the Wilcoxon rank test that detected heterozygous excess in comparison to the expected heterozygosity under mutation-drift equilibrium. There was also evidence that the WPEC population had undergone a recent bottleneck (Table 2.4). There was no evidence for extensive relatedness within or between populations (Table 2.5).

Table 2.3: Locus by locus AMOVA results over all 16 loci, with populations clustered in two groups, cultured and wild progenitor.

Source of Variation	Sum of Squares	Variance components	% of variation
Amongst groups	61.160	0.260	4.518
Amongst populations within groups	57.380	0.144	2.505
Within populations	1963.456	5.356	92.977
Total	2081.996	5.7610	
F_{ST}: 0.070	***P*: 0.000****		
F_{SC}: 0.026	***P*: 0.000****		
F_{CT}: 0.045	***P*: 0.001****		

*** statistical significance at the 1% nominal level.*

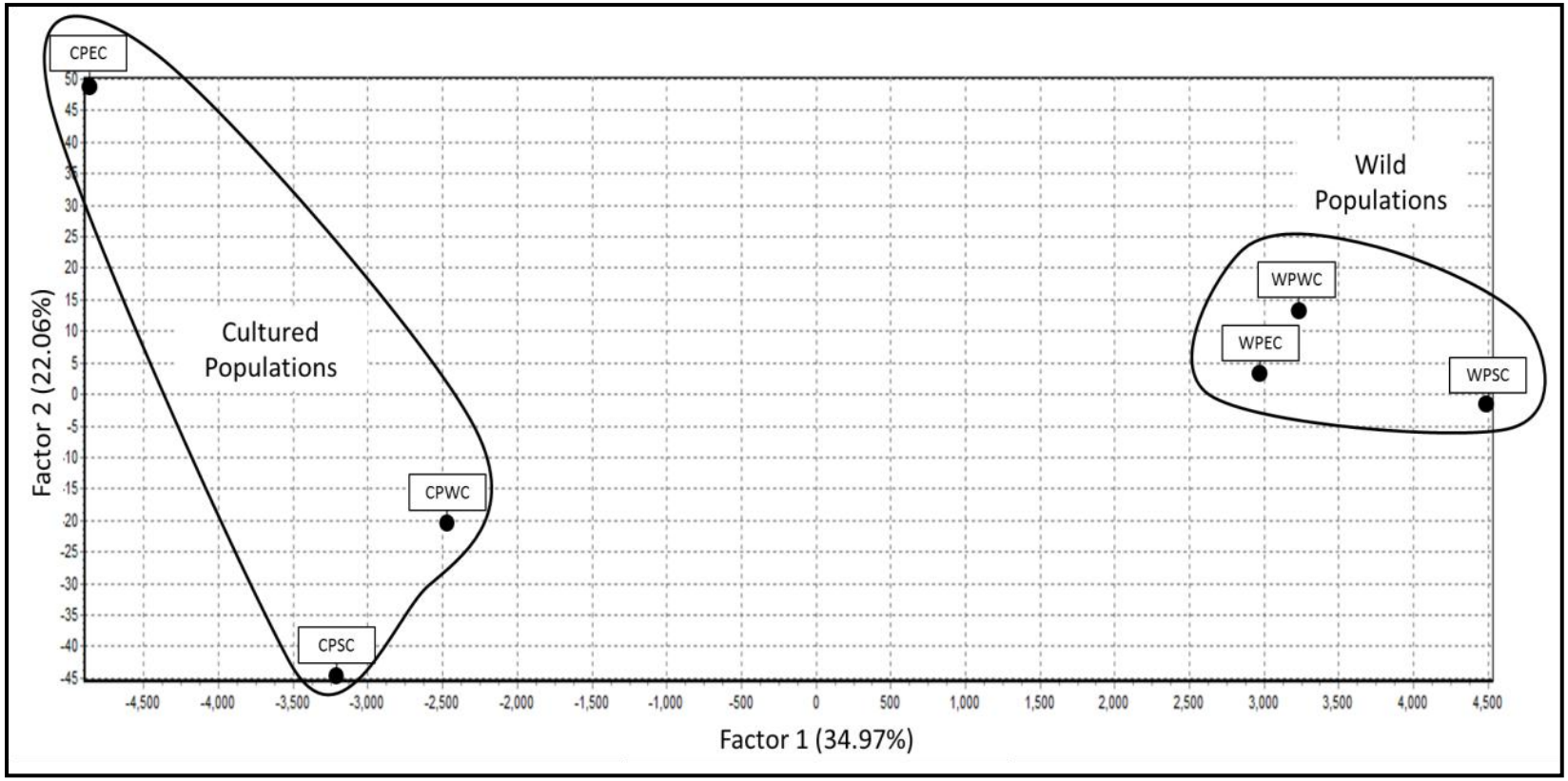

Figure 2.2: Factorial correspondence analysis, using 16 loci, showing two distinct population clusters grouped into wild and cultured populations.

Table 2.4: Various point estimates for effective population size (N_e) and a test for recent population bottleneck based on genomic-, EST-microsatellites and combined datasets.

	LD test for N_e (95% CI)			Heterozygosity excess test for N_e			Two temporal samples test for N_e (95% CI)			Wilcoxon's *P*-value for heterozygous excess in recent bottleneck		
	Genomic	**EST**	**Combined**	**Genomic**	**EST**	**Combined**	**Genomic**	**EST**	**Combined**	**Genomic**	**EST**	**Combined**
CPWC	288.4 (113-∞)	201.1 (56.9-∞)	185.1 (99.7-889.3)	∞	∞	∞	45.2 (21.4-212.3)	6.1 (3.6-10.2)	15.9 (10.8-24.4)	0.001**	0.273	0.019*
WPWC	∞ (1130.6-∞)	∞ (123.4-∞)	617.7 (179.7-∞)	∞	∞	∞	^	^	^	0.231	0.273	0.126
CPSC	199.3 (75.1-∞)	∞ (65.9-∞)	160.8 (87.3-737.3)	∞	∞	∞	42.6 (19.8-203.1)	4.1 (2.5-6.5)	10.9 (7.7-10.6)	0.002**	0.006**	0.000**
WPSC	∞ (500.1-∞)	∞ (273.3-∞)	∞ (719.9-∞)	∞	∞	∞	^	^	^	0.098	0.680	0.281
CPEC	41.9 (27.9-73.9)	37.4 (22.1-82.1)	57.9 (43.2-84.2)	∞	∞	∞	32.3 (16.9-87.2)	4.9 (3.0-7.8)	11.9 (8.4-17.3)	0.027*	0.231	0.029*
WPEC	78.5 (46.7-193.2)	96.7 (37.6-∞)	93.2 (63-162.8)	∞	2134.4	∞	^	^	^	0.014*	0.273	0.047*

**statistical significance at the 5% nominal level; ** statistical significance at the 1% nominal level. ^Note: Temporal estimate for wild populations could not be calculated, because temporal samples were not available.*

Table 2.5: Mean relatedness within and amongst populations, as estimated using genomic- and EST-microsatellites

	Genomic		EST	
	r	*P*-value	r	*P*-value
CPWC	0.032	0.497	0.024	0.493
WPWC	0.034	0.499	0.031	0.488
CPSC	0.033	0.509	0.034	0.538
WPSC	0.033	0.477	0.032	0.414
CPEC	0.033	0.506	0.028	0.470
WPEC	0.033	0.504	0.030	0.446
Over All	0.005	0.480	0.003	0.470

2.4. Discussion

2.4.1. Genetic Diversity: Genomic- vs. EST-Markers

Genetic diversity estimates revealed statistically significant differences between marker classes (genomic- and EST-microsatellites) in terms of number of alleles, allelic richness and expected heterozygosity. This is not a surprising result and conforms to previous studies demonstrating moderate EST-microsatellite polymorphism across various animal taxa (Zhan *et al.*, 2005; Wang *et al.*, 2007; Wang *et al.*, 2008; Zhan *et al.*, 2008; Kim *et al.*, 2009; Wang *et al.*, 2009). Because of the close proximity to coding regions, EST-microsatellites are often in linkage disequilibrium with functional genetic variants or may be itself a functional sequence. Therefore, EST-microsatellites are often under selective pressure, which in turn suppresses allelic variation (Li *et al.*, 2004). Evidence for this is present in the current marker-set, with three EST-microsatellites under possible positive differential selection between populations (Figure 2.1). It is therefore noteworthy that genomic- and EST-microsatellites might in fact represent two different sets of genetic diversity: neutral and adaptive genetic diversity, respectively (Sgrò *et al.*, 2011). A parallel analysis using sets of

both marker-types may therefore provide a more complete understanding of the evolutionary forces driving genetic diversity within and between populations.

2.4.2. Genetic Diversity: Wild vs. Cultured populations

Results indicated no significant loss of genetic diversity between wild and cultured populations of abalone based on estimates such as number of alleles, allelic richness or heterozygosity and in general were comparable amongst populations (Table S2.1, S2.2). This is in accord with previous findings for *Haliotis midae* (Slabbert *et al.*, 2009), estimates for the Pacific abalone (*H. discus hannai*; An *et al.*, 2011) and the blue abalone (*H. fulgens*; Gutierrez-Gonzalez and Perez-Enriquez, 2005) but contradicts findings for other studies on aquaculture species including abalone (Alarcón *et al.*, 2004; Evans *et al.*, 2004; Hara and Sekino 2007; Li *et al.*, 2007; Lind *et al.*, 2009; De la Cruz *et al.*, 2010). A similar investigation, comparing F_1-animals to wild populations, for *Haliotis midae* by Evans *et al.* (2004) was based on a single spawning event, with a particular spawning cohort; thus the population sample, in that study, was not representative of the total production population. The reported loss of genetic diversity could, therefore be considered an artefact of a specific spawning event in an isolated breeding group: Differential parental contributions are well documented for broadcast spawning molluscs including South African abalone (Slabbert *et al.*, 2009; Van den Berg and Roodt-Wilding, 2010). This may be for a number of reasons, including genetic fitness of particular individuals, but also stochastic variables, such as the condition (*e.g.* physiological stress because of disease) of an individual animal at any given spawning event. Contrary to previous studies, the present investigation sampled individuals across spawning events and groups and therefore provides population-wide estimates that can account for the observed maintenance of genetic diversity. Furthermore, the high levels of genetic diversity in cultured populations may be attributed to good management practice, by optimising the effective number of breeding individuals. It is noted that the cultured populations (in the present study) maintain comparatively large effective population sizes (57.9 – 185.1; combined LD estimate of N_e, Table 2.4). In comparison, estimates for, for example, cultured seabream (*Sparus aurata*; maximum N_e = 18; Brown *et al.*, 2005) and pearl oyster (*Pinctada maxima*; maximum N_e = 9.2; Lind *et al.*, 2009), reported

losses in genetic diversity. This is noteworthy, especially considering that mass-spawning is the primary means of production in all the aforementioned species.

The number of alleles observed per locus was significantly higher over all populations than within individual populations (Table S2.1, S2.2, S2.3) suggesting a number of population-specific alleles across these loci. Although this observation must be treated with caution due to the relatively small sample size used in the current study, a similar observation was made by An *et al.* (2011) for the Pacific abalone: This could be a result of founder effects that lead to a loss of rare alleles in cultured populations (Skaala *et al.*, 2004), noting that wild populations show the largest number of unique alleles, *e.g.* locus *HmAD102T, HmRS27T* and *HmRS80T* (Table S2.3). However, unique alleles also persists in the cultured population. This can be explained by random genetic drift or, alternatively, by selection of differentially favoured alleles in diverse heterogeneous environments. This holds particular reference to locus *Hm140858T* with unique alleles only in two cultured populations (CPSC and CPEC, Table S2.3) and evidence of differential selection between populations (Figure 2.1).

Within respective populations several loci demonstrated violation of Hardy-Weinberg equilibrium (Table S2.1, S2.2). This was mostly due to homozygous excess, based on F_{is}-estimates. Homozygous excess could be caused by a number of factors including the presence of null alleles and directional selection, with significant evidence for both at loci *HmAD102T* and *HmLCS72M*, therefore these markers were not excluded from analysis. A global analysis of Hardy-Weinberg equilibrium showed 11 of the 16 loci deviated from equilibrium expectations, as expected from mixing individual populations that differ significantly in allele frequencies (Table 2.1, 2.2).

2.4.3. Population Differentiation, Effective Population Size and Relatedness

The exact test for genotypic differentiation is often regarded as a more sensitive test for population differentiation if unique alleles persist in said populations (Goudet *et al.*, 1996; Balloux and Lugon-Moulin, 2002). The presence of such unique alleles could, in part, explain the significant genotypic differentiation of cultured populations from the wild progenitor

populations and amongst cultured populations (Table 2.1, S2.3) for the present abalone cohorts. A similar observation was made for captive bred Père David's deer (Zeng *et al.*, 2007).

Pairwise F_{st} estimates were generally highest between cultured and the wild progenitor populations (Table 2.2) and this agrees with the two distinct (cultured and wild) clusters obtained with the factorial correspondence analysis plot (Figure 2.2). However, there is evidence for further population differentiation within each cluster as demonstrated by the AMOVA and pairwise F_{st} results (Table 2.3). It is interesting to note that F_{st} estimates, based on EST-microsatellites were significant for almost all population pairs, whereas genomic estimates only reached significance for cultured populations (Table 2.2). As mentioned earlier, this is possibly an indication that selection is the major evolutionary force driving genetic differentiation at EST-loci (Figure 2.1). Population differentiation linked to adaptation to ecological niches or other environmental conditions is well documented in wild populations, even in high gene flow marine environments (Hemmer-Hansen *et al.*, 2007; Nielsen *et al.*, 2009).

Recently, Bester-van der Merwe *et al.* (2011) reported low, yet significant differentiation between wild *Haliotis midae* populations residing on the west and east coasts of South Africa, with a possible secondary contact zone around the Cape Agulhas region on the south coast. The present study also shows evidence of population heterogeneity between the west and east coasts and further indicates that population differentiation may be facilitated by adaptive processes. The development of ecotypes is known to occur in marine environments where environmental clines persist (Schmidt *et al.*, 2008). Such an environmental cline is consistent with the temperature gradient along the South African coast.

Differentiation of cultured populations from their progenitor populations and from one another is a common occurrence in aquaculture species with reports for salmon (Skaala *et al.*, 2004; Withler *et al.*, 2007), carp (Murakaeva *et al.*, 2003), shrimp (Dixon *et al.*, 2008), pearl oysters (Lind *et al.*, 2009) and various abalone species (Hara and Sekino, 2007; Li *et al.*, 2007; De la Cruz *et al.*, 2010; Praipue *et al.*, 2010; An *et al.*, 2011). Again it would seem as if selective pressures may be in part responsible for the observed patterns of genetic diversity;

considering that these cultured populations have been through the entire production system as well as grading procedures where inferior specimens were culled. The only pairwise F_{st} estimates based on EST-microsatellites that did not reach statistical significance was between the CPWC and CPSC as well as between the WPSC and WPEC populations and could be a result of possible convergent evolution. For the cultured populations this may be due to artificial selection of favourable production phenotypes or adaptation to similar aquaculture practices. Whereas for the wild populations this could be adaptation to similar environmental conditions, noting that both populations are located eastward from the major barrier to gene flow in the warm-temperate coastal regions (Bester-van der Merwe *et al.*, 2011). Convergent evolution has been described for scallops that inhabit analogous ecological niches and subsequently develop similar phenotypic characters (Alejandrino *et al.*, 2011).

Effective population size is an important parameter as it provides an indication of the rate of inbreeding and subsequent loss of genetic diversity over successive generations; however it is often difficult to estimate (Doyle and Talbot, 1986; Ryman and Laikre, 1991; Waples and Do, 2010). Waples and Do (2010) argued that the LD estimate of effective population size should provide greater precision than the temporal or heterozygous excess estimates, especially within a limited generational interval and small sample size. Based on the present data the temporal method suggests relatively low effective population sizes for the cultured populations (Table 2.4). Lind *et al.* (2009) reported effective population sizes to the same order for cultured pearl oysters, but the authors reported significant relatedness within the cultured populations. This is not the case for the current *H. midae* cultured populations that demonstrates no significant relatedness of individuals in respective populations and similar estimates across both wild and cultured populations (Table 2.5). The LD estimates are therefore a more reliable measure, considering that the Wilcoxon test results demonstrates evidence for a bottleneck event in the cultured populations. This result is mirrored by the linkage disequilibrium estimates and in accord with what is expected.

Previous estimates for long-term (N_e = 7 247.64 – 29 104.14) and contemporary (N_e = 62 496.875) effective population sizes, for the total wild population, as calculated by Bester-van der Merwe (2009) fall well within the confidence boundaries (as calculated by LDN_e) for the currently reported estimates of effective population size in the wild populations, WPWC

and WPSC (Table 2.4). Interestingly the WPEC population is the only wild population to show evidence of a possible recent bottleneck; further investigation is however needed before final conclusions can be drawn.

2.5. Conclusions

The presented data suggests that even though there is evidence of a population bottleneck, there is a sufficient number of breeding animals on respective aquaculture facilities to maintain levels of genetic diversity comparable with their wild progenitor populations. However, it is clear that the exact constitution of this genetic diversity is distinct from the wild populations and differs significantly between cultured populations from different facilities. On this account, it is therefore argued that there is evidence for genetically unique domesticated abalone strains produced by independent domestication events on respective aquaculture facilities. This initial phases of the domestication process seems to primarily driven by random genetic drift and possibly selection. Abalone and many other aquaculture species are in a unique position in that efforts to domesticate and conserve natural stocks will run in parallel. A thorough understanding of underlying genetic elements that contribute to the development of wild and aquaculture phenotypes is thus warranted for the effective management of the abalone resource.

References

Ajmone-Marsan P, The GLOBALDIV Consortium (2010) A global view of livestock biodiversity and conservation – GLOBALDIV. Anim Genet 41 (Suppl. 1): 1-5.

Alarcón JA, Magoulas A Georgakopoulos T, Zouros E, Alvarez MC (2004) Genetic comparison of wild and cultivated European populations of the gilthead sea bream (*Sparus aurata*). Aquaculture 230: 65-80.

Alejandrino A, Puslednik L, Serb JM (2011) Convergent and parallel evolution in life habit of the scallops (*Bivalvia*: *Pectinidae*). BMC Evol Biol 11: 164.

An HA, Wook J, Kim HC, Myeong J (2011) Genetic characterization of five hatchery populations of pacific abalone (*Haliotis discus hannai*) using microsatellite markers. Int J Mol Sci 12: 4836-4849.

Antao T, Lopes A, Lopes RL, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a F_{st}-outlier method. BMC Bioinformatics 9: 323.

Balloux FO, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. Mol Ecol 11: 155-165.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France). Available at http://kimura.univ-montp2.fr/genetix/

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

Bester AE, Roodt-Wilding R, Whitaker HA (2008) Discovery and evaluation of single nucleotide polymorphisms (SNPs) for *Haliotis midae*: a targeted EST approach. Anim Genet 39: 321-324.

Bester-Van der Merwe AE (2009) Population genetic structure and demographical history of South African abalone, *Haliotis midae*, in a conservation context. Unpublished PhD dissertation, Stellenbosch University, South Africa.

Bester-van der Merwe AE, Roodt-Wilding R, Volckaert FAM, D'Amato ME (2011) Historical isolation and hydrodynamically constrained gene flow in declining population of the South-African abalone, *Haliotis midae*. Conserv Genet 12: 543-555.

Blaauw S (2012) SNP screening and validation in *Haliotis midae*. Unpublished MSc thesis, Stellenbosch University, South Africa.

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

Brown RC, Woolliams JA, McAndrew BJ (2005) Factors influencing effective population size in commercial populations of gilthead seabream, *Sparus aurata*. Aquaculture 247: 219-225.

Day E, Branch GM (2000) Relationships between recruits of abalone *Haliotis midae*, encrusting corallines and the sea urchin *Parechinus angulosus*. S Afr J Mar Sci 22: 137-144.

De la Cruz F-L, Del Rio Portilla M, Gallardo-Escarate C (2010) Genetic variability of cultured populations of red abalone in Chile: An approach based on heterologous microsatellites. J Shellfish Res 29: 709-715.

Department of Agriculture, Forestry and Fisheries (2012) Aquaculture Annual Report 2011, South Africa.

Department of Agriculture, Forestry and Fisheries (2010) The conditional re-opening of commercial abalone fishery in the 2010 season - Progress report to cabinet. [online]. Available at: www.pmg.org.za/files/docs/110208reopening_0.ppt [accessed 28 August 2011].

Dixon TJ, Coman GJ, Arnold SJ, Sellars MJ, Lyons RE, Dierens L, Preston NP, Li Y (2008) Shifts in genetic diversity during domestication on Black Tiger shrimp, *Penaeus monodon*, monitored using two multiplexed microsatellite systems. Aquaculture 283: 1-6.

Doyle RW, Talbot AJ (1986) Effective population size and selection in variable aquaculture stocks. Aquaculture 57: 27-35.

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

Excoffier L, Lischer HEL (2010) Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10: 564-567.

FAO (2007) The State of the World's Animal Genetic Resources for Food and Agriculture [online]. Available at: www.fao.org/docrep/010/a1250e/a1250e00.htm [accessed 15 November 2011]

Franchini P, Van der Merwe M, Roodt-Wilding R (2011) Transcriptome characterization of the South African abalone *Haliotis midae* using sequencing-by-synthesis. BMC Research Notes 4: 59.

Gamborg C, Sandøe P (2005) Sustainability in farm animal breeding: a review. Livest Prod Sci 92: 221-231.

Groeneveld LF, Lenstra JA, Eding H, Toro MA, Scherf B, Pilling D, Negrini R, Finlay EK, Jianlin H., Groeneveld E, Weigend S, The GLOBALDIV Consortium (2010) Genetic diversity in farm animals – a review. Anim Genet 41 (Suppl. 1): 6-31.

Goudet J, Raymond M, De Meeiis T, Rousset F (1996) Testing Differentiation in Diploid Populations. Genetics 144: 1933-1940.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-Statistics. Journal of Heredity 86: 485-486.

Gutierrez-Gonzalez JL, Perez-Enriguez 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 (2007) Genetic differences between hatchery stocks and natural populations in Pacific Abalone (*Haliotis discus*) estimated using microsatellite DNA markers. Mar Biotechnol 9: 74-81.

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

Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). Heredity 99: 592-600.

Hepple J (2010) An integrated linkage map of Perlemoen (*Haliotis midae*). Unpublished MSc thesis, Stellenbosch University, South Africa.

Hoffmann I (2010) Climate change and the characterization, breeding and conservation of animal genetic resources. Anim Genet 41 (Suppl. 1): 32-46.

Jansen S (2012) Linkage mapping of gene linked markers in *Haliotis midae*. Unpublished MSc thesis, Stellenbosch University, South Africa.

Kim WJ, Kim YO, Nam BH, Kong JH, Park EM, Kang JH, Lee JH, Kim KK (2009) Development of 81 new polymorphic EST-derived microsatellite markers for the olive flounder, *Paralichthys olivaceus*. Conserv Genet 10: 1105-1111.

Konovalov DA, Manning C, Henshaw MT (2004) KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers. Mol Ecol Notes 4: 779-782.

Kucuktas H, Wang S, Li P, He C, Xu P, Sha Z, Liu H, Jiang Y, Baoprasertkul P, Somridhivej B, Wang Y, Abernathy J, Guo X, Liu L, Muir W, Liu Z (2009) Construction of genetic linkage maps and comparative genome analysis of catfish using gene-associated markers. Genetics 181: 1649-1660.

Li Q, Shu J, Yu R, Tian C (2007) Genetic variability of cultured populations of the Pacific abalone (*Haliotis discus hannai* Ino). Aquac Res 38: 981-990.

Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: Structure, function, and evolution. Mol Biol Evol 21: 991-1007.

Lind CE, Evans BS, Knauer J, Taylor JJU, Jerry DR (2009) Decreased genetic diversity and a reduced effective population size in cultured silver-lipped pearl oysters (*Pinctada maxima*). Aquaculture 286: 12-19.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89: 238-247.

Ma H, Ma Q, Ma C, Ma L (2011) Isolation and characterization of gene-derived single nucleotide polymorphism (SNP) markers in *Scylla paramamosain.* Biochem Syst Ecol 39: 419-424.

Medugorac I, Medugorac A, Russ I, Veit-Kensc CE, Taberlet HP, Luntz B, Mix HM, Forster M (2009). Genetic diversity of European cattle breeds highlights the conservation value of traditional unselected breeds with high effective population size. Mol Ecol 18: 3394-3410.

Murakaeva A, Kohlmann K, Kersten P, Kamilov B, Khabibullin D (2003) Genetic characterization of wild and domesticated common carp (*Cyprinus carpio* L.) populations from Uzbekistan. Aquaculture 218: 153-166.

Nielsen EE, Hemmer-Jansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18: 3128-3150.

Notter DR (1999) The importance of genetic diversity in livestock populations of the future. Journal of Ani Sci 77: 61-69.

Peel D, Ovenden JR, Peel SL (2004) NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in effective population size using allele frequency data. J Hered 90: 502-503.

Praipue P, Klinbunga S, Jarayabhand P (2010) Genetic diversity of wild and domesticated stocks of Thai abalone, *Haliotis asinina* (Haliotidae), analyzed by single-strand conformational polymorphism of AFLP-derived markers. Genet Mol Res 9: 1136-1152.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. Evolution 43: 258-275.

Raemaekers S, Hauck M, Bürgener M, Mackenzie A, Maharaj G, Plagányi EE, Britz PJ (2011) Review of the causes of the rise of the illegal South African abalone fishery and consequent closure of the rights-based fishery. Ocean Coast Manage 54: 433-445.

Rauw WM, Kanis EK, Noordhuizen-Stassen EN, Grommers FJ (1998) Undesirable side effects of selection for high production efficiency in farm animals: a review. Livest Prod Sci 56: 15-33.

Roodt-Wilding R (2007) Abalone ranching: a review on genetic considerations. Aquac Res 38: 1229-1241.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Res 8: 103-106.

Ryman N, Laikre L (1991) Effects of supportive breeding on the genetically effective population size. Conserv Biol 5: 352-329.

Saghai-Maroof MA, Solima KM, Jorgenson RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81: 8014-8018.

Schmidt PS, Serrão EA, Pearson GA, Riginos C, Rawson PD, Hilbish TJ, Brawley SH, Trussell GC, Carrington E, Wethey DS, Grahame JW, Bonhomme F, Rand DM (2008) Ecological genetics in the North Atlantic: Environmental gradients and adaptation at specific loci. Ecology 89: S91-S107.

Serapion J, Kucuktas H, Feng J, Liu Z (2004) Bioinformatic mining of type 1 microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*). Mar Biotechnol 6: 364-377.

Sgrò CM, Lowe AJ, Hoffmann AA (2011) Building evolutionary resilience for conserving biodiversity under climate change. Evol Appl 4: 326-337.

Skaala Ø, Høyheim B, Glover K, Dahle K (2004) Microsatellite analysis in domesticated and wild Atlantic salmon (*Salmo salar* L.): allelic diversity and identification of individuals. Aquaculture 240: 131-143.

Slabbert R, Ruivo NR, Van den Berg NC, Lizamore DL, Roodt-Wilding R (2008) Isolation and characterisation of 63 microsatellite loci for the abalone, *Haliotis midae*. J World Aquacult Soc 39: 429-435

Slabbert R, Bester AE, D'Amato ME (2009) Analyses of genetic diversity and parentage within a South African hatchery of the abalone *Haliotis midae* Linnaeus using microsatellite markers. J Shellfish Res 28: 369-375.

Slatkin M (1994) An exact test for neutrality based on the Ewens sampling distribution. Genet Res 64: 71-74.

Taberlet P, Valentini A, Rezaei HR, Naderi S, Pompanon F, Negrini R, Ajmone-Marsan P (2008) Are cattle, sheep, and goats endangered species? Mol Ecol 17: 275-284.

Van den Berg N-C, Roodt-Wilding R (2010) Parentage assignment in *Haliotis midae* L.: a precursor to future genetic enhancement programmes for South African abalone. Aquac Res 41: 1387-1395.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4: 535-538.

Wang D, Liao X, Cheng L, Yu X, Tong J (2007) Development of novel EST-SSR markers in common carp by data mining from public EST sequences. Aquaculture 271: 558-574.

Wang X, Song B, Qiu X, Meng X (2009) Development of EST-SSRs in scallop (*Patinopecten yessoensis*) from sequence database. Conserv Genet 10: 1129-1131.

Wang Y, Ren R, Yu Z (2008) Bioinformatic mining of EST-SSR loci in the Pacific oyster, *Crassostrea gigas*. Anim Genet 39: 287-289.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conserv Genet 7: 167-184.

Wapels RS, Do C (2010) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol Appl 3: 244-262.

Withler RE, Rundle T, Beacham TD (2007) Genetic identification of wild and domesticated strains of Chinook salmon (*Oncorhynchus tshawytscha*) in southern British Columbia, Canada. Aquaculture 272: 161-171.

Zeng Y, Jiang Z, Li C (2007) Genetic variability in relocated Pe`re David's deer (*Elaphurus davidianus*) populations - Implications to reintroduction program. Conserv Genet 8: 1051-1059.

Zhan AB, Bao ZM, Wang XL, Hu JJ (2005) Microsatellite markers derived from bay scallop *Argopecten irradians* expressed sequence tags. Fisheries Sci 71: 1341-1346.

Zhan A, Bao Z, Wang M, Chang D, Yaun J, Wang X, Hu X, Liang C, Hu J (2008) Development and characterization of microsatellite markers for the Pacific abalone (*Haliotis discus*) *via* EST database mining. Journal of Ocean University China 7: 219-222.

Zhang H, Scarpa J, Hare MP (2010) Differential fertilization success between two populations of Eastern oyster, *Crassostrea virginica*. Biol Bull 219: 142-150.

CHAPTER 3

Detection of Molecular Signatures of Selection at Microsatellite Loci in the South African Abalone (*Haliotis midae*)

Abstract

Identifying genomic regions that may be under selection is important for elucidating the genetic architecture of complex phenotypes underlying adaptation to heterogeneous environments. A population genomics approach, using a classical neutrality test and various F_{st}-outlier detection methods was employed to evaluate genome-wide polymorphism data in order to identify loci that may be candidates for selection amongst six populations (three cultured and three wild) of the South African abalone, *Haliotis midae*. Approximately 9% of the genome-wide microsatellite markers were subject to directional selection, whilst 6% to 18% of the genome is thought to be influenced by balancing selection. Genetic diversity estimates for candidate loci under directional selection was significantly reduced in comparison to candidate neutral loci, whilst candidate balancing selection loci demonstrated significantly higher levels of genetic diversity (Kruskal-Wallis test, $P < 0.05$). Pairwise F_{st} estimates based on candidate directional selection loci also demonstrated increased levels of differentiation between study populations. Various candidate loci under selection showed significant inter-chromosomal LD, suggesting possible gene-networks underling adaptive phenotypes. Furthermore, several loci had significant hits to known genes when performing BLAST searches to NCBI's non-redundant databases, whilst others are known to be derived from expressed sequences even though homology to a known gene could not be established. A number of loci also demonstrated relatively high similarity to transposable elements. The association of these loci to functional and genomically active sequences could in part explain the observed signatures of selection.

Keywords: Adaptation; F_{st}-outlier; Linkage Disequilibrium; Neutrality; Population Genomics; Selection

The results of this chapter has been submitted for peer review at the journal Marine Genomics
The results of this chapter was presented: The South African Genetics and Bioinformatics Society Conference. *Sept. 2012, Stellenbosch University, South Africa (oral presentation). - See Appendix D -*

3.1. Introduction

The neutral theory predicts that most molecular genetic variation will not have any fitness advantage and thus will be selectively neutral. However, in recent years increasing evidence suggests that a strict neutral model for molecular evolution is not tenable and it is now commonly accepted that evolutionary change is both a function of stochastic events and selection (Nielsen, 2005; Mitchell-Olds *et al.*, 2007; Nadeau and Jiggins, 2010). Understanding how selection shapes molecular diversity and how this diversity in turn facilitates the development of adaptive phenotypes, in heterogeneous environments, has thus become a key endeavour of modern evolutionary biology. Recently population genomics scans have become increasingly popular for detecting population divergence as a consequence of adaptation and identifying the underlying genetic architecture of complex divergent phenotypes (Black *et al.*, 2001; Luikart *et al.*, 2003; Storz, 2005; Biwas and Akey, 2006; Pavlidis *et al.*, 2008; Stinchcombe and Hoekstra, 2008; Nielsen *et al.*, 2009a; Nosil *et al.*, 2009).

The use of genome-wide polymorphism data allows for the partitioning of locus-specific effects such as recombination, mutation and selection from demographic effects (including: bottlenecks, founder effects, population stratification and migration etc.) (Luikart *et al.*, 2003; Stinchcombe and Hoekstra, 2008) and also provides for functional analyses of genetic polymorphisms to be extrapolated to a population level (Bonin, 2008). Unlike the more conventional linkage-based QTL analysis, population genomic scans do not rely on structured pedigree information and controlled breeding experiments, often impractical when working with natural populations or organisms with long generation times (Storz, 2005; Stinchcombe and Hoekstra, 2008). Furthermore, there is no dependence on *a priori* phenotypic information as is the case for association and conventional linkage-based studies. This is particularly advantageous when phenotypic traits are not readily observable or unknown, such as biochemical or physiological traits (Storz, 2005; Walsh, 2008).

A number of classical neutrality tests have been developed over the years, including the Ewens-Watterson test (Ewens, 1972; Watterson, 1978), Tajima's D test (Tajima, 1989), McDonald-Kreitman test (McDonald and Kreitman, 1991) and tests based on the relationship of synonymous and non-synonymous substitutions (Li *et al.*, 1985; Nei and

Gojobori, 1986). Although these tests are popular they are sensitive to demographic effects (*e.g.* Ewens-Watterson test, Tajima's D test) or rely on extensive gene sequence data that is not necessarily available for non-model species. Recently, F_{st}-outlier tests, first developed by Beaumont and Nichols (1996), became popular because it allowed for the simultaneous analysis of a large number of loci and both dominant and co-dominant marker data could be employed. This method assumes a simple island model, but Excoffier *et al.* (2009) argued that this simple model may not accurately reflect more multifaceted migration patterns and subsequently implemented a hierarchical island model in their execution of the method. Irrespectively, both of these frequentist methods are criticised for its inability to compensate for population and locus-specific effects. Pérez-Figueroa *et al.* (2010) and Narum and Hess (2011) for this reason argue that the reimplementation of the F_{st}-outlier method under a Bayesian statistical framework as in Foll and Gaggioti (2008) may provide more reliable results.

F_{st}-outlier tests have been used in a number of studies in order to detect loci that might be under selection (*e.g.* Bonin *et al.*, 2006; Paris *et al.*, 2010; Nunes *et al.*, 2011) including those for aquatic species such as Atlantic salmon (*Salmo salar*; Vasemägi *et al.*, 2005), lake whitefish (*Coregonus clupeaformis*; Campbell and Bernatchez, 2004), cod (*Gadus morhua*; Nielsen *et al.*, 2009b), Eastern oyster (*Crassostrea virginica*; Murry and Hare, 2006) and periwinkle snail (*Littorina saxatilis*; Wilding *et al.*, 2001) (for a review: Nielsen *et al.*, 2009a). Many of these studies used dominant AFLP markers due to the lack of genomic resources in non-model species. However, the use of microsatellite- and SNP-markers are now increasing (*e.g.* Vasemägi *et al.*, 2005; Nielsen *et al.*, 2009b; Willing *et al.*, 2010; Whiteley *et al.*, 2011).

The South African abalone, locally known as *perlemoen* (*Haliotis midae* Linnaeus, Gastropoda; Haliotidae), is an economically important marine mollusc. The species has a wide distribution along the cool to warm temperate regions of the South African coast ranging from west- (the Western Cape Province) through to the east coast (on the Eastern Cape Provincial seaboard). Although historically an important fisheries species, the commercial sector currently relies mainly on aquaculture due to the suspension of commercial fishery operations in 2008 for conservation purposes.

Previous studies, based mainly on neutral marker analysis, identified subtle population differentiation between wild populations on the west- and east coasts of South Africa coinciding with the major oceanographic characteristics of the South African coastline

(Evans *et al.*, 2004a; Bester-van der Merwe *et al.*, 2011). The current hypothesis suggests that an ancestral population was divided into two isolated populations during the last glacial maxima, approximately 20 000 years ago and a secondary contact zone was established on the south coast after glacial retreat (Bester-van der Merwe *et al.*, 2011). Furthermore, significant population differentiation has been reported between cultured and wild populations of *H. midae* after only one generation of breeding under artificial aquaculture settings. This population differentiation between wild and cultured populations is thought be a result of founder effects and selection for the new artificial aquaculture environment [Evans *et al.*, 2004b; Slabbert *et al.*, 2009; Rhode *et al.*, 2012 (Chapter 2)].

Demographic events that could lead to the observed patterns of population differentiation amongst various *H. midae* populations have been well studied. However, little is known about the contribution made by selection and the underlying genetic variation responsible for adaptation to particular environments. Aberrant patterns of genetic diversity at various loci in both wild (Evans *et al.*, 2004a) and cultured (Rhode *et al.*, 2012) populations have been ascribed to possible selection and it is therefore hypothesised that selection may play an important role in population divergence. Identifying adaptive genetic diversity may aid in defining management units in conservation and fisheries management. This has particular relevance for marine environments where populations often show little differentiation, based on neutral genetic diversity, because of few restrictions to migration and large effective population sizes. Thus the assumption of panmixia is often made. However, recent studies suggest population divergence due to adaptation even in cases where gene flow is prevalent (Hemmer-Hansen *et al.*, 2007; Nielsen *et al.*, 2009a, b; André *et al.*, 2011). These adaptations, leading to the development of ecotypes, are often correlated to heterogeneous or clinal environments across the species' natural distribution range (Schmidt *et al.*, 2008; Mariac *et al.*, 2011). In the case of the South African coast, such environmental heterogeneity is consistent with the temperature gradient along the cool temperate Atlantic Ocean on the west and the warm temperate Indian Ocean on the eastern seaboard of the country.

Furthermore, the domestication process of abalone in South Africa is in its initial stages, with some farms only now starting to retain F_1-generation animals for broodstock replacement. In other abalone species aquaculture has already produced the third and up to the sixth generation of animals in captivity (*e.g.* Li Q *et al.*, 2004; Praipue *et al.*, 2010).

Regardless of this initial phase of domestication, the elevated levels of population differentiation between wild and cultured populations of *H. midae* raise some questions: Which evolutionary forces are driving the domestication of abalone? What are the effects of different husbandry practises? What genes or loci are affected and how does this relate to the genetic constitution of cultured abalone? Are there any conservational concerns, implications for ranching or stock enhancement initiatives? Will this impact the future prospects for genetic improvement and management of commercial stocks in relation to conservation of wild populations?

The aim of this study, therefore, was to conduct a population genomic analysis of genome-wide microsatellite markers to identify candidate loci that may be under selection and that could elucidate the underlying adaptive genetic variation responsible for phenotypes (often cryptic) in wild and cultured abalone.

3.2. Materials and Methods

3.2.1. Study Populations, Specimens and Microsatellite Markers

Permits to collect and transport abalone for the purpose of research were obtained from the Department of Agriculture, Forestry and Fisheries of the Republic of South Africa. Thirty two wild animals from three localities (96 animals in total) representing the major geographical regions of the natural distribution of *H. midae* in South Africa were included: Saldanha Bay (geographic coordinates: 33°02'40.64''S; 17°56'00.53''E) on the west coast (WPWC); Witsand (geographic coordinates: 34°20'53.37''S; 19°01'39.75''E) on the south coast (WPSC) and Riet Point (geographic coordinates: 33°31'29.31''S; 27°06'51.18''E) on the east coast (WPEC) (populations as described in Bester-van der Merwe *et al.*, 2011). In total 96 F_1 cultured animals were collected from three aquaculture facilities (32 animals per facility), one each from the west- (CPWC) (geographic coordinates: 32°45'30.00''S; 18°01'40.00''E), south- (CPSC) (geographic coordinates: 34°35'04.00''S; 19°19'45.63''E) and east- (CPEC) (geographic coordinates: 32°45'43.20''S; 28°15'0.00''E) coast of South Africa. The respective aquaculture populations originated from wild broodstock animals collected from the corresponding geographic regions. Cultured animals were randomly selected,

across spawning cohorts, in order to attain a representative sample of the total F_1 population on each respective facility. All the wild abalone were adult, reproductively active animals. Cultured animals were juveniles of ages between three and four years and had gone through the entire production system, including several grading procedures according to each facility's specifications.

Muscle and/or gill tissues were collected from each individual and placed in 70% ethanol and stored at -20°C. DNA extraction was performed using the standard CTAB method of Saghai-Maroof *et al.* (1984). One-hundred-and-fifty microsatellite markers previously developed for *Haliotis midae* (Bester *et al.*, 2004; Slabbert *et al.*, 2008, 2010; Rhode *et al.*, 2012; Slabbert *et al.*, 2012) were selected providing an estimated average marker density of approximately 10 cM across the genome (estimated genome size: ~1400 cM; Franchini *et al.*, 2010). All individuals were genotyped for all markers. Polymerase chain reactions were done in a total volume of 10 µl using the Qiagen® multiplex kit, with all other conditions as described by the authors. This was followed by capillary electrophoreses and allele size scoring using GeneMapper® v.4 (Applied Biosystems) software.

3.2.2. Identifying Candidate Loci Under Selection

In order to identify candidate markers under selection, three F_{st}-outlier detection methods, as implemented in Lositan v.1.44 (Antao *et al.*, 2008), BayeScan v.2.01 (Foll and Gaggiotti, 2008) and Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010) was used as well as the classical Ewens-Watterson homozygosity test for neutrality *via* the exact test implementation of Slatkin (1994) in PyPop v.0.7 (Lancaster *et al.*, 2007). Separate analyses were run on wild and cultured population cohorts as well as a combined dataset containing all populations (wild and cultured) for computations in Lositan, BayesScan, and the Ewens-Watterson test in order to allow comparison with the Arlequin hierarchical results. Lositan parameters were as follow: 50 000 simulations, with a 95% confidence interval and a false discovery rate of 0.1, assuming the infinite alleles model. For BayeScan default parameters as set for co-dominant markers were used; statistical confidence levels were set according to the Jeffreys' scale: a Bayes factor (BF) greater than 10 [$\log_{10}(BF) > 1$] was interpreted as sufficient evidence for selection. The Ewens-Watterson test was run with 10 000 replicates and significance was set at $P < 0.05$. Outlier detection in Arlequin assumed the hierarchical

island model, populations were grouped as either wild or cultured, with 10 000 simulations (number of demes: 100, number of groups: 10), significance cut-off was set at $P < 0.05$.

For a locus to qualify as a candidate for selection it should have demonstrated congruent evidence for either directional or balancing selection in at least two of the tests done across any of the population cohorts investigated. The loci were then subdivided into three datasets: candidate neutral-, candidate directional selection-, and candidate balancing selection markers.

3.2.3. Genetic Diversity, Linkage Disequilibrium and Population Differentiation

Hardy-Weinberg equilibrium (exact probability test, 10 000 dememorisation, 500 batches, and 5000 iterations per batch) was computed using Genepop v.4.0 (Rousset, 2008). The following diversity statistics was calculated in GenALex v.6.41 (Peakall and Smouse, 2006): number of alleles, effective number of alleles, information index and heterozygosity. A Kruskal-Wallis test was performed to evaluate significant differences in number of alleles, effective number of alleles and heterozygosity between candidate markers under selection and neutral markers as well as between cultured and wild populations. To evaluate the relationship between LD and genetic distance (cM) two statistics were calculated: the LD parameter, D' [as defined by Hedrick (1987) for multi-allelic loci] and the standardised χ'^2 statistic (Zhao *et al.*, 2005) for each syntenic locus pair (pairs of loci on the same linkage group) of *H. midae* (linkage map, Vervalle *et al.*, in press). Furthermore, to test the LD due to functional associations, both statistics were calculated for each pair of candidate loci under selection. All pairwise LD statistics were computed using Pypop with significance testing by permutation (1000, $P < 0.05$ [for the χ'^2 Pypop calculates Cramer's V statistic (Cramer, 1946) equivalent to the square root of χ'^2]. In order to investigate the LD patterns that arise due to the domestication effect, but to maintain a cross-population comparison to better reflect the manner in which outlier loci were identified, the analyses were done for wild and cultured population respectively. To test whether LD amongst candidate selection loci was significantly different to the base LD, a Kruskal-Wallis test was performed.

Pairwise F_{st} (10 000 permutations at 5% significance level) between populations was estimated and a locus by locus molecular analysis of variance (AMOVA; 10 000 permutations) was also computed in Arlequin; populations were grouped as either cultured

or wild. Finally, dendrograms were constructed based on Nei's genetic distance (Nei, 1972) using the neighbour joining clustering algorithm in PowerMarker v.3.25, with 1000 bootstrap replicates (Liu and Muse, 2005).

3.2.4. Analyses for Possible Cause for Outlier Behaviour of Loci

Outlier loci were subjected to bioinformatic analysis, in order to identify possible association to known functional sequences. The bioinformatics protocol of Faber and Medrano (2003, 2004) was followed. In brief: repeat regions were masked using RepeatMasker v.3.3.0 (Smit *et al.*, 2011) to prevent superfluous hits due to microsatellite repeat motifs. The masked sequences were then used to conduct BLASTx and BLASTn (Altschul *et al.*, 1990) searches against the nr-protein and nr-nucleotide databases of NCBI. Masked sequences were also screened against the Repbase database (Jurka *et al.*, 2005) *via* the CENSOR v.4.2.27 program (Kohany *et al.*, 2006) to identify possible associations to dispersed repetitive elements. Hits with the smallest e-value (cut-off: e-value < $1e^{-04}$), the highest similarity and/or score were assumed to be the most likely homologue.

A case-control study design using a permutation-based distance test for genotypic differentiation (Prevosti distance, 10 000 permutations in PowerMarker) was used to evaluate the association of candidate loci under selection with domestication (wild vs. cultured) and a particular population (WPWC vs. WPSC vs. WPEC vs. CPWC vs. CPSC vs. CPEC). First significant genotypic differentiation at particular loci was tested between wild and cultured population cohorts; then each population in turn was compared to all other populations. Significance threshold was set to $P < 0.05$.

Environmental data from each geographic region was obtained from the South African Weather Service and the South African Data Centre for Oceanography. The average winter (SST-WinAve) and summer (SST-SumAve) sea surface temperatures, the average maximum summer (SST-SumMax) and minimum winter (SST-WinMin) sea surface temperatures as well as the average difference in the summer maximum and winter minimum sea surface temperature (ΔSST-SumWin) was calculated using data from the years 2000 to 2010. Average oxygenation and salinity levels were also calculated for the years 1992 to 2007 (Table 3.1). To evaluate the effect of these variables on population differentiation a regression analysis was performed with GESTE v.2.0 (Foll and Gaggiotti, 2006).

Table 3.1: Summary of environmental variables for each geographic region.

Environmental Variable	West Coast	South Coast	East Coast
SST-SumMax (°C)	22.10	27.00	24.90
SST-SumAve (°C)	18.90	21.82	20.46
SST-WinMin (°C)	9.50	12.00	13.00
SST-WinAve (°C)	14.16	14.87	15.87
ΔSST-SumWin (°C)	12.60	15.00	11.90
Dissolved O_2 (PPT)	4.16	4.66	4.59
Salinity (PPT)	35.00	35.17	35.04

°C – degrees Celsius
PPT – parts per thousand
SST – sea surface temperature

3.3. Results

3.3.1. Candidate Loci Under Selection

The highest number of candidate loci for selection was identified with Lositan, resulting in 48 candidate loci (14 under directional selection and 34 under balancing selection) across all populations. On the contrary, analysis with BayeScan resulted in the lowest number of candidate loci: 26 (across all populations), with 16 candidates for directional selection and 10 candidates for balancing selection. When considering wild- and cultured population cohorts separately, the number of loci identified by Lositan and BayeScan was reduced by almost half: total number of loci identified by Lositan across wild populations (23) and across cultured populations (26); total number of loci identified by BayeScan across wild populations (9) and across cultured populations (10). This reduction was not mirrored by the F_{st} (among populations) and F_{ct} (among groups of populations) estimates of Arlequin's hierarchical analysis or the Ewans-Watterson test. There were varying degrees of overlap in

the loci detected between the different methods, but overall 41 loci had congruent results with at least two of the detection methods employed: Fourteen loci were candidates for directional selection and 27 were candidates for balancing selection (Appendix B: Table S3.1).

3.3.2. Genetic Diversity, Linkage Disequilibrium and Population Differentiation

A summary of Hardy-Weinberg test statistics per locus per population are shown in the supplementary information (Appendix B: Table S3.2). A number of loci seem to violate the Hardy-Weinberg assumptions and demonstrate high F_{is} values indicating heterozygous deficiency. This is most likely due to the presence of null alleles, commonly found at microsatellite loci in abalone (*e.g*. Slabbert *et al*., 2008, 2010, 2012.). For the most part null alleles affect loci at random and across populations; because the population genomics approach takes into account stochastic genome-wide effects it is not anticipated to influence the results significantly. Therefore, these loci were not excluded from the analyses. Loci under selection are, also, expected to deviate from Hardy-Weinberg expectations. There were no overall significant differences ($P > 0.05$) in genetic diversity between wild and cultured populations, even though cultured populations demonstrated slightly reduced diversity across most measures. Diversity estimates across the three locus sets (for directional selection, balancing selection and neutral loci) however did reach significance ($P < 0.05$): Heterozygosity, number of alleles, effective number of alleles and the information index for candidate loci for balancing selection was significantly greater than that for neutral loci. On the contrary, candidate loci for directional selection demonstrated significantly less diversity as measured by number of alleles, effective number of alleles and the information index compared to neutral loci; heterozygosity between candidate directional selection loci and neutral loci failed to reach significance. Figure 3.1 summarises the various statistics across the six populations.

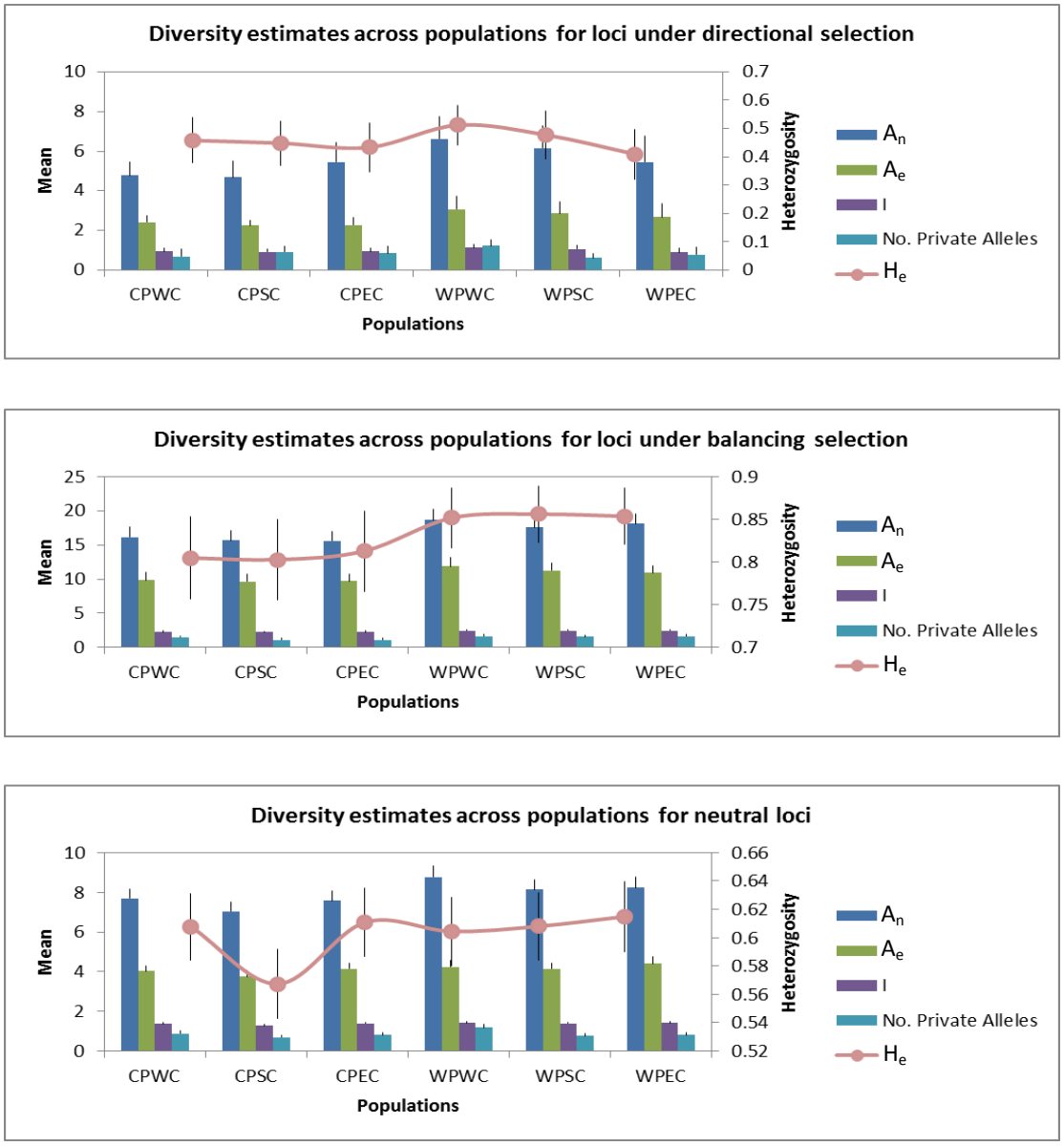

Figure 3.1: Summary statistics for mean diversity estimates, including number of alleles (A_n), effective number of alleles (A_e), information index (I), number of private alleles and heterozygosity (H_e) for candidate loci under selection and neutral loci per population.

Across wild populations approximately 10% of the syntenic locus pairs were in significant LD ($P < 0.05$); whilst across cultured populations this was more than double, at approximately 23% (Appendix B: Table S3.3). Genome-wide LD was relatively high, with similar mean D' and χ'^2 estimates for wild and cultured populations, although estimates for cultured populations were slightly elevated – D': 0.467 (± 0.0128) and χ'^2: 0.188 (± 0.0088) for wild populations and D': 0.473 (± 0.0147) and χ'^2: 0.2102 (± 0.0122) for cultured populations. D' estimates were inflated in comparison to the χ'^2 estimate; however both

statistics provided similar trends for LD. Wild populations seem to have reached the base level of LD among populations (noting the almost zero slope), whilst for cultured populations LD decayed as a function of genetic distance (cM) (Figure 3.2). For candidate loci under selection, 72 locus pairs reached statistically significant levels of LD ($P < 0.05$) for cultured populations and 63 locus pairs for wild populations (Appendix B: Table S3.4). The mean D'- and the χ'^2 statistic for pairs of candidate loci under selection was significantly higher (Kruskal-Wallis test, $P < 0.001$) than the corresponding estimate for pairs of syntenic loci: wild (D': 0.657±0.0189; χ'^2: 0.303±0.0167) and cultured (D' 0.689±0.0158; χ'^2: 0.317±0.0135). Linkage disequilibrium for these loci was not constrained to individual linkage groups, with significant LD spanning across putative chromosomes.

With the exception of the pairwise F_{st} estimates based on candidate loci for balancing selection, pairwise F_{st} estimates demonstrated evidence for population differentiation amongst almost all populations (Table 3.2). Interestingly, the pairwise F_{st} estimate based on directional selection loci failed to reach statistical significance for CPWC and CPSC. Furthermore, the only pairwise F_{st} estimate based on balancing selection loci to demonstrate significant population differentiation was between CPEC and CPSC. Similarly, the AMOVA results reflected the significant population differentiation as measured using the neutral and directional selection loci. For balancing selection loci average F-statistics remained low, but was significant for the F_{sc}- and F_{st} estimates (Table 3.3). The dendrograms corroborated the F-statistics, showing the close relationship between the CPWC and CPSC. Furthermore, the dendrograms also showed that selection is a major driving force for the development of the domestic phenotype as illustrated by the large genetic distance between the wild and cultured population clusters when using loci under positive selection (Figure 3.3).

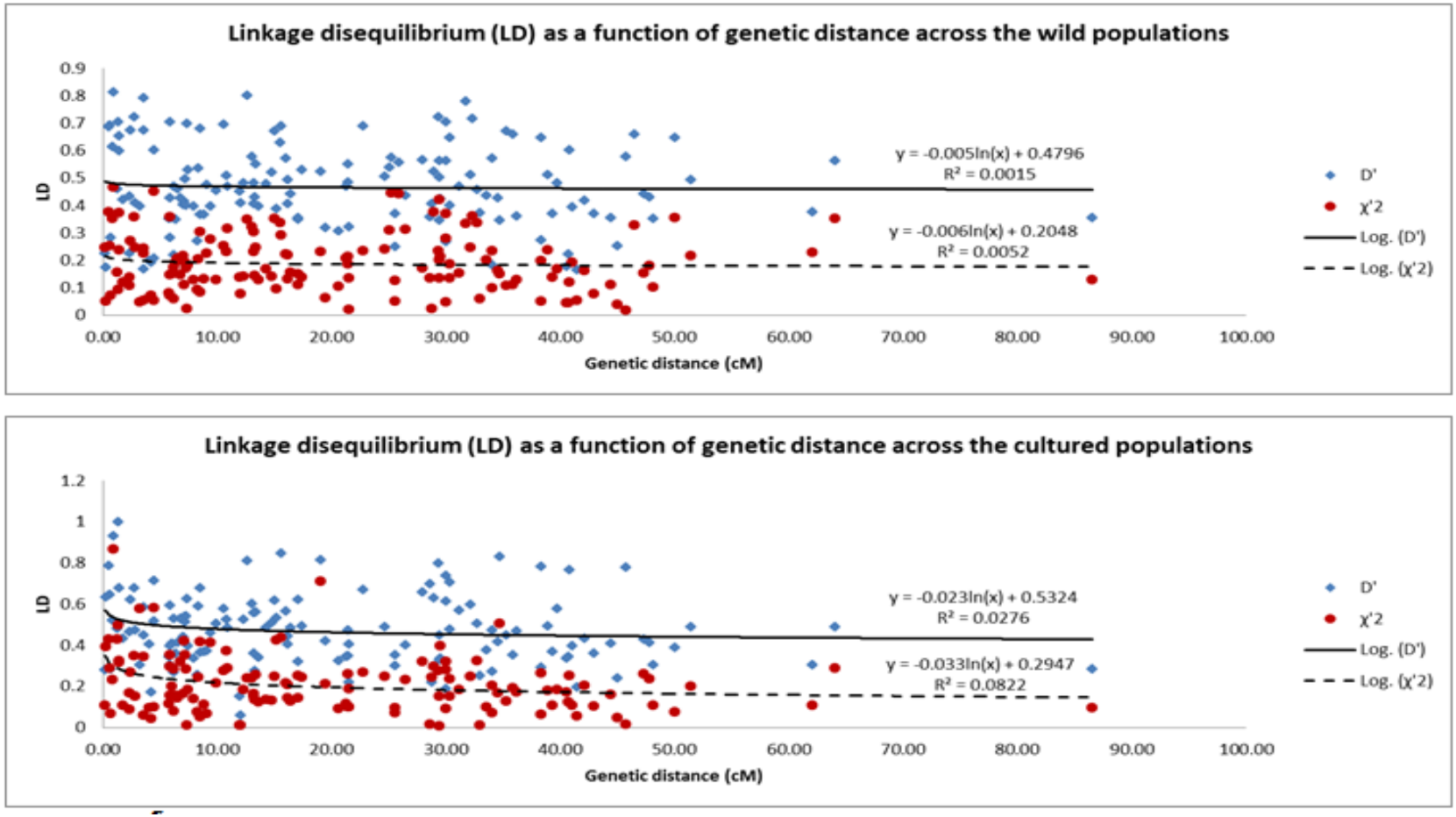

Figure 3.2: Linkage disequilibrium as measured by the D'- and χ'^2 statistic for each syntenic locus pair as a function of genetic distance (cM).

Table 3.2: Pairwise F_{st} estimates based on candidate loci for directional-, balancing selection and neutral loci.

Pairwise F_{st} estimates for candidate loci under directional selection						
	CPWC	**CPSC**	**CPEC**	**WPWC**	**WPSC**	**WPEC**
CPWC	-					
CPSC	-0.00566	-				
CPEC	0.09226*	0.14871*	-			
WPWC	0.21142*	0.23455*	0.25499*	-		
WPSC	0.21227*	0.24076*	0.29459*	0.04775*	-	
WPEC	0.21724*	0.23716*	0.30736*	0.02501*	0.05640*	-
Pairwise F_{st} estimates for candidate loci under balancing selection						
	CPWC	**CPSC**	**CPEC**	**WPWC**	**WPSC**	**WPEC**
CPWC	-					
CPSC	0.00313	-				
CPEC	0.00449	0.00722*	-			
WPWC	0.00001	-0.00054	0.00159	-		
WPSC	- 0.00372	-0.00269	0.00035	-0.00188	-	
WPEC	0.00203	0.00031	0.00248	0.00056	-0.00266	-
Pairwise F_{st} estimates for candidate neutral loci						
	CPWC	**CPSC**	**CPEC**	**WPWC**	**WPSC**	**WPEC**
CPWC	-					
CPSC	0.00851*	-				
CPEC	0.01815*	0.01476*	-			
WPWC	0.01486*	0.01131*	0.02012*	-		
WPSC	0.01459*	0.01889*	0.02194*	0.01627*	-	
WPEC	0.01443*	0.01428*	0.01332*	0.01005*	0.00809*	-

** Significant P-value at the 1% nominal level*

Table 3.3: AMOVA results based on candidate loci for directional-, balancing selection and neutral loci, populations grouped as cultured or wild.

Candidate loci for directional selection				
Source of Variation	**Sum of Squares**	**Variance Components**	**% Variation**	**Average F-statistic (* *P* < 0.05)**
Among groups	121.103	0.6634	17.70666	F_{CT} : 0.17707*
Among populations within groups	49.664	0.16054	4.28501	F_{SC} : 0.05207*
Within populations	951.015	2.92267	78.00834	F_{ST} : 0.21992*
Candidate loci for balancing selection				
Source of Variation	**Sum of Squares**	**Variance Components**	**% Variation**	**Average F-statistic (* *P* < 0.05)**
Among groups	15.329	0.0067	0.06191	F_{CT} : 0.00062
Among populations within groups	56.692	0.06123	0.56597	F_{SC} : 0.00566*
Within populations	3620.499	10.75047	99.37212	F_{ST} : 0.00628*
Candidate neutral loci				
Source of Variation	**Sum of Squares**	**Variance Components**	**% Variation**	**Average F-statistic (* *P* < 0.05)**
Among groups	89.638	0.15119	0.44726	F_{CT} : 0.00447*
Among populations within groups	249.302	0.54005	1.59758	F_{SC} : 0.01605*
Within populations	11092.658	33.1129	97.95516	F_{ST} : 0.02045*

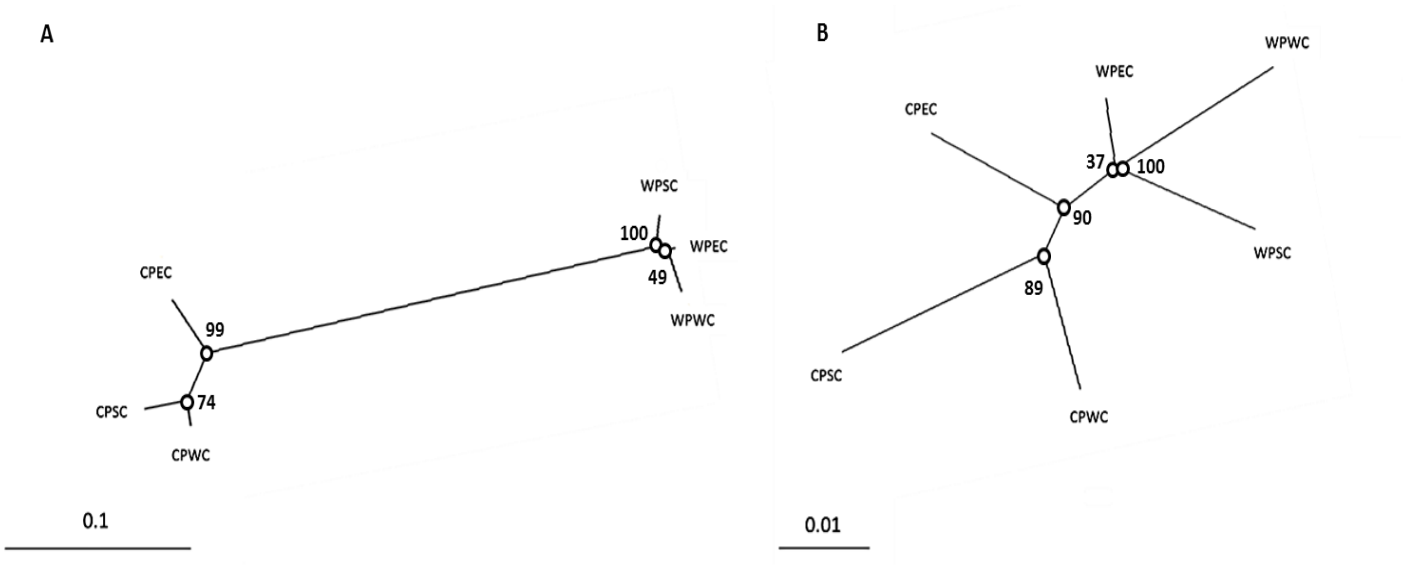

Figure 3.3: A: Dendrogram based on genetic distance (Nei, 1972) as calculated using candidate loci for directional selection and clustered *via* the Neighbour Joining algorithm. B: Dendrogram based on genetic distance (Nei, 1972) as calculated using candidate neutral loci and clustered *via* the Neighbour Joining algorithm. Nodal values: bootstrap replicates (in percentage) that supported the partitioning of branches.

3.3.3. Cause of Outlier Behaviour of Loci

No association between the environmental data and population differentiation could be found (highest probability model with P = 0.081, excluded all environmental variables as causative to population differentiation). However, a number of loci were associated with particular populations. Twelve of the 14 loci, suggested being under directional selection, demonstrated association with domestication. A few loci under balancing selection also demonstrated association to particular populations and domestication (Appendix B: Table S3.5). Seven loci had significant hits to known genic sequences (Table 3.4) and three additional loci (*HmidILL-146360*, *HmidILL-87955*, *HmidILL-118779*) that are known to be derived from expressed sequences were therefore *de facto* gene associated. Five loci had significant similarity to known transposable elements (Table 3.5).

Table 3.4: Candidate loci under selection with significant similarity to known genes.

Locus name	BLAST hit	e-Value	Identity/Positive (%)	Gene function
HmNR120	SH2 domain containing protein (*Danio rerio*) XP_687225.1	3.0e-05	74	Signal transduction
HmLCS5	ATPase α-subunit (*Haliotis rubra*) AY04305.1	7.0e-19	89	Energy metabolism
HmidILL-2192	14-3-3 protein zeta (*Bombyx mori*) NP_001040164.1	3.0e-102	78	Signal transduction
HmidILL-37506	Ras-related protein Rab-1a (*Haliotis discus*) ABO26625.1		100	Regulation of cellular vesicular transport
HmidILL-088398	LIM domain containing	3.0e-25	96	Cytoskeletal organisation

	protein 2-like (*Anolis carolinensis*) XP_003222427.1			and gene expression
HmidILL-64129	D-Lactate dehydrogenase (*Octopus vulgaris*) BAB33312.1	1.0e-75	79	Pyruvate metabolism
HmidILL-070036	Transaldolase (*Strongylocentrotus purpuratus*) XP_792583.2	3.0e-24	77	Reductive biosynthesis
HmidILL-076149	Adenosylhomocysteinase (*Pediculus humanus*) XP_002427522.1	3.0e-129	66	Methylation of biomolecules, including: phospholipids, proteins, DNA and RNA.

Table 3.5: Candidate loci under selection with significant similarity to known transposable elements.

Locus Name	Hit	Similarity (%)	Score
HmRS129	DNA transposon (EnSpm)	80.0	253
HmLCS48	DNA transposon (Polinton)	74.4	291
Hmid65	DNA transposon (Polinton)	68.3	323
HmidPS1.549	DNA transposon (hAT)	77.3	349
HmidPS1.559	Non-LTR retrotransposon	80.2	327

3.4. Discussion

This study investigated the effects of selection on genome-wide genetic diversity in the South African abalone, *Haliotis midae*. Several loci were identified as candidate regions under directional- or balancing selection. Although, the environmental variables investigated in this study could not explain the observed population differentiation, a number of candidate loci under selection were significantly associated to particular populations. The majority of loci under possible directional selection were also significantly associated to the domestication of abalone, *i.e.* loci that showed the most divergent genotypes between wild and cultured populations.

3.4.1. Candidate Loci Under Selection

This study employed four methods to identify candidate loci under selection. Three of these were based on detecting loci with excessively high or low estimates of locus-specific F_{st} among populations – F_{st}-outlier tests. Lositan implements the procedure based on FDist (Beaumont and Nichols, 1996). This is a frequentist method that uses coalescent simulations to determine the null distribution of F_{st} estimates under the neutral theory, assuming an island model – implicitly, assuming equal population sizes, -migration rates and -F_{st} variances amongst populations (Wright, 1931). Excoffier *et al.* (2009) proposed expanding this method by assuming a hierarchical island model (implemented in Arlequin) to compensate for more complex demographic scenarios, where migration rates may typically be higher amongst smaller sub-populations within a group than migration rates between groups. The third method used, as implemented in BayeScan, functions under a Bayesian statistical framework as described in Foll and Gaggioti (2008). This method makes no assumption on the equivalence of F_{st} variances amongst populations and allows for locus- and population-specific effects, but assumes a Dirichlet distribution for allele frequencies. Lastly, the Ewens-Watterson test assumes panmixia, whereby an excess or deficit of homozygotes could be explained by selection (Ewens, 1972; Watterson, 1977, 1978).

All these tests differ in the fundamental assumptions; albeit subtle in some cases, these varying assumptions lead to significant disparity in the resulting loci being detected as possible candidates for selection. This is reflected in the current data where the number of loci detected by each test differs. There is a fairly high degree of overlap in loci detected by the three F_{st}-outlier tests, but to a lesser extent with the results obtained by the Ewens-Watterson test (Appendix B: Table S3.1). A number of studies, comparing the Beaumont and Nichols (1996) method to a Bayesian method (BayeScan or analogous program), reported on the disparity in number of loci identified, with the latter method always resulting in a smaller set of candidate loci (Vasemägi *et al.*, 2005; Namroud *et al.*, 2008; Paris *et al.*, 2010; Nunes *et al.*, 2011). Two recent theoretical-simulation studies testing the robustness of various outlier detection methods, using dominant- (Pérez-Figueroa *et al.*, 2010) and co-dominant markers (Narum and Hess, 2011) concluded that the inability of the Beaumont and Nichols (1996) method to compensate for differing variances in F_{st} amongst populations, under a simple island model, may result in an increased type I error rate (false

positive) - especially with the detection of loci under balancing selection. Narum and Hess (2011) further argued that the hierarchical island model, as proposed by Excoffier *et al.* (2009), may further inflate both type I and type II error rates (false negative) if neutral genetic diversity contrasts adaptive genetic diversity.

The current data for *H. midae*, suggests a similar trend in results for the different outlier detection methods as do aforementioned investigations. Noting that BayeScan consistently showed the smallest set of loci as candidates for selection. In particularly, the inequality in the number of candidate balancing selection loci identified: 10 candidate balancing selection loci by BayeScan versus the 34 and 23 by Lositan and the Arlequin (hierarchical method) respectively, across all populations (Appendix B: Table S3.1). It is reasonable to suggest that many of these candidate balancing selection loci may thus be false positives. The apparent lack of congruence of the Ewens-Watterson test results to that of the outlier tests could be attributed to the rigidness of the panmixa assumption; it is known that this test may fail under complex demographic scenarios (Slatkin, 1982). Furthermore this test may have reduced power if high recombination rates persist in a population or samples sizes are small (Zhai *et al.*, 2009).

Based on the defined set of candidate loci under selection, 9.3% (14/150) of the *H. midae* genome maybe affected by divergent selection, whilst 18% (27/150) could be influenced by balancing selection; in total approximately 27% of genetic variation in the genome is presumably influenced by selection. This is an relatively large proportion with estimates from other organisms, including various species of plants (Scotti-Saintagne *et al.*, 2004; Herrera and Bazaga, 2008; Namroud *et al.*, 2008), a mosquito (Paris *et al.*, 2010), the common frog (Bonin *et al.*, 2006), the ocellated lizard (Nunes *et al.*, 2011), a number of fish species (Campbell and Bernatchez, 2004; Williams *et al.*, 2008; Nielsen *et al.*, 2009; Nielsen *et al.*, 2009b; Willing *et al.*, 2010; Whiteley *et al.*, 2011), wolves (Hagenblad *et al.*, 2009) and Holstein cattle (Qanbari *et al.*, 2010) ranging between 2.6% and 12%. In a bivalve mollusc (*Crassostrea virginica*), the estimate was even lower at 1.86% (Murray and Hare, 2006); however, for *Littorina saxatilis* (periwinkle snail; a gastropod mollusc) the estimate was slightly higher at 5% (Wilding *et al.*, 2001).

A more recent estimate for *L. saxatilis*, found that approximately 12% of the genome was putatively influenced by selection based on gene associated loci (Galindo *et al.*, 2010). This result is contrary to the previous molluscan and other aforementioned studies which used anonymous AFLP markers. A fairly high estimation was made for Atlantic salmon (24%, *Salmo salar*) based on 75 EST-associated microsatellites (Vasemägi *et al.*, 2005). Although it is known that many of the microsatellites used in the current study for *H. midae* were derived from expressed sequences, the majority are anonymous, derived from genomic fragments. However, a recent study by Rhode and Roodt-Wilding (2011) found that a significant number of *H. midae* microsatellites are located in/or close to genic sequences. Considering the high type I error rate of frequentist F_{st}-outlier methods (especially with the estimation of loci under balancing selection, which are particulary difficult to detect in general) using the more conservative BayeScan estimate for number of loci under balancing selection (10 loci) the percentage of loci under selection could be reduced from 27% to 16% (14 loci under directional selection, 10 loci under balancing selection out of 150 loci). Assuming a substantial proportion of these loci are gene-linked, 16% may therefore be a more realistic estimate.

Furthermore, the elevated level of directional selection could be explained by the use of aquaculture populations. These populations are F_1 generation animals; it is therefore expected that these animals will be under great selective pressure to adapt to the new artificial environment and may also be affected by artificial selection, noting that animals are subjected to various grade-and-cull procedures during production. Innan and Kim (2004) argued that comparisons of genetic polymorphism between domestic and wild populations could significantly increase the power for detecting loci underlying domestication. This observation is supported in the current study by the separate estimates for number of loci under selection for wild (6%) and cultured (6.67%) populations, as well as the association analysis that found all, but two, candidate loci for positive selection associated with domestication (Appendix B: Table S3.5). The estimates for genetic distance, as calculated using said marker cohort, in the dendrogram also provide evidence for this (Figure. 3.3). Similar estimates for genome-wide directional selection were also found for the domestic dog (±8.0%, Akey *et al.*, 2010) and European cattle (±9.5%, Medugorac *et al.*, 2009). Some caution with regards to the interpretation of the number of candidate loci for selection is,

however, warranted. The many loci deviating from Hardy-Weinberg expectations due to homozygous excess could inflate the number of candidate loci for directional selection. This abundance of homozygous excess could be attributed to the presence of null alleles, common for microsatellite loci (Campagne *et al*., 2012); however, a locus under directional will demonstrate a similar pattern of genetic diversity. Nonetheless, using a population genomics approach may compensate for possible overestimates if null alleles occur at random and affects all populations equally. If null alleles are, however, population specific it could cause a false positive result. Generally, population genetics approaches to detect null alleles use homozygous excess to estimate null allele frequencies, but this approach will assume all homozygous excess as a result of null alleles, despite the possibility of directional selection. Ideally, investigating patterns of allelic segregation using extensive pedigree data will provide an indication of the presence of null alleles, but is difficult to conduct in natural populations (Pompanon *et al*., 2005).

However, it is not always possible to draw direct comparisons among studies, because different studies use different cut-off values for significance, differing marker types and species biology must also be taken into account. It is expected, as demonstrated in the aforementioned comparisons that anonymous markers (such as AFLPs) will demonstrate lower number of loci under selection in comparison to EST-derived microsatellite- or SNP markers. Anonymous microsatellites in *H. midae* may show higher percentage of loci under selection than other studies, because species-specific dynamics of microsatellites allows for possible gene association - even if the gene is not known.

3.4.2. Patterns of Genetic Diversity, Population Differentiation and Linkage Disequilibrium

Previous investigations comparing genetic diversity between cultured and wild populations of the South African abalone were conducted using a limited number of molecular markers (Evans *et al.*, 2004b; Slabbert *et al.*, 2009; Rhode *et al.*, 2012). Results were often contradicting, with some cultured populations demonstrating reduced genetic diversity, whilst in others no significant decrease could be detected. There are various factors, including experimental design, biological processes, and mode of stock management that could explain such contradictory findings (Taniguchi, 2003; Slabbert *et al.*,

2009; Rhode *et al.*, 2012). The present genome-wide polymorphism data suggests that cultured populations have slightly decreased genetic diversity estimates based on allelic content at loci and/or heterozygosity levels. The noticeable decrease in genetic diversity is an expected consequence of the founder event, when breeding populations were established on respective aquaculture facilities, with some influence of stochastic factors (Taniguchi, 2003; Roodt-Wilding, 2007; Rhode *et al.*, 2012).

Expectantly, patterns of genetic diversity amongst the candidate loci under selection were in contrast to that of neutral loci. Candidates for balancing selection demonstrated a diversity excess as result of heterozygote advantage. Candidates for directional selection demonstrated significantly reduced diversity, most likely due to functional constraints (Li Y-C *et al.*, 2004). Interestingly, the candidate directional selection loci had similar heterozygosity estimates as neutral loci and could be explained as an artefact of the multi-allele microsatellite loci used in the current study. It is well-known that microsatellite-derived heterozygosity estimates are insensitive to the loss of low frequency alleles (Evans *et al.*, 2004b).

Pairwise F_{st} estimates based on both candidate neutral and -directional selection loci supported population differentiation. However, directional selection loci showed evidence for moderate to strong differentiation; indicating the effect of divergent selection, especially between population pairs with cultured populations. Candidate neutral loci, on the other hand supported subtle differentiation, by means of random genetic drift, in accord with previous estimates for wild populations (Bester-van der Merwe *et al.*, 2011). Pairwise F_{st} estimates based on candidate balancing selection loci showed, for the most part, a lack of population differentiation; demonstrating the maintenance of similar diversity across populations. The CPSC vs. CPWC pairwise F_{st} estimate (based on directional selection) failed to reach statistical significance; this could be explained by convergent evolution, i.e. adaptation to similar aquaculture environment/practices or artificial selection for favourable production traits. Such convergent evolution under domestication is well-known amongst crop and livestock species (Glémin and Bataillon, 2009; Wright *et al.*, 2010). Interestingly, the only pairwise F_{st} estimate based on candidate balancing selection loci that was significant was for CPSC vs. CPEC. Noting that these two populations show association to different balancing selection loci, this could explain the significant differentiation. It may be

that heterozygotes at these loci are favoured, but there could also be differential selection on particular alleles (allelic combinations within the heterozygous genotype) within each population. Other populations also show association with particular candidate loci under balancing selection and could in part explain the AMOVA result (based on balancing selection) showing significant differentiation amongst populations within groups and within populations (Table 3.3).

Amongst the wild populations it is noteworthy that the south coast (WPSC) population seems to be more distinct than the other wild population based on the adaptive diversity pairwise F_{st} values. Considering that this population represents a secondary contact zone between the west- and east coast populations (Bester-van der Merwe *et al.*, 2011), it may be that WPSC represents a unique arrangement of adaptive diversity. It has long been recognised that such hybridisation zones may represent "evolutionary test laboratories" where novel allelic combinations (within and across loci) may facilitate the development of new adaptations. Furthermore, Counterman *et al.* (2010) argued that these contact areas may be important as "population sieves" for adaptive variation, by allowing a degree of gene flow without compromising adaptability to local environmental conditions. Such gene flow between west- and east coast South African abalone populations has previously been reported with the south coast probably acting as a filtering corridor (Evans *et al.*, 2004a; Bester-van der Merwe *et al.*, 2011).

Linkage disequilibrium across wild and cultured populations of the South African abalone seem to be extensive, with relatively high D' and χ'^2 values. As expected the D' estimate was inflated in comparison to the χ'^2 estimate - it has been reported that D' may be sensitive to small sample size and the presence of low frequency alleles at highly polymorphic loci; leading to an overestimation of the level of LD (Slate and Pemberton, 2007). Nonetheless, both statistics provided similar trends for LD (Figure 3.2). The high estimates for LD would suggest inbred populations as found for the endangered Scandinavian wolf (Hagenbald *et al.*, 2009). However, the high diversity and previous estimates for relatedness for these populations [Rhode *et al.*, 2012 (Chapetr 2)] suggests the contrary. The significant population differentiation is more likely the cause of this significant LD. Populations sub-structuring results in a heterozygous deficit over the total population due to the Wahlund effect (Wahlund, 1928). Consequently the lack of recombinant genotypes leads to a

significant decrease in the rate of LD decay. The observation made for *H. midae* is supported by similar observations in other outbred structured populations (Slate and Pemberton, 2007; Meadows *et al.*, 2008; Li and Merila, 2010, 2011).

The apparent lack of LD decay (given by the slope of trend line, Figure 3.2), low R^2 value and small number of pairs of syntenic loci in significant linkage disequilibrium amongst the wild populations is probably because the populations have reached the base level of LD maintained by the population sub-division. The slightly increased level of LD amongst the cultured populations is likely due to the recent founder event (with the establishment of commercial broodstock populations) and most possibly also a result of the selective sweep due to the domestication effect: Noting that many of the locus pairs that demonstrated significant LD are with candidate loci under selection. As a result the cultured populations have not yet reached the base level of LD and thus explaining the more significant correlation between LD and distance (cM) (Figure 3.2).

3.4.3. Biological and Functional Interpretation of Outlier Loci

Locus pair *HmAD102* - *HmRS129* (inter-locus distance: 12.6 cM) on linkage group 6 and locus pair *HmAD102* - *HmRS129* (inter-locus distance: 0.6 cM) on linkage group 1 has a comparatively small inter-locus genetic distances on the respective chromosomes (Appendix B: Table S3.4). It is therefore not possible to discern at present whether the significant LD is due to functional linkage or because of hitch-hiking as a result of a recent selective sweep (Karasov *et al.*, 2010). Nonetheless, the significant increase in LD between pairs of loci under selection, even beyond intra-chromosomal associations, would suggest functional linkages amongst these loci and selection on multiple genes in a network that may contribute to the development of complex phenotypes. The higher number of significant locus pairs amongst cultured populations probably reflect new selection pressures due to domestication. It is noteworthy that many of the loci that showed significant similarity to known genes were genes involved in regulatory processes, such as signal transduction and gene expression (Table 3.4). This lends support to the idea of gene-networks being under selection; in fact Østman *et al.* (2012) argued that pleiotropic effects and epistasis play a vital role in adaptation. Furthermore, it has been suggested that only a few regulatory loci with

pleiotropic effects might be involved in the development of, for example, the domestic syndrome in traditional livestock species (Andersson and Georges, 2004; Mignon-Grasteau *et al.*, 2005; Dobney and Larson, 2006). As such, for example, locus *HmNR120* (on linkage group 10) may have a regulatory role in a signal transduction pathway with genes on linkage group 7, 9 and 18 (Table 3.4, Appendix B: S3.4).

A few loci also demonstrated similarity to transposable elements (Table 3.5). It is well-known that mobile elements, such as DNA transposons and LTR retrotransposons, are regularly found in genic regions, particularly in promoter areas and other regulatory motifs. In these sites, they have been found to actively alter gene expression by means of up-regulation and silencing (Bennetzen, 2000; Medstrand *et al.*, 2005; Feschotte and Pritham, 2007).

Nonetheless, as noted by Nikinmaa and Rytkönen (2012), the lack of genomic sequence information for many marine species, may hamper functional characterisation. It is therefore necessary for future supplementary studies to focus on correlating genotypes of outlier loci with particular phenotypes in association and/or QTL studies. This may be particularly useful in aquaculture populations to find loci that might be associated to particular economically important traits (*e.g.* Goddard and Hayes, 2009; Massault *et al.*, 2009; Lu *et al.*, 2012). Gene-expression analyses and protein functional assay investigations may also elucidate the functional roles of candidate loci under selection (*e.g.* Van der Merwe *et al.*, 2011). The present study could not find evidence for the association of particular environmental variables to the divergence of populations; however it should be noted that obtaining appropriate environmental data is not as simple. For example, regularly measured sea surface temperatures might provide a good estimate of sea surface temperature, but sea surface temperature does not necessarily reflect water temperature in lower strata where animals will be subject to selection. Furthermore, the data obtained reflects relatively stable climatic characteristics over large geographic regions that may not take into account local environmental fluctuations, such as an exceptionally warm year that may trigger a selective event. Nonetheless, clear evidence persists that particular loci are associated to particular populations; suggesting local environmental factors are probably driving diversity at these loci (Appendix B: Table S3.5).

3.5. Conclusions

This study identified several loci and genes in *H. midae*, where the patterns of genetic diversity could be explained by the effects of selection. It is postulated that these loci underlie the development of ecotypes in the wild and may contribute to the abalone domestic phenotype. There is evidence that these loci do not act singularly but form part of gene-networks, with significant LD amongst candidate loci beyond physical chromosomal constraints. The current data suggest that a substantial portion of the *H. midae* genome may be influenced by both directional and balancing selection. However, it is noted that these estimates may vary depending on the type of marker used (*e.g.* anonymous AFLP markers vs. EST-derived SNPs or -microsatellites) as seen when compared to other studies. And although the current analyses are based on the most extensive set of markers to date for abalone in terms of population genomics investigations, 150 markers do not necessarily provide complete genome coverage. Furthermore, this study supports previous finding of population differentiation between wild populations on the west-, south- and east coasts of South Africa and that this differentiation is, in part, due to adaptation. This result must be taken into consideration with regards to the management of wild populations. It has also been demonstrated that one generation of culture is sufficient to alter the genetic constitution of abalone significantly and that populations derived from independent domestication events will converge due to adaptation to an artificial environment or artificial selection.

References

Akey LM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madeoy J (2010) Tracking footprints of artificial selection in the dog genome. Proc Natl Acad Sci USA 107: 1160-1165.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403-410.

Andersson L, Georges M (2004) Domestic animal genomics: deciphering the genetics of complex traits. Nat Rev Genet 5: 202-212.

Andre´ C, Larsson LC, Laikre L, Bekkevold D, Brigham J, Carvalho GR, Dahlgren TG, Hutchinson WF, Mariani S, Mudde K, Ruzzante DE, Ryman N (2011) Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral *vs* putatively selected loci. Heredity 106: 270-280.

Antao T, Lopes A, Lopes RL, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9: 323.

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

Bester-van der Merwe AE, Roodt-Wilding R, Volckaert FAM, D'Amato ME (2011) Historical isolation and hydrodynamically constrained gene flow in declining populations of the South-African abalone, *Haliotis midae*. Conserv Genet 12: 543-555.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B 263: 1619-1626.

Bennetzen JF (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42: 151-169.

Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: Genome-wide sampling of insect populations. Annul Rev Entomol 46: 441-469.

Biswas S, Akay JM (2006) Genomic insights into positive selection. Trends Genet 22: 437-446.

Bonin A (2008) Population genomics; A new generation of genome scans to bridge the gap with functional genomics. Mol Ecol 17: 3583-3584.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). Mol Biol Evol 23: 773-783.

Campagne P, Smouse PE, Varouchas G, Silvain J-F, Leru B (2012) Comparing the van Oosterhout and Chybicki-Burczyk methods of estimating null allele frequencies for inbred populations. Mol Ecol Res 12: 975-982.

Campbell D, Bernatchez L (2004) Genetic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Mol Biol Evol 21: 945-956.

Counterman BA, Araujo-Perez F, Hines HM, Baxter SW, Morrison CM, Lindstrom DP, Papa R, Ferguson L, Joron M, French-Constant RH, Smith CP, Nielsen DM, Chen R, Jiggins CD, Reed RD, Halder G, Mallet J, McMillan WO (2010) Genomic hotspots for adaptation: The population genetics of Mullerian mimicry in *Heliconius erato*. PLoS Genet 6: e1000796.

Cramer H (1946) Mathematical Models of Statistics. Princeton University Press, Princeton, New Jersey.

Dobney K, Larson G (2006) Genetics and animal domestication: new windows on an elusive process. J Zool 269: 261-271.

Evans B, Sweijd N, Bowie RCK, Cook P, Elliott NG (2004a) Population genetic structure of the perlemoen *Haliotis midae* in South Arica: evidence of range expansion and founder events. Mar Ecol Prog Ser 270: 163-172.

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

Ewens WJ (1972) The sampling theory of selectively neutral alleles. Theor Popul Biol 3: 87-112.

Excoffier L, Lischer HEL (2010) Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10: 564-567.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103: 285-298.

Farber CR, Medrano JF (2003) Putative in silico mapping of DNA sequences to livestock genome maps using SSLP flanking sequences. Anim Genet 34: 11-18.

Farber CR, Medrano JF (2004) Identification of putative homology between horse microsatellite flanking sequences and cross-species ESTs, mRNAs and genomic sequences. Anim Genet 35: 28-33.

Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. Annu Rev Genet 41: 331-368.

Foll M, Gaggiotti OE (2006) Identifying the environmental factors that determine the genetic structure of populations. Genetics 174: 875-891.

Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180: 977-993.

Franchini P, Slabbert R, Van Der Merwe M, Roux A, Roodt-Wilding R (2010) Karyotype and genome size estimation of *Haliotis midae*: using these estimators to assist future studies on the evolutionary history of *Haliotidae*. J. Shellfish Res 29: 1-6.

Galindo K, Grahame JW, Bultin RK (2010) An EST-based genome scan using 454 sequencing in the marine snail *Littorina saxatilis*. J Evol Biol 23: 2004-2016.

Glémin S, Bataillon T (2009) A comparative view of the evolution of grasses under domestication. New Phytol 183: 273-290.

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10: 381-391.

Hagenblad J, Olsson M, Parker HG, Ostranders EA (2009) Population genomics of the inbred Scandinavian wolf. Mol Ecol 18: 1341-1351.

Hemmer-Hansen J, Nielsen EE, Frydenberg J, Loeschcke V (2007) Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). Heredity 99: 592-600.

Hedrick PW (1987) Gametic disequilibrium measures: proceed with caution. Genetics 41: 331-341.

Herrera CM, Bazaga P (2008) Population-genomics reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. Mol Ecol 17: 5378-5390.

Innan H, Kim Y (2004) Pattern of polymorphism after strong artificial selection in a domestication event. Proc Natl Acad Sci USA 101: 10667-10672.

Jurka JV, Kapitonov V, Pavlick A, Klownowski P, Kohany O, Walichiewicz J (2005) Repbase update, a database of eukaryotic repetitive elements. Cytogenet Genome Res 110: 462- 467.

Karasov T, Messer PW, Petrov DA (2010) Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. PLoS Genet 6: e1000924.

Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics 7: 474.

Lancaster AK, Single RM, Solberg OD, Nelson MP, Thomson G (2007) PyPop update - a software pipeline for large-scale multilocus population genomics. Tissue Antigens 69 (s1): 192-197.

Li W-H, Wu C-I, Luo C-C (1985) A new method for estimating synonymous and non-synonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2: 150-174.

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

Li Y-C, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: Structure, function, and evolution. Mol Biol Evol 21: 991-1007.

Li MH, Merila J (2010) Extensive linkage disequilibrium in a wild bird population. Heredity 104: 600-610.

Li MH, Merila J (2011) Population differences in levels of linkage disequilibrium in the wild. Mol Ecol 20: 2916-2928.

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128-2129.

Lu X, Wang H, Liu B, Xiang J (2012) Three EST-SSR markers associated with QTL for the growth of the clam *Meretrix meretrix* revealed by selective genotyping. Mar Biotechnol DOI: 10.1007/s10126-012-9453-4.

Luikart G, Englang PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4: 981-994.

Mariac, C., Jehin, L., Saidou, A.-A., Thuillet, A.C., Couderc, M., Sire, P., Jugdé H, Adam H, Bezancon G., Pham J-L, Vigouroux Y (2011) Genetic basis of pearl millet adaptation along an environmental gradient investigated by a combination of genome scan and association mapping. Mol Ecol 20: 80-91.

Massault C, Hellemans B, Louro B, Batargias C, Van Houdt JKJ, Canario A, Volckaert FAM, Bovenhuis H, Haley C, De Koning DJ (2009) QTL for body weight, morphometric traits and stress response in European sea bass *Dicentrarchus labrax*. Anim Genet 41: 337-345.

McDonald J, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature 351: 652-654.

Meadows JRS, Chan EKF, Kijas JW (2008) Linkage disequilibrium compared between five populations of domestic sheep. BMC Genet. 9: 61.

Medugorac I, Medugorac A, Russ I, Veit-Kensch CE, Taberlet P, Luntz B, Mix HM, Föster M (2009) Genetic diversity of European cattle breeds with high effective population size. Mol Ecol 19: 3394-3410.

Medstrand P, Van de Lagemaat LN, Dunn CA, Landry J-R, Svenback D, Mager DL (2005) Impact of transposable elements on the evolution of mammalian gene regulation. Cytogenet Genome Res 110: 342-352.

Mitchell-Olds T, Willis JH, Goldstein DB (2007) Which evolutionary processes influence natural genetic variation for phenotypic traits? Nat Rev Genet 8: 845-856.

Mignon-Grasteau S, Boissy A, Bouix J, Faure J-M, Fisher A, Hinch GN, Jensen P, Le Neindre P, Mormède P, Prunet P, Vandeputte M, Beaumont C (2005) Genetics of adaptation and domestication in livestock. Livest Prod Sci 93: 3-14.

Murray MC, Hare MP (2006) A genomic scan for divergent selection in a secondary contact zone between Atlantic and Gulf of Mexico oysters *Crassostrea virginica*. Mol Ecol 15: 4229-4242.

Namroud M-C, Beaulieu J, Juge N, LaRoche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in the white spruce. Mol Ecol 17: 3599-3613.

Narum SR, Hess JE (2011) Comparison of the F_{ST} outlier tests for SNP loci under selection. Mol Ecol Res 11 (Suppl 1): 184-194.

Nadeau NJ, Jiggins CD (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. Trends Genet 26: 484-492.

Nei M (1972) Genetic distance between populations. American Naturalist 106: 283-292.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and non-synonymous nucleotide substitutions. Mol Biol Evol 3: 418-426.

Nielsen EE, Hemmer-Hansen J, Larsen PF, Bekkevold D (2009a) Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18: 3128-3150.

Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer C, Taranger G-L, Ogden R, Carvalho GR (2009b) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC Evol Biol 9: 276.

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39: 197-218.

Nikinmaa M, Rytkönen KT (2010) From genomics to function in aquatic biology. Marine Genomics 5: 1-6.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18: 375-402.

Nunes VL, Beaumont MA, Butlin RK, Paulo OS (2011) Multiple approaches to detect outliers in a genome scan for selection in the ocellated lizards (*Lacerta lepida*) along an environmental gradient. Mol Ecol 20: 193-205.

Østman B, Hintze A, Adami C (2012) Impact of epistasis and pleiotropy on evolutionary adaptation. Proc R Soc B 279: 247-256.

Qanbari S, Pimentel ECG, Tetens J, Thaller P, Lichtner P, Sharifi AR, Simianer H (2010) A genome-wide scan for signatures of recent selection in Holstein cattle. Anim Genet 41: 377-389.

Paris M, Boyer S, Bonin A, Collado A, David J-P, Despres L (2010) Genome scan in the mosquito *Aedes rusticus*: population structure and detection of positive selection after insecticide treatment. Mol Ecol 19: 325-337.

Pavlidis P, Hutter S, Stephan W (2008) A population genetic approach to map recent positive selection in model species. Mol Ecol 17: 3585-3598.

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6: 288-295.

Pérez-Figueroa A, García-Pereira MJ, Saura M, Rolán-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. J Evol Biol 23: 2267-2276.

Praipue P, Klinbunga S, Jarayabhand P (2010) Genetic diversity of wild and domesticated stocks of Thai abalone, *Haliotis asinina* (Haliotidae), analyzed by single-strand conformational polymorphism of AFLP-derived markers. Genet Mol Res 9: 1136-1152.

Rhode C, Roodt-Wilding R (2011) Bioinformatic survey of *Haliotis midae* microsatellites reveals a non-random distribution of repeat motifs. Biol Bull 221: 147-154.

Rhode C, Hepple J, Jansen S, Davis T, Vervalle J, Bester-van der Merwe AE, Roodt-Wilding R (2012) A population genetic analysis of abalone domestication events in South Africa: Implications for the management of the abalone resource. Aquaculture 356-357: 235-242.

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6: 847-859.

Roodt-Wilding R (2007) Abalone ranching: a review on genetic considerations. Aquac Res 38: 1229-1241.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Res 8: 103-106.

Saghai-Maroof MA, Solima KM, Jorgenson RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81: 8014-8018.

Schmidt PS, Serrão EA, Pearson GA, Riginos C, Rawson PD, Hilbish TJ, Brawley SH, Trussell GC, Carrington E, Wethey DS, Grahame JW, Bonhomme F, Rand DM (2008) Ecological genetics in the North Atlantic: Environmental gradients and adaptation at specific loci. Ecology 89: S91-S107.

Scotti-Saintagne C, Marriete S, Porth I, Goicoechea PG, Barreneche T, Bodénès C, Burg K, Kremer A (2004) Genome scanning for the interspecific differentiation between two closely related oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. Genetics 168: 1615-1626.

Smit AFA, Hubley R, Green P (2011) Unpublished data current version: open-3.3.0 (RMLib: 20110920). Available: www.repeatmasker.org/ [Accessed 28 February 2012]

Slabbert R, Ruivo NR, Van den Berg NC, Lizamore DL, Roodt-Wilding R (2008) Isolation and characterisation of 63 microsatellite loci for the abalone, *Haliotis midae*. J World Aquacult Soc 39: 429-435.

Slabbert R, Bester AE, D'Amato ME (2009) Analyses of genetic diversity and parentage within a South African hatchery of the abalone *Haliotis midae* Linnaeus using microsatellite markers. J Shellfish Res 28: 369-375.

Slabbert R, Hepple J, Venter A, Nel S, Swart L, Van Den Berg NC, Roodt-Wilding R (2010) Isolation and segregation of 44 microsatellite loci in the South African abalone *Haliotis midae* L. Anim Genet 41: 332-333.

Slabbert R, Hepple J, Rhode C, Bester-Van der Merwe AE, Roodt-Wilding R (2012) New microsatellite markers for the abalone *Haliotis midae* developed by 454 pyrosequencing and *in silico* analyses. Genet Mol Res 11: 2769-2779.

Slate J, Pemberton JM (2007) Admixture and patterns of linkage disequilibrium in a free-living vertebrate population. J Evol Biol 20: 1415-1427.

Slatkin M (1982) Testing neutrality in subdivided populations. Genetics 100: 533-545.

Slatkin M (1994) An exact test for neutrality based on the Ewens sampling distribution. Genet Res 64: 71-74.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: Finding genes underlying ecological important traits. Heredity 100: 158-170.

Storz JF (2005) Using genome scans of DNA polymorphisms to infer adaptive population divergence. Mol Ecol 14: 671-688.

Tajima F (1983) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

Taniguchi N (2003) Genetic factors in broodstock management for seed production. Rev Fish Biol Fisher 13: 177-185.

Van der Merwe M, Franchini P, Roodt-Wilding R (2011) Differential growth-related gene expression in abalone (*Haliotis midae*). Mar Biotechnol 13: 1125-1139.

Vasemägi A, Nilsson J, Primmer CR (2005) Express Sequence Tag-linked microsatellite as a source for gene-associated polymorphisms for detecting signatures of selection in Atlantic Salmon (*Salmo salar* L.). Mol Biol Evol 22: 1067-1076.

Wahlund S (1928) Zusammensetzung von populationen und korrelation-serscheinungen von standpunkt der verebungslehre aus betrachtet. Hereditas 11: 65-106.

Walsh B (2008) Using molecular markers for detecting domestication, improvement, and adaptation genes. Euphytica 161: 1-17.

Watterson A (1977) Heterosis or neutrality? Genetics 85: 789-814.

Watterson A (1978) The homozygosity test of neutrality. Genetics 88: 405-417.

Wilding CS, Bultin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorena saxatilis* detected using AFLP markers. J Evol Biol 14: 611-619.

Williams LM, Olecksiak MF (2008) Signatures of selection in natural populations adapted to chronic pollution. BMC Evol Biol 8: 282.

Willing E-M, Bentzen P, Van Oosterhout C, Hoffmann M, Cable J, Breden F, Weigel D, Dreyer C (2010) Genome-wide single nucleotide polymorphisms reveal populations history and adaptive divergence in the wild guppies. Mol Ecol 19: 968-984.

Whiteley AR, Bhat A, Martins EP, Mayden RL, Arunachalam M, Uusi-Heikkilä S, Ahmad ATA, Shrestha J, Clark M, Stemple D, Bernatchez L (2011) Population genomics of wild and laboratory zebrafish (*Danio rerio*). Mol Ecol 20: 4259-4276.

Wright D, Rubin C-J, Martinez Barrio A, Schutz K, Kerje S, Brändström H, Kindmark A, Jensen P, Andersson L (2010) The genetic architecture of domestication in the chicken: effects of pleiotropy and linkage. Mol Ecol 19: 5140-5156.

Wright S (1931) Evolution in Mendelian populations. Genetics 16: 97-159.

Zhai W, Nielsen R, Slatkin M (2009). An investigation of the statistical power of neutrality tests based on comparative and population genetic data. Mol Biol Evol 26: 273-283.

Zhao H, Nettleton D, Soller M, Deckers JCM (2005) Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between markers and QTL. Genet Res 86: 77-87.

CHAPTER 4

Spatio-Temporal Assessment of Genetic Variation in the South African Abalone (*Haliotis midae*) using SNP Loci.

Abstract

Organisms live in dynamic environments that are subject to change and generally respond to such environmental stressors by means of genetic adaptation in the long-term. The South African abalone is a gastropod mollusc of economic importance. In recent years natural populations have come under considerable pressure due to overharvesting and ecological shifts. The spatial genetic structure of abalone has been determined; however the degree to which adaptive evolutionary forces in contrast to neutral mechanisms maintain this population subdivision has not yet been established. Furthermore, to date there has not been a temporal assessment of abalone population dynamics. Using a population genomics approach this study aimed to assess fluctuations in genetic diversity among wild and cultured abalone populations through time and space. Various estimates of genetic diversity and population differentiation were calculated using EST-derived SNP markers. All populations seemed to possess comparable levels of genetic diversity and long-term effective population size appears to be sufficiently large for the wild populations, despite evidence for recent bottlenecks. Population differentiation was for the most part geographically correlated, with spatial genetic structure maintained across temporal samples. Significant genetic differentiation was however detected among temporal samples taken from the same locality. This temporal heterogeneity could be caused by changes in selection pressures over time. There was also evidence for comparatively small short-term effective population sizes that could explain large changes in allele frequencies due to stochastic effects.

Keywords: Effective Population Size; *Haliotis midae*; Spatio-temporal Genetic Variation; Selection; Population Differentiation; SNPs

This chapter is in preparation for submission to a peer-review journal.

4.1. Introduction

Organisms function in dynamic environments where a number of variables often fluctuate through time and space, for example, the emergence of new pathogens (Parker and Gillbert, 2004; Antunes *et al.*, 2008) or the introduction of new species that may alter the dynamics of predator-prey interactions or inter-species competition for available resources (Mack *et al.*, 2000) and changes in climatic conditions (Bradshaw and Holzapfel, 2008; Karell *et al.*, 2011). Recently, however, environmental changes due to anthropogenic effects have received considerable attention (Smith and Bernatchez, 2008): Organisms are faced with increasing levels of pollution (Anderson *et al.*, 1994; Williams and Olecksiak, 2008), overharvesting (fishing and hunting) of wild populations (Allendorf *et al.*, 2008; Coltman, 2008) and arguably the most topical issue; global warming and the emission of greenhouse gasses (Kerr, 2007; Visser, 2008; Moss, 2010).

Changes in the environment subject organisms to novel selection regimes, and as such, they can respond to these pressures in three ways (Davis *et al.*, 2005; Hansen *et al.*, 2012): Firstly, by means of range-shifting, whereby a population's distribution range changes through the gradual migration of individuals to more suitable habitats (Walther *et al.*, 2002; Karban and Strauss, 2004; Perry *et al.*, 2005; Parmesan, 2006). Secondly, phenotypic plasticity allows for physiological acclimation, by means of genotype-environmental interactions (Przybylo *et al.*, 2000; De Jong, 2005; Cheviron *et al.*, 2008) and lastly by means of genetic (micro-evolutionary) adaptation (Stockwell *et al.*, 2003; Davis *et al.*, 2005). How important these coping mechanisms are, relative to one another, is a function of many factors, including: life-history and dispersal ability of the particular species, timescale and extent of the environmental permutation and the availability of suitable habitats (Gienapp *et al.*, 2008). In reality it is more likely that more than one mechanism is at work simultaneously (Davis and Shaw, 2001). Nonetheless, in the short-term organisms will generally rely on phenotypic plasticity for a rapid response to any environmental alteration that might subsequently allow for a sufficient period of time to migrate to more appropriate habitats. In the absence of suitable habitats, limited dispersal ability of particular organisms or under continuing directional environmental change, phenotypic plasticity is however

unlikely to provide a long-term solution: There are limits to the extent to which plastic traits can buffer the effects of extreme conditions and beyond that threshold phenotypic plasticity will not be able to mitigate a loss in fitness (DeWitt *et al.*, 1998; De Jong, 2005; Visser, 2008). Therefore, ultimately, the development of a "stable", adapted phenotype will depend on micro-evolutionary processes, *i.e.* the interplay between the generation of genetic variation, random drift and selection (Stockwell *et al.*, 2003; Davis *et al.*, 2005).

Elucidating such micro-evolutionary processes is important for several reasons: Firstly, it allows for assessing the ability of a population to "launch" an adaptive response and consequently predicting if said population could persist in an altered environment. Within a conservation context it may provide an indication of the necessity to take action, such as, whether genetic rescue should be performed (Richards, 2000; Ingvarsson, 2001). If an adaptive response is detected it could serve as an environmental "flag" for either locally adapted populations or possible stressors that are affecting a population if the adaptation signal is recent (Hansen *et al.*, 2012). In order to identify adaptive responses it is necessary to monitor genetic diversity on a temporal scale, however this is not always easy or practical, especially if organisms have long generation intervals. Consequently, many studies focus on investigating spatial variation across environmental gradients (*e.g.* Bonin *et al.*, 2006; Byars *et al.*, 2007; Cheviron *et al.*, 2008; Jensen *et al.*, 2008; Riba *et al.*, 2009). Although this approach has proved useful, populations often have distinct evolutionary histories and consequently may have different demographic characteristics, such as degrees of gene flow and effective population sizes. Therefore, populations may differ in the standing genetic variation that may be of adaptive value and thus have unique genetic architectures for loci underlying adaptation (Hansen *et al.*, 2012).

The South African abalone, *Haliotis midae*, as an economically important species has been under great pressure in the wild due to overharvesting, particularly poaching (Raemaekers *et al.*, 2011). Furthermore, due to poorly understood and unknown factors there has been a major range-shift of the South African west coast rock lobster, *Jasus lalandii*, to the main regions generally associated with the commercial abalone fishery. The rock lobster is a major predator of the abalone commensal symbiont urchin (*Parechinus angulosus*) that provides shelter to juvenile animals (Mayfield and George, 2000; Cockcroft *et al.*, 2008). Aquaculture and the recent implementation of selective breeding for

production traits are also exerting new selective pressures on abalone. This study therefore aims to assess both spatial and temporal patterns of genetic diversity, using a newly developed SNP assay, and to identify the micro-evolutionary (selection and genetic drift) forces that may influence such variation by mean of a population genomics approach. This will be the first study to attempt spatio-temporal genetic assessment for abalone.

4.2. Materials and Methods

4.2.1. Study Populations and Specimens

Permits to collect and transport abalone for research purposes were obtained from the Department of Agriculture, Forestry and Fisheries (Republic of South Africa). Specimens were collected across the geographical range of *H. midae*, including wild animals from the west- [Saldanha Bay (SD), Gansbaai (GB)], south- [Witsand (WS)] and east coast [Cape Recife (CR), Riet Point (RP)] of the country (Table 4.1). Temporal samples for the wild populations were taken at an interval of six to eleven years representing approximately one to two generations. Cultured specimens were obtained from two aquaculture facilities; one on the west coast (Atlantic Sea Farm, AS) and the other on the east coast (Wild Coast Abalone, WC). Where applicable, temporal samples for cultured animals were based on the generation under culture, *i.e.* F_1- or F_2- generation (Table 4.1). Cultured populations were founded by wild animals originating from the respective geographic regions. F_1's were generated by means of random mating under semi-natural conditions, whilst F_2's were a result of random mating amongst individuals phenotypically selected for superior growth rate. Muscle and/or gill tissues were collected from each individual and preserved in 70% ethanol and stored at -20°C until DNA extraction could be performed *via* the standard CTAB method (Saghai-Maroof *et al.*, 1984).

Table 4.1: Study populations and sample sizes with geographic coordinates and indication of temporal separation (Wild population as per year of collection and cultured populations as per generation under culture)[#].

Geographic region	Geographic coordinates	Study population name (generation/year collected)	Number of specimens
West coast	32°45'30.00''S; 18°01'40.00''E	AS(F_1)[#]	29
	33°02'40.64''S; 17°56'00.53''E	SD(2004)	48
		SD(2010)	31
	34°35'04.00''S; 19°19'45.63''E	GB(2003)	24
South coast	34°20'53.37''S; 19°01'39.75''E	WS(2004)	33
		WS(2011)	9
East coast	34°02'15.98''S; 25°42'17.64''E	CR(2000)	19
		CR(2011)	30
	33°31'29.31''S; 27°06'51.18''E	RP(2003)	51
		RP(2011)	32
	32°45'43.20''S; 28°15'0.00''E	WC(F_1)[#]	21
		WC(F_2)	49

[#]*Originally 48 specimens for each of the F_1 cultured populations [AS(F_1) and WC(F_1)] as well as 48 specimens for a south coast cultured population was included for analysis; however, due to an unexpected technical failure at the genotyping facility this plate was lost. These specimens where thus excluded from further analysis and discussion as no conclusions could be drawn.*

4.2.2. SNP Assay (Illumina® BeadXpress®) Development and Genotyping

Animals collected during 2010/2011 were transported on ice in an oxygenated container to minimise transportation stress to the animals. From these population cohorts six wild specimens (three male and three female) were selected at random from each of the three geographic regions (SD, WS, RP) and six cultured specimens (three each of the aquaculture facilities; AS and WC) – 24 individuals in total. Tissue was collected from the following five organs from each individual: muscle (from the epipoduim), ganglion, hepato-pancreas, gonad and gill. Biopsied tissues were immediately placed in RNA*Later*® (Ambion®) solution and stored at -20°C until RNA extraction could be performed. RNA was extracted from 1-2 g

of tissue from each organ and individual, separately. The extraction protocol as described in Van der Merwe (2010) was used. Denaturing agarose gel electrophoresis [2% agarose; 1X 3-(N-morpholino) propanesulfonic acid (MOPS) buffer] was performed to determine RNA integrity and absorbance measurements were taken using the NanoDrop® ND-1000 spectrophotometer to assess RNA purity and concentration. The SuperScript® Double-Stranded cDNA Synthesis Kit (Invitrogen) was used according to the manufacturer's specifications to generate double stranded cDNA. Oligo (dT) primers were used to selectively amplify poly-adenylated mRNA. The cDNA from each population group (three wild and one cultured group) was pooled individually in equal-Molar amounts to provide a final pooled sample for each population group containing 2 µg of cDNA at a concentration of 100 ng/µl. Pooled cDNA was sent to Inqaba Biotechnical Industries (Pty) Ltd (Pretoria, South Africa) for pyro-(454)-sequencing, according to the manufacturer's protocol. Pyro-sequencing was conducted on a full-plate of the 454 GS FLX platform (Roche); effectively a quarter of a plate for each of the population cohorts.

The CLC genomics workbench software v.4.9 (CLC bio) was employed to conduct *de novo* sequence assembly using the default parameters for "long reads". Putative SNP calling was also done using the CLC genomics workbench SNP detection module. The following parameters were set for SNP identification: Quality score: 20; minimum coverage: 8X; minor allele frequency: 0.25. Annotation of the contigs was done using the Blast2Go® software. Putative SNPs were then selected for the Illumina® BeadXpress® assay (Illumina® GoldenGate™ Genotyping Assay with VeraCode Technology; Fan *et al.*, 2006) based on homogenous sequences of at least 60 bp flanking the putative SNP and a designability score greater than 0.75 (Illumina® assay design tool). Single nucleotide polymorphisms located in know genes were preferentially selected. A final set of 142 putative SNPs from the 454-sequence data was selected and 50 SNPs previously confirmed and mapped to the *H. midae* linkage map was also included (Du Plessis, 2012; Vervalle *et al.*, in press) in a 192 plex BeadXpress® SNP assay (Appendix C, Table S4.1).

Genotyping reactions were performed by the National Health Laboratory Service of South Africa (Johannesburg, South Africa) on the BeadXpress® platform in 96-well plates. DNA for each sample was standardised at 0.5 µg (50 ng/µl). As internal positive controls, two individuals with known genotypes (for previously developed SNPs) were included on each

plate. The GenomeStudio™ genotyping module v.1.0 (Illumina®) was used to assign genotypes to individuals. Loci with a GenCall score (a quality metric that provides an indication of the reliability of assigned genotype) less than 0.25, a GenTrain score (indication of cluster separation among the genotypes) less than 0.35 and a call rate (number of individuals assigned a genotype for any particular locus) less that 0.7 were excluded from further analysis. Furthermore, loci that had a minor allele frequency lower that 0.01 across all populations were deemed monomorphic and were also excluded from subsequent analyses.

4.2.3. Population Genetic Data Analysis

Hardy-Weinberg equilibrium was computed *via* the exact probability test (10 000 dememorisations, 500 batches, and 5000 iterations per batch) using Genepop v.4.0 (Rousset, 2008). The following diversity statistics were also calculated in Genepop: minor allele frequency, observed- and expected heterozygosity and locus specific F_{is}. A Kruskal-Wallis test was performed to evaluate significant differences in heterozygosity and minor allele frequency among populations. Two F_{st}-outlier detection methods, as implemented in Lositan v.1.44 (Antao *et al.*, 2008) and BayeScan v.2.1 (Foll and Gaggiotti, 2008) were used to identify potential loci under selection. Lositan parameters were as follows: 50 000 simulations, with a 95% confidence interval and a false discovery rate of 0.1, assuming the infinite alleles model. For BayeScan default parameters as set for co-dominant markers were used; statistical confidence levels were set according to the Jeffreys' scale: a posterior probability (PO) greater than 10 [$\log_{10}(PO) > 1$] was interpreted as sufficient evidence for selection. Outlier analyses was first done across all populations through time and space, and then rerun for population cohorts as follows: wild populations as sampled between 2000 and 2004; wild populations as sampled between 2010 and 2011; F_1 cultured populations; and finally as pairwise comparisons among temporal samples for each population. Cultured populations were also compared with both "historic" and contemporary wild populations of origin. To reduce false positives, loci should have demonstrated congruent evidence for either directional or balancing selection in both methods across any of the population cohorts investigated before it was deemed a candidate locus under selection. A Kruskal-

Wallis test was performed to evaluate the differences in diversity statistics among selection- and neutral loci.

Estimates for population differentiation were calculated for two data sets: all loci and excluding loci under selection. Pairwise F_{st} (10 000 permutations at 5% significance level) between populations was estimated and a locus-by-locus hierarchical molecular analysis of variance (AMOVA; 10 000 permutations) was also computed in Arlequin v.3.5.1.2 (Excoffier and Lischer, 2010). For the multivariate analysis populations were grouped as follows: cultured population from the west coast [AS(F_1)]; cultured populations from the east coast [WC(F_1), WC(F_2)]; wild populations from the west coast [SD(2004), SD(2010), GB(2003)]; wild populations from the south coast [WS(2004), WS(2011)]; wild populations from the east coast, Riet point [RP(2003), RP(2011)] and Cape Recife [CR(2000), CR(2011)] – five groups in total. Dendrograms were constructed based on Nei's genetic distance (Nei, 1972) using the neighbour joining clustering algorithm in PowerMarker v.3.25, with 1000 bootstrap replicates (Liu and Muse, 2005). Finally, a factorial correspondence analysis plot was drawn in Genetix v. 4.05.2 (Belkhir *et al.*, 2004).

Effective population sizes were calculated using the heterozygous excess test and the moment-based temporal test (where applicable) in NeEstimator v.1.3 (Peel *et al.*, 2004); as well as a LD test (minimum allele frequency, 0.01) in LDNe v.1.0 (Waples, 2006). To further investigate the occurrence of recent bottlenecks, the Wilcoxon signed rank test (Cornuet and Luikart, 1997) and the mode shift indicator test (Luikart *et al.*, 1998), assuming the infinite alleles model, in Bottleneck v.1.2.02 (Piry *et al.*, 1999) was used. Population size estimates and bottleneck detection was performed on a dataset containing neutral loci only.

4.3. Results

4.3.1. SNP Assay Development and Genotyping

The pyro-sequencing generated 606 102 sequence reads with a mean read length of 441 bp. Of these 344 650 reads assembled into 50 378 contig sequences, ranging in size from 232 bp to 3 396 bp. Average coverage was 4.1X. Only 7 862 of the 50 378 contigs (15.61%)

had significant BLAST hits to known genes of which only approximately half could be fully annotated in terms of functionality (Appendix C: Figure S4.1, S4.2, S4.3). Seven-thousand-and-eighty-four putative SNPs could be identified of which 573 adhered to the requirements for inclusion on the GoldenGate SNP assay. Only a subset of these were included in the 192 plex SNP assay, which consisted of 142 putative SNPs (identified from the pyro-sequencing experiment) and 50 confirmed SNP markers previously identified. The internal positive controls were all assigned genotypes that correlated to the expected genotypes based on a previous study (Du Plessis, 2012). Thirty-three loci were excluded from analysis due to low quality genotyping scores, including GenCall-, GenTrain scores and call rates; providing an assay success rate of 82.8% (159/192). A further 43 monomorphic loci [of which the majority (39) were from the putative SNPs, identified from the pyro-sequencing run] were excluded, leaving 116 polymorphic loci available for analysis and a conversion rate of 60.42% (116/192; number of polymorphic SNPs divided by the total number of SNPs assayed; Fan *et al.*, 2003). All individuals were assigned genotypes at the majority of loci (call rate > 0.7) and therefore none were excluded from analyses (NOTE: #Table 4.1).

4.3.2. Genetic Diversity

The majority of loci conformed to Hardy-Weinberg expectation within populations, but demonstrated mild heterozygous excess, with generally negative F_{is} values. Across all populations mean F_{is} values were low ranging from -0.04 to 0.037; mean observed heterozygosity (H_o) per population ranged from 0.285 to 0.327 and mean expected heterozygosity (H_e) ranged from 0.281 to 0.346. Mean minor allele frequencies (MAF) were moderate ranging from 0.197 to 0.239. Across all populations no difference in heterozygosity or MAF could be detected (Kruskal-Wallis test, $P > 0.05$). Figure 4.1 provides a general overview of diversity statistics (for a detailed treatment refer to Appendix C: Table S4.2). Amongst the confirmed polymorphic loci the transition to transversion ratio was 1.7:1.

4.3.3. Outlier Loci

Across all populations (spatial and temporal) Lositan detected nine loci under directional selection and 48 loci under balancing selection. BayeScan identified 12 loci under directional selection and only one locus to be under balancing selection. Among the wild populations sampled between 2000 and 2004, Lositan and BayeScan detected eight and nine loci under directional selection respectively; 33 loci were found to be under balancing selection by Lositan, whilst BayeScan failed to detect any loci under balancing selection. For comparison among contemporary wild populations (samples 2010/11), Lositan identified ten loci that may be influenced by directional selection and 33 loci under balancing selection; whilst BayeScan only detected four loci under directional selection. Among the F_1 generation cultured abalone populations Lositan identified 26 loci putatively under selection (five under directional selection and 21 under balancing selection); BayeScan could only detect four loci influenced by directional selection. Among the pairwise population comparisons BayeScan failed to detect any outlier loci, whilst Lositan detected a number of loci ranging from eight to 29 depending on the population pair (Appendix C: Table S4.3).

Based on the set criteria, however, only 13 loci had congruent results as determined by both test methods; 12 loci were candidates for directional selection and one candidate for balancing selection. Among this final set of loci under selection, two SNP markers were located in the 3' UTR, three in the 5' UTR and eight in the exonic regions of the respective genes. Of the SNPs located in the coding regions, five were synonymous- and three were non-synonymous substitutions. Two of the contig sequences had no BLAST hits, whilst one contig had no known homologue in any other species, it did however, contain a conserved functional protein domain (Table 4.2). The transition to transversion ratio among loci under selection was less than that estimated for the total dataset (1.16). Statistically significant differences in minor allele frequencies and expected heterozygosity could not be found (Kruskal-Wallis test, $P > 0.05$), but loci under selection demonstrated significantly reduced observed heterozygosity (Kruskal-Wallis test, $P < 0.05$) (Figure 4.2).

4.3.4. Population Differentiation and Effective Population Size

Mean pairwise F_{st} using all loci (including candidate loci under selection) was almost three times higher than the mean pairwise F_{st} excluding loci under selection at 0.062 (range: 0.003-0.132) and 0.022 (range: -0.001-0.080), respectively; indicating low to moderate population differentiation in general. Differentiation amongst the temporal samples were low to moderate, with the Riet Point samples (RP) reaching the highest F_{st} estimates for its temporal comparisons at 0.110 (using all markers). Subtle population differentiation between populations of the major geographic regions (west-, south- and east coast) is also supported, with significant divergence of the cultured populations from wild progenitor populations (Table 4.3). This population differentiation is further supported by the AMOVA results that show significant differentiation amongst groups of populations, amongst populations within groups and within populations. The percentage of variation explained among groups of populations increase from 1.420 to 5.674% when using all loci, including the loci under selection (Table 4.4).

Both the factorial correspondence analysis and dendrograms show clear patterns of population clustering according to geographic distributions. There is also evidence for a temporal arrangement with temporal samples occupying distinct branches on the dendrograms. The WC(F_1) population however demonstrated atypical positioning on both factorial correspondence analysis and dendrograms, failing to cluster with its progenitor population [WC(F_1)] and geographic origin in general. Bootstrap values for the dendrograms were generally lower when using the dataset excluding the loci under selection. Among the west coast populations the separation of the culture population [AS(F_1)] was particularly supported by the complete dataset (bootstrap value = 100%), but to a lesser extent by the dataset excluding the loci under selection. The east coast cultured populations [WC(F_1)] showed the inverse relationship, with the exception for the second generation cultured population [WC(F_2)] for which strong differentiation was supported by both datasets (Figure 4.3, 4.4).

Point estimates for effective population size varied considerably amongst the methods used. The heterozygous excess methods estimated an infinitely large effective population size for all populations with the exception of the two cultured populations. Temporal

estimates were generally lower than LD estimates for effective population sizes, but a degree of overlap considering the 95% confidence intervals at some populations was observed. In general cultured populations had lower effective population sizes than wild populations and effective population sizes seemed to be fairly stable across temporal estimates among wild populations, based on heterozygosity excess and LD tests. However, for the SD- and RP populations the 2003/4 ("historic") samples demonstrated a lower effective population size than the 2010/11 (contemporary) samples. There was evidence for recent bottlenecks at all populations based on the Wilcoxon signed rank test for heterozygous excess, but this was only supported by the mode-shift indicator test at seven populations (Table 4.5).

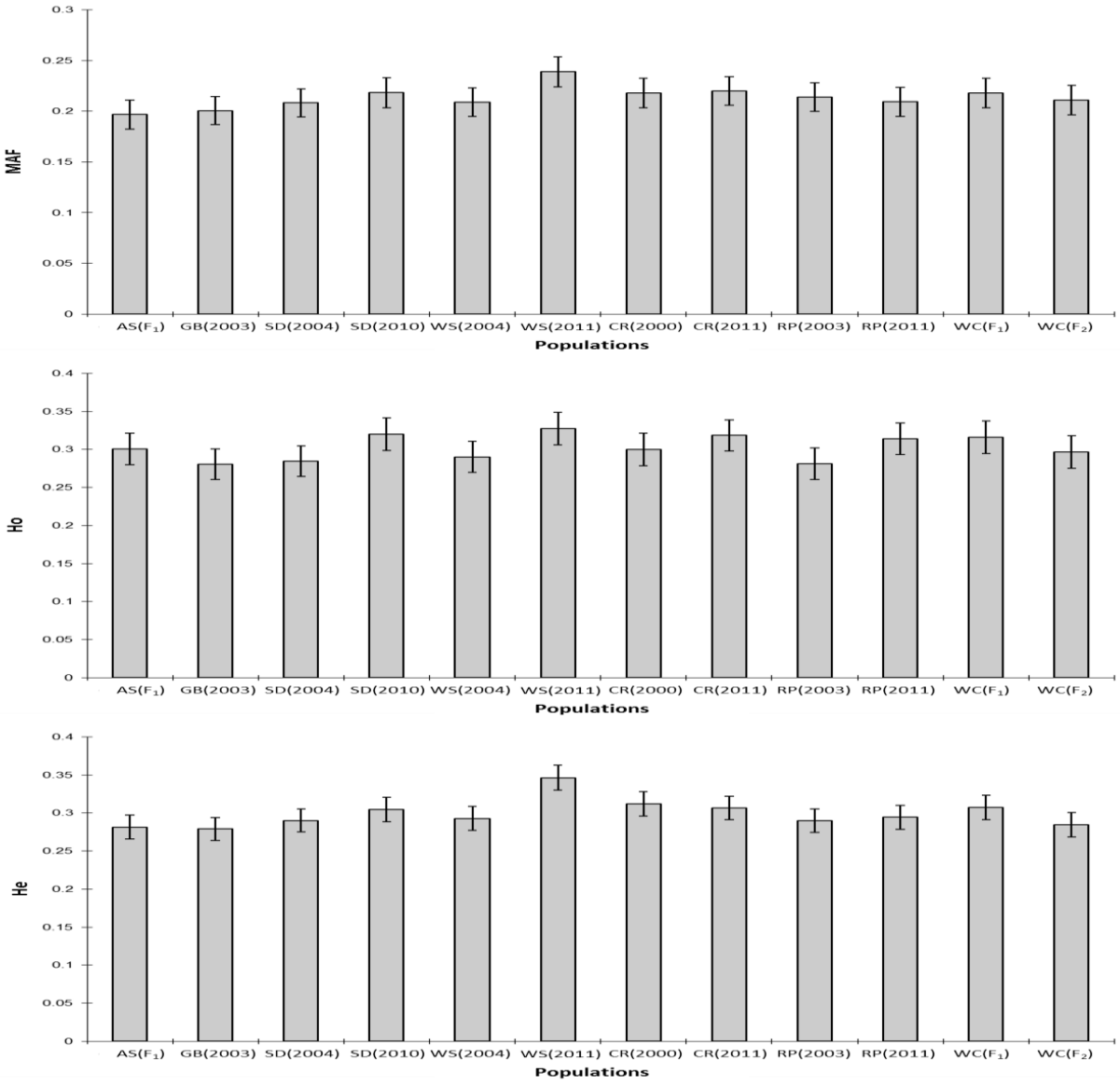

Figure 4.1: Summary of diversity statistics, minor allele frequency (MAF), expected heterozygosity (H_e) and observed heterozygosity (H_o) per population.

Table 4.2: Candidate loci under the influence of selection, most likely gene of origin and position of SNP marker in respective genes.

Locus	Selection (D/B)[a]	Gene	E-value	Similarity (%)	Gene location of variant	Type of Substitution (S/NS)[b]
PS_C23591_200_[T/C]	D	TFG protein (RNA promoter binding)	2.00E-09	76	3' UTR	N/A
PS_C34501_77_[T/G]	D	NADH dehydrogenase subunit (oxidative phosphorylation, electron transport)	0.00	95	5' UTR	N/A
PS_C34501_638_[A/G]	D				exonic	S
PS_C23075_525_[G/C]	D	Unkown protein with EF-hand, calcium binding motif	9.52E-03	65	exonic	S
PS_C1652_228_[A/C]	D	NADH dehydrogenase subunit (oxidative phosphorylation, electron transport)	8.00E-21	91	5' UTR	N/A
PS_C25083_285_[A/G]	D	Transportase (endonuclease activity)	7.00E-14	65	exonic	S
PS_C35977_153_[T/C]	D	NADH dehydrogenase subunit (oxidative phosphorylation, electron transport)	8.00E-48	98	exonic	NS (Pro>Ser)
PS_C36237_70_[C/G]	D	Unkown protein (no BLAST hit)	N/A	N/A	exonic	NS (Ala>Gly)
PS_C28810_290_[A/C]	D	Hypothetical protein (endonuclease activity)	5.00E-17	62	exonic	NS (Asn>His)
PS_C47375_253_[A/G]	D	NADH dehydrogenase subunit (oxidative phosphorylation, electron transport)	1.00E-30	95	5' UTR	N/A
PS_C47340_198_[A/G]	D	Unkown protein (no BLAST hit)	N/A	N/A	5' UTR	N/A
ILL_C6061_1289_[T/G]	D	Phosphoglycerate mutase (energy metabolism - glycolysis)	7.00E-126	82	3' UTR	N/A
PS_C38608_168_[T/C]	B	Reverse transcriptase	1.00E-25	71	exonic	S

a – Directional selection(D) /Balancing selection (B)

b – Synonymous (S)/Non-synonymous substitution (NS) (indication of amino acid substitution)

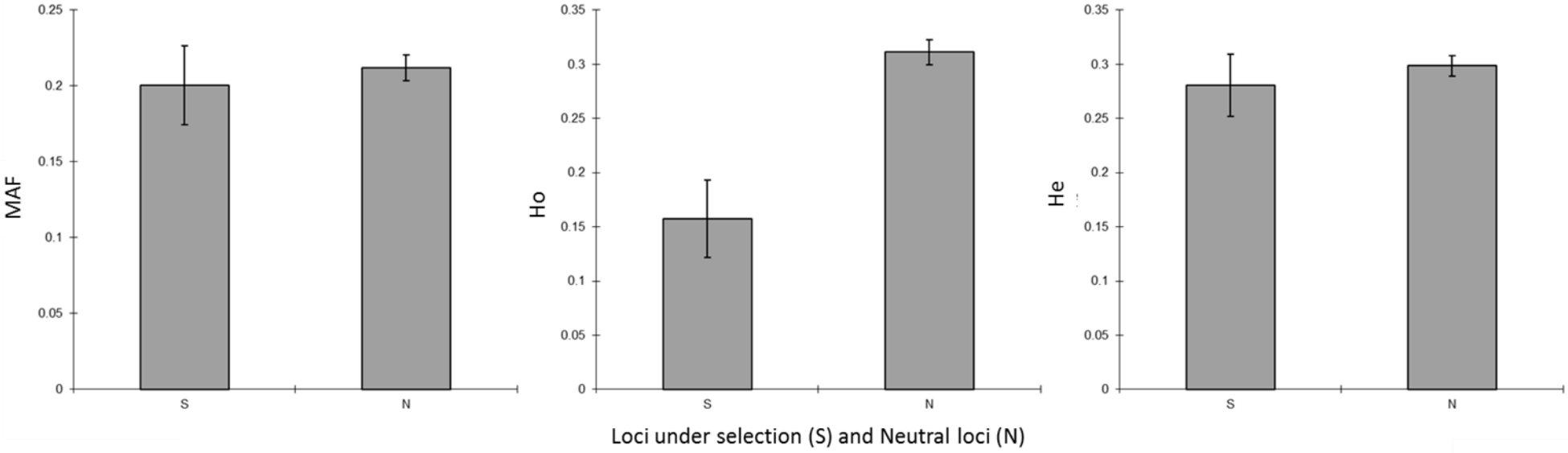

Figure 4.2: Summary of mean diversity statistics among loci under selection and neutral loci.

Table 4.3: Pairwise F_{st} estimates based on all markers (including loci under selection, lower diagonal) and excluding loci under selection (upper diagonal), shaded values highlights the pairwise F_{st} estimates among temporal samples.

	AS(F_1)	CR(2000)	CR(2011)	GB(2003)	RP(2003)	RP(2011)	SD(2004)	SD(2010)	WC(F_1)	WC(F_2)	WS(2004)	WS(2011)
AS(F_1)	-	0.018**	0.033**	0.029**	0.027**	0.010**	0.014**	0.008*	0.024**	0.069**	0.030**	0.037**
CR(2000)	0.119**	-	0.003	0.003	0.004	0.012**	0.010**	0.009*	0.002	0.055**	0.006	0.002
CR(2011)	0.127**	0.007*	-	0.016**	0.004*	0.029**	0.027**	0.027**	0.005	0.058**	0.023**	0.006
GB(2003)	0.120**	0.004	0.020**	-	0.008	0.020**	0.016**	0.012**	0.002	0.065**	0.007**	0.008
RP(2003)	0.119**	0.006*	0.003	0.011**	-	0.019**	0.019**	0.015**	0.008*	0.056**	0.011**	-0.001
RP(2011)	0.127**	0.097**	0.112**	0.097**	0.110**	-	0.004*	0.004	0.014**	0.056**	0.015**	0.017**
SD(2004)	0.013**	0.119**	0.132**	0.114**	0.132**	0.005*	-	0.002	0.024**	0.056**	0.014**	0.015**
SD(2010)	0.010**	0.092**	0.107**	0.090**	0.103**	0.003	0.005*	-	0.011**	0.070**	0.010**	0.016**
WC(F_1)	0.110**	0.004	0.003	0.003	0.006*	0.089**	0.117**	0.083**	-	0.060**	0.009*	0.011*
WC(F_2)	0.092**	0.075**	0.081**	0.080**	0.078	0.073**	0.087**	0.088**	0.070**	-	0.080**	0.065**
WS(2004)	0.112**	0.012**	0.035**	0.014**	0.021	0.087**	0.100**	0.079**	0.016**	0.091**	-	0.010*
WS(2011)	0.103**	0.006	0.016*	0.010	0.008	0.071**	0.092**	0.068**	0.013	0.066**	0.014*	-

*** statistically significant at the 0.01 nominal level*

** statistically significant at the 0.05 nominal level*

Table 4.4: AMOVA results using all loci (including loci under selection) and excluding loci under selection.

AMOVA over all SNP loci (including loci under selection)				
Source of Variation	**Sum of Squares**	**Variance Components**	**% Variation**	**Fixation Indices**
Among groups	666.544	0.960	5.674	F_{CT} = 0.056**
Among populations within groups	222.993	0.295	1.742	F_{CS} = 0.018**
Within populations	10993.941	15.666	92.584	F_{ST} = 0.074**
Total	11883.478	16.921		
AMOVA over SNP loci (excluding loci under selection)				
Source of Variation	**Sum of Squares**	**Variance Components**	**% Variation**	**Fixation Indices**
Among groups	223.176	0.209	1.420	F_{CT} = 0.014**
Among populations within groups	179.892	0.206	1.404	F_{CS} = 0.014**
Within populations	10021.205	14.276	97.176	F_{ST} = 0.028**
Total	10424.273	14.691		

*** statistically significant at the 0.01 nominal level*

** statistically significant at the 0.05 nominal level*

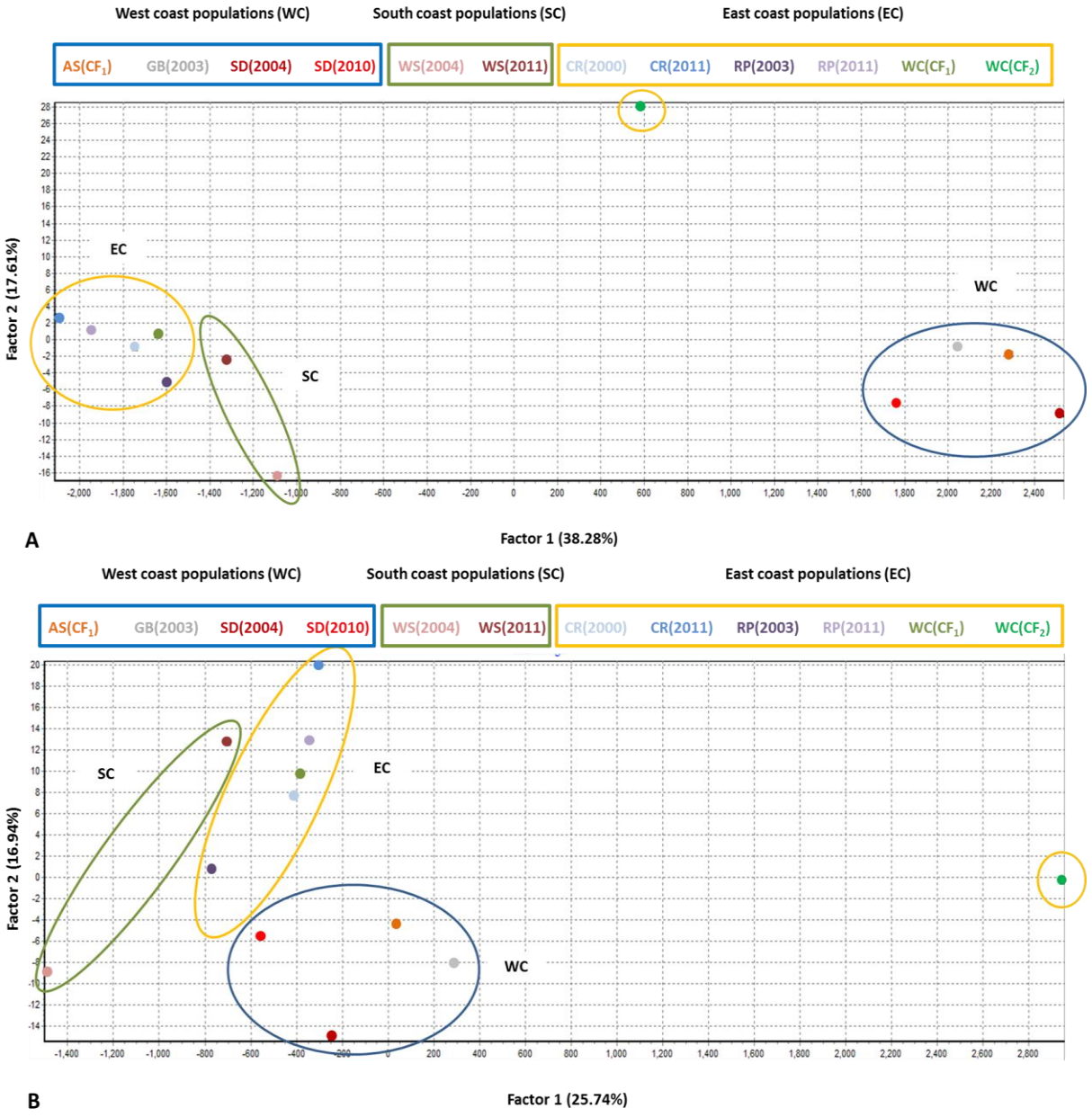

Figure 4.3: Factorial correspondence analysis plot constructed using all genotype data (including loci under selection) (A) and excluding loci under selection (B).

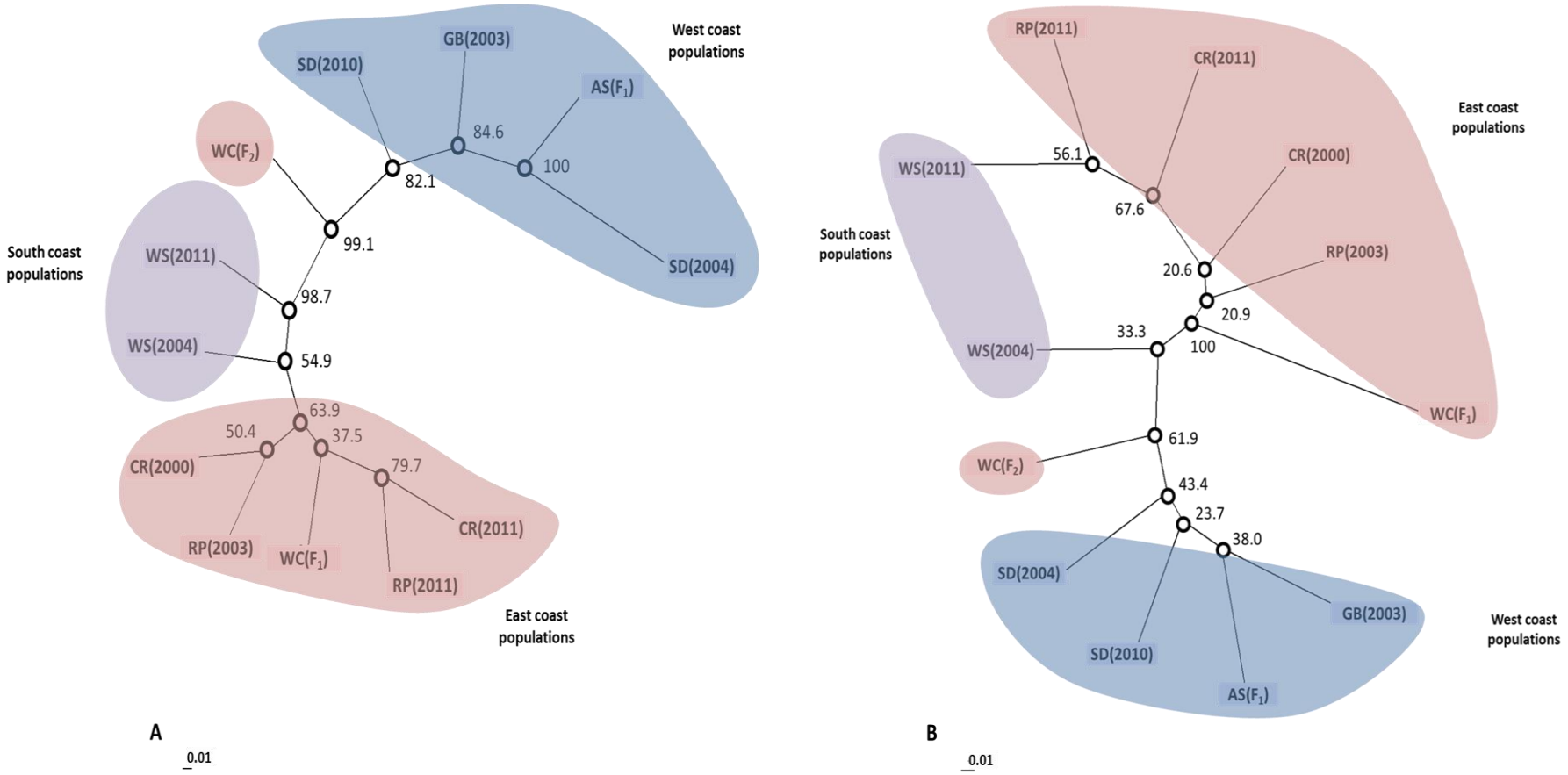

Figure 4.4: Dendrograms based on Nei's genetic distance (Nei, 1972) constructed using the neighbour joining algorithm with 1000 bootstrap replicates (nodal values). (A) Dendrogram based on all loci, including loci under selection. (B) Dendrogram excluding loci under selection.

Table 4.5: Estimates for effective population size based on three methods, heterozygosity excess, LD and temporal and two methods for the detection of recent bottlenecks, Wilcoxon singed rank test for heterozygosity excess and mode-shift indicator test.

	Heterozygosity excess test	LD test	Temporal test	Wilcoxon *P*-value for recent bottleneck	Mode shift test for recent bottleneck
AS(F_1)	19	137.9 (75.2-597.5)	21.1 (12.0-42.3)	0.00	normal L-shaped distribution
SD(2004)	∞	84.1 (60.7-130.0)	-	0.00	normal L-shaped distribution
SD(2010)	∞	179.3 (88.4-3681.6)	41.0 (19.4-173.1)	0.00	shifted mode
GB(2003)	∞	∞(242.3-∞)	-	0.00	normal L-shaped distribution
WS(2004)	∞	∞ (-207.2-∞)	-	0.00	shifted mode
WS(2011)	∞	∞ (-44-∞)	24.0 (9.3-∞)	0.00	shifted mode
CR(2000)	∞	75.0 (40.6-321.0)	-	0.00	normal L-shaped distribution
CR(2011)	∞	41.2 (31.7-56.9)	44.8 (16.3-∞)	0.00	shifted mode
RP(2004)	∞	160.3 (104.3-318.3)	-	0.00	shifted mode
RP(2011)	∞	179 (86.0-27539.3)	24.0 (13.6-49.8)	0.00	shifted mode
WC(F_1)	∞	44.6 (29.3-83.9)	26.0 (13.0-77.8)	0.00	shifted mode
WC(F_2)	19.1	32.1 (27.1-37.7)	9.0 (6.8-13.2)	0.00	normal L-shaped distribution

4.4. Discussion

4.4.1. Transcriptome Sequencing, Evaluation of SNP Assay Performance and Genotyping Success

The use of next-generation sequencing technology has become common and has already been employed in a number of species (see Deschamps and Campbell, 2010; Davey *et al.*, 2011 for reviews). Transcriptome sequencing in particular is popular because it provides the opportunity for identifying genetic variants that have direct association to genes and facilitates easier *de novo* assembly in the absence of genomic reference sequences (*e.g.* Barbazuk *et al.*, 2007; Novaes *et al.*, 2008; Parchman *et al.*, 2010; Renaut *et al.*, 2010). This study is not the first to generate transcriptomic sequence data for *H. midae* by means of next-generation sequencing. Franchini *et al.* (2011) used the Illumina® sequence-by-synthesis method to generate an initial transcriptome sequencing experiment using a limited number of full-sib individuals. Transcriptome sequencing has only been done for two other haliotids (using conventional Sanger sequencing and Illumina® HiSeq 2000, respectively), *H. diversicolor* (Jiang J-Z *et al.*, 2011) and *H. rufescens* (De Wit and Palumbi, 2012). What is evident from these studies is the relatively small portion of contig sequences that have known homologous genes in other species as evaluated by BLAST searches to the public databases, such as NCBI's nr databases for protein and DNA sequences. Jiang J-Z *et al.* (2011) found homologues for 60.6% of the contigs, but Franchini *et al.* (2011) could only find homologues for 16.8% of the *H. midae* contigs. Furthermore, only 13.14% of these *H. midae* contigs could be functionally annotated using Blast2Go®. In the present study 15.6% of the contig sequences had homologues among the protein nr database of NCBI and approximately half of these could be annotated by Blast2Go®. The quantity of novel sequences among the abalone transcripts suggests that there are a substantial number of unique abalone genes that are not present in animals, such as mammals that are overrepresented in public databases. Furthermore, although homologues could be found, many of these genes were denoted as "hypothetical" and were only postulated by virtue of genomic sequence open reading frame predictions. The presence of such transcripts in abalone could serve as evidence that these hypothetical genes are in fact transcribed and therefore might be functionally involved in biological processes, however the functionalities

of these genes remain unknown. Consequently, there is much scope for further investigation of the abalone transcriptomic profile that could elucidate the characteristics of such anomalous gene products.

Previous studies suggest that SNPs might be frequent in the *H. midae* genome, with estimates of one SNP every 100 to 500 bp having been reported depending on the genomic region under investigation (Bester *et al.*, 2008; Rhode *et al.*, 2008; Franchini *et al.*, 2011). Although this may provide an almost limitless source of genetic variation that could be investigated, the high frequency of SNPs might create a technical challenge for the development of SNP assays *via* the Illumina® GoldenGate™ platforms. The allelic discrimination is based on the hybridisation of allele-specific oligonucleotide probes to the target DNA and as such the flanking sequences to the SNP of interest must be homogeneous with no other polymorphisms in the 60 bp on either side. If additional polymorphism did occur, it could influence the annealing ability of the hybridisation probes. Although much attention was given to prevent this, undetected polymorphisms could explain some of the genotyping failures observed at putative SNP loci in the current study. A preliminary assessment of the BeadXpress® platform for SNP genotyping in abalone (specifically, in *H. midae*) found that loci known to be in hyper-variable genomic regions are prone to genotypic failures (Blaauw, 2012). The use of cDNA as the original template DNA for polymorphism detection might also influence hybridisation probe annealing if SNPs of interest are located close to exon-intron boundaries: The presence of intronic sequences in the gDNA during the genotyping reaction could interfere with probe hybridisation (Wang *et al.*, 2008). Nonetheless, the assay success rate (82.80%) was comparable with a number of studies reporting SNP assay success (ranging from 66.90% to 92.0%), using GoldenGate™ platforms, for a variety of species (Pavey *et al.*, 2008; Wang *et al.*, 2008; Eckert *et al.*, 2009; Hyten *et al.*, 2010; Yan *et al.*, 2010; Khan *et al.*, 2012). Blaauw (2012) reported a success rate of 85.40% using mostly previously validated SNP loci for *H. midae*, whilst Du Plessis (2012) reports 76.34% success using predominantly putative markers identified *in silico*. A success rate within this range is therefore expected for this study.

A comparatively high number (22.40%) of the putative SNPs were found to be monomorphic. Sequence coverage was generally low, but not unusual for 454-pyro-squencing. However, it does complicate the identification of sequence variants especially

considering that 454-pyro-sequencing generally has one of the highest sequence error rates among the next-generation sequencing technologies (Gilles *et al.*, 2011). Harismendy *et al.* (2009) compared 454-pyrosequecing SNP call accuracy with other next-generation sequencing platforms (Illumina® Genome Analyzer (GA), ABI SOLiD™) and conventional Sanger sequencing using Illumina's Hap550 BeadChip and found 454-pyrosequencing had the lowest accuracy (97.4%) compared to, for example Illumina® GA (100%) and ABI SOLiD™ (99.7%). Wang *et al.* (2008) found that coverage (sequence depth, number of reads in a contig) and MAF were significant factors in determining the conversion rate of a SNP assay. To compensate for this, minimum coverage was set to 8X and MAF to 0.25 in the present study; thus at least two observations of the alternative variant was necessary to call a putative SNP. In general, however, is not unexpected for SNPs identified *in silico* to have lower conversion rates than SNPs first validated *in vitro* (Lepoittevin *et al.*, 2010) Furthermore, when using cDNA the occurrence of RNA editing, the post-transcriptional modification of nucleotide base pairs within mRNA cannot be excluded; thus the nucleotide diversity of the mRNA may not necessarily reflect genomic diversity (*e.g.* Seiwert and Stuart, 1994). Consequently, the conversion rate for the this assay (60.42%) was marginally lower than the previous developed assays for *H. midae* (64.5%, Blaauw, 2012; 68.82%, Du Plessis, 2012); but was higher in comparison to a number of other studies that reported estimates as low as 40.63% (Wang *et al.*, 2008; Hyten *et al.*, 2010; Khan *et al.*, 2012). The success- and conversion rates of a SNP assay might be improved by validating *in silico* SNPs through Sanger sequencing before incorporating a particular locus into the assay (*e.g.* Seeb *et al.,* 2011). In this way compensations could be made for intronic sequences and sequencing errors, however this will significantly increase marker development costs.

Single nucleotide polymorphisms might be susceptible to ascertainment bias: the discrepancies in allele frequencies that arise due to biased sampling (often unintentional) and detection methodologies employed to find putative markers (Brumfield *et al.*, 2003; Morin *et al.*, 2004; Seddon *et al.*, 2005; Helyar *et al.*, 2011). In this study ascertainment bias was minimised by aiming to capture as much of the genetic variation as possible. This was facilitated by using specimens for the initial cDNA construction from across the natural distribution range of abalone (west-, south-, and east coast of South Africa) as well as including wild and cultured individuals. Furthermore, tissue samples were taken from the

major organ groups and pooled in equal molar concentrations in order to provide a "more complete" gene representation. Finally, by incorporating SNPs that had been previously identified and utilised, any methodological bias could at least be partially negated. Nonetheless, biases may still persist: firstly, using only genic sequence potentially introduces unequal representation of different genomic regions (although this could be perceived as a favourable bias) and secondly, tissues for RNA extraction were only available for 2010/2011 specimens (introducing a temporal bias in population samples). A degree of caution must thus be taken with regards to the interpretation of the data.

4.4.2. Genetic Diversity and Effective Population Size

In general, for most animal taxa, the transition (purine to purine or pyrimidine to pyrimidine substitution) to transversion (purine to pyrimidine or pyrimidine to purine substitution) ratio is generally in favour of transitions. It is postulated that cytosine nucleotides are more readily subjected to methylation, especially in CpG like repeat units. The 5-methyl cytosine is mutationally unstable and during spontaneous deamination can transition to a thymidine nucleotide; consequently resulting in a transition excess (Brookes, 1999; Vignal *et al.*, 2002). Transition to transversion ratios in mammals are reported to be between 1.4:1 and 1.7:1 (Collins and Jukes, 1994; Picoult-Newberg *et al.*, 1999), while birds show higher ratios, 2.3:1 to 4.0:1 (Smith *et al.*, 2001; Vignal *et al.*, 2002). Amongst the invertebrates, including molluscs, the general trend seems to prevail with estimates for the silkworm (*Bombyx mori*) at 1.66:1 (Cheng *et al.*, 2004); 1.3:1 for Eastern oysters (*Crassostrea virginica*; Quilang *et al.*, 2007) and 2.4:1 for weathervane scallops (*Patinopecten caurinus;* Elfstrom *et al.,* 2005). To date, 25 SNPs have been confirmed for *H. discus hannai*. Based on these loci the transition to transversion ratio was fairly high at 3.6:1 (Qi *et al.*, 2008, 2009). Estimates for confirmed loci for *H. midae* are contradictory to the general expectations with Rhode *et al.* (2008) reporting a 1:1 ratio (12 SNPs), whilst Bester *et al.* (2008) even reported a transversion excess (1:1.5) (12 SNPs). Based on the current estimate of 116 confirmed SNPs, the ratio of 1.7:1 is within the expected range for most animal taxa. Based on putative SNP loci Franchini *et al.* (2011) also estimated the transition to transversion ration for *H.*

midae to be with this range at 1.6:1. The initial estimates (Bester *et al*., 2008; Rhode *et al*., 2008) are most likely an artefact of the small number of loci investigated

Single nucleotide polymorphism heterozygosity estimates are expected to be considerably less than estimates based on microsatellite markers (*e.g.* Chapter 2 and 3), because of the biallelic nature of the maker type. The current estimates for heterozygosity and MAF are comparable with observations for a number of mammalian species with estimates ranging from 0.05 to 0.50 for heterozygosity and 0.02 to 0.50 for MAF (*e.g.* Brouillette and Venta, 2002; Seddon *et al.*, 2005; Pariset *et al.*, 2006; Cappuccio *et al.*, 2006). Diversity statistics also correlated well with estimates for other broadcast spawning molluscs, *e.g.* various species of scallop and oyster, with estimates for MAF ranging from 0.01 to 0.50 and heterozygosity ranging from 0.11 to 0.87 (Elfstorm *et al.*, 2005; Arias *et al.*, 2009; Li *et al.*, 2009; Varney *et al.*, 2009; Jiang G *et al.*, 2011). Estimates for the Pacific abalone (*H. discus hannai*) was slightly higher (Qi *et al.*, 2008, 2009) and Bester-van der Merwe *et al.* (2011) also reported SNP heterozygosities for a number of wild populations of *H. midae* marginally higher than the presently reported estimates. It must however be noted that these estimates were based on a limited number of loci; therefore the current diversity estimates most likely provides a more realistic genome-wide estimate of genetic diversity at SNP loci in *H. midae.* There was no evidence for genetic diversity disparities between the populations (spatial and temporal) even for the cultured populations; which conforms to the microsatellite based estimates [Rhode *et al.*, 2012 (Chapter 2); Chapter 3]. Regardless of the fact that South African abalone populations have been particularly impacted by overharvesting in the past this has not translated into the expected decline in genetic diversity. The complex life-history of abalone, with overlapping generations could act as a buffer against the loss of genetic diversity due to animals from different age classes contributing during any given spawning event (Heath *et al.*, 2002; Riccioni *et al.*, 2010).

Overall, mean F_{is} values for the populations under investigation were low and even demonstrated heterozygous excess for many of the populations (Appendix C: Table S4.2), suggesting relatively low levels of inbreeding. The majority of loci conformed to Hardy-Weinberg expectations, however with the Wilcoxon signed rank test significant heterozygous excess at all populations was detected. This could be an indication of recent population bottlenecks (Cornuet and Luikart, 1996); however, the population bottleneck

could not be supported by the mode-shift indicator test for all populations (Table 4.5). Evidence for a population bottleneck amongst the cultured populations could be expected as it is known that these populations are derived from a limited number of broodstock collected from the wild (Rhode *et al.*, 2012; Chapter 2). Bester-van der Merwe *et al.* (2011) could, however not find evidence for a historical bottleneck for wild South African abalone populations based on heterozygosity; this seems to be supported by the heterozygous excess estimate for effective population size based on the present data. However, if a population has undergone a recent bottleneck the accompanying medium- to long-term reduction in heterozygosity will not be evident (Leberg, 1992; Cornuet and Luikart, 1996) and thus the heterozygosity method may overestimate the effective population size. The LD method might thus provide a more accurate indication of medium- to long-term effective population size because it takes into consideration historical recombination events (Waples and Do, 2010). When the effective population size is reduced it generally increases the genome-wide LD. This could explain why the SD(2004) and RP(2003) samples demonstrate a reduced point estimate for effective population size in comparison to the 2010/11 sample populations: If these populations did undergo a recent bottleneck, which is likely given the overharvesting of abalone (Raemaekers *et al.*, 2011), recent ecosystem shifts (Mayfield and George, 2000; Cockcroft *et al.*, 2008) and estimates of stock biomass (Dichmont *et al.*, 2000; Plagányi *et al.*, 2001; Proudfoot *et al.*, 2006; Plagányi and Butterworth, 2010), the bottleneck probably precedes 2003/4. The increased estimate in 2010/11 could therefore be ascribed to rapid LD decay post-bottleneck. The temporal estimate for effective population size measures the fluctuation in allele frequencies across successive generations and therefore provides a short-term estimate. The effective population size point estimate based on the temporal method was substantially less for most of the populations, but there was sufficient overlap in the confidence intervals. Nonetheless, being a more contemporary estimate it is not surprising that it is less than the long-term estimate, especially in broadcast spawning animals. Under aquaculture conditions, considerable variation in parental contributions have been observed in broadcast spawning molluscs, including abalone, at any given spawning event (Slabbert *et al.*, 2009; Van den Berg and Roodt-Wilding, 2010). It is comprehendible that this phenomenon is also replicated in the wild. Such differential spawning contributions could thus lead to skewed allele frequencies from

generation to generation and consequently reduce the short-term effective population size, especially in bottlenecked populations.

4.4.3. Candidate Loci Under Selection

Thirteen (approximately 11.2%) loci demonstrating evidence of being under selection based on the set criteria were identified. This is somewhat less than the percentage that was identified for the microsatellite loci in Chapter 3. However, this seems to be due to the relatively few loci under balancing selection amongst the current marker set, only one locus. With regards to the number of loci under directional selection the microsatellite and SNP data correlates well; approximately 9% and 10%, respectively. The disparity in loci under balancing selection might be due to the genomic distribution differentials of the microsatellite- and SNP loci used. Estimates for SNP loci under selection is considerably less for a number of other species with estimates ranging from 3.9% to 7.9% (Namroud *et al.*, 2008; Narum *et al.*, 2010; Willing *et al.*, 2010; Renaut *et al.*, 2011; Whiteley *et al.*, 2011), but were similar to estimates for the Atlantic cod (±10%; *Gadus morhua*; Nielsen *et al.*, 2009) and the periwinkle snail (7-12%; *Littorina saxatilis*; Galindo *et al.*, 2010). A recent study found that 5.23% of EST-SNP loci in wild red abalone (*H. rufescens*) were candidates for divergent selection amongst populations (De Wit and Palumbi, 2012).

Five loci are associated to genes involved in energy metabolism (Table 4.2). Energy metabolism genes appear to be frequently under selection in a variety of organisms (*e.g.* Namroud *et al.*, 2008; Galindo *et al.*, 2010; Guatier and Naves, 2011; Whiteley *et al.*, 2011). This is however not surprising as energy metabolism is directly correlated to various environmental stressors such as temperature, oxygenation and availability of food sources. An interesting observation is that loci associated to transposable element genes, such as reverse transcriptase and endonucleases are also under selection. Galindo *et al.* (2010) made a similar observation for the gastropod mollusc *L. saxatilis*. Transposable elements seem to be plentiful in the *H. midae* genome and are transcriptionally active (Rhode and Roodt-Wilding, 2011; *H. midae* unpublished transcriptome data). Transposable elements are known to alter gene functions and facilitate genome evolution (Bennetzen, 2000; Kidwell, 2002; Medstrand *et al.*, 2005; Gogvadze and Buzdin, 2009).

Single nucleotide polymorphisms that locate in coding region may influence peptide composition. Seven of the outlier loci were in fact found in exonic regions, of which three produced non-synonymous substitutions (Table 4.2). The non-synonymous loci might be direct targets for selection as changing the peptide sequence could alter protein structure and function. Such a direct causation cannot necessarily be made for synonymous loci or loci in the UTRs. Nonetheless, loci in UTRs may exert functional effects when present in regulatory elements or motifs such as transcription binding sites (Majewski and Ott, 2002). Furthermore, codon usage biases may result in selection differentials on synonymous substitutions (Williams and Hurst, 2000; Chamary and Hurst, 2004). The extent of such codon usage biases in *H. midae* remains to be investigated, but work done on the Pacific oyster suggest that codon usage might be under strong selective pressure in molluscs (Sauvage *et al.*, 2007). Studies have also suggested that synonymous substitutions may have a variety of effects on protein functionality and avalibility, with effects on mRNA structure, RNA processing, post-transcriptional regulation and translation (Sauna and Kimchi-Sarfaty, 2011). Genetic hitch-hiking, due to close physical linkage to a causal variant can, however, not be excluded (Nosil *et al.*, 2009; Karasov *et al.*, 2010).

Observed heterozygosity was significantly reduced at loci under selection compared to neutral loci which is expected given that most loci were under directional selection (Figure 4.2). The lack of significant outliers (based on the BayeScan estimate) among the pairwise temporal samples suggest that selective pressures were relatively stable across the given timeframe. One locus in particular, however, suggests that some shifts in selective pressure did occur over time - locus PS_C23591_200_[T/C] (Appendix C: Table S4.3). This locus is identified by the joint spatio-temporal (across all populations) outlier analysis by both Lositan and BayeScan, but fails to reach significance when temporal samples are analysed independently. This locus is also detected as an outlier by Lositan amongst the pairwise temporal comparisons between the WS, CR and RP populations. This locus is associated to a transcription factor gene (TFG) thus giving credence to the hypothesis that regulatory variation will be the first to respond to selection and that in general variation in the regulation of gene expression might be an important mechanism in phenotypic evolution (Purugganan, 2000; Barrier *et al.*, 2001; Hoekstra and Coyne, 2007). The agent that is

responsible for this altered selection regime remains unknown; however could possibly be ascribed to anthropogenic factors and/or ecological shifts.

4.4.4. Population Genetic Structure

The geographic correlation with population genetic structuring of *H. midae* around the South African coast has been well documented. The major barrier to gene flow is most likely the retroflection of the Agulhas current at Cape Agulhas, dividing the population into two major reproductive stocks (Bester-van der Merwe *et al.*, 2011). This genetic structuring is supported by the present data, with both the factorial correspondence analysis and the dendrograms (Figure 4.3, 4.4) demonstrating a clear clustering pattern of populations in groups of geographic origin. This geographic genetic structuring is particularly supported when the loci under selection are included in the analyses, with the factorial correspondence analysis showing distinct population clusters west and east of Cape Agulhas and higher bootstrap values at nodal junctions on the dendrogram. Furthermore, when including loci under selection the pairwise F_{st} values are higher and the percentage of among group genetic variation increases almost three fold (AMOVA, Table 4.4). This could be interpreted as an indication of selection facilitating adaptation to the local environmental conditions and will give credence to the biogeographical vicariance hypothesis of Bester van der-Merwe *et al.* (2012) for population divergence and ultimately speciation of abalone along the South African coast due to ecological adaptation to the three biogeographical provinces.

The only population that does not seem to conform to the geographic clustering patterns is the WC(F_2) cultured population, which fails to group with any of the general population groupings, irrespective of the data used (including or excluding loci under selection). This is probably because of the pronounced effects of the domestication process. The WC(F_2) population is a selected second generation aquaculture population and is for all practical purposes an isolated population with a small effective population size (also supported by unpublished microsatellite data). In chapter 3, when using the microsatellite markers under selection there was a clear wild-/cultured population separation, irrespective of the geographic origin of a particular population. This pattern is not replicated with the SNP

markers. For the most part the SNP loci used in this analysis represents structural variation, whereas the microsatellite loci represent regulatory variation (Rhode and Roodt-Wilding, 2011). As discussed previously, regulatory variation will more readily respond to novel selection pressures; in this case domestication (Purugganan, 2000; Barrier *et al.*, 2001; Mignon-Grasteau *et al.*, 2005; Dobney and Larson, 2006; Hoekstra and Coyne, 2007). This is not to say that the SNP loci under selection do not contribute to the development of the domestic phenotype; on the contrary when including these loci, the partitioning of the AS population from the wild populations is particularly supported in the population dendrogram. Random drift on the other hand seems to be a more important factor in the divergence of the WC cultured populations (Figure 4.4).

The relatively large long-term effective population sizes and high rates of gene flow (Bester-van der Merwe *et al.*, 2011) suggests that genetic variation in the wild *H. midae* populations should be fairly stable across temporal scales (Hansen *et al.*, 2002; Palm *et al.*, 2003; Hoffman *et al.*, 2004; Lee and Boulding, 2009). Nonetheless, there is sufficient evidence to demonstrate that this is not the case [some caution is however heeded due to unbalanced samples that could reduce statistical power, Goudet *et al.*, 1996]; which is surprising given that the populations are at most removed by only two generations. Genetic differentiation of temporal populations inhabiting a particular region is not uncommon. Such differentiation is generally associated with environmental instabilities creating differential selection regimes, but can also be caused by frequent population extinction and recolonisation events within a meta-population structure (Vandewoestijne *et al.*, 1999; Østergaard *et al.*, 2003; Hoffman *et al.*, 2004). Furthermore, an increasing number of studies are reporting temporal genetic heterogeneity, whilst spatial genetic structure remains stable in a number of marine organisms with planktonic larval stages (*e.g.* Chapman *et al.*, 2002; Robainas *et al.*, 2005; Florin and Höglund, 2007; Lee and Boulding, 2007, 2009). The temporal fluctuation in the selection on locus PS_C23591_200_[T/C] suggests that changing selection regimes may result in the genetic differentiation of temporal samples. However, the estimates based on neutral loci only, also demonstrates significant differentiation. A strict extinction-recolonisation scenario seems unlikely given the limited number of generations under investigation and the life-history characteristics of the South African abalone. However, gene flow by means of larval settlement from other genetically distinct

populations may provide an explanation: Larvae from distinct populations, whom may too vary in genetic composition through time and space, may settle in a divergent population. Because there are generational overlaps animals that differ in age may contribute to any spawning event leading to fluctuation in temporal genetic diversity due to such larval recruitment regimes (Johnson and Wernham, 1999; Moberg and Burton, 2000). This is a likely scenario given the migration patterns of abalone larvae along the South African coast (Bester van der Merwe *et al.*, 2011). Another plausible explanation for the temporal observation could be ascribed to Hedgecock's (1994) "sweepstakes hypothesis", noting the potential disparity in long-term and short-term effective population size (Table 4.5) and the broadcast spawning mode of reproduction of *H. midae*. Under this hypothesis the short-term effective population size is small enough for significant fluctuations in allele frequencies to occur by chance, *i.e.* random genetic drift. Thus, due to abalone's high fecundity and low rate of planktonic larval survival only a small portion of adult animals will contribute at any given generational interval, which reduces the effective population size and consequently exacerbates the effects of random drift; leading to significant temporal fluctuations in genetic variation. Lee and Boulding (2009) made a similar observation for two Pacific littorinid gastropod molluscs. In reality it is probable that combinations of the aforementioned processes are simultaneously at play. When considering only neutral loci the WS population also demonstrate some deviation in the general geographic grouping of populations. The south coast of South Africa has been postulated to be a secondary contact zone for abalone (Bester-van der Merwe *et al.*, 2011). The WS population is thus in all likelihood an "corridor" population connecting the major reproductive stocks on the west and east, and consequently particularly susceptible to variations in genetic diversity over time. However, it must be noted that the WS(2011) population sample only consisted of nine specimens and thus the effects of sampling error must be taken into consideration.

4.5. Conclusions

This study investigated the genetic diversity at both spatial and temporal scales of the economically important South African abalone, *H. midae*, using SNP markers. Spatial

diversity conformed to expectations as previously reported, with possible adaptation to local geographical environmental conditions maintaining this spatial genetic structuring. Nonetheless, evidence suggests significant population differentiation among temporal samples collected from the same locality. These temporal fluctuations are thought to be in response to ecosystems shifts (possibly in response to global warming) and anthropological effects (overfishing and poaching), although the generational interval at present is too limited to draw definitive conclusions. However, differential reproductive performance leading to small short-term effective population sizes and high gene flow between differentiated populations may also lead to substantial variations in the genetic constitution of abalone populations along the South African coast. Thus, these populations might be more dynamic than previously thought and could bear significance on strategies for conservation and fisheries management.

References

Allendorf FW, England PR, Luikart G, Ritchie PA, Ryman N (2008) Genetic effects of harvest on wild animal populations. Trends Ecol Evol 23: 327-337.

Anderson S, Sadinski W, Shugart L, Brussard P, Depledge M, Ford T, Hose J, Stegeman J, Suk W, Wirgin I, Wogan G (1994) Genetic and molecular ecotoxicology – a research framework. Environ Health Persp 102: 3-8.

Antao T, Lopes A, Lopes RL, Beja-Pereira A, Luikart G (2008) LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9: 323.

Antunes A, Troyer JL, Roelke ME, Pecon-Slattery J, Packer C, Winterbach C, Winterbach H, Hemson G, Frank L, Stander P, Siefert L, Driciru M, Funston PJ, Alexander KA, Prager KC, Mills G, Wildt D, Bush M, O'Brien SJ, Johnson WE (2008) The evolutionary dynamics of the lion *Panthera leo* revealed by host and viral population genomics. PLoS Genet 4: e1000251.

Arias A, Freire R, Boudry P, Heurtebise S, Méndez J, Insua A (2009) Single nucleotide polymorphism for population studies in the scallops *Aequipecten opercularis* and *Mimachlamys varia*. Conserv Genet 10: 1491-1495.

Barbazuk WB, Emrich SJ, Chen HD, Li L, Schnable PS (2007) SNP discovery via 454 transcriptome sequencing. Plant J 51: 910-918.

Barrier M, Robichaux RH, Purugganan MD (2001) Accelerated regulatory gene evolution in an adaptive radiation. Proc Natl Acad Sci USA 98: 10208-10213.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France). Available at http://kimura.univ-montp2.fr/genetix/

Bennetzen JL (2000) Transposable element contribution to plant gene and genome evolution. Plant Mol Biol 42: 251-269.

Bester AE, Roodt-Wilding R, Whitaker HA (2008) Discovery and evaluation of single nucleotide polymorphisms (SNPs) for *Haliotis midae*: a targeted EST approach. Anim Genet 39: 321-324.

Bester-van der Merwe AE, Roodt-Wilding R, Volckaert FAM, D'Amato ME (2011) Historical isolation and hydrodynamically constrained gene flow in declining populations of the South-African abalone, *Haliotis midae*. Conserv Genet 12: 543-555.

Blaauw S (2012) SNP screening and validation in *Haliotis midae.* Unpublished M.Sc. thesis. Stellenbosch University, South Africa.

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). Mol Biol Evol 23: 773-783.

Bradshaw WE, Holzapfel CM (2008) Genetic response to rapid climate change: it's seasonal timing that matters. Mol Ecol 17: 157-166.

Brookes AJ (1999) The essence of SNPs. Gene 234: 177-186.

Brouillette JA, Venta PJ (2002) Within-breed heterozygosity of canine single nucleotide polymorphisms identified by across-breed comparison. Anim Genet 33: 464-467.

Brumfield RT, Beerli P, Nickers DA, Edwards SV (2003) The utility of single nucleotide polymorphism in inferences of population history. Trends Ecol Evol 18: 249-256.

Byars SG, Papst W, Hoffmann AA (2007) Local adaptation and cogradient selection in the alpine plant, *Poa hiemata*, along a narrow altitudinal gradient. Evolution 61: 2925-2941.

Cappuccio L, Pariset P, Ajmone-Marsan P, Dunner S, Cortes O, Erhardt G, Lühken G, Gutscher K, Joost S, Nijman J, Lenstra JA, England PR, Zundel S, Obexer-Ruff G, Beja-Pereira A, Valentini A, The ECONOGENE Consortium (2006) Allele frequencies and diversity parameters of 27 single nucleotide polymorphisms within and across goat breeds. Mol Ecol Notes 6: 992-997.

Chamary JV, Hurst LD (2004) Similar rates but different modes of sequence evolution in introns and at exonic silent sites in rodents: evidence for selectively driven codon usage. Mol Biol Evol 21: 1014-1023.

Chapman RW, Ball AO, Mash LR (2002) Spatial homogeneity and temporal heterogeneity of red drum (*Sciaenops ocellatus*) microsatellites: effective population sizes and management implications. Mar Biotechnol 4: 589-603.

Cheng T-C, Xia Q-Y, Qian J-F, Liu QC, Lin Y, Zha XF, Xiang Z-H (2004) Mining single nucleotide polymorphisms from EST data of silkworm, *Bombyx mori*, inbred strain Dazao. Insect Biochem Molec 34: 523-530.

Cheviron ZA, Whitehead A, Brumfield RT (2008) Transcriptomic variation and plasticity in rufous-collared sparrows (*Zonotrichia capensis*) along an altitudinal gradient. Mol Ecol 17: 4556-4569.

Cockcroft AC, Van Zyl D, Hutchings L (2008) Large-scale changes in the spatial distribution of South African West Coast rock lobsters: an overview. Afr J Mar Sci 30: 149-159.

Collins DW, Jukes TH (1994) Rate of transition and tranversions in coding sequences since the human-rodent divergence. Genomics 20: 386-396.

Coltman DW (2008) Molecular ecological approaches to studying the evolutionary impact of selective harvesting in wildlife. Mol Ecol 17: 221-235.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144: 2001-2014.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate change. Science 292: 673-679.

Davis MB, Shaw RG, Etterson JR (2005) Evolutionary responses to changing climate. Ecology 86: 1704-1714.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12: 499-510.

De Jong G (2005) Evolution of phenotypic plasticity: pattern of plasticity and the emergence of ecotypes. New Phytol 166: 101-108.

Deschamps S, Campbell MA (2010) Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. Mol Breeding 25: 553-570.

De Wit P, Palumbi SR (2012) Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. Mol Ecol doi: 10.1111/mec.12081.

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. Trends Ecol Evol 13: 77-81.

Dichmont CM, Butterworth DS, Cochrane KL (2000) Towards adaptive approaches to management of the South African abalone *Haliotis midae* fishery. S Afr J Mar 22: 33-42.

Dobney K, Larson G (2006) Genetics and animal domestication: new windows on an elusive process. J Zool 269: 261-271.

Du Plessis (2012) Medium-throughput SNP genotyping and linkage mapping in *Haliotis midae*. Unpublished MSc thesis, Stellenbosch University, South Africa.

Eckert AJ, Pande B, Ersoz ES, Wright MH, Rashbrook VK, Nicolet CM, Neale DB (2009). High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda* L.). Tree Genet Genomes 5: 225-234.

Elfstrom CM, Gaffney PM, Smith CT, Seeb JE (2005) Characterization of 12 nucleotide polymorphisms in weathervane scallop. Mol Ecol Notes 5: 406-409.

Excoffier L, Lischer HEL (2010) Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Res 10: 564-567.

Franchini P, Van der Merwe M, Roodt-Wilding R (2011) Transcriptome characterization of the South African abalone *Haliotis midae* using sequencing-by-synthesis. BMC Research Notes 4: 59.

Fan J-B, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL, Hansen M, Steemers F, Butler SL, Deloukas P, Galver L, Hunt S, McBride C, Bibikova M, Rubano T, Chen J, Wickham E, Doucet D, Chang W, Campbell D, Zhang, B, Kruglyak S, Bentley D, Haas J, Rigault P, Zhou L, Stuelpnagel J, Chee MS (2003). Highly parallel SNP genotyping. Cold SH Q B 68: 69-78.

Fan J-B, Gunderson KL, Bibikova M, Yeakley JM, Chen J, Garcia EW, Lebruska LL, Laurent M, Shen R, Barker D (2006) Illumina Universal Bead Arrays. Method Enzymol 410: 57-73.

Florin A, Höglund J (2007) Absence of population structure of turbot (*Psetta maxima*) in the Baltic Sea. Mol Ecol 16: 115-126.

Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180: 977-993.

Galindo K, Grahame JW, Bultin RK (2010) An EST-based genome scan using 454 sequencing in the marine snail *Littorina saxatilis*. J Evol Biol 23: 2004-2016.

Gautier M, Naves M (2011) Footprints of selection in the ancestral admixture of a new world Creole cattle breed. Mol Ecol 20: 3128-3143.

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution: disentangling environmental and genetic responses. Mol Ecol 17: 167-178.

Gilles A, Meglécz E, Pech N, Ferreira S, Malausa T, Martin J-F (2011) Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. BMC Genomics 12: 245.

Gogvadze E, Buzdin A (2009) Retroelements and their impact on genome evolution and functioning. Cell Mol Life Sci 66: 3727-3742.

Goudet J, Raymond M, de Meeiis T, Rousset F (1996) Testing Differentiation in Diploid Populations. Genetics 144: 1933-1940.

Hansen, MM, Ruzzante DE, Nielsen EE, Bekkevold D, Mensberg K-L (2002) Long-term effective population sizes, temporal stability of genetic composition and potential for local adaptation in anadromous brown trout (*Salmo trutta*) populations. Mol Ecol 11: 2523-2535.

Hansen MM, Olivieri I, Waller DM, Nielsen EE, The GeM Working Group (2012) Monitoring adaptive genetic responses to environmental change. Mol Ecol 21: 1311-1329.

Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA (2009) Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome Biology 10: R32.

Heath DD, Busch C, Kelly J, Atagi DY (2002) Temporal change in genetic structure and effective population size in steelhead trout (*Oncorhynchus mykiss*). Mol Ecol 11: 197-214.

Hedgecock D (1994) Does variance in reproductive success limit effective population sizes of marine organisms? In: Genetic and Evolution of Aquatic Organisms (ed. Beaumont AR), pp. 122-134.

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol Ecol Res 11 (S1): 123-136.

Hoekstra H, Coyne JA (2007) The locus of evolution: Evo Devo and the genetics of adaptation. Evolution 61: 995-1016.

Hoffman EA, Schueler FW, Blouin MS (2004) Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. Evolution 58: 2536-2545.

Hyten DL, Song Q, Fickus EW, Quigley CV, Lim JS, Choi IY, Hwang EY, Pastor-Corrales M, Cregan PB (2010). High-throughput SNP discovery and assay development in common bean. BMC Genomics 11: 475.

Ingvarsson PK (2001) Restoration of genetic variation lost – the genetic rescue hypothesis. Trends Ecol Evol 16: 62-63.

Jensen LF, Hansen MM, Pertoldi C, Holdensgaard G, Mensberg K-L, Loeschcke V (2008) Local adaptation in brown trout early life-history traits: implications for climate change adaptability. P Roy Soc Lond B Bio 275: 2859-2868.

Jiang G, Li J, Li L, Zhang L, Bao Z (2011) Development of 44 gene-based SNP markers in the Zhikong scallop, *Clamys farreri*. Conserv Genet Res 3: 659-663.

Jiang J-Z, Zhang W, Guo Z-X, Cai C-C, Su Y-L, Wang R-X, Wang J-Y (2011) Functional annotation of an expressed sequence tag library from *Haliotis diversicolor* and analysis of its plant-like sequences. Marine Genomics 4: 189-196.

Johnson MS, Wernham J (1999) Temporal variation of recruits as a basis of ephemeral genetic heterogeneity in the western rock lobster *Panulirus cygnus*. Mar Biol 135: 133-139.

Karasov T, Messer PW, Petrov DA (2010) Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. PLoS Genet 6: e1000924.

Karban R, Strauss SY (2004) Physiological tolerance, climate change, and a northward range shift in the spittlebug, *Philaenus spumarius*. Ecol Entomol 29: 251-254.

Karell P, Ahola K, Karstinen T, Valkama J, Brommer JE (2011) Climate change drives microevolution in a wild bird. Nat Comm 2: 208.

Kerr RA (2007) Global warming is changing the world. Science 316: 188-190.

Khan MA, Han Y, Zhao YF, Korban SS (2012) A high-throughput apple SNP genotyping platform using the GoldenGate™ assay. Gene 494: 196-201.

Kidwell MG (2002) Transposable elements and the evolution of genome size in eukaryotes. Genetica 115: 49-63.

Leberg PL (1992) Effects of population bottlenecks on genetic diversity as measured by allozyme electrophoresis. Evolution 46: 477-494.

Lee HJ, Boulding EG (2007) Mitochondrial DNA variation in space and time in the northeastern Pacific gastropod, *Littorina keenae*. Mol Ecol 16: 3084-3103.

Lee HJE, Boulding EG (2009) Spatial and temporal population genetic structure of four northeastern Pacific littorinid gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. Mol Ecol 18: 2165-2184.

Lepoittevin C, Frigerio J-M, Garnier-Géré P, Salin F, Cervera M-T, Vornam B, Harvengt L, Plomion C (2010). *In vitro vs in silico* detected SNPs for the development of a genotyping array: What can we learn from a non-model species? PLos One 5: 1-9.

Li R, Li Q, Kong L (2009) Characterization of expressed sequence tag-derived single-nucleotide polymorphisms in the bay scallop *Argopecten irradians irradians*. Fish Sci 75: 1389-1400.

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21: 2128-2129.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89: 238-247.

Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA (2000) Biotic invasions: causes, epidemiology, global consequences, and control. Ecol Appl 10: 689-710.

Majewski J, Ott J (2002) Distribution and characterization of regulatory elements in the human genome. Genome Res 12: 1827-1836.

Mayfield S, George M (2000) Branch Interrelations among rock lobsters, sea urchins, and juvenile abalone: implications for community management. Can J Fish Aquat Sci 57: 2175-2185.

Medstrand P, Van de Lagemaat LN, Dunn CA, Landry J-R, Svenback D, Mager DL (2005) Impact of transposable elements on the evolution of mammalian gene regulation. Cytogenet Genome Res 110: 342-352.

Mignon-Grasteau S, Boissy A, Bouix J, Faure J-M, Fisher AD, Hinch GN, Jensen P, Le Neindre P, Morméde P, Prunet P, Vandeputte M, Beaumont C (2005) Genetics of adaptation and domestication in livestock. Livest Prod Sci 93: 3-14.

Moberg PE, Burton RS (2000) Genetic heterogeneity among adult and recruit red sea urchins, *Strongylocentrotus franciscanus*. Mar Biol 136: 773-784.

Morin PA, Luikart G, Wayne RK, The SNP workshop group (2004) SNPs in ecology, evolution and conservation. Trends Genet 19: 208-216.

Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, Van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riah K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010) The next generation of scenarios for climate change research and assessment. Nature 463: 747-756.

Namroud M-C, Beaulieu J, Juge N, Laroche J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Mol Ecol 17: 3599-3613.

Narum SR, Campbell NR, Kozfkay CC, Meyer KA (2010) Adaptation of redband trout in desert and montane environments. Mol Ecol 19: 4622-4637.

Nei M (1972) Genetic distance between populations. Am Nat 106: 283-292.

Nielsen EE, Hemmer-Hansen J, Poulsen NA, Loeschcke V, Moen T, Johansen T, Mittelholzer C, Taranger G-L, Ogden R, Carvalho GR (2009) Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). BMC Evol Biol 9: 276.

Nosil P, Funk DJ, Ortiz-Marrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18: 375-402.

Novaes E, Drost DR, Farmerie WG, Pappas GJ, Grattapaglia D, Sederoff RR, Kirst M (2008) High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. BMC Genomics 9: 312.

Palm S, Laikre L, Jorde PE, Ryman N (2003) Effective population size and temporal genetic change in stream resident brown trout (*Salmo trutta*, L.). Conserv Genet 4: 249-264.

Parchman TL, Geist KS, Grahnen JS, Benkman CW, Buerkle CA (2010) Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. BMC Genomics 11: 180.

Pariset L, Cappuccio I, Joost S, D'Andrea M, Marletta D, Ajmone-Marsan P, Valentini A, The ECONOGENE Consortium (2006) Characterization of single nucleotide polymorphisms in sheep and their variation as evidence of selection. Anim Genet 37: 290-292.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Annu Rev Ecol Evol Syst 37: 637-669.

Parker IM, Gilbert GS (2004) The evolutionary ecology of novel plant-pathogen interactions. Annu Rev Ecol Evol Syst 35: 675-700.

Pavy N, Pelgas B, Beauseigle S, Blais S, Gagnon F, Gesselin I., Lamothe M, Isabel N, Bousquet J (2008). Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and black spruce.BMC Genomics 9: 21.

Peel D, Ovenden JR, Peel SL (2004) NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.

Perry AL, Low PJ, Ellis JR, Reynolds JD (2005) Climate change and distribution shifts in marine fishes. Science 308: 1912-1915.

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M (1999) Mining SNPs From EST databases. Genome Res 9: 167-174.

Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in effective population size using allele frequency data. J Hered 90: 502-503.

Plagányi EE, Butterworth DS (2010) A spatial- and age-structured assessment model to estimate the impact of illegal fishing and ecosystem change on the South African abalone *Haliotis midae* resource. Afr J Mar Sci 32: 207-236.

Plagányi EE, Butterworth DS, Brandão A (2001) Toward assessing the South African abalone *Haliotis midae* stock using an age-structured production model. J Shellfish Res 20: 813-827.

Proudfoot L, Kaehler S, McGarry DK, Uppink PK, Aereboe M, Morris KM (2006) Exploitation status of infralittoral abalone (*Haliotis midae*) and alikreukel (*Turbo sarmaticus*) in the southern section of the Eastern Cape coast, South Africa. S Afr J Sci 102: 162-168.

Przybylo R, Sheldon BC, Merilä J (2000) Climatic effects on breeding and morphology: Evidence for phenotypic plasticity. J Anim Ecol 69: 395-403.

Purugganan MD (2000) The molecular population genetics of regulatory genes. Mol Ecol 9: 1451-1461.

Qi H, Liu Z, Zhang G-F (2008) Characterization of 12 single nucleotide polymorphisms (SNPs) in Pacific abalone, *Haliotis discus hannai.* Mol Ecol Res 8: 974-976.

Qi H, Liu X, Zhang G, Wu F (2009) Mining expressed sequences for single nucleotide polymorphisms in the Pacific abalone *Haliotis discus hannai*. Aquac Res 40: 1661-1667.

Quilang J, Wang S, Li P, Abernathy J, Peatman E, Wang Y, Wang L, Shi Y, Wallace R, Guo X, Liu Z (2007) Generation and analysis of ESTs from the eastern oyster *Crassostrea virginica* Gmelin and identification of microsatellite and SNP markers. BMC Genomics 8: 157-167.

Raemaekers S, Hauck M, Bürgener M, Mackenzie A, Maharaj G, Plagányi EE, Britz PJ (2011) Review of the causes of the rise of the illegal South African abalone fishery and consequent closure of the rights-based fishery. Ocean Coast Manage 54: 433-445.

Renaut S,Nolte AW, Bernatchez L (2010) Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae) Mol Ecol 19: 115-131.

Renaut S, Nolte AW, Rogers SM, Derome N, Bernatchez L (2011) SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). Mol Ecol 20: 545-559.

Rhode C, Roodt-Wilding R (2011) Bioinformatic survey of *Haliotis midae* microsatellites reveals a non-random distribution of repeat motifs. Biol Bull 221: 147-154.

Rhode C, Slabbert R, Roodt-Wilding R (2008) Microsatellite flanking regions: a snp mine in south african abalone (*Haliotis midae*). Anim Genet 39: 329.

Rhode C, Hepple J, Jansen S, Davis T, Vervalle J, Bester-van der Merwe AE, Roodt-Wilding R (2012) A population genetic analysis of abalone domestication events in South Africa: implications for the management of the abalone resource. Aquaculture 356-357: 235-242.

Riba M, Mayol M, Giles BE, Ronce O, Imbert E, Van der Velde M, Chauvet S, Ericson L, Bijlsma R, Vosman B, Smulders MJM, Olivieri I (2009) Darwin's wind hypothesis: does it work for plant dispersal in fragmented habitats? New Phytol 183: 667-677.

Riccioni G, Landi M, Ferrara G, Milano I, Cariani A, Zane L, Sella M, Barbujani G, Tinti F (2010) Spatio-temporal population structuring and genetic diversity retention in depleted Atlantic Bluefin tuna of the Mediterranean Sea. Proc Natl Acad Sci USA 5: 2102-2107.

Richards CM (2000) Inbreeding depression and genetic rescue in a plant metapopulation. Am Nat 155: 383-394.

Robainas BA, López GE, Hernández D, García-Machado E (2005) Temporal variation of the population structure and genetic diversity of *Farfantepenaeus notialis* assessed by allozyme loci. Mol Ecol 14: 2933-2942.

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Res 8: 103-106.

Saghai-Maroof MA, Solima KM, Jorgenson RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. Proc Natl Acad Sci USA 81: 8014-8018.

Sauna ZE, Kimchi-Sarfaty C (2011) Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet 12: 683-691.

Sauvage C, Bierne N, Lapègue S, Boudry P (2007) Single nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas.* Gene 406: 13-22.

Seddon JM, Parker HG, Ostrander EA, Ellegren H (2005) SNPs in ecology and conservation studies: a test in the Scandinavian wolf population. Mol Ecol 14: 503-511.

Seeb JE, Pascal CE, Grau ED, Seeb LW, Templin WD, Harkins T, Roberts SB (2011) Transcriptome sequencing and high-resolution melt analysis advance single nucleotide polymorphism discovery in duplicated salmonids. Mol Ecol Res 11: 335-348.

Seiwert SD, Stuart K (1994) RNA editing: transfer of genetic information from gRNA to precursor mRNA *in vitro*. Science 266: 114-117.

Slabbert R, Bester AE, D'Amato ME (2009) Analyses of genetic diversity and parentage within a South African hatchery of the abalone *Haliotis midae* Linnaeus using microsatellite markers. J Shellfish Res 28: 369-375.

Smith EJ, Shi L, Drummond P, Rodriguez L, Hamilton R, Ramlal S, Smith G, Pierce K, Foster J (2001) Expressed sequence tags for the chicken genome from a 10-day-old White Leghorn whole embryo cDNA library: DNA sequence characterization and linkage analysis. J Hered 92: 1-8.

Smith TB, Bernatchez L (2008) Evolutionary change in human-altered environments. Mol Ecol 17: 1-8.

Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. Trends Ecol Evol 18: 94-101.

Van den Berg N-C, Roodt-Wilding R (2010) Parentage assignment in *Haliotis midae* L.: a precursor to future genetic enhancement programmes for South African abalone. Aquac Res 41: 1387-1395.

Van der Merwe M (2010) Growth-related gene expression in *Haliotis midae*. Unpublished PhD dissertation, Stellenbosch University, South Africa.

Vandewoestijne S, Nève G, Baguette M (1999) Spatial and temporal population genetic structure of the butterfly *Aglais urticae* L. (Lepidoptera, Nymphalidae). Mol Ecol 8: 1539-1543.

Varney RL, Galindo-Sánchez CE, Cruz P, Gaffney PM (2009) population genetics of the eastern oyster *Crassostrea virginica* (Gmelin, 1791) in the Gulf of Mexico. J Shellfish Res 28: 855-864.

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other molecular markers and their use in animal genetics. Genet Sel Evol 34: 275-305.

Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. P R Soc B 275: 649-659.

Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O, Bairlein F (2002) Ecological responses to recent climate change. Nature 416: 389-395.

Wang S, Sha Z, Sonstegard TS, Liu H, Xu P, Somridhivej B, Peatman E, Kucuta H, Liu Z (2008).Quality assessment parameters for EST-derived SNPs for catfish. BMC Genomics 9: 450.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conserv Genet 7: 167-184.

Wapels RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evol Appl 3: 244-262.

Whiteley AR, Bhat A, Martins EP, Mayden RL, Arunachalam M, Uusi-Heikkilä S, Ahmad ATA, Shrestha J, Clark M, Stemple D, Bernatchez L (2011) Population genomics of wild and laboratory zebrafish (*Danio rerio*). Mol Ecol 20: 4259-4276.

Williams EJ, Hurst LD (2000) The proteins of linked genes evolve at similar rates. Nature 407: 900-903.

Williams LM, Olecksiak MF (2008) Signatures of selection in natural populations adapted to chronic pollution. BMC Evol Biol 8: 282.

Willing E-M, Bentzen P, Van Oosterhout C, Hoffmann M, Cable J, Breden F, Weigel D, Dreyer C (2010) Genome-wide single nucleotide polymorphisms reveal populations history and adaptive divergence in the wild guppies. Mol Ecol 19: 968-984.

Yan J, Yang X, Shah T, Sánchez-Villeda H, Li J, Warburton M, Zhou Y, Crouch JH, Xu Y (2010) High-throughput SNP genotyping with the GoldenGate assay in maize. Mol Breeding 25: 441-451.

CHAPTER 5

Synopsis: Summarising Discussion and Conclusions

5.1. Overview of the Research Endeavour

Recent developments, particularly the advancement of DNA sequencing technologies, have led to an explosion in biological research. It has become possible to investigate such questions pertaining to the complexities of genome structure, -function and -evolution and how this impacts on the development of phenotypes. Moreover, these studies were generally restricted to model organisms, but it has now become possible to readily investigate species that to date had limited genomic resources (see reviews on genomic approaches in marine species: Hauser and Seeb, 2008; Nielsen *et al.*, 2009; Oleksiak, 2010). As such, there has been a renewed interest in evolutionary biology, perhaps more specifically – evolutionary genetics – with a predominantly genomics focus (Nadeau and Jiggins, 2010). The use of population genomics to identify loci that are putatively influenced by selection and could therefore explain the development of adapted phenotypes to heterogeneous environments has become particularly popular (*e.g.* Campbell and Bernatchez, 2004; Namroud *et al.*, 2008; Willing *et al.*, 2010; Nunes *et al.*, 2011).

The South African abalone, *Haliotis midae* (a marine gastropod mollusc), is one of the most sought-after marine organisms in the world. Historically, the abalone fishery was South Africa's most lucrative and currently abalone culture is the largest and economically the most important sector within the South African aquaculture industry (Department of Agriculture, Forestry and Fisheries, 2012). Since the commencement of abalone aquaculture endeavours in South Africa, a number of research efforts were launched to better understand species biology (*e.g.* Mackay and Coyne, 2005; Vosloo and Vosloo, 2006; Roux *et al.*, 2008; Mackay *et al.*, 2011; Mouton and Gummow, 2011; Bester-van der Merwe *et al.*,

2011). With increasing interest in the genetic improvement of *perlemoen* a substantial number of genetic/genomic resources for *H. midae* have also been developed, including numerous molecular markers (especially microsatellites and SNPs) and ESTs (*e.g.* Bester *et al.*, 2008; Rhode *et al.*, 2008; Slabbert *et al.*, 2008, 2010; Franchini *et al.*, 2011; Rhode *et al.*, 2012; Slabbert *et al.*, 2012). In this study these genetics resources are exploited to gain a greater understanding of the micro-evolutionary processes, in particularly the role of selection, affecting both wild and cultured populations of *H. midae.* The South African abalone is in a unique position in that both efforts to domesticate abalone for commercial gains and conservation initiatives for preserving wild populations are running in parallel.

In Chapter 2, a standard population genetic analysis was performed to investigate the general genetic properties of cultured populations in relation to their wild progenitor populations. A comparison was also drawn between the use of anonymous genomic microsatellites and EST-derived microsatellites. Although at least two previous studies had assessed genetic diversity in cultured *perlemoen* and compared it to wild populations, these studies were limited in terms of the population cohort investigated (Evans *et al.*, 2004a; Slabbert *et al.*, 2009). The results presented in Chapter 2 are the first to provide a thorough population genetic assessment taking into account commercial processes that might influence the genetic constitution of these particular cultured populations. The data generated in Chapter 3 expands on the work done in Chapter 2 by providing a more genomic evaluation of genetic diversity at microsatellite loci using a population genomics approach. In particular the contribution of selection was evaluated in the development of the domestic phenotype. In Chapter 4 a temporal assessment of genetic diversity and factors influencing this diversity was done using SNP markers. The aim of Chapter 4 was to evaluate how stable the genetic structure of *perlemoen* populations were over time; given that both anthropogenic and changing ecological conditions were most likely affecting these populations. The work presented in Chapter 3 and 4 is one of the first of this nature for any haliotid species.

5.2. Synthesis of the Biological Findings

5.2.1. Molecular Markers, Outlier Loci and Evidence for Selection

Microsatellites and SNPs are currently the most used molecular markers in animal-, ecological- and population genetics. Biologically, these molecular markers exhibit a number of characteristics that have led to their popularity, including co-dominant mode of inheritance, high genomic frequency (especially SNPs) and high information content (particularly microsatellites due to its multi-allelic nature) (Beuzen *et al.*, 2000; Brumfield *et al.*, 2003; Lui and Cordes, 2004; Morin and McCarthy, 2007; Pérez-Enciso and Ferretti, 2010; Liu *et al.*, 2011). This study is the first to utilise such an extensive marker set (150 microsatellite- and 116 SNP markers) to evaluate the population genetic properties of South African abalone populations. Previous investigations (*e.g.* Evans *et al*., 2004a, b; Slabbert *et al.*, 2009; Bester van der Merwe *et al.*, 2011) used a limited number of markers ranging from three microsatellite loci to 12 SNP loci. This extensive marker set allows for a more representative assessment of genome-wide polymorphism. Given that the *H. midae* genome size is estimated at approximately 1400 cM (Franchini *et al.*, 2010, Vervalle *et al.*, in press), the average genome coverage for the microsatellite markers was approximately one marker every ±10 cM and for SNP makers, one marker every ±12 cM. Massault *et al.* (2008) recommends an average marker interval of 10 cM for conventional QTL mapping; however, considering that conventional QTL mapping exploits within family recombination rates it might be possible to use a lower marker density. When using population data (*i.e.* individuals selected at random, assuming no familial relationship among individuals) the number of markers necessary might increase depending on the level of LD maintained in the population under investigation. Meadows *et al.* (2008) recommends a marker interval of 0.1 - 2.5 cM in populations with short range LD (0 - 5 cM). Thus, although the current study used an extensive marker set, in all likelihood only a fraction of the *H. midae* genome was surveyed.

For the most part both the microsatellite- and SNP data provided congruent results, *e.g.* both marker types detected no differences in the levels of genetic diversity across populations (wild and cultured); both marker sets demonstrated evidence for population

bottlenecks and supported genetic differentiation amongst the wild- and cultured populations. The two marker types did, however, demonstrate contradictory results with regards to the grouping of aquaculture populations especially when loci assumed to be under selection was incorporated in the analysis: Microsatellite loci grouped aquaculture populations separately from wild populations, whilst SNP loci clustered aquaculture populations with their respective geographically correlated wild progenitor populations. In general, because microsatellite loci are multi-allelic (in comparison to SNPs that are predominantly biallelic) they demonstrate higher information content and subsequently have higher resolving power (Vignal *et al.*, 2002; Lui and Cordes, 2004). It might therefore be that the microsatellite loci are more sensitive to the recent effects of domestication than the SNP loci. Furthermore, the microsatellite loci are probably also more representative of genome-wide variation in general, because the microsatellite loci were developed from a variety of sources, including anonymous genomic fragments and expressed sequences (*e.g.* Slabbert *et al.*, 2008, 2010; Rhode *et al.*, 2012; Slabbert *et al.*, 2012). On the contrary, the SNPs used in this study were solely developed from ESTs and thus only represents genic variation of which the majority relate to peptide structural variation. Many microsatellite loci that were found to be outlier loci may in fact also represent genic variation (Rhode and Roodt-Wilding, 2011). These were, however, in the UTR or intronic sequences where they may be closely associated with gene regulatory motifs (Li *et al.*, 2004). This could explain why there is a marked clustering of cultured- vs. wild populations when using microsatellite loci putatively under directional selection: Rapid evolution is often attributed to selection on regulatory variation altering pleiotropic interaction within gene-networks (Purugganan, 2000; Barrier *et al.*, 2001; Mignon-Grasteau *et al.*, 2005; Dobney and Larson, 2006; Hoekstra and Coyne, 2007; Østman *et al.,* 2012). The domestication effect is most likely creating the selection pressure driving the divergence of culture populations from the wild populations at these loci. This evidence for gene-networks is supported by the significantly higher LD amongst microsatellite loci under selection in compsrison to syntenic LD.

This study identified a relatively high number of outlier loci (27% for microsatellites and 11% for SNPs); nonetheless it remains comparable with previous studies that report 0.4% to 26% for a variety of species (Nosil *et al.*, 2009 for a review). Both the microsatellite- and the SNP loci demonstrated a similar number of loci under directional selection (±10%), whilst

microsatellites had more loci under balancing selection than SNPs. Again this disparity might be related to the genomic locality and particular characteristics (*e.g.* bi- vs. multi-allelic) of the respective loci. The simplest explanation for outlier loci is heterogeneous genomic divergence due to selection (Nosil *et al.*, 2009; Bierne *et al.*, 2011). However various factors, other than selection, have been reported to influence heterogeneous divergence of genomic regions. Firstly, complex demographic scenarios, including co-ancestry correlations amongst subpopulations and hierarchically structured populations may increase neutral variance of genetic differentiation beyond the null distribution assumed by an outlier test (Excoffier *et al.*, 2009; Bonhomme *et al.*, 2010). This could easily be overcome by using multiple outlier tests with differing assumptions on population demography, for example BayeScan (Foll and Gaggiotti, 2008) makes no *a priori* assumption on population history and takes into account population- and allele-specific effects. Secondly, "gene surfing" (whereby the frequency of novel, neutral mutations increase in the wake of an expanding population) could mimic a signature of selection (Klopfstein *et al.* 2006; Hofer *et al.* 2009). Although there is evidence of an historic population expansion for *H. midae* (Bester-van der Merwe *et al.*, 2011), the data presented in this study suggests that the populations are in decline; and therefore "gene surfing" is improbable, but cannot be entirely excluded. Thirdly, when effective population size is small, background selection against deleterious mutations could be expected to increase population divergence and thus to inflate allelic variance at loci with differential recombination rates (Charlesworth *et al.*, 1997; Bierne *et al.*, 2002). Current data and previous work (Bester-van der Merwe *et al.*, 2011) suggests that long-term effective population size for *H. midae* is sufficiently large, which will make the aforementioned scenario unlikely; however the small short-term effective population sizes may have some influence. Lastly, Bierne *et al.* (2011) propose that "endogenous genetic barriers" (*i.e.* reproductive isolation or -incompatibilities) might restrict neutral gene flow and thus inflate neutral genetic differentiation and produce outlier loci that are not due to ecological adaptation. Estimates of genetic differentiation among putatively neutral loci show evidence of low to moderate population differentiation indicating that gene flow in general is still high amongst wild populations (Bester-van der Merwe *et al.*, 2011). Furthermore, most of the aquaculture populations (with the exception of one F_2-generation population) were directly descended from wild broodstock. Taking into account the aforementioned, heterogeneous genomic divergence in *H. midae* is most likely a

consequence of selection; however gene flow amongst populations may still persist (Nosil *et al*., 2009). The high number of outlier loci could possibly be explained by incorporating aquaculture populations in the analyses. Aquaculture populations are under selective pressures to adapt to a new environment and also artificial selection for production traits (animals go through regular grade-and-cull procedures during the production grow-out phases) and therefore, the domestication event is likely resulting in a selective sweep (Innan and Kim, 2004; Bierne, 2010; Ralph and Coop, 2010).

5.2.2. Insights into Population Dynamics in the Wild

As adults, abalone are benthic sessile animals and thus gene flow is mostly dependent on the brief pelagic larval stages, when the planktonic larvae are particularly susceptible to ocean currents by which they then disperse. Previously, Bester-van der Merwe *et al.* (2011) reported that the retroflection of the Agulhas current at Cape Agulhas on the south coast of South Africa is most likely a major barrier to gene flow, dividing the *H. midae* population into two main reproductive stocks on the west- and east coast. A less prominent, secondary barrier was also postulated at the thermal front in the Algoa Bay region that could subdivide the populations into three stocks corresponding with the known geographical, marine biomes around the South African coast. The presently presented genome-wide microsatellite- and -SNP data supports the subtle population differentiation of wild populations on the west-, south-, and east coast of South Africa as previously reported (Bester-van der Merwe *et al.*, 2011) when using loci that are presumed to be selectively neutral. When incorporating loci that are putatively under the influence of directional selection the population differentiation becomes noticeably higher; this is particularly evident when using the SNP data (Chapter 4). This geographic correlation of population structure is in accordance with the biogeographical vicariance hypothesis for the origin of abalone around the South African coast (Bester-van der Merwe *et al.*, 2012). Under this hypothesis an ancestral abalone species migrated in a south-easterly direction from what is today the Mediterranean Sea. As such, population divergence and ultimately speciation is a product of adaptation to the environmental conditions, which within the South African context relates to the three biogeographically provinces around the country's coast: cool-

temperate on the west-, warm-temperate on the south- and subtropical on the east coast (Emanuel *et al.*, 1992).

Although this study did not investigate mitochondrial genetic diversity *per se*, some conclusions could be made to explain the observed decrease in mitochondrial diversity as observed by Evans *et al.* (2004b). Evans *et al.* (2004b) maintain that the observed decrease in mitochondrial genetic diversity in east coast populations is consistent with a founder event from the west coast. However, this study [and the study by Bester-van der Merwe *et al.* (2011)] could not support the decrease in genetic diversity at nuclear loci expected if a founder event did in fact occur. Furthermore, the results of this study suggest that adaptation to local environmental conditions might play an important role in maintaining the geographically correlated population structure. This lends credence to the biogeographical vicariance hypothesis, which is in contrast to the west-east colonisation hypothesis of Evans *et al.* (2004b). In an extensive investigation among 3000 animal species Bazin *et al.* (2006) could not detect any correlation between mitochondrial genetic diversity and population size. They consequently question whether it is appropriate to deduce demographic history of a population from mitochondrial data. The authors continue to explain that the variations in mitochondrial genetic diversity is more like due to selective events and that this is particularly true for marine fauna with relatively large effective population sizes. The mitochondrion is a vital cellular organelle that functions in cellular respiration, therefore it is not surprising that is under selective constraints. Various loci associated to genes operating in the respiratory mechanism (*e.g.* NADH dehydrogenase, ATPase, Phosphoglycerate mutase) have been identified as outlier loci in this study. It is therefore comprehendible that the mitochondrial genome itself has been subject to a selective sweep reducing genetic diversity in east coast populations, particularly considering the fairly large long-term effective population size of abalone.

In recent years, natural abalone populations have come under considerable pressure due to overharvesting and changing ecological circumstances, consequently there is evidence to suggest that a recent bottleneck has occurred. Furthermore, it would seem that these populations are responding to the permutations by genetic adaptation. Nonetheless, the reproductive strategy of abalone, broadcast spawning, and high migration rates seem to be the major contributing factors in temporal genetic variation. Overlapping generations, high

fecundity and low rates of larval survival seem to generate substantial differences in the genetic constitution of populations through time, even if samples are only one generation removed, leading to comparatively small effective population sizes in the short-term. Overall, the population dynamics of abalone appears to be more complex than previously thought.

5.2.3. Population Dynamics under Aquaculture Conditions

Abalone aquaculture populations seem to differentiate from their wild progenitor populations at a rapid rate. This phenomenon has been observed for a number of abalone species (Hara and Sekino, 2007; Li *et al.*, 2007; De la Cruz *et al.*, 2010; Praipue *et al.*, 2010; An *et al.*, 2011). Irrespective of this the *H. midae* aquaculture populations seem to maintain levels of genetic diversity equivalent to the wild populations, which is consistent with previous estimates for genetic diversity in a number of abalone species, including *H. midae* (Gutierrez-Gonzalez and Perez-Enriquez, 2005; Slabbert *et al.*, 2009; An *et al.*, 2011). The domestication event is generally accompanied by a population bottleneck produced by the founder effect and will generally lead to an increase in the rate of inbreeding and substantial stochastic fluctuations in allele frequencies (Mignon-Grasteau *et al.*, 2005). Although there is evidence for population bottlenecks in the aquaculture populations, in the current study, the effective population sizes have not been reduced to a level that genetic diversity is lost. The significant fluctuation at neutral loci can thus be attributed to the differential broodstock contributions, high fecundity and low larval survival rates (Lind *et al.*, 2009; Slabbert *et al.*, 2009).

With regards to selection, the domestication effect is expected to create significant selective pressures on the aquaculture populations and is most likely resulting in a selective sweep. Evidence for this is the high number of candidate loci under selection and possible evolutionary convergence of aquaculture populations due to similar aquaculture practices and/or artificial selection (Chapter 2, 3). Loci under selection seem to be involved in regulatory processes of gene expression rather than structural variation, which is expected (see discussion above). At present there is no formal breeding programme at most abalone aquaculture facilities and production rarely extends beyond the F_1-generation. However, there is increasing interest to develop such programmes for the development of animals

that show superior production traits. It is therefore expected that as selective breeding programmes advance aquaculture populations will increasingly become reproductively isolated from the wild populations and diverge even more, effective population sizes will decrease and the rate of inbreeding will increase. The pronounced genetic effects of the domestication process can already be observed in the F_2-generation with this population failing to group with any other population (Chapter 4).

5.3. Managerial Considerations for the South African Abalone Resource

5.3.1. Preservation of Wild Populations

As an economically important marine animal the effective management of the abalone resources is of the utmost importance. *Haliotis midae* has come under considerable pressure due to overfishing and poaching and population recovery in abalone species globally has proven to be more complex than expected (Tegner, 2000). These pressures on the natural populations are now starting to reflect on the genetic constitution of these populations with evidence of population contractions and changes in selection regimes. At present the South African Marine Living Resource act (Republic of South Africa, 1998) regulates the usage and preservation of the marine *biota* within the country's coastal waters. Although an ecosystems approach to reserve management was adopted, it has been previously argued that at present the marine protected areas around South Africa might not necessarily reflect the population genetic structure of many of the marine organisms (Von der Heyden, 2009). A population genetic assessment of any population provides an understanding of the evolutionary forces that shape population/species diversity; in turn this can be used to predict the long-term trends in population viability and robustness (Moritz, 2002; Palumbi, 2003).

Historically the *H. midae* fishery was regulated by means of minimum size restrictions of harvested animals, closed seasons, annual quotas, total allowable catch (TAC) per fishing zone and reserves where harvesting is prohibited (Tarr, 2000). In order to better align managerial strategies for the abalone resource with the genetic structure of this species,

Bester-van der Merwe *et al.* (2011) proposed that as a minimum precaution the populations on the west- and east coast (with the transition point at Cape Agulhas - the major barrier to gene flow) should be managed as separate stocks. Nonetheless, noting the secondary barrier to gene flow and the propensity of South African marine fauna populations to structure in accord to the biogeographical provinces (*e.g.* Ridgway *et al.*, 1998; Teske *et al.*, 2006, 2007; Zardi *et al.*, 2007; Teske *et al*., 2008; Von der Heyden *et al.*, 2008), Bester-van der Merwe *et al.* (2011) argues that adaptive diversity could be lost if the transition zone on the south coast is not recognised as an independent management unit (Crandall *et al.*, 2000; Fraser and Bernatchez, 2001; Counterman *et al.*, 2010). To date, most endeavours in conservation genetics have not explicitly formulated strategies incorporating adaptive diversity mostly because the majority of studies are based on a small number of markers that are assumed to be neutral. With a genomics approach to conservation genetics it has become easier to identify adaptive variation at the molecular level that will certainly allow for the refinement of conservation strategies (Wenne *et al.*, 2007; Allendorf *et al*., 2010; Ouborg *et al.*, 2010). As such, the current study provides evidence that adaptation to local environments may indeed be an important determining factor for the maintenance of spatial population structure of wild populations of *H. midae.* The recognition of south coast populations as an independent management unit may thus warranted. Furthermore, the apparent temporal instabilities in genetic variation, created by the differential spawning and high larval mortalities that lead to a decrease in short-term effective population sizes may be a point of concern. Long-term effective population sizes may be sufficiently large, however if short-term effective population sizes decline, the number of new recruits will decrease with successive generations. There has already been reports of juvenile recruitment failure in a number of abalone fishing zones around South Africa (Day and Branch, 2000; Plagányi and Butterworth, 2010). The Allee effect, (the direct relationship between spawner density and successful fertilisation in sessile broadcast spawning animals, Allee *et al.*, 1949) might therefore be much more prominent than previously thought and could contribute to the slow and troublesome recovery of abalone populations seen world-wide (Tegner, 2000). Strategically placed reserves, taking into account larval dispersal capabilities, where a high density of broodstock animals could be maintained in order to maximise short-term effective population size may thus be vital for the long-term preservation of abalone (Hobday *et al.*, 2001).

5.3.2. Breeding Objectives and Implications for Commercial Stocks

The presented data suggests that the respective aquaculture facilities hold a sufficient number of breeding animals to maintain levels of genetic diversity comparable with their wild progenitor populations. However, there is clear evidence that the exact constitution of this genetic diversity is distinct from the wild populations and differs significantly between cultured populations from different facilities. It is therefore argued that there is evidence for genetically unique domesticated abalone strains produced by independent domestication events on respective aquaculture facilities. For this reason, farmers should be careful when translocating animals between farms: The production value of animals, adapted to specific conditions, may change when placed in a different environment, as a consequence of genotype by environment interactions; such interactions have been reported for salmon (Evans *et al.*, 2010), shrimp (Ibarra and Famula, 2008) and mussels (Shields *et al.*, 2008).

Considering that cultured populations mostly consist of F_1-generation animals (one F_2 population), sufficient time has not yet lapsed for noticeable decrease in genetic variation to occur. Nonetheless, with the implementation of selective breeding programmes it is anticipated that inbreeding and relatedness will increase, due to a further decrease in effective population size. At present, however, the relatively high effective population sizes and low relatedness in cultured cohorts, means that producers could select broodstock from these F_1-animals with little deleterious consequences. However, it should be done with an air of caution, noting the dramatic decrease in effective population size of the F_2 population. Inbreeding depression has been reported for the Pacific abalone (*H. discus hannai*) after only one generation of full-sib-mating, with significant decreases in the survival rate of offspring (Kobayashi and Kijima, 2010). This could have grave consequences for the profitability of abalone production. Hayes *et al.* (2006) suggested several methods for maximising genetic diversity for aquaculture selective breeding programmes, including random mating, minimising kinship and maximising heterozygosity; as genetic diversity remains vital for continued and long-term genetic gains in variable environments. This fact was often overlooked by traditional animal breeders (Notter, 1999; Ajmone-Marsan *et al.*, 2010; Groeneveld *et al.*, 2010) and subsequently concerns led to the adoption of the “global

plan of action for conserving indigenous farm animal genetic resources, FAnGR" (FAO, 2007).

The genetic distinctness of cultured populations also warrants regulations to be put in place in order to limit escapees from cultured populations into the wild. It has been demonstrated in salmon that such escapes, when interbreeding with wild animals, could lead to maladapted individuals due to outbreeding depression, potentially causing the collapse of the natural population (Naylor *et al.*, 2005). Similar findings were reported for oysters (Camara and Adopalas, 2009; Zhang *et al.*, 2010) and mussels (Jones *et al.*, 2006). This is of further importance, when considering natural stock enhancement or ranching initiatives using culture derived seed (Roodt-Wilding, 2007; Hara *et al.*, 2008). Such an initiative for red abalone (*Haliotis rufescens*) demonstrated the possible adverse effects on the genetic integrity of wild populations (Gaffney *et al.*, 1996). Furthermore, Gutierrez-Gonzalez and Perez-Enriquez (2005) found no loss of genetic diversity between cultured and wild blue abalone (*Haliotis fulgens*); however recapture of ranched animals was low. Pilot studies conducted in South Africa for *H. midae*, also demonstrated differential survival rates at various sites (Sweijd *et al.*, 1998; De Waal *et al.*, 2003), with similar reports for the Pacific abalone (Hamasaki and Kitada, 2008) and greenlip abalone (*Haliotis laevigata*) (Dixon *et al.*, 2006). There are many compounding factors that contribute to recapture success rate of ranched animals. In the past much attention was given to simply maintaining neutral genetic diversity; however consideration of adaptive diversity/potential may be an important determinant.

In summary, from a genetic management perspective, aquaculture facilities should define their long-term breeding objectives under at least one of two broad aims: 1) Implement a selective breeding programme to enhance favourable production traits. This is the traditional animal production route, but maintaining sufficient genetic diversity to ensure sustainable breeding should be a key imperative (perhaps as a national breeding objective, to enable individual farms to develop specialised strains, but conserving genetic diversity throughout a national breeding structure). Furthermore, putting in place measures to prevent interbreeding of cultured and wild animals should also be taken. 2) If the facility's focus will be on ranching or stock enhancement, standard conservation genetic practices

should be implemented to maximise survival rates of seeded animals and minimise possible adverse effects of introducing cultured animals into the wild.

5.4. Shortcomings and Perspectives on Future Undertakings

This study is, to date, the most comprehensive survey of genome-wide genetic variation within and amongst wild and cultured populations of *H. midae* and abalone in general. Nonetheless, in all likelihood only a fraction of the genome was investigated and many more markers will likely be necessary to gain a complete understanding of the evolutionary forces that shape genome-wide genetic variation within these populations. A preliminary assessment of genome-wide LD, over populations was done in order to evaluate the possible co-segregation of outlier loci (presumed to be under selection) that could indicate functional linkages within gene-networks. However, to assess the number of markers needed to provide a more comprehensive genome coverage, it is necessary to evaluate the extent of population-specific LD. This will also allow for the assessment of particular outlier loci within the specific genomic contexts of individual populations. In turn, this will aid a more thorough investigation into the history of selection events, *i.e.* historic vs. recent selection or hard- vs. soft selective sweeps (*e.g.* Karasov *et al.*, 2010), which at present can only be speculated on. The recent completion of the *H. midae* linkage map (Vervalle *et al.*, in press) makes such an investigation possible. However, not all the markers used in this study have yet been successfully mapped, whilst some markers currently on the linkage map have not been included in this study. It is therefore necessary to genotype all mapped markers in the study populations for sufficient genome coverage. In future it might also be advantageous to run a combined analysis of both SNP and microsatellite markers that could increase the precision of population genetic estimates and marker coverage across the genome (Liu *et al.*, 2005; Ryynanen *et al.*, 2007; Narum *et al.*, 2008). It may also be necessary to include more sampling sites within each of the three geographic regions to identify possible fine-scale population structure due to adaptation to cryptic local environmental conditions. Furthermore, at present the sample sizes form each population was limited. It is known that when sample sizes are small (less than 50) it may create sampling errors that could bias estimates of genetic diversity and –differentiation. This is

particularly relevant when using microsatellite markers: due to its high polymorphism a large sample is needed to accurately reflect allele frequencies for especially rare alleles. Nonetheless, sampling variance can be partially negated by standardising sample sizes (as has been done for all microsatellite analyses) and using a large number of marker loci (as in the case of a genomic approach) (Ruzzante, 1998).

Although many of the outlier loci in this study could be associated to genic regions, they remain for the most part putatively influenced by selection. If these loci are indeed under selection it is expected that they will demonstrate some functional activity (or other closely linked causal variant) and thus be associated to particular phenotypes. Additional analyses such as QTL- and association mapping or gene expression /transcriptome profiling will thus be extremely useful in confirming that the identified loci are indeed functionally active and therefore under selection (*e.g.* Stinchcombe and Hoekstra, 2008; Larsen *et al.*, 2011). During this study an extensive transcriptome resource was created and could be used to further characterise the unique properties of the respective populations with regards to gene expression. From an aquaculture perspective, many of these loci might be associated with economically important production traits and could be used in future marker assisted breeding programmes. On the other hand, if these loci are found to be associated to adaptive phenotypes for specific environmental conditions in the wild it could aid in refining long-term conservation strategies through highlighting environmental stressors that could solicit a selective pressure or by predicting how populations will adapt to various climate scenarios. This is particularly relevant in the light of global climate change, where sea temperatures are expected to increase (Roessig *et al.*, 2004; Cheung *et al*., 2009). In South Africa the residual effects of water temperature increases has already been noted on corals (Riegl, 2003). A study by Hobday and Tegner (2002) suggest that water temperature might be an important determining factor for the natural distribution of abalone populations. With regards to *H. midae* a range shift might be observed in future. Warming sea temperatures could result in a population contraction on the east coast, whilst the range might expand northwardly on the west coast of South Africa. Populations may also respond adaptively that could lead to reductions in genetic diversity at particular loci, *i.e.* produces new signatures of selection (Parmesan, 2006). It is therefore important to continue monitoring both environmental and biological (including genetic) parameters (Clark, 2006)

5.5. Concluding Remarks

The South African abalone, *Haliotis midae*, is an economically important, large gastropod marine mollusc. In recent years this species has been particularly threatened by both human activities and ecological changes. This study represents one of first attempt to quantify genome-wide genetic variation for any haliotid species and to assess the micro-evolutionary forces, especially selection that can account for the observed patterns in genetic diversity. By using a population genomics approach it was ascertained that spatial population structure of wild abalone along the South African coast seems to be stable; however the mode of reproduction, long lifespan and high gene flow leads to significant fluctuation in genetic diversity through time. The spatial stability on the other hand is probably maintained by adaptation to local environmental conditions. Furthermore, the data suggests that the recent domestication of abalone is a major selective agent driving the divergence of cultured populations from their progenitor populations in the wild. As such, a relatively large percentage of the *H. midae* genome might be under the influence of selection at present - 10% to 27%. Continued monitoring of both wild and cultured populations are essential in order to manage genetic resources in such a manner as to ensure the integrity of the wild populations and to sustainably expand the aquaculture industry.

References

Ajmone-Marsan P, The GLOBALDIV Consortium (2010) A global view of livestock biodiversity and conservation – GLOBALDIV. Anim Genet 41 (Suppl. 1): 1-5.

An HA, Wook J, Kim HC, Myeong J (2011) Genetic characterization of five hatchery populations of pacific abalone (Haliotis discus hannai) using microsatellite markers. Int J Mol Sci 12: 4836-4849.

Allee WC, Emerson AE, Park O, Partk T, Schmidt KP (1949) Principles in Animal Ecology. Saunders Publishing Company, Philadelphia, PN.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11: 697-710.

Barrier M, Robichaux RH, Purugganan MD (2001) Accelerated regulatory gene evolution in an adaptive radiation. Proc Natl Acad Sci USA 98: 10208-10213.

Bazin E, Glémin S, Galtier N (2006) Population size does not influence mitochondrial genetic diversity in animals. Science 312: 570-572.

Bester AE, Roodt-Wilding R, Whitaker HA (2008) Discovery and evaluation of single nucleotide polymorphisms (SNPs) for *Haliotis midae*: a targeted EST approach. Anim Genet 39: 321-324.

Bester-van der Merwe AE, Roodt-Wilding R, Volckaert FAM, D'Amato ME (2011) Historical isolation and hydrodynamically constrained gene flow in declining populations of the South-African abalone, *Haliotis midae*. Conserv Genet 12: 543-555.

Bester-van der Merwe AE, D'Amato ME, Swart BL, Roodt-Wilding R (2012) Molecular phylogeny of South African abalone, its origin and evolution as revealed by two genes. Mar Biol Res 8: 727-736.

Bierne N (2010) The distinctive footprints of local hitchhiking in a varied environment and global hitchhiking in a subdivided population. Evolution 64: 3254-3272.

Bierne N, Lenormand T, Bonhomme F, David P (2002) Deleterious mutations in a hybrid zone: can mutational load decrease the barrier to gene flow? Genet Res 80: 197-204.

Bierne N, Welch J, Loire E, Bonhomme FO, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. Mol Ecol 20: 2044-2072.

Black WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: Genome-wide sampling of insect populations. Annul Rev Entomol 46: 441-469.

Bonhomme M, Chevalet C, Servin B et al. (2010) Detecting selection in population trees: the Lewontin and Krakauer test extended. Genetics 186: 241-262.

Brumfield RT, Beerli P, Nickers DA, Edwards SV (2003) The utility of single nucleotide polymorphism in inferences of population history. Trends Ecol Evol 18: 249-256.

Beuzen ND, Stear MJ, Chang KC (2000) Molecular markers and their use in animal breeding. Vet J 160: 42-52.

Camara MD, Vadopalas B (2009) Genetic aspects of restoring olympia oysters and other native bivalves: balancing the need for action, good intentions, and the risks of making things worse. J Shellfish Res 28: 121-145.

Campbell D, Bernatchez L (2004) Genetic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Mol Biol Evol 21: 945-956.

Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet Res 70: 155-174.

Cheung WWL, Lam VWY, Sarmiento JL, Kearney K, Watson R, Pauly D (2009) Projecting global marine biodiversity impacts under climate change scenarios. Fish Fish 10: 235-251.

Clark BM (2006) Climate change: A looming challenge for fisheries management in southern Africa. Mar Policy 30: 84-95.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. Trends Ecol Evol 15: 290-295.

Counterman BA, Araujo-Perez F, Hines HM, Baxter SW, Morrison CM, Lindstrom DP, Papa R, Ferguson L, Joron M, French-Constant RH, Smith CP, Nielsen DM, Chen R, Jiggins CD, Reed RD, Halder G, Mallet J, McMillan WO (2010) Genomic hotspots for adaptation: The population genetics of Mullerian mimicry in *Heliconius erato*. PLOS Genet 6: e1000796.

Day E, Branch GM (2000) Evidence for a positive relationship between juvenile abalone *Haliotis midae* and the Sea urchin *Parechinus angulosus* in the South-Western Cape, South Africa. S Afr J Mar Sci 22: 145-156.

De la Cruz F-L, Del Rio Portilla M, Gallardo-Escarate C (2010) Genetic variability of cultured populations of red abalone in Chile: An approach based on heterologous microsatellites. J Shellfish Res 29: 709-715.

Department of Agriculture, Forestry and Fisheries (2012) Aquaculture Annual Report 2011, South Africa.

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

Dixon CD, Daya RW, Huchette SMH, Shepherd SA (2006) Successful seeding of hatchery-produced juvenile greenlip abalone to restore wild stocks. Fish Res 78: 179-185.

Dobney K, Larson G (2006) Genetics and animal domestication: new windows on an elusive process. J Zool 269: 261-271.

Emanuel BP, Bustamante RH, Branch GM, Eekhout S, Odendaal FJ (1992) A zoogeographic and functional approach to the selection of marine reserves on the West coast of South Africa. S Afri J Mar Sci 12: 341-354.

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

Evans BS, Sweijd NA, Bowie RCK, Cook PA, Elliott NG (2004b) Population genetics structure of the perlemoen *Haliotis midae* in South Africa: evidence of range expansion and founder events. Mar Ecol Prog Ser 270: 163-172.

Evans ML, Neff BD, Heath DD (2010) Quantitative genetic and translocation experiments reveal genotype-by-environment effects on juvenile life-history traits in two populations of Chinook salmon (*Oncorhynchus tshawytscha*). J Evol Biol 23: 687-698.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103: 285-298.

FAO (2007) The State of the World's Animal Genetic Resources for Food and Agriculture [online]. Available at: www.fao.org/docrep/010/a1250e/a1250e00.htm [accessed 15 November 2011]

Foll M, Gaggiotti OE (2008) A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. Genetics 180: 977-993.

Franchini P, Slabbert R, Van der Merwe M, Roux A, Roodt-Wilding R (2010) Karyotype and genome size estimation in *Haliotis midae*: Estimators to assist future studies on the evolutionary history of the Haliotidae. J Shellfish Res 29: 945-950.

Franchini P, Van der Merwe M, Roodt-Wilding R (2011) Transcriptome characterization of the South African abalone *Haliotis midae* using sequencing-by-synthesis. BMC Research Notes 4: 59.

Fraser DJ, Bernatchez L (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Mol Ecol 10: 2741-2752.

Gaffney PM, Rubin VP, Hedgecock D, Powers DA, Morris G, Hereford L (1996) Genetic effects of artificial propagation: signals from wild and hatchery populations of red abalone in California. Aquaculture 143: 257-266.

Groeneveld LF, Lenstra JA, Eding H, Toro MA, Scherf B, Pilling D, Negrini R, Finlay EK, Jianlin H., Groeneveld E, Weigend S, The GLOBALDIV Consortium (2010) Genetic diversity in farm animals – a review. Anim Genet 41 (Suppl. 1): 6-31.

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.

Hamasaki K, Kitada S (2008) The enhancement of abalone stocks: lessons from Japanese case studies. Fish Fish 9: 243-260.

Hara M, Sekino M (2007) Genetic differences between hatchery stocks and natural populations in Pacific Abalone (Haliotis discus) estimated using microsatellite DNA markers. Mar Biotechnol 9: 74-81.

Hara, M., Onoue, S., Taniguchi, N., 2008. Assessing the impact of releasing exogenous hatchery-reared juveniles of Pacific Abalone, *Haliotis discus*. Reviews in Fisheries Science 16, 278-284.

Hauser L, Seeb JE (2009) Advances in molecular technology and their impact on fisheries genetics. Fish Fish 9: 473-486.

Hayes B, He J, Moen T, Bennewits J (2006) Use of molecular markers to maximise diversity of founder populations for aquaculture breeding programs. Aquaculture 255: 573-578.

Hobday AJ, Tegner MJ (2002) The warm and the could: Influence temperature and fishing on local population dynamics of red abalone. CalCOFI Rep 43: 74-96.

Hobday AJ, Tegner MJ, Haaker PL (2001) Over-exploitation of a broadcast spawning marine invertebrate: Decline of the white abalone. Rev Fish Biol Fisher 10: 493-514.

Hoekstra H, Coyne JA (2007) The locus of evolution: Evo Devo and the genetics of adaptation. Evolution 61: 995-1016.

Hofer T, Ray N, Wegmann D, Excoffier L (2009) Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. Ann Hum Genet 73: 95-108.

Ibarra AM, Famula TR (2008) Genotype by environment interaction for adult body weights of shrimp *Penaeus vannamei* when grown at low and high densities. Genet Sel Evol 40: 541-551.

Innan H, Kim Y (2004) Pattern of polymorphism after strong artificial selection in a domestication event. Proc Natl Acad Sci USA 101: 10667-10672.

Jones JW, Hallerman EM, Neves RJ (2006) Genetic management guidelines for captive propagation of freshwater mussels (Unionoidea). J Shellfish Res 25: 527-535.

Larsen PF, Schultes PM, Nielsen EE (2011) Gene expression analysis for the identification of selection and local adaptation in fishes. J Fish Biol 78: 1-22.

Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: Structure, function, and evolution. Mol Biol Evol 21: 991-1007.

Li Q, Shu J, Yu R, Tian C (2007) Genetic variability of cultured populations of the Pacific abalone (Haliotis discus hannai Ino). Aquac Res 38: 981-990.

Lind CE, Evans BS, Knauer J, Taylor JJU, Jerry DR (2009) Decreased genetic diversity and a reduced effective population size in cultured silver-lipped pearl oysters (*Pinctada maxima*). Aquaculture 286: 12-19.

Liu ZJ, Cordes JF (2004) DNA marker technologies and their application in aquaculture genetics. Aquaculture 238: 1-37.

Liu N, Chen L, Wang S, Oh C, Zhao H (2005) Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure. BMC Genet 6: S26.

Liu S, Zhou Z, Lu J, Sun F, Wang S, Liu H, Jiang Y, Kucuktas H, Kaltenboeck L, Peatman E, Liu Z (2011) Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. BMC Genomics 12: 53.

Karasov T, Messer PW, Petrov DA (2010) Evidence that adaptation in *Drosophila* is not limited by mutation at single sites. PLoS Genet 6: e1000924.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. Mol Biol Evol 23: 482-490.

Kobayashi T, Kijima A (2010) Effects of inbreeding depression in the Pacific abalone *Haliotis discus hannai*. J Shellfish Res 29: 643-649.

Macey BM, Coyne VE (2005) Improved growth rate and disease resistance in farmed Haliotis midae through probiotic treatment. Aquaculture 245: 249-261.

Macey BM, Christison KW, Mouton A (2011) *Halioticida noduliformans* isolated from cultured abalone (*Haliotis midae*) in South Africa. Aquaculture 315: 187-195.

Massault C, Bovenhuis H, Haley C, De Koning D. (2008) QTL mapping designs for aquaculture. Aquaculture 285: 23-29

Meadows JRS, Chan EKF, Kijas JW (2008) Linkage disequilibrium compared between five populations of domestic sheep. BMC Genet 9: 61.

Mignon-Grasteau S, Boissy A, Bouix J, Faure J-M, Fisher AD, Hinch GN, Jensen P, Le Neindre P, Morméde P, Prunet P, Vandeputte M, Beaumont C (2005) Genetics of adaptation and domestication in livestock. Livest Prod Sci 93: 3-14.

Morin PA, McCarthy M (2007) Highly accurate SNP genotyping from historical and low-quality samples. Mol Ecol Notes 7: 937-946.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. Systematic Biol 51: 238-254.

Mouton A, Gummow B (2011) The occurrence of gut associated parasites in the South African abalone, *Haliotis midae*, in Western Cape aquaculture facilities. Aquaculture 313: 1-6.

Nadeau NJ, Jiggins CD (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. Trends Genet 26: 484-492.

Namroud M-C, Beaulieu J, Juge N, LaRoche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in the white spruce. Mol Ecol 17: 3599-3613.

Narum SR, Banks M, Beacham TD, Bellinger MR, Campbell MR, Dekoning J, Elz A, Guthrie CM, Kozfkay C, Miller KM, Moran P, Phillips R, Seeb LW, Smith CT, Warheit K, Young SF, Garza CJ (2008) Differentiating salmon populations at broad and fine geographical scales with microsatellites and single nucleotide polymorphisms. Mol Ecol 17: 3464-3477.

Naylor R, Hindar K, Fleming IA, Goldburg R, Williams S, Volpe J, Whoriskey F, Eagle J, Kelso D, Mangel M (2005) Fugitive salmon: Assessing the risks of escaped fish from net-pen aquaculture. BioScience 55: 427-437.

Nielsen EE, Hemmer-Jansen J, Larsen PF, Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. Mol Ecol 18: 3128-3150.

Notter DR (1999) The importance of genetic diversity in livestock populations of the future. Journal of Ani Sci 77: 61-69.

Nosil P, Funk DJ, Ortiz-Marrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18: 375-402.

Nunes VL, Beaumont MA, Butlin RK, Paulo OS (2011) Multiple approaches to detect outliers in a genome scan for selection in the ocellated lizards (Lacerta lepida) along an environmental gradient. Mol Ecol 20: 193-205.

Oleksiak MP (2010) Genomic approaches with natural fish populations. J Fish Biol 76: 1067-1093.

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma R, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. Trend Genet 26: 117-187.

Østman B, Hintze A, Adami C (2012) Impact of epistasis and pleiotropy on evolutionary adaptation. Proc R Soc B 279: 247-256.

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. Ecol Appl 13: 146-158.

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Annu Rev Ecol Evol Syst 37: 637-669.

Pérez-Enciso M, Ferretti L (2010) Massive parallel sequencing in animal genetics: wherefroms and wheretos. Anim Genet 41: 561-569.

Plagányi EE, Butterworth DS (2010) A spatial- and age-structured assessment model to estimate the impact of illegal fishing and ecosystem change on the South African abalone Haliotis midae resource. Afr J Mar Sci 32: 207-236.

Praipue P, Klinbunga S, Jarayabhand P (2010) Genetic diversity of wild and domesticated stocks of Thai abalone, *Haliotis asinina* (Haliotidae), analyzed by single-strand conformational polymorphism of AFLP-derived markers. Genet Mol Res 9: 1136-1152.

Purugganan MD (2000) The molecular population genetics of regulatory genes. Mol Ecol 9: 1451-1461.

Ralph P, Coop G (2010) Parallel adaptation: One or many waves of advance of an advantageous allele? Genetics 186: 647-668.

Riegl B (2003) Climate change and coral reefs: different effects in two high-latitude areas (Arabian Gulf, South Africa) Coral Reefs 22: 433-446.

Republic of South Africa (1998) Marine Living Resources Act no. 18 of 1998.

Rhode C, Roodt-Wilding R (2011) Bioinformatic survey of Haliotis midae microsatellites reveals a non-random distribution of repeat motifs. Biol Bull 221: 147-154.

Rhode C, Slabbert R, Roodt-Wilding R (2008) Microsatellite flanking regions: A SNP mine in South African abalone (*Haliotis midae*). Anim Genet 39: 329.

Rhode C, Hepple J, Jansen S, Davis T, Vervalle J, Bester-van der Merwe AE, Roodt-Wilding R (2012) A population genetic analysis of abalone domestication events in South Africa: Implications for the management of the abalone resource. Aquaculture 356-357: 235-242.

Roessig MJ, Woodley CM, Cech JJ, Hansen LJ (2004) Effects of global climate change on marine and estuarine fishes and fisheries. Rev Fish Biol Fisher 14: 251-275.

Roodt-Wilding R (2007) Abalone ranching: a review on genetic considerations. Aquaculture Research 38: 1229-1241.

Roux A, Sandenbergh L, Roodt-Wilding R (2008) Preliminary investigation to determine the cytotoxicity of various cryoprotectants on southern African abalone (*Haliotis midae*) embryos. Cryobiology 57: 308-311.

Ruzzante DE (1998) A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. Can J Fish Aquat Sci 55: 1-14.

Ryynanen HJ, Tonteri A, Vasemagi A, Primmer CR (2007)A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). J Hered 98: 692-704.

Tarr RJQ (2000) The South African abalone (*Haliotis midae*) fishery: a decade of challenges and change. Can Spec Publ Fish Aquat Sci 130: 32-40.

Tegner MJ (2000) Abalone (*Haliotis* spp.) enhancement in California: what we've learned and where we go from here. Can Spec Publ Fish Aquat Sci 130:61-71.

Teske PR, McQuaid CD, Froneman CD, Barker NP (2006) Impacts of marine biogeographic boundaries on phylogeographic patterns of three South African estuarine crustaceans. Mar Ecol Prog Ser 314: 283-293.

Teske PR, Froneman CD, Barker NP, McQuaid CD (2007) Phylogeographic structure of the caridean shrimp *Palaemon peringueyi* in South Africa: further evidence for intraspecific genetic units associated with marine biogeographic provinces. Afr J Mar Sci 29: 253-258.

Teske PR, Papadopoulos I, Newman BK, Dworschak P, McQuaid C, Barker NG (2008) Oceanic dispersal barriers, adaptation and larval retention: an interdisciplinary assessment of potential factors maintaining a phylogeographic break between sister lineages of an African prawn. BMC Evol Biol 8: 341.

Shields JL, Barnes P, Heath DD (2008) Growth and survival differences among native, introduced and hybrid blue mussels (*Mytilus* spp.): genotype, environment and interaction effects. Mar Biol 154: 919-928.

Slabbert R, Ruivo NR, Van den Berg NC, Lizamore DL, Roodt-Wilding R (2008) Isolation and characterisation of 63 microsatellite loci for the abalone, *Haliotis midae*. J World Aquacult Soc 39: 429-435.

Slabbert R, Bester AE, D'Amato ME (2009) Analysis of genetic diversity and parentage of the abalone Haliotis midae Linnaeus using microsatellite markers. J Shellfish Res 28: 1-7.

Slabbert R, Hepple J, Venter A, Nel S, Swart L, Van den Berg NC, Roodt-Wilding R (2010) Isolation and segregation of 44 microsatellite loci in the South African abalone, *Haliotis midae* L. Anim Genet 41: 332-333.

Slabbert R, Hepple J, Rhode C, Bester-van der Merwe AE, Roodt-Wilding R (2012). New microsatellite markers for the abalone *Haliotis midae* developed by 454 pyrosequencing and *in silico* analyses. Genet Mol Res 11: 2769-2779.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: Finding genes underlying ecological important traits. Heredity 100: 158-170.

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

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other molecular markers and their use in animal genetics. Genet Sel Evol 34: 275-305.

Von der Heyden S (2009) Why do we need to integrate population genetics into South African marine protected area planning? Afr J Mar Sci 31: 263-269.

Von der Heyden S, Prochazka K, Bowie RCK (2008) Significant population structure and asymmetric gene flow patterns amidst expanding populations of *Clinus cottoides* (Perciformes, Clinidae): application of molecular data to marine conservation planning in South Africa. Mol Ecol 17: 4812-4826.

Vosloo A, Vosloo D (2006) Routes of water loss in South African abalone (*Haliotis midae*) during aerial exposure. Aquaculture 261: 670-677.

Wenne R, Boubry P, Hemmer-Hansen J, Lubieniecki KP, Was A, Kuase A (2007) What role for genomics in fisheries management and aquaculture? Aquat Livingn Resour 20:241-255.

Willing E-M, Bentzen P, Van Oosterhout C, Hoffmann M, Cable J, Breden F, Weigel D, Dreyer C (2010) Genome-wide single nucleotide polymorphisms reveal populations history and adaptive divergence in the wild guppies. Mol Ecol 19: 968-984.

Zardi GI, McQuaid CD, Teske PR, Barker NP (2007) Unexpected genetic structure of indigenous (*Perna perna*) and invasive (*Mytilus galloprovincialis*) mussel populations in South Africa.Mar Ecol Prog Ser 337: 135-144.

Zhang H, Scarpa J, Hare MP (2010) Differential fertilization success between two populations of Eastern oyster, *Crassostrea virginica*. Biol Bull 219: 142-150.

Appendix A

Supplementary Information for Chapter 2

Table S2.1: Diversity estimates for genomic-microsatellites.

Table S2.2: Diversity estimates for EST-microsatellites.

Table S2.3: Number of unique alleles per locus per population.

Table S2.1: Diversity estimates for genomic-microsatellites.

Population	Locus	A_n	R_s	H_o	H_e	F_{is}	HWE *P*-value	fr_{null}	Slatkin's *P*-value
CPWC	*HmAD102T*	27	24.968	0.688	0.962	0.289	0.0**	0.133*	0.024*
	HmLCS1T	7	6.614	0.688	0.692	0.007	0.216	-0.004	0.524
	HmLCS67T	5	4.625	0.500	0.524	0.046	0.884	0.010	0.697
	HmNS6T	6	5.674	0.581	0.622	0.067	0.600	0.019	0.660
	HmRS27T	28	24.71	0.906	0.950	0.047	0.150	0.015	0.777
	HmRS36T	5	4.934	0.719	0.613	-0.176	0.264	-0.072	0.428
	HmRS80T	16	15.418	0.875	0.922	0.051	0.072	0.017	0.026*
	HmLCS72M	8	7.812	0.313	0.831	0.628	0.000**	0.278*	0.020*
	Average	**12.7500**	**11.844**	**0.659**	**0.764**	**0.200**	-	-	-
WPWC	*HmAD102T*	28	28.881	0.633	0.974	0.354	0.000**	0.166*	0.049*
	HmLCS1T	7	7.716	0.500	0.564	0.116	0.010**	0.035	0.727
	HmLCS67T	5	4.690	0.241	0.432	0.138	0.001**	0.129*	0.922
	HmNS6T	7	7.736	0.688	0.772	0.108	0.211	0.041	0.362
	HmRS27T	28	26.343	0.969	0.973	-0.003	0.757	-0.006	0.000**
	HmRS36T	4	5.733	0.700	0.662	-0.054	0.267	-0.030	0.552
	HmRS80T	19	19.588	0.969	0.941	-0.018	0.527	-0.022	0.056
	HmLCS72M	7	7.742	0.438	0.685	0.321	0.005**	0.141*	0.388
	Average	**13.125**	**13.554**	**0.642**	**0.750**	**0.120**	-	-	-
CPSC	*HmAD102T*	26	24.383	0.710	0.963	0.266	0.000**	0.122*	0.0180*
	HmLCS1T	8	7.868	0.719	0.733	0.019	0.583	0.001	0.308
	HmLCS67T	4	3.733	0.367	0.505	0.278	0.159	0.087	0.708
	HmNS6T	7	7	0.654	0.725	0.100	0.372	0.033	0.289
	HmRS27T	26	24.253	0.936	0.965	0.031	0.649	0.007	0.014*
	HmRS36T	6	5.614	0.563	0.614	0.085	0.864	0.026	0.746
	HmRS80T	18	17.042	0.875	0.915	0.045	0.190	0.017	0.107
	HmLCS72M	8	7.733	0.400	0.821	0.517	0.000**	0.225*	0.124
	Average	**12.875**	**12.203**	**0.653**	**0.780**	**0.168**	-	-	-
WPSC	*HmAD102T*	28	26.136	0.548	0.966	0.436	0.000**	0.206*	0.021*
	HmLCS1T	7	6.624	0.469	0.570	0.180	0.217	0.059	0.630
	HmLCS67T	5	4.747	0.469	0.492	0.047	0.112	0.010	0.777
	HmNS6T	7	6.792	0.759	0.767	0.010	0.064	-0.003	0.349
	HmRS27T	28	26.496	0.966	0.966	0.001	0.303	-0.008	0.153
	HmRS36T	4	3.897	0.724	0.591	-0.230	0.279	-0.091	0.360
	HmRS80T	19	17.984	0.867	0.921	0.060	0.694	0.020	0.341
	HmLCS72M	7	6.968	0.133	0.666	0.802	0.000**	0.315*	0.250
	Average	**13.125**	**12.456**	**0.617**	**0.742**	**0.163**	-	-	-
CPEC	*HmAD102T*	26	23.782	0.625	0.960	0.353	0.000**	0.164*	0.086
	HmLCS1T	8	7.736	0.625	0.736	0.153	0.026	0.058	0.397
	HmLCS67T	3	3.00	0.355	0.503	0.299	0.055	0.094	0.249
	HmNS6T	7	6.674	0.613	0.753	0.188	0.068	0.073	0.379
	HmRS27T	23	21.945	0.966	0.956	-0.010	0.136	-0.014	0.092
	HmRS36T	6	5.513	0.452	0.463	0.026	0.862	0.003	0.932

	HmRS80T	17	16.149	0.800	0.919	0.131	0.038*	0.054*	0.234
	HmLCS72M	7	6.998	0.200	0.788	0.749	0.000**	0.324*	0.029*
	Average	**12.125**	**11.475**	**0.579**	**0.760**	**0.236**	-	-	-
WPEC	*HmAD102T*	28	25.747	0.742	0.962	0.232	0.000**	0.105*	0.167
	HmLCS1T	8	7.747	0.688	0.699	0.017	0.468	0.000	0.32
	HmLCS67T	5	4.620	0.375	0.477	0.217	0.140	0.065	0.783
	HmNS6T	8	7.620	0.656	0.794	0.176	0.108	0.070	0.241
	HmRS27T	32	28.433	0.969	0.966	-0.003	0.611	-0.009	0.401
	HmRS36T	5	4.812	0.594	0.609	0.026	0.292	0.004	0.326
	HmRS80T	22	19.646	0.906	0.925	0.020	0.762	0.002	0.788
	HmLCS72M	8	7.973	0.290	0.827	0.653	0.000**	0.289*	0.022*
	Average	**14.500**	**13.325**	**0.653**	**0.782**	**0.167**	-	-	-
Over all populations	*HmAD102T*	53	27.589	0.658	0.969	0.322	0.000**	-	-
	HmLCS1T	11	7.393	0.616	0.682	0.097	0.014*	-	-
	HmLCS67T	8	4.350	0.387	0.488	0.207	0.000**	-	-
	HmNS6T	10	7.553	0.658	0.747	0.120	0.001**	-	-
	HmRS27T	52	26.673	0.951	0.965	0.014	0.705	-	-
	HmRS36T	10	5.105	0.624	0.625	0.003	0.736	-	-
	HmRS80T	38	18.320	0.883	0.928	0.049	0.150	-	-
	HmLCS72M	9	7.904	0.297	0.798	0.628	0.000**	-	-
	Average	**23.875**	**13.111**	**0.634**	**0.775**	**0.180**	-	-	-

*A_n: Number of observed alleles; R_s: Alellic richness; H_o: Observed heterozygosity; H_e: Expected heterozygosity; F_{is}: Inbreeding coefficient; HWE: Hardy-Weinberg Equilibruim; fr_{null}: Null allele frequency. *statistical significance at the 5% nominal level; ** statistical significance at the 1% nominal level.*

Table S2.2: Diversity estimates for EST-microsatellites.

Population	Locus	A_n	R_s	H_o	H_e	F_{is}	HWE *P*-value	fr_{null}	Slatkin's *P*-value
CPWC	*HmidILL-128551T*	4	3.737	0.438	0.369	-0.191	0.807	-0.055	0.703
	HmidILL-140858T	5	4.749	0.563	0.538	-0.046	0.650	-0.021	0.448
	HmidILL-002192T	2	2.000	0.063	0.347	0.822	0.000**	0.208*	0.287
	HmidILL-047613T	17	14.984	0.875	0.852	-0.027	0.629	-0.020	0.823
	HmidILL-006622T	3	2.750	0.344	0.335	-0.027	1.000	-0.011	0.673
	HmidILL-071359P	4	3.997	0.594	0.550	-0.081	0.651	-0.034	0.162
	HmidILL-084787T	7	6.937	0.656	0.821	0.237	0.012*	0.084*	0.020*
	HmidILL-087955T	9	8.627	0.969	0.814	-0.194	0.152	-0.093	0.144
	Average	**6.375**	**5.973**	**0.563**	**0.578**	**0.0618**	-	-	-
WPWC	*HmidILL-128551T*	5	4.250	0.500	0.402	-0.069	0.675	-0.075	0.929
	HmidILL-140858T	5	4.794	0.533	0.468	-0.072	0.862	-0.050	0.575
	HmidILL-002192T	4	3.737	0.219	0.470	0.224	0.001**	0.167*	0.619
	HmidILL-047613T	14	13.293	0.906	0.901	-0.031	0.931	-0.010	0.023*
	HmidILL-006622T	3	2.750	0.250	0.388	0.202	0.066	0.096	0.632
	HmidILL-071359P	5	4.736	0.563	0.500	-0.076	0.651	-0.049	0.545
	HmidILL-084787T	8	7.720	0.625	0.777	0.148	0.007**	0.079*	0.121
	HmidILL-087955T	10	9.484	0.969	0.871	-0.059	0.003**	-0.060	0.040*
	Average	**6.750**	**6.346**	**0.571**	**0.597**	**0.034**	-	-	-
CPSC	*HmidILL-128551T*	4	3.500	0.594	0.442	-0.353	0.129	-0.111	0.835
	HmidILL-140858T	6	6.000	0.500	0.637	0.219	0.089	0.076	0.507
	HmidILL-002192T	2	2.000	0.094	0.246	0.622	0.005**	0.119*	0.373
	HmidILL-047613T	14	13.031	0.906	0.884	-0.026	0.971	-0.019	0.168
	HmidILL-	2	2.000	0.375	0.310	-0.216	0.557	-0.054	0.323

	006622T								
	HmidILL-071359P	5	4.765	0.710	0.694	-0.023	0.009**	-0.016	0.295
	HmidILL-084787T	7	6.935	0.813	0.798	-0.019	0.203	-0.015	0.038*
	HmidILL-087955T	7	6.631	0.656	0.748	0.124	0.238	0.046	0.323
	Average	**5.875**	**5.608**	**0.581**	**0.595**	**0.041**	-	-	-
WPSC	*HmidILL-128551T*	2	2.00	0.226	0.204	-0.111	1.000	-0.021	0.456
	HmidILL-140858T	7	6.488	0.500	0.448	-0.118	0.699	-0.041	0.917
	HmidILL-002192T	4	3.655	0.207	0.250	0.174	0.424	0.031	0.922
	HmidILL-047613T	15	13.906	0.862	0.870	0.009	0.628	-0.004	0.584
	HmidILL-006622T	3	2.973	0.345	0.424	0.188	0.024*	0.050	0.416
	HmidILL-071359P	5	4.539	0.516	0.457	-0.133	0.911	-0.046	0.805
	HmidILL-084787T	7	6.827	0.483	0.816	0.4123	0.000**	0.177*	0.064
	HmidILL-087955T	9	8.356	0.484	0.747	0.356	0.000**	0.145*	0.661
	Average	**6.500**	**6.093**	**0.453**	**0.527**	**0.097**	-	-	-
CPEC	*HmidILL-128551T*	5	4.690	0.686	0.530	-0.305	0.385	-0.100	0.547
	HmidILL-140858T	6	5.537	0.419	0.457	0.083	0.556	0.021	0.798
	HmidILL-002192T	4	3.548	0.065	0.405	0.843	0.000**	0.239*	0.851
	HmidILL-047613T	14	13.177	0.936	0.912	-0.026	0.302	-0.020	0.044
	HmidILL-006622T	2	2.000	0.194	0.229	0.155	0.402	0.026	0.431
	HmidILL-071359P	5	4.927	0.781	0.692	-0.131	0.183	-0.060	0.192
	HmidILL-084787T	7	6.865	0.688	0.718	0.043	0.054	0.011	0.211
	HmidILL-087955T	8	7.731	0.750	0.786	0.046	0.568	0.013	0.092
	Average	**6.375**	**6.059**	**0.565**	**0.591**	**0.088**	-	-	-
WPEC	*HmidILL-128551T*	3	2.952	0.355	0.306	-0.162	1.000	-0.041	0.494
	HmidILL-140858T	7	6.356	0.467	0.447	-0.044	0.367	-0.019	0.946
	HmidILL-002192T	2	2.000	0.226	0.204	-0.111	1.000	-0.021	0.462

	HmidILL-047613T	16	14.883	0.839	0.890	0.059	0.041*	0.020	0.221
	HmidILL-006622T	3	2.952	0.484	0.450	-0.041	1.000	-0.029	0.384
	HmidILL-071359P	6	5.688	0.533	0.566	0.058	0.527	0.015	0.768
	HmidILL-084787T	8	7.726	0.774	0.823	0.060	0.015*	0.020	0.062
	HmidILL-087955T	11	10.761	0.724	0.870	0.170	0.021*	0.070*	0.037*
	Average	**7**	**6.665**	**0.550**	**0.569**	**-0.001**	-	-	-
Over all populations	*HmidILL-128551T*	7	3.826	0.468	0.381	-0.229	0.044*	-	-
	HmidILL-140858T	13	7.156	0.497	0.685	0.275	0.000**	-	-
	HmidILL-002192T	5	2.868	0.144	0.419	0.656	0.000**	-	-
	HmidILL-047613T	22	14.585	0.888	0.899	0.012	0.628	-	-
	HmidILL-006622T	3	2.564	0.332	0.356	0.069	0.040*	-	-
	HmidILL-071359P	8	4.930	0.617	0.592	-0.043	0.402	-	-
	HmidILL-084787T	10	7.688	0.676	0.812	0.169	0.000**	-	-
	HmidILL-087955T	11	9.830	0.761	0.869	0.125	0.000**	-	-
	Average	**9.875**	**6.681**	**0.548**	**0.627**	**0.129**	-	-	-

*A_n: Number of observed alleles; R_s: Alellic richness; H_o: Observed heterozygosity; H_e: Expected heterozygosity; F_{is}: Inbreeding coefficient; HWE: Hardy-Weinberg Equilibruim; fr_{null}: Null allele frequency. *statistical significance at the 5% nominal level; ** statistical significance at the 1% nominal level.*

Table S2.3: Number of unique alleles per locus per population.

	CPWC	WPWC	CPSC	WPSC	CPEC	WPEC
HmAD102T	2	6	2	2	2	1
HmLCS1T	-	1	-	1	-	-
HmLCS67T	-	-	-	1	-	1
HmNS6T	-	-	-	1	-	-
HmRS27T	3	4	2	2	-	4
HmRS36T	-	1	2	-	1	1
HmRS80T	1	3	2	4	-	5
HmLCS72M	-	-	1	-	-	-
HmidILL-128551T	-	-	1	-	1	-
HmidILL-140858T	-	-	1	-	1	2
HmidILL-002192T	-	-	-	1	-	-
HmidILL-047613T	-	-	-	1	-	-
HmidILL-006622T	-	-	-	-	-	-
HmidILL-071359P	-	-	-	-	-	-
HmidILL-084787T	-	-	-	-	-	-
HmidILL-087955T	-	-	-	-	-	-

Appendix B

Supplementary Information for Chapter 3

Table S3.1: Candidate loci for selection as identified by the various Fst-outlier methods and the Ewens-Watterson test. Candidates for balancing selection identified by individual programs are underlined, with the final candidate loci (conforming to the set criteria) marked as "B". Candidates for directional selection identified by individual programs are highlighted in bold with final candidate loci (conforming to the set criteria) marked as "D".

Table S3.2: Hardy-Weinberg statistics per locus per population and reference to the marker information.

Table S3.3: Pairwise linkage disequilibrium statistics for syntenic markers across Wild and Cultured populations.

Table S3.4: Linkage disequilibrium (based on D' and $\chi'2$) estimates for candidate locus pairs under selection, with significance tested by means of 1000 simulations. For loci mapped to the *H. midae* linkage map (Vervalle et al., in press), linkage group allocation is given in parenthesis.

Table S3.5: Distance-based association analysis of candidate loci under selection with domestication and particular population.

Table S3.1: Candidate loci for selection as identified by the various F_{st}-outlier methods and the Ewens-Watterson test. Candidates for balancing selection identified by individual programs are underlined, with the final candidate loci (conforming to the set criteria) marked as "B". Candidates for directional selection identified by individual programs are highlighted in bold with final candidate loci (conforming to the set criteria) marked as "D".

Locus	Across all populations			Across cultured populations			Across wild populations			Hierarchical Analysis (2 Groups)		Final Candidate Loci
	Lositan	BayScan	EWH	Lositan	BayScan	EWH	Lositan	BayScan	EWH			
	P-value*	log10(PO)	*P*-value	*P*-value*	log10(PO)	*P*-value	*P*-value*	log10(PO)	*P*-value	Fst *P*-value	Fct *P*-value	
HmLCS72	ns	ns	ns	Ns	ns	0.02	ns	ns	ns	ns	ns	
HmLCS47	0.04	ns	ns	0.04	ns	ns	ns	ns	ns	ns	ns	
HmLCS48	**1.00**	**1000.00**	ns	Ns	ns	ns	ns	ns	ns	**0.00**	**0.00**	D
HmLCS63	ns	ns	ns	Ns	ns	**0.99**	ns	ns	ns	0.02	ns	
HmLCS7	ns	ns	**1.00**	Ns	ns	**1.00**	ns	ns	**1.00**	ns	0.01	
HmRS38	ns	ns	**1.00**	Ns	ns	**0.99**	ns	ns	**0.99**	ns	ns	
HmRS83	ns	ns	ns	Ns	ns	ns	ns	ns	ns	ns	0.01	
HmIF33	**1.00**	ns	ns	**1.00**	ns	0.01	ns	ns	ns	ns	ns	
HmLCS18	0.00	ns	ns	0.01	ns	ns	ns	ns	ns	ns	ns	
HmLCS37	0.00	ns	0.02	Ns	ns	ns	ns	ns	0.02	ns	ns	B
HmLCS5	**1.00**	**1000.00**	**0.99**	**1.00**	**1000.00**	ns	ns	ns	**0.99**	ns	ns	D
HmRS117	0.00	ns	ns	Ns	ns	ns	ns	ns	ns	0.05	ns	B
HmAD102	0.00	3.40	0.00	0.05	1.29	0.00	ns	2.70	0.01	0.04	ns	B
HmRS36	ns	ns	ns	Ns	ns	ns	**0.99**	ns	ns	ns	0.03	
HmG16	ns	ns	0.01	Ns	ns	0.01	ns	ns	ns	ns	ns	
HmLCS73	ns	ns	0.00	Ns	ns	ns	**0.98**	ns	0.00	ns	ns	
HmRS37	ns	ns	**0.98**	Ns	ns	ns	ns	ns	ns	ns	ns	

HmRS27	0.00	1000.00	0.01	0.02	1.50	0.01	0.01	1000.00	0.02	0.02	0.00	B
HmRS80	0.00	1.08	ns	Ns	ns	ns	ns	ns	ns	0.04	0.05	B
HmD55	ns	ns	ns	Ns	ns	ns	ns	ns	ns	ns	0.04	
HmD59	0.01	-0.61	0.01	0.05	ns	0.00	ns	ns	ns	ns	ns	B
HmRS129	0.00	1.61	0.02	0.02	2.40	ns	ns	ns	ns	ns	ns	B
HmidPS1.870	0.00	-0.71	ns	Ns	ns	ns	ns	ns	ns	0.05	ns	B
HmidNR120	0.00	2.74	0.00	0.00	2.01	0.00	ns	1.06	0.00	0.05	ns	B
HmidPS1.305	ns	ns	ns	Ns	ns	0.01	ns	ns	ns	ns	ns	
HmidPS1.818	ns	ns	0.02	Ns	ns	0.01	ns	ns	ns	ns	0.05	
HmDL34b	0.02	ns	ns	Ns	ns	ns	ns	ns	ns	ns	ns	
HmG53	0.00	1000.00	0.00	0.02	1000.00	0.00	0.01	1000.00	0.00	0.03	ns	B
HmRS62	0.00	ns	ns	0.04	ns	ns	ns	ns	ns	0.03	0.03	
HmLCS55	ns	**1.03**	ns	Ns	ns	ns	**0.96**	ns	ns	ns	ns	
HmNR191	ns	ns	ns	Ns	ns	ns	**1.00**	ns	ns	ns	ns	
HmNR106	**1.00**	**1000.00**	ns	Ns	ns	ns	**1.00**	**1000.00**	ns	**0.00**	**0.00**	D
HmNR185	0.00	ns	ns	0.03	ns	ns	ns	ns	ns	ns	ns	
HmNR258	ns	ns	ns	Ns	ns	0.01	ns	ns	ns	ns	0.01	
HmDL207	**0.97**	ns	ns	Ns	ns	ns	ns	ns	ns	ns	ns	
HmDL214	ns	ns	**1.00**	Ns	ns	1.00	ns	ns	**1.00**	0.04	ns	
HmDL50	0.00	ns	0.02	Ns	ns	ns	ns	ns	ns	ns	ns	B
HmNR224	0.00	3.40	0.00	0.01	3.22	0.00	ns	ns	0.01	0.03	ns	B
HmNR54	0.00	ns	ns	0.04	ns	ns	0.02	ns	ns	ns	ns	
HmNR180	ns	ns	ns	0.01	ns	ns	**1.00**	ns	ns	ns	ns	
HmNR20	0.02	ns	ns	Ns	ns	ns	ns	ns	ns	ns	0.04	
HmNS17b	ns	ns	0.01	Ns	ns	0.02	ns	ns	0.02	ns	0.01	B
HmNS56D	ns	ns	1.00	Ns	ns	**0.99**	ns	ns	**1.00**	ns	ns	
HmNS14	ns	ns	ns	Ns	ns	**0.98**	ns	ns	ns	ns	0.00	
HmNS19	0.01	ns	0.00	0.02	ns	0.00	ns	ns	ns	ns	0.04	B

HmNS31	ns	ns	ns	Ns	ns	ns	**0.99**	ns	ns	ns	ns	
HmNSS1H	**1.00**	**3.70**	ns	Ns	ns	**0.98**	ns	ns	ns	**0.00**	**0.00**	D
HmNR136	ns	ns	1.00	Ns	ns	**1.00**	ns	ns	**1.00**	ns	ns	
HmNR281	ns	ns	1.00	Ns	ns	ns	ns	ns	ns	ns	ns	
HmNS18	**1.00**	**1000.00**	ns	Ns	ns	ns	ns	ns	ns	**0.00**	**0.00**	D
HmNS58	0.00	ns	ns	0.01	ns	0.02	ns	ns	ns	0.01	ns	B
HmidlLL-140858	**1.00**	**1000.00**	ns	Ns	ns	ns	ns	ns	ns	**0.00**	**0.00**	D
HmidlLL-2192	**1.00**	**1000.00**	ns	**1.00**	**1000.00**	ns	ns	ns	ns	**0.00**	ns	D
HmidlLL-47613	ns	ns	ns	Ns	ns	ns	**0.96**	ns	ns	ns	ns	
HmidlLL-84787	ns	ns	ns	Ns	ns	ns	**0.98**	ns	ns	ns	0.01	
HmidlLL-87955	**1.00**	**1000.00**	ns	Ns	ns	ns	**1.00**	ns	0.02	**0.00**	**0.00**	D
HmidlLL-118779	0.03	ns	ns	Ns	ns	ns	ns	ns	ns	ns	0.01	B
HmidlLL-70036	0.01	ns	ns	Ns	ns	ns	ns	ns	ns	0.04	ns	B
HmidlLL-76149	0.02	ns	ns	Ns	ns	ns	ns	ns	ns	0.00	ns	B
HmidlLL-39227	ns	ns	ns	Ns	ns	ns	**0.95**	ns	ns	ns	0.04	
HmidlLL-126949	0.02	ns	ns	Ns	ns	ns	ns	ns	ns	ns	ns	
HmidlLL-128607	ns	ns	ns	Ns	ns	ns	ns	ns	ns	ns	0.03	
HmidlLL-112066	ns	**1.21**	ns	Ns	ns	ns	ns	ns	ns	ns	ns	
HmidlLL-98293	ns	ns	ns	Ns	ns	ns	**0.97**	ns	ns	ns	ns	
HmidlLL-87955	ns	ns	0.00	Ns	ns	ns	ns	ns	ns	ns	ns	
HmidlLL-8738	ns	ns	ns	Ns	ns	ns	**0.95**	ns	ns	ns	ns	
HdhSSR60b	ns	ns	ns	Ns	ns	ns	ns	ns	ns	0.01	ns	
HmidPS1.147	0.00	ns	0.01	0.00	ns	ns	ns	ns	0.00	0.04	ns	B
HmidPS1.559	**1.00**	**1000.00**	**0.99**	**1.00**	ns	ns	ns	ns	**1.00**	**0.00**	**0.00**	D
Hmid310	0.03	ns	ns	0.00	ns	ns	ns	ns	ns	ns	ns	
Hmid563	0.00	ns	ns	Ns	ns	ns	ns	ns	ns	0.05	ns	B

HmNS56	0.05	ns	**1.00**	Ns	ns	**0.99**	ns	ns	**1.00**	ns	0.02	B
HmidPS1.629	0.01	ns	ns	Ns	ns	ns	ns	ns	ns	0.04	0.00	B
HmidPS1.247	0.00	ns	ns	Ns	ns	ns	ns	ns	ns	ns	ns	
HmDL110	ns	ns	ns	**0.99**	ns	ns	ns	ns	ns	ns	ns	
HmLCS388	0.00	ns	ns	0.00	ns	ns	ns	ns	ns	ns	ns	
HmidPS1.1012	ns	ns	ns	**0.96**	ns	ns	ns	ns	0.00	ns	ns	
HmidPS1.228	ns	ns	ns	Ns	ns	ns	ns	ns	ns	0.04	ns	
HmD61	0.00	ns	ns	Ns	ns	ns	ns	ns	ns	0.00	ns	B
Hmid007	**1.00**	ns	ns	**1.00**	ns	ns	**1.00**	ns	ns	ns	0.03	
Hmid553	ns	ns	ns	Ns	ns	ns	**0.98**	ns	ns	ns	ns	
Hmid610	0.02	ns	ns	Ns	ns	ns	ns	ns	ns	0.01	ns	B
Hmid321	ns	ns	ns	Ns	ns	ns	**0.97**	ns	ns	ns	0.04	
Hmid36	ns	ns	ns	Ns	ns	ns	ns	ns	ns	0.04	ns	
HmidPS1.379	ns	ns	**0.99**	Ns	ns	ns	ns	**1.44**	ns	ns	ns	
HmidPS1.561	**0.99**	**2.15**	**1.00**	**1.00**	**2.47**	**0.98**	0.00	**1.85**	ns	**0.00**	ns	D
HmidPS1.859	ns	ns	ns	Ns	ns	ns	**0.97**	ns	ns	ns	ns	
HmidPS1.549	0.01	ns	ns	Ns	ns	ns	ns	ns	ns	0.03	ns	B
HmidPS1.874	0.00	1.83	ns	0.01	2.12	ns	ns	ns	0.01	ns	ns	B
Hmid65	0.00	ns	0.00	Ns	ns	ns	ns	ns	0.00	ns	0.04	B
HmidPS1.227	ns	ns	ns	0.04	ns	ns	ns	ns	ns	ns	ns	
HmidPS1.692	ns	ns	**0.99**	Ns	ns	**0.98**	ns	ns	ns	ns	ns	
HmidPS1.840	ns	ns	ns	Ns	ns	ns	**1.00**	ns	ns	ns	ns	
HmidILL-140027	ns	ns	**1.00**	Ns	ns	ns	ns	ns	**1.00**	ns	0.00	
HmidILL-72605	ns	ns	ns	Ns	ns	ns	ns	ns	**0.99**	ns	ns	
HmidILL-37506	ns	**1.84**	ns	Ns	ns	ns	ns	ns	**1.00**	ns	ns	D
HmidILL-146360	ns	**1.08**	**1.00**	Ns	ns	ns	**0.99**	**1.20**	ns	ns	ns	D

HmidILL-88398	ns	**1.66**	ns	Ns	ns	ns	**0.99**	**1000.00**	ns	ns	ns	D
HmidILL-64192	**0.99**	**2.01**	ns	Ns	ns	ns	**1.00**	ns	ns	**0.00**	ns	D
Total for directional selection	14	16	12	7	3	10	20	5	11	10	7	**14**
Total for balancing selection	34	10	19	19	7	14	3	4	13	23	23	**27**

Bold: Statistically significant for directional selection at the 5% nominal level.

Underlined: Statistically significant for balancing selection at the 5% nominal level.

ns: Not significant.

**Simulated F_{st} < Sample F_{st}*

Table S3.2: Hardy-Weinberg statistics per locus per population and reference to the marker information.

Candidate Neutral Loci													
Locus	**CPWC**		**CPSC**		**WPEC**		**WPWC**		**CPEC**		**WPSC**		**Marker Reference**
	***P*-value**	**F_{is}**	***P*-value**	**Fis**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	
HmLCS67	0.88	0.05	0.16	0.28	0.14	0.22	0.00	0.45	0.05	0.30	0.12	0.05	Slabbert *et al*., 2008
HmLCS72	0.00	0.63	0.00	0.52	0.00	0.65	0.00	0.36	0.00	0.75	0.00	0.80	Slabbert *et al*., 2008
HmLCS47	0.02	0.12	0.93	-0.12	0.05	0.19	0.85	0.08	0.86	-0.01	0.11	0.19	Slabbert *et al*., 2008
HmLCS63	N/A	N/A	0.03	0.38	1.00	-0.02	N/A	N/A	0.02	0.50	0.05	0.66	Slabbert *et al*., 2008
HmLCS9	0.02	0.22	0.01	0.36	0.00	0.58	0.00	0.48	0.01	0.34	0.00	0.51	Slabbert *et al*., 2008
HmLCS7	0.00	0.53	0.05	0.16	0.16	0.22	0.00	0.36	0.09	0.24	0.03	0.25	Slabbert *et al*., 2008
HmRS38	0.24	0.09	0.47	0.02	1.00	-0.03	1.00	-0.04	0.17	0.22	0.10	0.25	Slabbert *et al*., 2008
HmRS83	0.00	0.21	0.10	0.09	0.00	0.25	0.00	0.22	0.06	0.00	0.00	0.49	Slabbert *et al*., 2008
HmRS88	0.01	0.33	0.08	0.23	0.06	0.24	0.24	0.16	0.03	0.24	0.75	0.03	Slabbert *et al*., 2008
HmIF33	0.01	0.75	0.01	1.00	0.00	0.21	0.00	0.27	0.01	0.75	0.00	0.34	Slabbert *et al*., 2008
HmLCS18	0.02	0.40	N/A	N/A	0.99	-0.07	1.00	-0.09	0.55	0.08	0.78	-0.06	Slabbert *et al*., 2008

HmRS90	1.00	0.12	0.03	0.75	0.05	0.36	0.01	0.45	0.22	0.49	0.04	0.40	Slabbert *et al*., 2008
HmLCS1	0.22	0.01	0.57	0.02	0.46	0.02	0.01	0.12	0.02	0.15	0.22	0.18	Slabbert *et al*., 2008
HmRS36	0.27	-0.18	0.86	0.09	0.29	0.03	0.27	-0.06	0.86	0.03	0.28	-0.23	Slabbert *et al*., 2008
HmG16	0.76	-0.05	0.31	-0.07	0.04	-0.01	0.06	-0.24	0.47	0.04	0.02	0.12	Slabbert *et al*., 2008
HmLCS73	0.00	0.26	0.00	0.41	0.00	0.74	0.00	0.62	0.00	0.37	0.00	0.70	Slabbert *et al*., 2008
HmRS37	0.00	0.71	0.00	0.79	0.00	0.50	0.00	0.70	0.00	0.46	0.00	0.59	Slabbert *et al*., 2008
HmD55	0.00	0.34	0.01	0.29	0.06	0.19	0.03	0.23	0.11	0.19	0.16	0.22	Bester *et al*., 2004
HmidPS1.967	0.03	-0.27	0.26	-0.14	0.23	0.04	0.11	-0.01	0.20	-0.10	0.38	-0.01	Slabbert *et al*., 2012
HmidPS1.305	0.01	0.21	0.15	0.10	0.30	0.12	0.12	0.05	0.10	0.21	0.31	0.05	Slabbert *et al*., 2012
HmidPS1.818	0.25	0.07	0.04	0.01	0.26	0.12	0.48	0.05	0.04	0.27	0.03	0.17	Slabbert *et al*., 2012
HmDL34b	0.00	0.29	0.00	0.33	0.08	0.15	0.00	0.39	0.00	0.44	0.00	0.44	Slabbert *et al*., 2008
HmRS62D	0.02	0.19	0.04	0.14	0.00	0.28	0.00	0.26	0.00	0.23	0.01	0.22	Slabbert *et al*., 2008
HmLCS55	0.00	0.54	0.10	0.17	0.00	0.44	0.00	0.52	0.00	0.69	0.02	0.42	Slabbert *et al*., 2008
HmNR191	0.00	0.33	0.00	0.31	0.00	0.45	0.00	0.17	0.00	0.25	0.15	0.16	Slabbert *et al*., 2010
HmNR258	0.22	-0.10	0.45	0.20	0.46	-0.01	0.16	0.07	0.81	-0.01	0.36	0.12	Slabbert *et al*., 2010
HmNR289	1.00	-0.01	1.00	-0.07	0.30	0.23	0.43	0.16	0.49	0.12	0.02	0.66	Slabbert *et al*., 2010
HmDL207	0.00	0.45	0.00	0.33	0.00	0.49	0.01	0.28	0.00	0.44	0.00	0.37	Slabbert *et al*., 2010
HmDL214	0.12	0.13	0.57	0.07	0.00	0.19	0.22	-0.03	0.09	-0.10	0.43	-0.05	Slabbert *et al*., 2010

HmNR54	0.30	-0.02	0.29	-0.12	0.45	0.00	0.51	-0.03	0.60	-0.03	0.18	-0.05	Slabbert *et al*., 2010
HmLCS67	0.88	0.05	0.16	0.28	0.14	0.22	0.00	0.45	0.05	0.30	0.11	0.05	Slabbert *et al*., 2010
HmNS6	0.60	0.07	0.38	0.10	0.10	0.18	0.21	0.11	0.07	0.19	0.06	0.01	Slabbert *et al*., 2010
HmNR180	0.00	0.48	0.00	0.46	0.00	0.69	0.00	0.62	0.00	0.60	0.00	0.63	Slabbert *et al*., 2010
HmNR20	0.01	0.18	0.94	-0.07	0.02	0.11	0.60	-0.01	0.00	0.24	0.50	0.02	Slabbert *et al*., 2010
HmNS56	0.34	0.04	0.09	0.23	0.96	-0.05	0.23	0.08	0.44	0.12	0.20	0.13	Slabbert *et al*., 2010
HmNS14	1.00	-0.01	1.00	-0.03	1.00	-0.02	1.00	-0.06	1.00	-0.03	1.00	-0.06	Slabbert *et al*., 2010
HmNS31	0.00	0.73	0.02	0.28	0.00	0.27	0.00	0.63	0.00	0.57	0.00	0.29	Slabbert *et al*., 2010
HmNR136	0.52	0.02	0.33	0.09	0.97	-0.13	0.12	0.07	0.78	0.07	0.16	0.15	Slabbert *et al*., 2010
HmNR281	0.00	0.49	0.03	0.22	0.00	0.36	0.00	0.41	0.00	0.41	0.00	0.48	Slabbert *et al*., 2010
HmidILL-128551	0.81	-0.19	0.13	-0.35	1.00	-0.16	0.68	-0.25	0.39	-0.31	1.00	-0.11	Rhode *et al*., 2012
HmidILL-47613	0.61	-0.03	0.97	-0.03	0.03	0.06	0.93	-0.01	0.29	-0.03	0.59	0.01	Rhode *et al*., 2012
HmidILL-6622	1.00	-0.03	0.56	-0.22	1.00	-0.08	0.07	0.36	0.40	0.15	0.02	0.19	Rhode *et al*., 2012
HmidILL-71359	0.65	-0.08	0.01	-0.02	0.53	0.06	0.65	-0.13	0.19	-0.13	0.91	-0.13	Rhode *et al*., 2012
HmidILL-84787	0.01	0.20	0.20	-0.02	0.01	0.06	0.01	0.20	0.05	0.04	0.00	0.41	Rhode *et al*., 2012
HmidILL1-140027	0.71	0.07	0.72	-0.10	1.00	0.00	0.15	0.31	0.71	-0.09	0.93	0.04	Rhode *et al*., 2012
HmidILL-38396	1.00	-0.02	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Rhode *et al*., 2012

HmidILL-64307	0.04	-0.24	0.04	-0.23	0.29	0.03	0.13	-0.20	0.02	0.17	0.25	-0.03	Rhode *et al*., 2012
HmidILL-29450	1.00	-0.08	1.00	-0.06	0.31	0.23	0.31	0.21	0.44	0.16	1.00	-0.10	Rhode *et al*., 2012
HmidILL-39227	0.60	0.12	0.56	-0.19	0.08	-0.35	0.02	0.53	0.58	0.14	0.71	0.13	Rhode *et al*., 2012
HmidILL-60863	1.00	0.03	0.04	0.26	1.00	-0.08	1.00	-0.02	0.73	-0.15	0.63	-0.16	Rhode *et al*., 2012
HmidILL-64121	1.00	0.03	0.67	0.12	0.61	-0.05	0.07	0.19	0.31	0.18	0.15	0.25	Rhode *et al*., 2012
HmidILL-6458	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Rhode *et al*., 2012
HmidILL-66010	0.39	-0.19	1.00	-0.12	0.03	0.06	1.00	-0.13	1.00	-0.12	1.00	-0.05	Rhode *et al*., 2012
HmidILL-97931a	0.14	-0.33	0.03	-0.45	0.30	-0.28	0.03	-0.40	0.01	-0.56	0.01	-0.48	Rhode *et al*., 2012
HmidILL-126949R	0.25	0.09	0.62	0.03	0.69	-0.05	0.10	0.00	0.23	0.05	0.06	0.14	Rhode *et al*., 2012
HmidILL-128607	0.04	0.29	1.00	-0.09	1.00	0.00	0.04	0.16	1.00	-0.13	1.00	-0.18	Rhode *et al*., 2012
HmidILL-112066	1.00	-0.10	1.00	-0.11	1.00	-0.17	0.00	0.36	1.00	-0.04	0.19	-0.01	Rhode *et al*., 2012
HmidILL-98293	0.68	0.02	0.12	0.06	0.36	0.07	0.49	-0.06	0.00	0.34	0.93	-0.09	Rhode *et al*., 2012
HmidILL-87955	0.90	-0.01	0.61	0.16	0.49	0.15	0.00	0.23	0.42	0.07	0.46	0.15	Rhode *et al*., 2012
HmidILL-8738	0.04	-0.17	0.37	-0.06	0.27	0.17	0.00	0.24	0.41	0.03	0.22	-0.16	Rhode *et al*., 2012
HaSSRgd842	0.40	-0.13	0.04	0.34	0.34	0.11	0.09	-0.08	0.66	-0.03	0.45	-0.16	Rhode, 2010
HdSSRex495	0.00	0.42	0.00	0.51	0.02	0.09	0.05	0.06	0.01	0.17	0.00	0.28	Rhode, 2010
HmSSRex489b	0.94	-0.07	1.00	0.03	0.87	-0.08	0.74	0.09	0.82	-0.15	0.01	0.12	Rhode, 2010

HaSSRdw239	0.00	0.37	0.00	0.44	0.00	0.37	0.00	0.59	0.00	0.62	0.01	0.24	Rhode, 2010
HdhSSR60b	0.36	-0.07	0.72	-0.22	0.38	-0.05	0.47	-0.23	0.59	-0.04	0.77	-0.16	Rhode, 2010
HmSSRex489a	0.83	0.06	0.27	0.12	0.97	-0.05	0.76	0.13	1.00	-0.04	0.00	0.03	Rhode, 2010
HmidPS1.1018	0.02	0.15	0.01	0.15	0.04	-0.16	0.04	-0.16	0.09	0.20	0.00	0.05	Slabbert *et al*., 2012
HmidPS1.1063	0.00	0.32	0.00	0.52	0.01	0.10	0.56	-0.01	0.27	0.23	0.00	0.20	Slabbert *et al*., 2012
HmidPS1.138	0.00	0.58	0.04	0.29	0.07	0.12	0.64	0.12	0.00	0.43	0.00	0.33	Slabbert *et al*., 2012
HmidPS1.332	0.22	0.13	0.05	0.11	0.15	0.18	0.57	-0.06	0.00	0.25	0.19	0.02	Slabbert *et al*., 2012
Hmid310	0.00	0.22	0.00	0.25	0.19	0.17	0.10	0.17	0.21	0.05	0.71	0.05	Slabbert *et al*., 2010
HmidILL-62675	1.00	-0.13	0.00	-0.05	1.00	-0.11	0.37	-0.23	0.00	0.25	0.06	-0.10	Rhode *et al*., 2012
HmDL131	1.00	-0.03	0.84	-0.09	0.78	-0.10	0.99	-0.21	0.04	-0.02	0.88	-0.19	Slabbert *et al*., 2008
HmidPS1.370	0.03	-0.44	0.59	-0.05	0.00	-0.53	0.03	-0.46	1.00	-0.21	0.28	-0.34	Slabbert *et al*., 2012
HmidPS1.487	1.00	-0.06	1.00	-0.03	1.00	-0.17	1.00	-0.10	1.00	-0.07	1.00	-0.15	Slabbert *et al*., 2012
HmidPS1.551	0.73	-0.12	1.00	-0.11	0.05	0.13	0.71	0.02	0.84	0.10	1.00	-0.16	Slabbert *et al*., 2012
Hmid2044	0.00	0.60	0.00	0.42	0.00	0.35	0.00	0.43	0.00	0.39	0.00	0.37	Slabbert *et al*., 2010
Hmid558	0.28	0.00	0.07	-0.07	0.13	0.01	0.72	0.09	0.21	-0.08	0.33	-0.05	Slabbert *et al*., 2010
HmidPS1.247	0.63	0.11	0.72	-0.10	0.10	0.08	0.18	-0.18	0.60	-0.01	0.33	-0.25	Slabbert *et al*., 2012
HmidPS1.860	0.96	-0.10	0.01	0.38	0.98	-0.09	0.36	-0.03	0.59	0.07	0.63	0.04	Slabbert *et al*., 2012
HmDL110	0.34	-0.40	N/A	N/A	0.08	0.21	0.20	0.22	0.01	0.58	0.00	0.39	Slabbert *et al*., 2008

HmLCS388	1.00	-0.19	0.20	1.00	1.00	-0.02	1.00	-0.01	0.62	0.36	1.00	-0.12	Slabbert *et al*., 2010
HmidPS1.1012	0.15	0.08	0.60	0.20	0.94	0.03	0.09	0.16	0.66	0.15	0.18	0.20	Slabbert *et al*., 2012
HmidPS1.228	1.00	-0.11	0.20	0.32	0.13	0.07	0.50	0.20	0.23	0.02	0.26	0.08	Slabbert *et al*., 2012
Hmid007	0.04	0.31	N/A	N/A	0.00	0.61	0.17	0.34	0.10	0.30	0.09	0.00	Slabbert *et al*., 2010
Hmid553	0.00	0.06	0.27	0.33	0.31	0.16	1.00	-0.09	0.80	-0.14	0.38	0.18	Slabbert *et al*., 2010
Hmid221	0.00	0.22	0.59	-0.11	0.84	-0.10	0.57	-0.09	0.46	-0.06	0.84	0.00	Slabbert *et al*., 2010
Hmid321	0.00	0.30	0.37	0.06	0.06	-0.11	0.77	-0.02	0.44	0.05	0.03	0.31	Slabbert *et al*., 2010
HmidPS1.457	0.00	0.39	0.02	0.07	0.00	0.42	0.00	0.48	0.00	0.50	0.00	0.39	Slabbert *et al*., 2012
HmidPS1.711	0.00	0.24	0.01	-0.48	0.39	-0.28	0.01	-0.15	0.47	-0.13	0.37	-0.28	Slabbert *et al*., 2012
Hmid136	0.00	0.24	0.57	-0.10	0.95	0.01	0.97	-0.02	0.14	0.12	0.32	-0.24	Slabbert *et al*., 2010
HmidPS1.379	0.00	0.23	0.75	-0.07	0.38	0.14	0.42	0.06	0.00	0.51	0.00	0.12	Slabbert *et al*., 2012
HmidPS1.859	0.27	0.00	0.67	-0.10	0.04	0.24	0.43	0.16	0.00	0.36	0.00	0.31	Slabbert *et al*., 2012
Hmid315	0.07	0.19	0.03	0.32	0.37	-0.04	0.17	0.07	0.13	-0.23	0.97	-0.07	Slabbert *et al*., 2010
Hmid4009	1.00	-0.10	1.00	-0.09	1.00	-0.06	1.00	-0.04	1.00	-0.07	1.00	-0.03	Slabbert *et al*., 2010
HmidPS1.1026	0.01	0.13	0.00	0.44	0.00	0.29	0.03	0.24	0.00	0.52	0.01	0.11	Slabbert *et al*., 2012
HmidPS1.1058	0.91	-0.14	0.43	0.11	0.23	0.05	0.21	0.03	0.44	-0.01	0.00	0.20	Slabbert *et al*., 2012
HmidPS1.227	0.03	-0.42	0.46	-0.28	0.51	-0.23	0.40	-0.23	0.07	-0.19	1.00	-0.12	Slabbert *et al*., 2012
Hmid4010	0.32	0.24	0.24	0.17	0.12	0.03	0.20	0.21	0.33	0.14	0.05	0.30	Slabbert *et al*., 2010

HmidPS1.193	0.84	0.04	0.48	0.13	0.07	0.23	0.00	0.57	0.00	0.30	0.01	0.37	Slabbert *et al*., 2012
HmidPS1.692	1.00	-0.02	1.00	-0.02	N/A	N/A	N/A	N/A	0.05	0.66	N/A	N/A	Slabbert *et al*., 2012
HmidPS1.840	0.27	0.25	0.82	0.03	0.01	0.18	0.04	0.31	0.18	0.27	0.00	0.41	Slabbert *et al*., 2012
HmidPS1.95	0.10	0.49	0.02	0.40	1.00	-0.01	0.16	0.30	0.05	0.66	0.02	0.63	Slabbert *et al*., 2012
HmidILL-46948	0.85	-0.20	0.39	0.22	0.35	-0.27	1.00	-0.17	0.03	0.27	0.36	0.14	Rhode *et al*., 2012
HmidILL-140027	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1.00	-0.02	Rhode *et al*., 2012
HmidILL-72605	N/A	N/A	N/A	N/A	N/A	N/A	1.00	-0.03	0.02	1.00	N/A	N/A	Rhode *et al*., 2012
HmNR185	0.98	-0.05	0.01	0.21	0.29	0.13	0.02	0.24	0.00	0.18	0.00	0.59	Slabbert *et al*., 2007
HmidILL-7898	0.16	0.00	0.27	-0.02	0.12	0.16	0.00	0.26	0.11	0.16	0.04	0.05	Rhode *et al*., 2012

Candidate Directional Selection Loci

Locus	**CPWC**		**CPSC**		**WPEC**		**WPWC**		**CPEC**		**WPSC**		**Marker Reference**
	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	
HmLCS48	0.62	-0.14	0.77	-0.23	0.01	0.28	0.23	0.18	0.02	0.00	0.02	0.15	Slabbert *et al*., 2007
HmLCS5	N/A	N/A	0.00	0.83	0.02	1.00	0.00	0.84	N/A	N/A	0.00	0.79	Slabbert *et al*., 2008
HmNR106	0.01	-0.33	0.61	-0.16	0.00	0.61	0.00	0.16	0.69	-0.22	0.00	0.56	Slabbert *et al*., 2009
HmNSS1H	0.00	0.45	0.01	0.55	0.00	-0.04	0.12	0.02	0.00	0.30	0.12	-0.04	Slabbert *et al*., 2010
HmNS18	0.67	-0.12	0.08	0.34	0.00	-0.54	0.01	-0.44	0.00	0.10	0.05	-0.39	Slabbert *et al*., 2010

HmidILL-140858	0.64	-0.05	0.10	0.22	0.38	-0.04	0.87	-0.14	0.57	0.08	0.73	-0.12	Rhode *et al*., 2012
HmidILL-2192	0.00	0.82	0.01	0.62	1.00	-0.11	0.00	0.54	0.00	0.84	0.43	0.17	Rhode *et al*., 2012
HmidILL-87955	0.15	-0.19	0.25	0.12	0.02	0.17	0.00	-0.11	0.58	0.05	0.00	0.36	Rhode *et al*., 2012
HmidPS1.559	0.02	0.35	0.00	0.46	0.31	0.09	0.09	0.27	0.04	0.28	0.01	0.19	Slabbert *et al*., 2012
HmidPS1.561	0.00	0.70	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Slabbert *et al*., 2012
HmidILL-146360	N/A	N/A	N/A	N/A	N/A	N/A	1.00	-0.11	0.02	1.00	N/A	N/A	Rhode *et al*., 2012
HmidILL-88398	0.70	-0.08	0.59	0.15	0.00	-1.00	0.00	-1.00	0.83	-0.17	0.49	0.13	Rhode *et al*., 2012
HmidILL-64192	N/A	N/A	N/A	N/A	N/A	N/A	0.72	-0.22	N/A	N/A	0.02	0.66	Rhode *et al*., 2012
HmidILL-37506	0.00	0.15	0.23	0.15	0.16	0.09	0.72	-0.01	0.11	-0.01	0.01	0.08	Rhode *et al*., 2012

Candidate Balancing Selection Loci

Locus	**CPWC**		**CPSC**		**WPEC**		**WPWC**		**CPEC**		**WPSC**		**Marker Reference**
	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	***P*-value**	**F_{is}**	
HmLCS37	N/A	N/A	N/A	N/A	0.01	0.06	0.00	0.13	N/A	N/A	0.18	0.06	Slabbert *et al*., 2007
HmRS117	0.18	0.08	0.00	0.34	0.02	0.03	0.58	0.04	0.02	0.15	0.05	0.15	Slabbert *et al*., 2007
HmAD102	0.00	0.29	0.00	0.27	0.00	0.23	0.00	0.35	0.00	0.35	0.00	0.44	Slabbert *et al*., 2007
HmRS27	0.04	0.05	0.86	0.03	0.50	0.00	0.66	0.00	0.07	-0.01	0.36	0.00	Slabbert *et al*., 2007
HmRS80	0.07	0.05	0.18	0.04	0.74	0.02	0.54	-0.03	0.04	0.13	0.69	0.06	Slabbert *et al*., 2007

HmD59	0.09	0.06	0.64	0.01	0.27	0.07	0.05	0.12	0.23	0.10	0.84	-0.02	Bester *et al*., 2004
HmRS129	0.00	0.23	0.00	0.16	0.00	0.38	0.00	0.42	0.00	0.39	0.00	0.30	Slabbert *et al*., 2007
HmidPS1.870	0.96	-0.03	0.83	0.01	0.43	-0.03	0.81	-0.08	0.08	0.02	0.77	0.05	Slabbert *et al*., 2012
HmNR120	0.06	0.08	0.79	-0.03	0.37	0.07	0.68	-0.01	0.12	-0.02	0.13	0.02	Slabbert *et al*., 2007
HmG53	0.00	0.10	0.80	0.02	0.86	-0.03	0.61	-0.03	0.85	-0.04	0.00	0.14	Slabbert *et al*., 2007
HmDL50	0.11	0.15	0.00	0.35	0.07	0.14	0.52	0.03	0.03	0.09	0.01	0.28	Slabbert *et al*., 2007
HmNR224	0.03	0.00	0.38	-0.04	0.00	0.09	0.66	-0.04	0.09	0.06	0.88	-0.01	Slabbert *et al*., 2007
HmNS17b	0.00	0.28	0.00	0.28	0.00	0.24	0.81	0.05	0.00	0.42	0.03	0.11	Slabbert *et al*., 2007
HmNS19	0.00	0.23	0.00	0.07	0.00	0.18	0.85	-0.01	0.00	0.31	0.02	0.08	Slabbert *et al*., 2007
HmNS58	0.00	0.33	0.11	0.12	0.00	0.30	0.07	0.16	0.08	0.08	0.00	0.16	Slabbert *et al*., 2007
HmidILL-118779	N/A	N/A	1.00	-0.02	1.00	-0.02	1.00	-0.01	1.00	-0.02	1.00	-0.02	Rhode *et al*., 2012
HmidILL-70036	0.23	0.12	0.00	0.22	0.75	-0.02	0.07	0.02	0.22	0.07	0.06	0.18	Rhode *et al*., 2012
HmNS56	0.22	0.06	0.17	0.10	0.99	-0.16	0.77	-0.02	0.00	0.19	0.10	0.14	Slabbert *et al*., 2007
HmidILL-76149	0.94	-0.09	0.16	0.22	0.18	-0.05	0.00	0.29	0.00	0.31	0.48	-0.06	Rhode *et al*., 2012
HmPS1.147	0.00	0.12	0.03	0.19	0.45	0.09	0.23	-0.08	0.01	0.09	0.92	-0.09	Slabbert *et al*., 2012
Hmid563	0.07	0.09	0.22	-0.07	0.09	0.03	0.53	0.01	0.76	0.06	0.11	0.09	Slabbert *et al*., 2010
HmidPS1.629	0.19	-0.20	0.09	-0.31	0.04	-0.29	0.02	-0.31	0.04	-0.23	0.11	0.01	Slabbert *et al*., 2012
HmD61	0.89	-0.01	0.00	0.50	0.00	0.38	0.02	0.24	0.00	0.28	0.00	0.52	Bester *et al*., 2004

Hmid610	0.16	0.27	0.94	0.08	0.54	0.08	0.02	0.18	0.01	0.24	0.64	-0.01	Slabbert *et al*., 2010
HmidPS1.549	0.44	0.11	0.05	0.08	0.12	0.24	0.00	0.25	0.14	0.15	0.67	0.02	Slabbert *et al*., 2012
HmidPS1.874	0.19	0.11	0.34	0.05	0.22	-0.03	0.00	0.25	0.07	-0.03	0.21	0.06	Slabbert *et al*., 2012
Hmid65	0.00	0.17	0.01	0.19	0.08	0.05	0.00	0.24	0.00	0.12	0.69	0.01	Slabbert *et al*., 2010

Table S3.3: Pairwise linkage disequilibrium statistics for syntenic markers across Wild and Cultured populations.

Estimates of linkage disequilibrium for pairwise syntenic loci across wild populations				
	Distance (cM)	**D'**	**χ'2**	***P*-value**
Locus pairs on LG_1				
HmidPS1.332_HmidPS1.859	6.1960	0.4686	0.1743	0.3493
HmidPS1.332_HmidPS1.227	9.4020	0.3979	0.2774	0.4724
HmidPS1.332_HmNS19	38.3320	0.6478	0.1979	0.4855
HmidPS1.332_HmNS56	38.9300	0.5115	0.2377	0.8639
HmidPS1.332_HmNR54	39.7450	0.4801	0.1661	0.3804
HmidPS1.859_HmidPS1.227	3.2060	0.3974	0.0484	0.9820
HmidPS1.859_HmNS19	32.1360	0.5105	0.2462	0.9530
HmidPS1.859_HmNS56	32.7340	0.4579	0.3354	0.4164
HmidPS1.859_HmNR54	33.5490	0.4354	0.2017	0.1593
HmidPS1.227_HmNS19	28.9300	0.5226	0.3761	0.6597
HmidPS1.227_HmNS56	29.5280	0.4228	0.2139	0.8128
HmidPS1.227_HmNR54	30.3430	0.3983	0.1345	0.5220
HmNS19_HmNS56	0.5980	0.6932	0.2504	0.2112
HmNS19_HmNR54	1.4130	0.6534	0.2365	0.0120*
HmNS56_HmNR54	0.8150	0.6125	0.3517	0.0070*
Locus pairs on LG_2				
HmidPS1.138_HmidILL-40027	40.7860	0.6018	0.1199	0.4865
HmidPS1.138_HmD61	42.1310	0.4182	0.1601	0.6246
HmidPS1.138_HmidILL-76149	48.1550	0.3511	0.1004	0.9840
HmidPS1.138_HmidILL-8738	86.5290	0.3544	0.1281	0.3253
HmidILL-40027_HmidD61	1.3450	0.7028	0.0915	0.7277
HmidILL-40027_HmidILL-76149	7.3690	0.6978	0.0224	0.6727
HmidILL-40027_HmidILL-8738	45.7430	0.5787	0.0178	0.9960
HmD61_HmidILL-76149	6.0240	0.3556	0.1484	0.2678
HmD61_HmidILL-8738	44.3980	0.3541	0.1094	0.9349
HmidILL-76149_HmidILL-8738	38.3740	0.2739	0.0500	0.2818
Locus pairs on LG_3				

HmidPS1.967_Hmid65	13.3800	0.5493	0.2458	0.4655
HmidPS1.967_HmNR185	36.1590	0.3604	0.1274	0.0761
Hmid65_HmNR185	22.7790	0.6889	0.2342	0.0220*
Locus pairs on LG_4				
HmidPS1.1058_HmRS38	21.4370	0.5518	0.2119	0.3013
HmidPS1.1058_HmRS27	29.9720	0.7045	0.3682	0.6527
HmidPS1.1058_HmLCS67	42.9680	0.3689	0.0757	0.8297
HmRS38_HmRS27	8.5350	0.6807	0.3018	0.5395
HmRS38_HmLCS67	21.5310	0.1864	0.0192	0.5265
HmRS27_HmLCS67	12.9960	0.5787	0.3217	0.2953
Locus pairs on LG_5				
HmidILL-2192_HmidPS1.228	6.1610	0.2197	0.0599	0.3023
HmidILL1-2192_HmNR281	10.5830	0.6964	0.2524	0.2543
HmidILL-2192_HmidPS1.551	12.0040	0.4078	0.0769	0.4289
HmidILL-2192_HmidILL.47613	13.2860	0.4304	0.1426	0.0501
HmidILL-2192_Hmid221	25.5820	0.2499	0.0504	0.8829
HmidPS1.228_HmNR281	4.4220	0.6018	0.4513	0.0490*
HmidPS1.228_HmidPS1.551	5.8430	0.2827	0.0707	0.1962
HmidPS1.228_HmidILL-47613	7.1250	0.4144	0.1115	0.6847
HmidPS1.228_Hmid221	19.4210	0.3172	0.0627	0.8246
HmNR281_HmidPS1.551	1.4210	0.5987	0.3713	0.5185
HmNR281_HmidILL-47613	2.7030	0.7234	0.3586	0.5355
HmNR281_Hmid221	14.9990	0.6697	0.3500	0.2082
HmidPS1.551_HmidILL-47613	1.2820	0.4588	0.1563	0.2132
HmidPS1.551_Hmid221	13.5780	0.3952	0.1290	0.1743
HmidILL-47613_Hmid221	12.2960	0.4815	0.1417	0.8198
Locus pairs on LG_6				
HmidILL-64121_HmLCS9	30.0030	0.2662	0.0472	0.4665
HmidILL-64121_Hmid321	34.0430	0.1824	0.0972	0.4795
HmidILL-64121_HmRS129	51.4040	0.4927	0.2160	0.1001
HmidILL-64121_HmAD102	64.0160	0.5610	0.3527	0.4705
HmLCS9_Hmid321	4.0400	0.1973	0.0608	0.7498
HmLCS9_HmRS129	21.4010	0.4851	0.1868	0.1301

HmLCS9_HmAD102	34.0130	0.5705	0.2341	0.1982
Hmid321_HmRS129	17.3610	0.5299	0.1372	0.2573
Hmid321_HmAD102	29.9730	0.5623	0.2799	0.5135
HmRS129_HmAD102	12.6120	0.8019	0.3489	0.0010*
Locus pairs on LG_7				
HmLCS388_HmidPS1.860	8.2440	0.2688	0.0920	0.1123
HmLCS388_HmNS17b	19.0650	0.5222	0.2297	0.7097
HmLCS388_Hmid310	34.6690	0.3442	0.1505	0.7698
HmidPS1.860_HmNS17b	10.8210	0.5066	0.2320	0.7277
HmidPS1.860_Hmid310	26.4250	0.4363	0.3116	0.0321*
HmNS17b_Hmid310	15.6040	0.6899	0.2911	0.1762
Locus pairs on LG_8				
HmNR191_HmRS62	14.2350	0.4785	0.1662	0.8669
HmNR191_HmD59	14.7540	0.5197	0.1398	0.6166
HmNR191_HmILL-71359	17.0650	0.3482	0.1096	0.2613
HmNR191_HmSSRex489a	20.6000	0.3056	0.1045	0.7355
HmNR191_HmLCS1	21.2850	0.4700	0.2059	0.0901
HmNR191_HmSSRex489b	21.4830	0.3212	0.1352	0.5946
HmNR191_HmLCS37	46.5150	0.6593	0.3270	0.7227
HmNR191_HmidILL-72605	50.0650	0.6457	0.3537	0.0410*
HmNR191_HmNR258	62.0440	0.3745	0.2285	0.7117
HmRS62_HmD59	0.5190	0.6862	0.3747	0.0000*
HmRS62_HmidILL-71359	2.8300	0.4078	0.2455	0.7227
HmRS62_HmSSRex489a	6.3650	0.3495	0.1918	0.1493
HmRS62_HmLCS1	7.0500	0.4029	0.2165	0.0240*
HmRS62_HmSSRex489b	7.2480	0.3993	0.1692	0.1212
HmRS62_HmLCS37	32.2800	0.7146	0.3618	0.6076
HmRS62_HmidILL-72605	35.8300	0.6586	0.1116	0.4545
HmRS62_HmidNR258	47.8090	0.4300	0.1791	0.1061
HmD59_HmidILL-71359	2.3110	0.4325	0.1383	0.8327
HmD59_HmSSRex489a	5.8460	0.4281	0.1468	0.9438
HmD59_HmidLCS1	6.5310	0.4585	0.2063	0.1792
HmD59_HmSSRex489b	6.7290	0.4280	0.1467	0.8529

HmD59_HmLCS37	31.7610	0.7793	0.3335	0.0120*
HmD59_HmidILL-72605	35.3110	0.6717	0.1072	0.7337
HmD59_HmNR258	47.2900	0.4403	0.1522	0.3854
HmidILL2.71359_HmSSRex489a	3.5350	0.1684	0.0527	0.5000
HmidILL2.71359_HmLCS1	4.2200	0.1925	0.0716	0.4443
HmidILL2.71359_HmSSRex489b	4.4180	0.2053	0.0520	0.5886
HmidILL2.71359_HmLCS37	29.4500	0.5623	0.4210	0.7698
HmidILL2.71359_HmidILL-72605	33.0000	0.3718	0.0576	0.2513
HmidILL2.71359_HmNR258	44.9790	0.2519	0.0369	0.8878
HmSSRex489a_HmLCS1	0.6850	0.2826	0.0717	0.2568
HmSSRex489a_HmSSRex489b	0.8830	0.8138	0.4650	0.0000*
HmSSRex489a_HmLCS37	25.9150	0.5570	0.4417	0.4785
HmSSRex489a_HmidILL-72605	29.4650	0.3455	0.1357	0.1421
HmSSRex489a_HmNR258	41.4440	0.1650	0.0543	0.7505
HmLCS1_HmSSRex489b	0.1980	0.2240	0.2473	0.1294
HmLCS1_HmLCS37	25.2300	0.5744	0.4436	0.5385
HmLCS1_HmidILL-72605	28.7800	0.4055	0.0226	0.8398
HmLCS1_HmNR258	40.7590	0.2218	0.0446	0.9509
HmSSRex489b_HmLCS37	25.0320	0.5396	0.3079	0.7986
HmSSRex489b_HmidILL-72605	28.5820	0.3559	0.1356	0.1191
HmSSRex489b_HmNR258	40.5610	0.1766	0.0443	0.7477
HmLCS37_HmidILL-72605	3.5500	0.7925	0.2434	0.8028
HmLCS37_HmNR258	15.5290	0.6291	0.3373	0.4344
HmidILL1-72605_HmNR258	11.9790	0.4520	0.1363	0.3514
Locus pairs on LG_9				
HmLCS48_HmNR180	9.8860	0.4546	0.1273	0.2422
HmLCS48_HmPS1.549	39.3390	0.3681	0.1364	0.5866
HmLCS48_HmNS58	41.0610	0.3937	0.1903	0.2412
HmNR180_HmPS1.549	29.4530	0.5025	0.2006	0.0040*
HmNR180_HmNS58	31.1750	0.4694	0.1519	0.0911
HmPS1.549_HmNS58	1.7220	0.4221	0.1184	0.8118
Locus pairs on LG_10				
HmRS117_HmNR120	29.3150	0.7215	0.2330	0.6486

Locus pairs on LG_12				
HmNR20_Hmid553	17.0920	0.3555	0.1508	0.2402
HmNR20_Hmid610	27.9440	0.5647	0.1710	0.6406
HmNR20_HmPS1.874	30.3030	0.6474	0.1865	0.6296
Hmid553_Hmid610	10.8520	0.4691	0.3163	0.2525
Hmid553_HmPS1.874	13.2110	0.4068	0.2322	0.6356
Hmid610_HmPS1.874	2.3590	0.6735	0.2696	0.0230*
Locus pairs on LG_13				
Hmid4010_Hmid563	16.1360	0.4925	0.2199	0.2282
Locus pairs on LG_14				
HmidPS1.1063_HmidPS1.818	8.2920	0.5350	0.2037	0.0120*
HmidPS1.1063_HmidPS1.247	16.1780	0.4056	0.1300	0.7738
HmidPS1.1063_HmidPS1.370	24.6690	0.5040	0.2394	0.6196
HmidPS1.818_HmidPS1.247	7.8860	0.3972	0.1276	0.0661
HmidPS1.818_HmidPS1.370	16.3770	0.4405	0.1550	0.4985
HmidPS1.247_HmidPS1.370	8.4910	0.3655	0.0838	0.2863
Locus pairs on LG_15				
HmidPS1.305_HmidILL-87955	8.8280	0.3649	0.1311	0.1301
HmidPS1.305_HmDL50	15.9870	0.5707	0.2212	0.4234
HmidILL-87955_HmDL50	7.1590	0.4949	0.1914	0.2162
Locus pairs on LG_17				
HmidPS1.1012_HmLCS7	15.2110	0.3873	0.0960	0.6907
Locus pairs on LG_18A				
HmNS6_HmDL110	0.2290	0.1742	0.0500	0.8848
Locus pairs on LG_18B				
HmNS6_HmDL214	25.5440	0.3687	0.1265	0.4985
HmNS6_HmDL34	34.5820	0.4253	0.1624	0.2292
HmDL214_HmDL34	9.0380	0.4749	0.2259	0.0371*
Locus pairs on LG_18C				
Hmid2044_HmidPS1.559	7.4340	0.5290	0.1822	0.5586
Hmid2044_HmidPS1.193	13.1700	0.4805	0.3043	0.1171
PS1.559_HmidPS1.193	5.7360	0.3510	0.0813	0.8186
Locus pairs on LG_18D				

HmidILL-66010a_HmPS1.559	2.3100	0.2244	0.1078	0.3884
HmidILL-66010a_HmG53	5.8770	0.7038	0.3587	0.2365
HmidPS1.559_HmG53	3.5670	0.6747	0.2262	0.8739

Estimates of linkage disequilibrium for pairwise syntenic loci across cultured populations

	Distance (cM)	D'	χ'2	p-value
Locus pairs on LG_1				
HmidPS1.332_HmidPS1.859	6.1960	0.5276	0.2853	0.0000*
HmidPS1.332_HmidPS1.227	9.4020	0.4602	0.4119	0.1914
HmidPS1.332_HmidNS19	38.3320	0.7828	0.2645	0.0450*
HmidPS1.332_HmidNS56	38.9300	0.4912	0.1793	0.6767
HmidPS1.332_HmidNR54	39.7450	0.5752	0.1832	0.0270*
HmidPS1.859_HmidPS1.227	3.2060	0.3020	0.5776	0.0020*
HmidPS1.859_HmNS19	32.1360	0.5969	0.2462	0.2182
HmidPS1.859_HmNS56	32.7340	0.5039	0.3254	0.0010*
HmidPS1.859_HmNR54	33.5490	0.3766	0.0974	0.3938
HmidPS1.227_HmNS19	28.9300	0.6299	0.2956	0.1692
HmidPS1.227_HmNS56	29.5280	0.3368	0.3955	0.0190*
HmidPS1.227_HmNR54	30.3430	0.4775	0.1508	0.3287
HmNS19_HmNS56	0.5980	0.6458	0.2887	0.2082
HmNS19_HmNR54	1.4130	0.6764	0.3233	0.0000*
HmNS56_HmNR54	0.8150	0.5196	0.2297	0.1101
Locus pairs on LG_2				
HmidPS1.138_HmidILL-40027	40.7860	0.7657	0.2507	0.3223
HmidPS1.138_HmD61	42.1310	0.4331	0.2024	0.5055
HmidPS1.138_HmidILL-76149	48.1550	0.3046	0.1062	0.2352
HmidPS1.138_HmidILL-8738	86.5290	0.2826	0.0949	0.3904
HmidILL-40027_HmD61	1.3450	1.0000	0.4967	0.2252
HmidILL-40027_HmidILL-76149	7.3690	0.6263	0.0113	0.8969
HmidILL-40027_HmidILL-8738	45.7430	0.7769	0.0151	0.7137
HmD61_HmidILL-76149	6.0240	0.4067	0.2004	0.8659
HmD61_HmILL-8738	44.3980	0.4094	0.1608	0.8148
HmILL-76149_HmILL-8738	38.3740	0.2912	0.0623	0.3894
Locus pairs on LG_3				

HmidPS1.967_Hmid65	13.3800	0.5616	0.2562	0.0460*
HmidPS1.967_HmNR185	36.1590	0.4691	0.1720	0.1151
Hmid65_HmNR185	22.7790	0.6714	0.2656	0.0110*
Locus pairs on LG_4				
HmidPS1.1058_HmRS38	21.4370	0.3483	0.0999	0.9249
HmidPS1.1058_HmRS27	29.9720	0.7391	0.2802	0.0120*
HmidPS1.1058_HmLCS67	42.9680	0.3615	0.1024	0.1774
HmRS38_HmRS27	8.5350	0.6779	0.4149	0.0020*
HmRS38_HmLCS67	21.5310	0.2174	0.0994	0.0752
HmRS27_HmLCS67	12.9960	0.6024	0.2398	0.0290*
Locus pairs on LG_5				
HmidILL-2192_HmidPS1.228	6.1610	0.2660	0.0781	0.3524
HmidILL-2192_HmNR281	10.5830	0.5782	0.2748	0.0380*
HmidILL-2192_HmidPS1.551	12.0040	0.0600	0.0120	0.8989
HmidILL-2192_HmILL-47613	13.2860	0.2766	0.1351	0.0871
HmidILL-2192_Hmid221	25.5820	0.3522	0.0687	0.0120*
HmidPS1.228_HmNR281	4.4220	0.7117	0.5818	0.0691
HmidPS1.228_HmidPS1.551	5.8430	0.3046	0.3504	0.0201*
HmidPS1.228_HmidILL-47613	7.1250	0.5380	0.4199	0.4615
HmidPS1.228_Hmid221	19.4210	0.4222	0.2105	0.4194
HmNR281_HmidPS1.551	1.4210	0.5064	0.3152	0.6617
HmNR281_HmILL-47613	2.7030	0.6792	0.3481	0.3814
HmNR281_Hmid221	14.9990	0.6185	0.2486	0.6196
HmidPS1.551_HmidILL-47613	1.2820	0.4787	0.4283	0.0020*
HmidPS1.551_Hmid221	13.5780	0.3413	0.1233	0.0954
HmidILL-47613_Hmid221	12.2960	0.5246	0.1843	0.0350*
Locus pairs on LG_6				
HmidILL-64121_HmLCS9	30.0030	0.1871	0.0909	0.1383
HmidILL-64121_Hmid321	34.0430	0.2699	0.0709	0.6066
HmidILL-64121_HmRS129	51.4040	0.4894	0.2005	0.3724
HmidILL-64121_HmAD102	64.0160	0.4901	0.2883	0.3347
HmLCS9_Hmid321	4.0400	0.4041	0.0964	0.0791
HmLCS9_HmRS129	21.4010	0.4721	0.2605	0.3303

HmLCS9_HmAD102	34.0130	0.4741	0.2028	0.9540
Hmid321_HmRS129	17.3610	0.4927	0.2427	0.8348
Hmid321_HmAD102	29.9730	0.6123	0.3178	0.0460*
HmRS129_HmAD102	12.6120	0.8102	0.2393	0.0190*
Locus pairs on LG_7				
HmLCS388_HmPS1.860	8.2440	0.2561	0.0765	0.9039
HmLCS388_HmNS17b	19.0650	0.8135	0.7093	0.2152
HmLCS388_Hmid310	34.6690	0.8305	0.5046	0.2683
HmPS1.860_HmNS17b	10.8210	0.5238	0.3726	0.4004
HmPS1.860_Hmid310	26.4250	0.4015	0.2302	0.4605
HmNS17b_Hmid310	15.6040	0.8455	0.4370	0.0040*
Locus pairs on LG_8				
HmNR191_HmRS62	14.2350	0.4826	0.1350	0.7928
HmNR191_HmD59	14.7540	0.5096	0.1326	0.3213
HmNR191_HmidILL-71359	17.0650	0.3182	0.1435	0.2234
HmNR191_HmSSRex489a	20.6000	0.3238	0.0891	0.8388
HmNR191_HmLCS1	21.2850	0.3455	0.1154	0.3083
HmNR191_HmSSRex489b	21.4830	0.4023	0.1858	0.2876
HmNR191_HmILL-72605	50.0650	0.3888	0.0750	0.7988
HmNR191_HmNR258	62.0440	0.3037	0.1048	0.5936
HmRS62_HmD59	0.5190	0.7852	0.4287	0.0000*
HmRS62_HmidILL-71359	2.8300	0.4724	0.1493	0.2886
HmRS62_HmSSRex489a	6.3650	0.3615	0.1554	0.2823
HmRS62_HmLCS1	7.0500	0.4419	0.1614	0.0460*
HmRS62_HmSSRex489b	7.2480	0.5420	0.2821	0.1471
HmRS62_HmidILL-72605	35.8300	0.3528	0.1896	0.8579
HmRS62_HmNR258	47.8090	0.4107	0.2336	0.0040*
HmD59_HmidILL-71359	2.3110	0.4627	0.1673	0.4668
HmD59_HmSSRex489a	5.8460	0.3981	0.1529	0.0873
HmD59_HmLCS1	6.5310	0.4121	0.1429	0.5065
HmD59_HmSSRex489b	6.7290	0.5295	0.3196	0.0440*
HmD59_HmidILL-72605	35.3110	0.4497	0.1281	0.3814
HmD59_HmNR258	47.2900	0.4280	0.2583	0.0010*

HmidILL-71359_HmSSRex489a	3.5350	0.4499	0.0569	0.1074
HmidILL-71359_HmLCS1	4.2200	0.1728	0.0440	0.8858
HmidILL-71359_HmSSRex489b	4.4180	0.5171	0.0994	0.2643
HmidILL-71359_HmidILL-72605	33.0000	0.2534	0.0083	0.7347
HmidILL-71359_HmNR258	44.9790	0.2390	0.0475	0.3202
HmSSRex489a_HmLCS1	0.6850	0.2816	0.0646	0.4674
HmSSRex489a_HmSSRex489b	0.8830	0.9300	0.8657	0.0000*
HmSSRex489a_HmidILL-72605	29.4650	0.3271	0.0080	0.7257
HmSSRex489a_HmNR258	41.4440	0.1944	0.0530	0.1043
HmLCS1_HmSSRex489b	0.1980	0.2810	0.1052	0.7108
HmLCS1_HmidILL-72605	28.7800	0.2205	0.2433	0.0320*
HmLCS1_HmNR258	40.7590	0.3419	0.1214	0.0271*
HmSSRex489b_HmidILL-72605	28.5820	0.6959	0.0125	0.7437
HmSSRex489b_HmNR258	40.5610	0.3364	0.1699	0.0060*
HmidILL-72605_HmNR258	11.9790	0.1500	0.0083	0.9920
Locus pairs on LG_9				
HmLCS48_HmNR180	9.8860	0.5053	0.2166	0.4634
HmLCS48_HmPS1.549	39.3390	0.3674	0.1063	0.5281
HmLCS48_HmNS58	41.0610	0.3954	0.1053	0.9389
HmNR180_HmPS1.549	29.4530	0.4476	0.1514	0.9640
HmNR180_HmNS58	31.1750	0.5688	0.1849	0.2873
HmidPS1.549_HmNS58	1.7220	0.4320	0.1060	0.8038
Locus pairs on LG_10				
HmRS117_HmNR120	29.3150	0.7965	0.2756	0.4444
Locus pairs on LG_12			0.0000	
HmNR20_Hmid553	17.0920	0.6206	0.2505	0.7603
HmNR20_Hmid610	27.9440	0.6590	0.3181	0.1572
HmNR20_HmidPS1.874	30.3030	0.7048	0.2336	0.8238
Hmid553_Hmid610	10.8520	0.4826	0.2868	0.1263
Hmid553_HmidPS1.874	13.2110	0.5549	0.1627	0.8937
Hmid610_HmidPS1.874	2.3590	0.6203	0.2664	0.5936
Locus pairs on LG_13				
Hmid4010_Hmid563	16.1360	0.4054	0.1420	0.5872

Locus pairs on LG_14				
HmidPS1.1063_HmidPS1.818	8.2920	0.5892	0.2448	0.2212
HmidPS1.1063_HmidPS1.247	16.1780	0.4452	0.2073	0.4114
HmidPS1.1063_HmidPS1.370	24.6690	0.4875	0.2455	0.2375
HmidPS1.818_HmidPS1.247	7.8860	0.3350	0.1378	0.4253
HmidPS1.818_HmidPS1.370	16.3770	0.4831	0.1276	0.1984
HmidPS1.247_HmidPS1.370	8.4910	0.3656	0.0496	0.8404
Locus pairs on LG_15				
HmidPS1.305_HmidILL-87955	8.8280	0.3692	0.1088	0.7257
HmidPS1.305_HmDL50	15.9870	0.5646	0.2144	0.3023
HmidILL-87955_HmDL50	7.1590	0.5118	0.3507	0.0440*
Locus pairs on LG_17				
HmidPS1.1012_HmLCS7	15.2110	0.5311	0.4231	0.3834
Locus pairs on LG_18A			0.0000	
HmidNS6_HmDL110	0.2290	0.6350	0.3937	0.0741
Locus pairs on LG_18B				
HmNS6_HmDL214	25.5440	0.2988	0.0944	0.8298
HmNS6_HmDL34	34.5820	0.4178	0.1688	0.2234
HmDL214_HmDL34	9.0380	0.3709	0.0683	1.0000
Locus pairs on LG_18C				
Hmid2044_HmPS1.559	7.4340	0.3967	0.1814	0.2873
Hmid2044_HmPS1.193	13.1700	0.3592	0.1511	0.5521
HmidPS1.559_HmPS1.193	5.7360	0.3455	0.1139	0.8539
Locus pairs on LG_18D			0.0000	
HmidILL-66010a_HmPS1.559	2.3100	0.2663	0.0880	0.8529
HmidILL-66010a_HmG53	5.8770	0.5935	0.2966	0.6066
HmidPS1.559_HmG53	3.5670	0.5838	0.3456	0.2222

** Statistically significant at the 5% nominal level.*

Table S3.4: Linkage disequilibrium (based on D' and χ'2) estimates for candidate locus pairs under selection, with significance tested by means of 1000 simulations. For loci mapped to the *H. midae* linkage map (Vervalle *et al*., in press), linkage group allocation is given in parenthesis.

Cultured Populations				Wild Populations			
Locus pair	**D'**	***χ'2***	***P*-value**	**Locus pair**	**D'**	***χ'2***	***P*-value**
HmLCS48 (LG_9)_HmRS129 (LG_6)	0.5304	0.3019	0.0310*	*HmLCS48 (LG_9)_HmDL50*	0.5351	0.2392	0.0390*
HmLCS48 (LG_9)_HmNS17b (LG_7)	0.6119	0.3706	0.0480*	*HmLCS48 (LG_9)_HmidILL-88398*	0.2574	0.1843	0.0410*
HmLCS5_HmRS117 (LG_10)	0.7328	0.6364	0.0080*	*HmLCS37 (LG_8)_HmD59 (LG_8)*	0.7769	0.3342	0.0120*
HmRS117 (LG_10)_HmNR106	0.6356	0.2825	0.0460*	*HmLCS37 (LG_8)_HmidILL-87955 (LG_15)*	0.6709	0.3224	0.0440*
HmRS117 (LG_10)_HmidILL-140858	0.6092	0.3395	0.0320*	*HmLCS37 (LG_8)_HmD61 (LG_2)*	0.7067	0.3464	0.0190*
HmRS117 (LG_10)_HmD61 (LG_2)	0.7464	0.2918	0.0050*	*HmLCS37 (LG_8)_HmidPS1.561*	0.9934	0.7483	0.0470*
HmRS117 (LG_10)_Hmid610 (LG_12)	0.6769	0.3265	0.0480*	*HmLCS5_HmNR120 (LG_10)*	0.6509	0.3731	0.0220*
HmRS117 (LG_10)_Hmid65 (LG_3)	0.7857	0.3272	0.0090*	*HmLCS5_HmidILL-146360*	0.5604	0.3015	0.0040*
HmAD102 (LG_6)_HmRS129 (LG_6)	0.8068	0.2644	0.0140*	*HmRS117 (LG_10)_HmRS27 (LG_4)*	0.8140	0.2909	0.0320*
HmAD102 (LG_6)_HmidPS1.870	0.7261	0.2568	0.0450*	*HmRS117 (LG_10)_HmidPS1.559 (LG_18)*	0.6297	0.2400	0.0190*
HmAD102 (LG_6)_HmNR106	0.6385	0.3225	0.0250*	*HmRS117 (LG_10)_HmD61 (LG_2)*	0.6617	0.2553	0.0390*
HmAD102 (LG_6)_HmidPS1.559 (LG_18)	0.5910	0.4143	0.0090*	*HmRS117 (LG_10)_HmidILL-88398*	0.6504	0.3582	0.0100*
HmAD102 (LG_6)_HmD61 (LG_2)	0.7392	0.3194	0.0000*	*HmAD102 (LG_6)_HmRS129 (LG_6)*	0.7999	0.3521	0.0070*
HmRS27 (LG_4)_HmNR120 (LG_10)	0.8208	0.3172	0.0060*	*HmAD102 (LG_6)_HmNS19L (LG_1)*	0.8138	0.3004	0.0040*

HmRS27 (LG_4)_HmG53 (LG_18)	0.8527	0.2626	0.0050*	*HmAD102 (LG_6)_HmidILL-87955 (LG_15)*	0.62560	0.3010	0.0390*
HmRS27 (LG_4)_HmNS17b (LG_7)	0.8697	0.3860	0.0290*	*HmAD102 (LG_6)_HmidILL-70036*	0.7184	0.2538	0.0090*
HmRS27 (LG_4)_HmNS19 (LG_1)	0.8585	0.2493	0.0010*	*HmAD102 (LG_6)_HmidILL-146360*	0.8391	0.3318	0.0440*
HmRS27 (LG_4)_HmD61 (LG_2)	0.7174	0.2985	0.0140*	*HmRS27 (LG_4)_HmNR120 (LG_10)*	0.8222	0.2977	0.0380*
HmD59 (LG_8)_HmDL50	0.6488	0.1883	0.0120*	*HmRS27 (LG_4)_HmidPS1.561*	0.9937	0.7484	0.0380*
HmD59 (LG_8)_HmNS17b (LG_7)	0.7263	0.3058	0.0110*	*HmRS27 (LG_4)_HmidPS1.874 (LG_12)*	0.8426	0.3328	0.0110*
HmD59 (LG_8)_HmidPS1.559 (LG_18)	0.4937	0.1387	0.0350*	*HmD59 (LG_8)_HmidPS1.559 (LG_18)*	0.5228	0.1951	0.0390*
HmD59 (LG_8)_HmidILL-146360	0.9163	0.2204	0.0350*	*HmD59 (LG_8)_HmidPS1.561*	0.9870	0.6226	0.0300*
HmRS129 (LG_6)_HmNS19 (LG_1)	0.8123	0.2701	0.0020*	*HmRS129 (LG_6)_HmNS17b (LG_7)*	0.7596	0.2809	0.0020*
HmRS129 (LG_6)_HmidILL-2192 (LG_5)	0.4907	0.2526	0.0080*	*HmRS129 (LG_6)_HmidPS1.147 (LG_5)*	0.6147	0.2519	0.0210*
HmidPS1.870_HmidILL-2192 (LG_5)	0.3627	0.1366	0.0200*	*HmRS129 (LG_6)_Hmid610 (LG_12)*	0.6486	0.2533	0.0000*
HmidPS1.870_HmNS56 (LG_1)	0.5124	0.1867	0.0040*	*HmidPS1.870_HmNR120 (LG_10)*	0.70959	0.2425	0.0330*
HmidPS1.870_Hmid610 (LG_12)	0.6237	0.2009	0.0460*	*HmidPS1.870_HmidILL-87955 (LG_15)*	0.5144	0.2053	0.0200*
HmidPS1.870_Hmid65 (LG_3)	0.6759	0.2218	0.0180*	*HmidPS1.870_HmidPS1.549 (LG_9)*	0.5269	0.2677	0.0470*
HmNR120 (LG_10)_HmG53 (LG_18)	0.8260	0.2983	0.0080*	*HmNR120 (LG_10)_HmNS18M*	0.6227	0.3079	0.0340*
HmNR120 (LG_10)_HmNS17b (LG_7)	0.8219	0.3495	0.0150*	*HmNR120(LG_10)_HmidPS1.874*	0.7634	0.2242	0.0280*
HmNR120 (LG_10)_HmNS18M	0.7287	0.5173	0.0270*	*HmG53 (LG_18)_HmidPS1.549 (LG_9)*	0.7349	0.3365	0.0150*
HmNR120 (LG_10)_HmidPS1.549 (LG_9)	0.6105	0.2064	0.0220*	*HmG53 (LG_18)_Hmid65 (LG_3)*	0.8589	0.3448	0.0340*
HmG53 (LG_18)_HmNS58 (LG_9)	0.7245	0.3165	0.0220*	*HmNR106_HmidPS1.147 (LG_5)*	0.4992	0.2343	0.0390*

HmG53 (LG_18)_HmidPS1.629	0.6639	0.2955	0.0090*	*HmNR106_HmidILL-146360*	0.7771	0.7264	0.0000*
HmG53 (LG_18)_Hmid65 (LG_3)	0.7896	0.3149	0.0350*	*HmNR106_HmidILL-64192*	0.6580	0.4849	0.0130*
HmNR106_HmNR224	0.7000	0.4410	0.0140*	*HmDL50_HmidILL-140858*	0.5754	0.2528	0.0110*
HmDL50_HmNS17b (LG_7)	0.8000	0.3361	0.0020*	*HmDL50_HmidPS1.147 (LG_5)*	0.6272	0.2147	0.0220*
HmDL50_HmNS58 (LG_9)	0.643	0.2302	0.0360*	*HmNR224_HmidPS1.549 (LG_9)*	0.6687	0.4269	0.0410*
HmDL50_HmidILL-140858	0.5066	0.2238	0.0391*	*HmNR224_HmidILL-88398*	0.6832	0.4667	0.0190*
HmDL50_HmidPS1.147 (LG_5)	0.6899	0.2585	0.0220*	*HmNS17b (LG_7)_HmidILL-140858*	0.6012	0.3103	0.0310*
HmDL50_HmidPS1.559 (LG_18)	0.52283	0.2082	0.0480*	*HmNS17b (LG_7)_HmidILL-87955 (LG_15)*	0.6103	0.2583	0.0160*
HmNR224_HmNS17b (LG_7)	0.8811	0.3523	0.0470*	*HmNS17b (LG_7)_HmidPS1.147 (LG_5)*	0.7088	0.3174	0.0480*
HmNR224_HmidILL-140858	0.6572	0.3934	0.0210*	*HmNS17b (LG_7)_HmidPS1.549 (LG_9)*	0.6324	0.3117	0.0140*
HmNR224_HmidILL-87955 (LG_15)	0.6853	0.3779	0.0050*	*HmNS19 (LG_1)_HmNS58 (LG_9)*	0.6355	0.2254	0.0190*
HmNR224_HmidPS1.561	0.6754	0.5679	0.0320*	*HmNS19 (LG_1)_HmidILL-87955 (LG_15)*	0.5929	0.2133	0.0220*
HmNS17b (LG_7)_HmNS18M	0.7318	0.3878	0.0080*	*HmNS19 (LG_1)_HmNS56 (LG_1)*	0.6977	0.2417	0.0160*
HmNS17b (LG_7)_HmNS58 (LG_9)	0.7459	0.3315	0.0000*	*HmNS19 (LG_1)_HmidPS1.874*	0.7641	0.2912	0.0290*
HmNS17b (LG_7)_HmidILL-87955 (LG_15)	0.6501	0.4566	0.0050*	*HmNSS1H_HmidPS1.147 (LG_5)*	0.6084	0.1820	0.0250*
HmNS17b (LG_7)_HmidPS1.147 (LG_5)	0.8071	0.4596	0.0030*	*HmNS18_HmidPS1.561*	0.9251	0.2716	0.0160*
HmNS17b (LG_7)_HmD61 (LG_2)	0.7207	0.3484	0.0300*	*HmidILL-140858_HmidILL-76149 (LG_2)*	0.2884	0.1168	0.0060*
HmNS17b (LG_7)_Hmid610 (LG_12)	0.7491	0.3908	0.0100*	*HmidILL-140858_HmNS56 (LG_1)*	0.5701	0.2883	0.0010*
HmNS17b (LG_7)_HmidPS1.561	0.9122	0.4430	0.0090*	*HmidILL-2192(LG_5)_HmidILL-87955 (LG_15)*	0.4078	0.0852	0.0190*

HmNS17b (LG_7)_Hmid65 (LG_3)	0.8276	0.2808	0.0310*	*HmidILL-70036_HmidILL-76149 (LG_2)*	0.4244	0.1590	0.0471*
HmNS19 (LG_1)_HmidILL-146360	0.9870	0.7467	0.0130*	*HmidILL-70036_HmD61 (LG_2)*	0.5112	0.1969	0.0170*
HmNSS1H_HmidPS1.561	0.7266	0.3975	0.0380*	*HmidILL-70036_HmidILL-88398*	0.4802	0.2167	0.0230*
HmNSS1H_HmidPS1.549 (LG_9)	0.6008	0.3521	0.0380*	*HmidPS1.147 (LG_5)_HmD61 (LG_2)*	0.5442	0.1850	0.0330*
HmNS18_HmNS56 (LG_1)	0.5978	0.2169	0.0310*	*HmidPS1.147 (LG_5)_HmidILL-146360*	0.7927	0.4116	0.0360*
HmNS58 (LG_9)_HmidILL-87955 (LG_15)	0.4522	0.1608	0.0070*	*HmidPS1.147 (LG_5)_HmidILL-88398*	0.5379	0.2753	0.0310*
HmidILL-140858_HmidILL-87955 (LG_15)	0.4418	0.2762	0.0000*	*HmidPS1.559_HmidILL-88398*	0.4668	0.2017	0.0060*
HmidILL-140858_HmD61 (LG_2)	0.5012	0.3239	0.0230*	*HmNS56 (LG_1)_Hmid65 (LG_3)*	0.6613	0.2886	0.0120*
HmidILL-2192 (LG_5)_HmidPS1.559 (LG_18)	0.4114	0.0929	0.0130*	*HmidPS1.629_HmidILL-146360*	0.5412	0.1359	0.0333*
HmidILL-87955 (LG_15)_Hmid65 (LG_3)	0.5932	0.2981	0.0120*	*Hmid610 (LG_12)_HmidPS1.874*	0.6745	0.2612	0.0250*
HmidILL-118779_HmNS56 (LG_1)	0.6816	0.2864	0.0170*	*HmidILL-146360_HmidILL-64192*	0.6112	0.3982	0.0000*
HmidILL-70036_Hmid65 (LG_3)	0.6948	0.2555	0.0010*				
HmidILL-70036_HmidILL-88398	0.5059	0.2161	0.0050*				
HmidPS1.147 (LG_5)_Hmid65 (LG_3)	0.7243	0.2745	0.0040*				
HmidPS1.559 (LG_18)_HmidPS1.561	0.8735	0.2389	0.0290*				
HmNS56 (LG_1)_Hmid65 (LG_3)	0.6019	0.3198	0.0050*				
HmidPS1.629_HmidPS1.874	0.5758	0.3934	0.0230*				
HmD61 (LG_2)_Hmid65 (LG_3)	0.6411	0.2871	0.0410*				
HmD61 (LG_2)_HmidILL-146360	0.9868	0.6201	0.0390*				

HmidPS1.874_Hmid65 (LG_3)	0.7467	0.2124	0.0350*

** Statistically significant at the 5% nominal level.*

Table S3.5: Distance-based association analysis of candidate loci under selection with domestication and particular population.

	Marker	Wild/Cultured Domestication		CPWC		CPSC		CPEC		WPWC		WPSC		WPEC	
		Prevosti's Distance	*P*-value	Prevosti's Distance	*P*-value	Prevosti's Distance	*P*-value	Prevosti's Distance	*P*-value	Prevosti's Distance	*P*-value	Prevosti's Distance	*P*-value	Prevosti's Distance	*P*-value
Candidates for directional selection	*HmLCS48*	0.99	0.00	0.64	0.00	0.67	0.00	0.62	0.00	0.71	0.00	0.70	0.00	0.75	0.00
	HmLCS5	0.18	ns	0.26	0.02	0.87	0.00	0.00	ns	0.17	ns	0.19	ns	0.21	ns
	HmNR106	0.98	0.00	0.60	0.00	0.63	0.00	0.60	0.00	0.81	0.00	0.78	0.00	0.85	0.00
	HmNSS1H	0.78	0.00	0.74	ns	0.91	0.01	0.76	ns	0.63	ns	0.63	ns	0.65	0.04
	HmNS18	0.72	0.00	0.61	0.01	0.88	0.00	0.56	0.00	0.34	ns	0.41	0.02	0.43	0.02
	HmidILL-140858	0.89	0.00	0.58	0.00	0.62	0.00	0.62	0.00	0.56	0.00	0.57	0.00	0.64	0.00
	HmidILL-2192	0.26	0.00	0.12	ns	0.19	ns	0.67	0.00	0.11	ns	0.23	ns	0.29	0.00
	HmidILL-87955	0.73	0.00	0.64	0.00	0.64	0.00	0.56	0.03	0.76	0.00	0.79	0.00	0.44	ns
	HmidPS1.559	0.58	0.00	0.72	0.00	0.53	0.01	0.33	ns	0.56	0.00	0.49	0.03	0.53	0.01
	HmidPS1.561	0.09	ns	0.25	0.00	0.07	ns	0.07	ns	0.06	ns	0.07	ns	0.06	ns
	HmidILL-146360	0.09	0.01	0.07	ns	0.06	ns	0.06	ns	0.25	0.00	0.06	ns	0.07	ns
	HmidILL-	0.43	0.00	0.31	0.02	0.34	0.02	0.31	0.01	0.52	0.00	0.52	0.00	0.52	0.00

	88398													
	HmidILL-64192	0.17	0.00	0.10	ns	0.10	ns	0.10	ns	0.43	0.00	0.08	ns	0.10	ns
	HmidILL-37506	0.46	0.00	0.42	0.027	0.40	0.046	0.36	ns	0.35	ns	0.42	0.08	0.41	0.04
Candidates for balancing selection	*HmLCS37*	0.00	ns	0.00	ns	0.00	ns	0.00	ns	0.95	ns	0.94	ns	0.98	0.01
	HmNS56	0.58	0.00	0.50	ns	0.58	ns	0.60	ns	0.58	ns	0.56	ns	0.60	ns
	HmRS117	0.76	ns	0.88	ns	0.91	ns	0.81	ns	0.79	ns	0.83	ns	0.78	ns
	HmAD102	0.74	ns	0.85	ns	0.85	ns	0.90	ns	0.87	ns	0.83	ns	0.87	ns
	HmRS27	0.90	0.04	0.92	ns	0.95	ns	0.96	ns	0.96	ns	0.94	ns	0.97	0.04
	HmRS80	0.70	ns	0.76	ns	0.80	ns	0.75	ns	0.79	ns	0.78	ns	0.75	ns
	HmD59	0.64	ns	0.86	0.00	0.76	ns	0.69	ns	0.81	0.05	0.74	ns	0.73	ns
	HmRS129	0.70	ns	0.78	ns	0.81	ns	0.79	ns	0.77	ns	0.84	ns	0.76	ns
	HmPS1.870	0.61	ns	0.65	ns	0.79	0.04	0.76	ns	0.73	ns	0.73	ns	0.73	ns
	HmNR120	0.74	ns	0.85	ns	0.94	0.00	0.91	ns	0.88	ns	0.91	ns	0.90	ns
	HmG53	0.88	ns	0.95	ns	0.94	ns	0.93	ns	0.95	ns	0.95	ns	0.95	ns
	HmDL50	0.69	ns	0.81	ns	0.79	ns	0.81	ns	0.79	ns	0.73	ns	0.80	ns
	HmNR224	0.81	ns	0.93	ns	0.92	ns	0.93	ns	0.95	0.05	0.95	ns	0.93	ns
	HmNS17b	0.81	ns	0.90	ns	0.91	ns	0.94	0.05	0.85	ns	0.92	ns	0.89	ns
	HmNS19	0.83	0.03	0.91	ns	0.93	ns	0.93	ns	0.91	ns	0.85	ns	0.81	ns
	HmNS58	0.51	ns	0.58	ns	0.63	ns	0.68	ns	0.62	ns	0.63	ns	0.59	ns
	HmidILL-118779	0.04	ns	0.05	ns	0.04	ns	0.04	ns	0.03	ns	0.04	ns	0.08	ns

HmidILL-70036	0.50	ns	0.68	ns	0.66	ns	0.64	ns	0.64	ns	0.57	ns	0.59	ns
HmidILL-76149	0.24	ns	0.31	ns	0.16	ns	0.34	ns	0.30	ns	0.22	ns	0.33	ns
HmidPS1.147	0.62	ns	0.78	ns	0.74	ns	0.80	ns	0.79	0.04	0.76	ns	0.73	ns
Hmid563	0.63	ns	0.71	ns	0.80	0.03	0.77	ns	0.72	ns	0.77	ns	0.68	ns
HmidPS1.629	0.25	ns	0.34	ns	0.24	ns	0.23	ns	0.22	ns	0.43	0.01	0.18	ns
HmD61	0.49	ns	0.66	ns	0.60	ns	0.68	0.04	0.59	ns	0.60	ns	0.59	ns
Hmid610	0.49	ns	0.57	ns	0.56	ns	0.54	ns	0.65	ns	0.55	ns	0.50	ns
HmidPS1.549	0.42	ns	0.39	ns	0.55	ns	0.48	ns	0.53	ns	0.48	ns	0.51	ns
HmidPS1.874	0.80	0.00	0.79	ns	0.82	ns	0.86	ns	0.87	ns	0.85	ns	0.85	ns
Hmid65	0.79	ns	0.90	ns	0.85	ns	0.94	0.03	0.88	ns	0.96	0.02	0.89	ns

ns: Not significant.

Appendix C

Supplementary Information for Chapter 4

Figure S4.1: Distribution of contigs with regards to biological process.

Figure S4.2: Distribution of contigs with regards to cellular components.

Figure S4.3: Distribution of contigs with regards to molecular function.

Table S4.1: Nucleotide probes used in the 192-plex BeadXpress® SNP assay.

Table S4.2: Diversity statistics, minor allele frequency (MAF), observed- (H_o) and expected (H_e) heterozygosity and F_{is} values, including Hardy-Weinberg equilibrium *P*-value per locus per population.

Table S4.3: Summary of outlier results, numerical values highlighted denotes loci putatively under selection as indicated by the respective tests. Where the locus name has been highlighted it shows a locus that has been identified as a locus under selection by both test methods across any of the population cohorts (Bold: locus under directional selection; Underlined: locus under balancing selection).

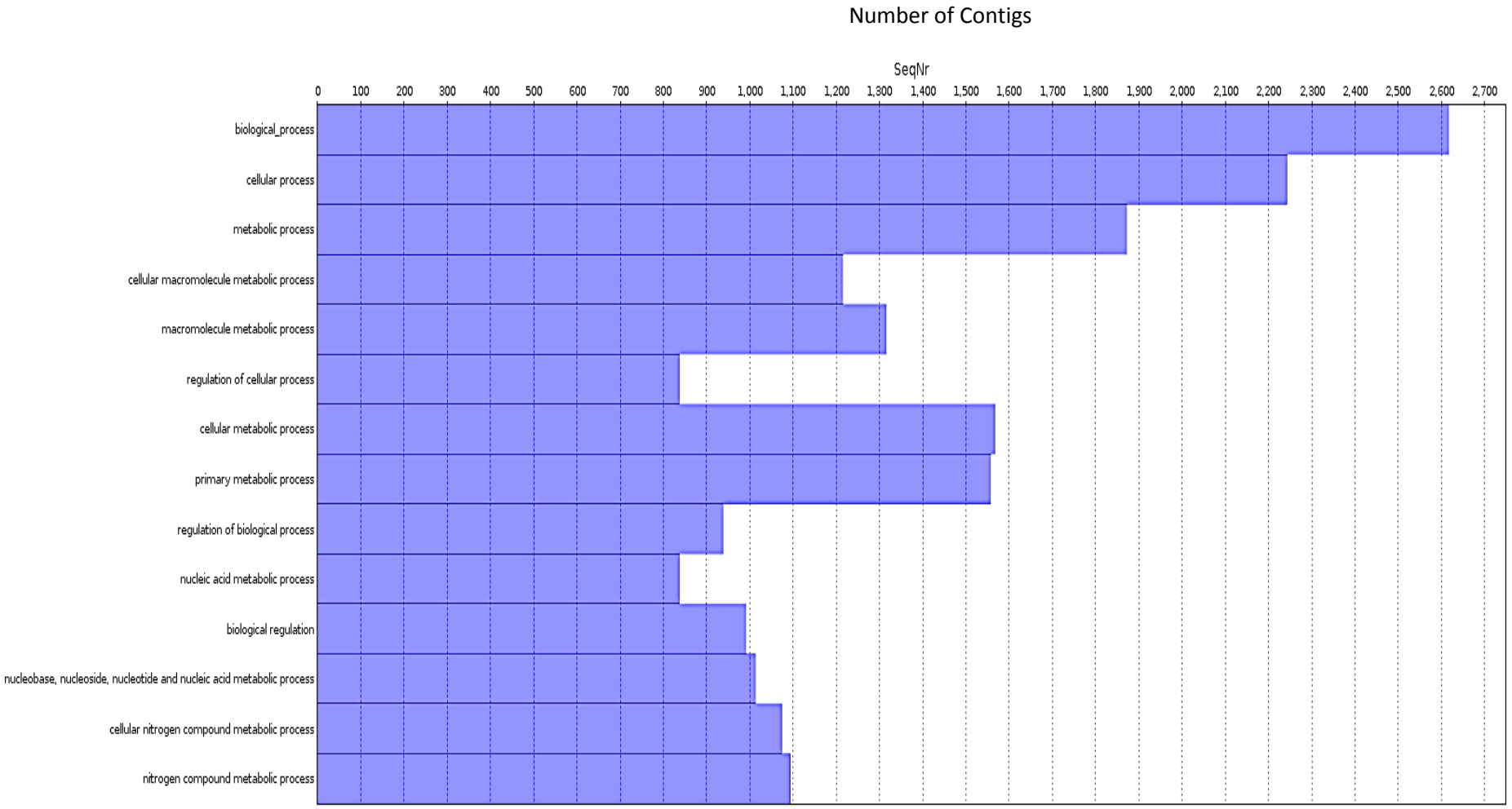

Figure S4.1: Distribution of contigs with regards to biological process.

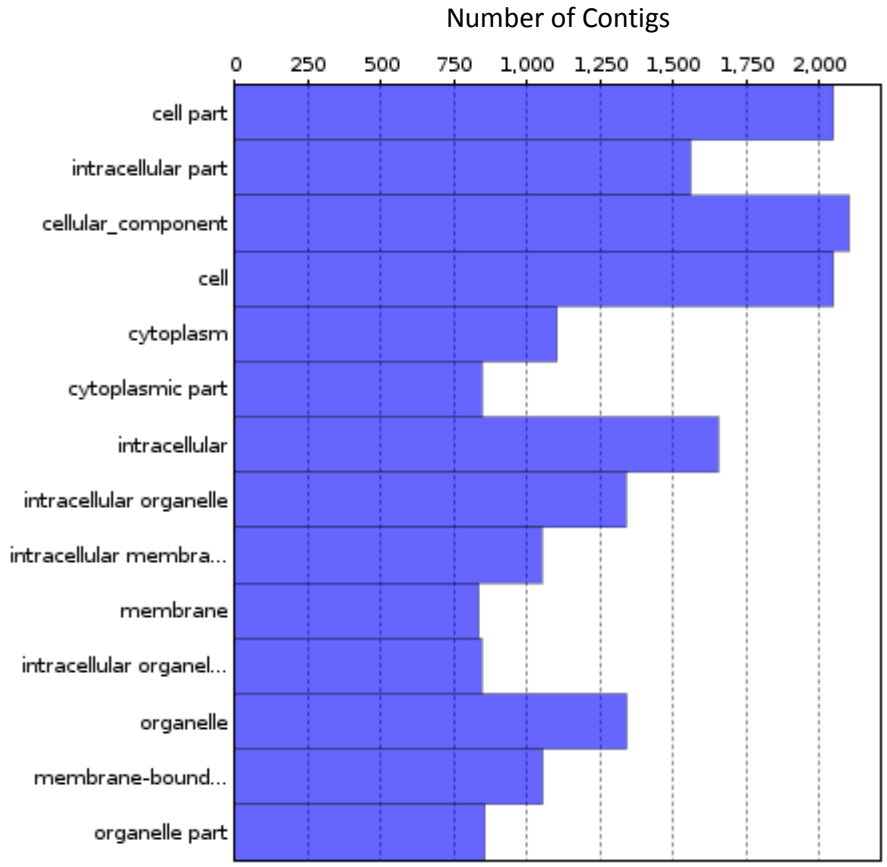

Figure S4.2: Distribution of contigs with regards to cellular components.

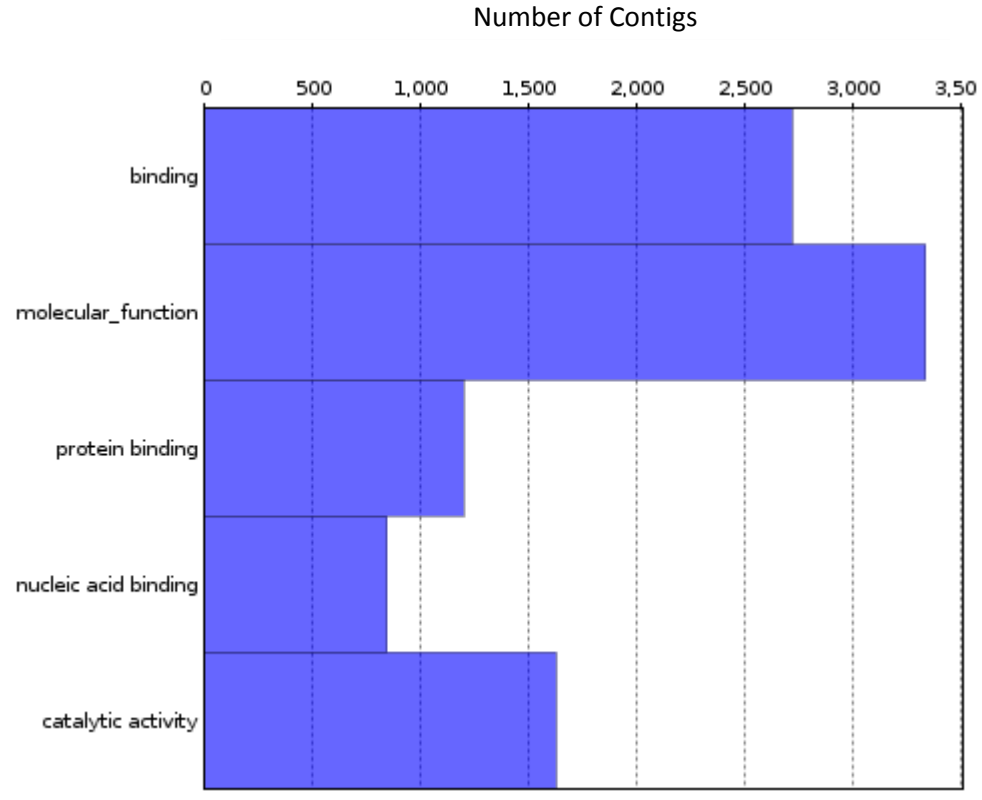

Figure S4. 3: Distribution of contigs with regards to molecular function

Table S4. 1: Nucleotide probes used in the 192-plex BeadXpress® SNP assay.

Locus	Nucleotide probe sequence
ILL_C2122_257_[A/G]	GACCGCGAGTACTATGACAGACTGTACCTCCACAACCTGAATCTGGAGTACCGTTTGAAC[T/C]TGCGCACCCGCAACTGTACCATCACCACTCTCACCCGTCCCTGGATTCCATTCGGAGTCCCC
PS_C15402_271_[T/A]	TTATGGATAAACGTGAATGCTATGAACGTCATTGAACATTTTCTACGTGTTTCTCTGTAT[A/T]TCAATTTGCACGTCTGTGTATAATGTGTAATCTTTATAAGCCATATCCGAAAGAAAATGC
ILL_C253_1545_[A/C]	TATATCCATCTCTCTAATCAGTAACTTACCACAGATATCCTCTTGCTCATGAGAGTAATT[T/G]TACTGGATGGATATTTTAAGCACCGCAGTTAGTCAACTCTTTATGAAGTCATATTACTCTA
ILL_C428_225_[A/G]	AGGCCACACACAGAACAGTCGACAACAGTACTTGAACAGTACCAGGCCTTACAAGATGTC[T/C]GCAGACACTCATTTGAAGCTGTCATCTATATTTATTCGTAAGATTGCAATTATTTCATATT
PS_C23094_578_[G/C]	AGTGTCCGAGCAGCGCAAGACTGACCAGAAGGAGGTTGACAAGCAGCTGCTGGATGTCCT[C/G]AGGAAGAGTCCCGAGAAGAAGCTCCTGTTTGGATATCTCGGCTCCATGTTCTCACTCCGG
ILL_C20682_843_[A/T]	AAGCAGAGGAGATGAATACTGTGTAATGTACAAGAGTCATGTCAGTTGAAGCTACACGAAA[A/T]AGACATTTTGTTACCATGGTAACATCTCTGGCTACACTACAAGTGTCATCTTTGTACAAAGT
PS_C34490_403_[A/G]	TCTGCAGGACCAAAGTGGAGCTTTGATTAAAGTATTCAGTGACCAGGCAGAAGGTGGAAG[A/G]GTTCGGATCCAGTTGTCTGGAGAGAGTCAACACAGACAAGCAGCAGAAGTGTTGATCAGG
PS_C25208 _377_[T/G]	TTCGGCAGGTGAGTTGTTACACACTCCTTAGCGGGTTCCGACTTCCATGGCCACCGTCCT[T/G]CTGTCTATATCGACCAACACCTTTTATGGTCTCTAACTGTATGTACGTCTAGTCGTGTTA
ILL_C1783_492_[T/G]	GGCGAAGACTAGGCAACAATAGTATGTACAATGTCTTTAAACATCGGTCAGCCTTTTTCGAA[T/G]CAGCACCAGTCTCTTATCACAGTGTAGCCTGGAGGCGATTTGCATATGATTTCTTTCCGTGTA
ILL_C929_734_[T/C]	GCCACAGTCAGAACAGTGCACCAATTTGATCAATACCGGTAACAAATGTTTGTTCTTCGAAC[A/G]CTTCAACACACAAACGACGCTCTACGTCTACTAGAATCACAAGTGTTTCTGGGTACTGACAGA
PS_C11665_287_[A/G]	ACTTTGTCGAAACTCTTTAGCAATTTCCTTCTTCCTGTCCGCTGTTTCTGGCTTGTTTTC[A/G]ATCATACAAACAACCCTCCACGACGATCTTCTAGCGCCGACAACATTCTTAAATGCCACT
ILL_C2141_350_[T/A]	GTGGAGGGATGCTCCCCAGCTTGTCCAGCATCATCCTGGCTTCCTTAATCTGTCCCTGAAG[A/T]ACTCCAACGAAAGCTTTACTTTCAGTATCAGCCTTCTTCTCCTGTGCATCTACCTCCTGAA
PS_C23591_200_[T/C]	GCCTTCTACACAGTATTCACGCCCTCTTACTCATCATAACTTCTATTTGCTATGATGGTT[A/G]AGTATCGGTATCACGCTTAGCGCCATCAATTTTCAGGGCTGGTTAATTCGGCTGGTGAGT
PS_C15689_162_[A/T]	CTTACCTGTGTTCCTCCAGGAAGTGTAATGGTTCTTCGTCTTGTAATCACGTGACCGGGA[A/T]GTGTGAACGTGGCTGTGTACCCGGATTTAAGCATGTTGATTGCGCAGAAGCATGCATACA
ILL_C229_2772_[A/G]	TCTGCTGGGGAACTACCATCTTGGGTACAACCAAAGATACTTTCATGACATCACAGCTAAAA[T/C]GGCTTTGCATGGATTGTTATTATCCTGTGTAGTCAGTGTGTCATTGCTGTTGCCCATTGTTG
ILL_C22574_507_[A/C]	ACAAGATCCTGTCTGAGATCAAGAACGTCGTCAATCTTCACGACATCTTCCAGAACTCCCTT[T/G]TGGAGACCCTCATTAAGAACCCAGGACAGTTCCTCATCAAGGCTGCCGAGGGTCTTATCAGCA
PS_C15088_268_[A/G]	TGGTGGATAGAGGGGTACTCGAAGATGATGATGATACTAATGATGATGATGACGAATGTC[T/C]GTGCGTGTAGGTTGTGTGACTGGCAGGCACGGCACCCCCCTCTGTTCACTGTTGTTTGGA
ILL_C2406_641_[T/C]	GTGCAACACTTTTACGGGGATATCTGGGAGCTCTGGGCAACTTTATGGTCTTAGGCCTCCT[A/G]AATTTGACGGAAGTCCTTGTTTTCCTGTTTCTTTTATCATGGACACCACGTAGGACGGCTT
PS_C12925_666_[A/G]	GGAGATGGACTAGACACACAGAGTCAGGACTGACATCACCAGTATTGAAATAAGACTTGT[T/C]AAATGTGTGTGAAATATTGTATGGATGTTGATGGATACATGACAAATCTATAATTTTGTG
PS_C11984_159_[T/G]	TTTTCTCTGCTGCTGATTTAGGAGATGTTGTTGGATGTCTCTTGGCATCTGTGGATGATG[T/G]TTGTTTATCATTATGCTTCTTCTCTGGGCGTCTAGCTTCTCTTCTAGCACGTTCGTCATG
ILL_C22449_261_[A/G]	CAGGGAAGATTGACGTCAACAGCTTTATTAGGAACCAGAGGAATTCTGGGCAGTGGTCCTC[A/G]TCTTCTTCACTGTCAGACAGTCGGGATCTGTCCAGCGACCCTGAGATGGGGAGTGTTGAATCT
PS_C11970_157_[A/G]	TGAACAGCAGTTAAAGGAGCTCGAAGAACGACTCTTGGATGAAGAAGATGCAAATGCTGA[T/C]GCGTCTGAGAGGAGGAAGAAGATGGAGGGCGAAATCGACGATCTCAAGAAAGACTGTGAG
PS_C23051_368_[A/G]	CTGAACACGACTGAACAAGTGTACAGATCTGTACACCGACCACAACAAAAATGCCATTCA[A/G]TAAGTTTTTAAGTCAACATCACAGTGGCGGAATAATCAATTCCAGTCAATTGAGTAAAGA
ILL_C6012_280_[T/C]	TAGCCCCTGTCAACAACAATGTGTCATCATTTATCTCACAGATCCGAACATTTCAAACAAG[T/C]GCTGTACAACGTGATATTGATCAGGCCGCCAAATACATTGGTGCTGGAGCTGCCACAGTAGG
ILL_C2903_1043_[T/A]	GAGCCAGGAAGCCATCTTGGTTTTTGACAGTGTTATTGTAAGCTATCAAGTATTCGTTTTA[A/T]AAAAAGGAATGGCTAGAAAGCCCAAAACATGGTAATAGTCCACAAGAATTCTCCCAGTGACT

PS_C34501_638_[A/G]	GATGACTAAAAGGGCTCAGGTTCCGTTCTCTAGGTGGTTGCCAGCTGCTATGGCGGCGCC[A/G]ACTCCCGTTTCTGCTTTAGTTCATTCGTCTACCCTGGTTACTGCTGGAGTGTTTCTCTTT
PS_C12352_527_[T/C]	CTGGCTTCACTAGTTCTTCCTCTTTACCGGAACTGACATTGTCTGGCTTCTCTAGTTCTT[T/C]CTCTACTGATTCTTCTTCCTTAGGTTTAAGCTCTATATGTTCCTCTTTGCTCCCTCGATC
PS_C34725_229_[T/C]	TTCGAATGTATATTCTCTCATGATGGCCTCCAGACCAGCTCTGGCCACTCGATCTTCTCG[T/C]TTTCCATCCTCATCGAGTTCATCATCATCTTCCAGAGGTTTGGTTTCTTCTTCTCCCTCT
PS_C15230_93_[A/G]	TGGTCTTACAAACTTCCTCTACTTGTGCTACCTTCCGAAGACGCTGGAGCCTGACCAGGG[A/G]GAGCCGACATGTGTGTTGCTACGTATTTATGGACACATTGCCAAGAGCAGTAAGGAGTTT
PS_C23075_525_[G/C]	CAATAAGGATGGCGTAATTGACCAGTTGGAAGCCGAGAAAGTTCTTGATGAGCGCGCAGT[C/G]AAGGACTTTCTTCCTCTTGTGGATGAAGACGGCAATTCTCAAGTGAGCGTTGAAGAGTTC
ILL_C3835_411_[A/G]	ATCGGACCATCACCTTGACGGCTACCCCGCTCGCCGTTGTCTGCAGAGCTTCAAGAGACTT[T/C]TCCAGTGGGAATCGGTGGGTGATGAGAGGCTTTACGTTCACCTTTCCACTGGCGACCAGAG
PS_C24743_123_[T/G]	TGTTCCACTTCCGCCGTAGCCTCCATTGCCTCCGCCGTAGCCTCCATTGCCTCCACCGCG[T/G]CCTCCATTGCCTCCTGTACTACCACCACTTCCATGTTTCCACCTTGTTCCTCTTCCACCT
ILL_C5634_234_[A/G]	AGAAAATGACAAGTGTAGTTTCAGCTTCTTGTTGGGTGTTCCCTCTAGTTCCGAGTGTGGC[A/G]CCGGTTTACACTGTAGCTTGAAGACCGTGACTTGTGTCAAAACTGTCAAAAGACAAGTGGC
PS_C23630_237_[T/C]	GGCTGACTTGATGCATTTATTGGAAAATGTCGACACTAGGATATACATGAACTTAAATCT[A/G]TGAATGAAAATGCATGCCTTTTACTTATGAGTGTGCTCATGATAAAAGCATACTTCTTGA
ILL_C1813_300_[T/G]	ATAGGGTCCATCTTGTGATTAATTCTGATACACCTGTAGCCACTCCCTTTGAAGGGCTTTTC[T/G]GGGAACCAGTGATTTTGATAGTGGTTGCACAGAAGATTTTGCAGATGGGTGCGAAAAGAGTC
PS_C12218_188_[A/G]	GTGCTCCTCACGTTCTTTCTCCAGCCTAGCAAGTTCTGCATCTGGATCAGGATGTCCTTC[T/C]TCCATGAGCTTCCTTCGCAGTTCTTCATATTCTTCCTTTTCCCTCTGTCTATCTCTGGAG
PS_C36136_185_[T/C]	GAATGTCCCATTCAGAGTCCGTGTCCGCCTTGCCCGTAAGAGGAACGAAGATGAAGACTC[A/G]ACACACAGACTGTACACCCTCGTCACCTATGTTCCCTGTGTGGAATTTAAAGGTAAACAG
PS_C23070_1364_[A/G]	ACCTCCAGTGTGCAAAATAAAGGTTAATGTCACCTGATTTCAACGCTGTTATTTACCCTG[T/C]TTGCGGTAGCGACAGATATGTGTAGTGTAATGCATGGTCAGCTTGGGCTTGGTGTCCGCG
ILL_C618_116_[A/G]	AGAGTGAGAATTCAACCATATTGCAGCTGTTCACATACATACAGTCTCAATCACACAAGG[T/C]GTCACAGCAGTGACCAACAGTGAGGTATTGCCCCCATGGTGATGATGAGGACGACGAGAG
PS_C1652_228_[A/C]	GCTAGGTTGGGGTTGGCAGTTCTTGTTTCTCTGATTCGGACTCATGGGAATGATTACGTT[A/C]ACAGGTTTAATGGGTATAAGTGTTAGGGCTTTTGTCTGTTGGTTTGACTTTGTCGGTTGG
PS_C25083_285_[A/G]	CGAAAGCTTGAATGTGGTGTTATTGCGAGCTATGAAATTTTTACAGTTTCCCCAGATCAA[T/C]TCGATTGGATTTAGCTCACAATGTCTGACAGGTGTTCTCAAAACAGCGTGGCCATGATTT
ILL_C394_1510_[A/G]	CTTGTGTAGCTTTATGGTAGCTGTGGTCAGTTAAAGAATAGAGAGTGTTCGAGAAGTTCA[T/C]ACACAGGTTGTGCAATGTTTGCATCTCTGATACATGCTCAGGTATATTTTGATCACGCCA
PS_C24267_71_[A/G]	TGTGTCTGAGTCTCTGAGTGTGTGGGTGGTGGTGCGTGGCCGTTCTTGGTTGGTGGGGTG[A/G]TTTGTCTGGTTAATTCCGATAACGAACGAGACTCTAGCCTGCTAACTAGTTCGCCGACAG
PS_C34507_1191_[T/C]	GAAGAAACTTGAACAGAAGGGTTCAAATTCCTCTGCTAGAATTGCAGAGCTTGAAAAGCA[T/C]GAGAAGGAACTGATGGAAAAAGTGTCTCAATATGCTAAGACAGATTCCACATCAAAACAG
PS_C36273_73_[A/C]	ACACCAAAGCTTGGCTTCGTCGGAACTGCATTCAGGTAGTGCAGTTTCCGGCCGCTTCAC[A/C]AGATCTGAACCCCATCGAAAATGTCTGGCAAATCATCAAGGACAAAGTGGAAAAGTTGGA
PS_C35977_153_[T/C]	AGGCAAAGACGCAGTAATTGTATAAAGCATCATATACATCCCAGCCTGCAATCGCTCAGG[T/C]TGATACCCCCATCCTAAAATCAAAATAAAAGTAGGCACTAAAGACGCCTCAAAAAAGATA
PS_C16093_142_[T/C]	TGGACGCCTGATTGCTGATGGCGTTGGTGTGCAGTACAGACCCAACAAGGGACCTCTGAA[T/C]AAGTGGAGGAAGGACCAGACCAACCTCAGGGCCTAGACACTGTAGGAACAGACATGGTCA
PS_C16031_147_[A/G]	TATATCGTCCTCATCAATCTGTTTCCATTTCTCTGGAAGGATGTCCCACAGATCGTCCTT[A/G]ACATCGCCTTGGATGACTATCTCATCCTCGCCCATCACTGACGACCCACAGGAAAACTTG
ILL_C5106_273_[T/C]	CTTCTTTCAACTCCATTAGCTGTTTGTCAGCATTCTCCAAATGTGTTACTTCAACCTCCAA[A/G]TTTCGTAGACTGTCAAGTGCGTTGGCACGCATTATTTCTGTCCTATCACATTCCTCCTTCCA
PS_C00512_245_[T/C]	CACAGATTCACGGACTCTATCTGGTTTCGCCACCCAGGCGTACGATCAACAGAACAATCC[A/G]ACACCAGCCCAGACAAGTCAGCACACCAGCAACAGCATAGCAGAACAGCAACAGCGTTAT
ILL_C140_2421_[A/G]	GGTATTGCTGGAAGTGACGTCAGTAAACAAGCTGCTGACATGATCTTGCTGGATGATAACTT[T/C]GCCTCCATTGTCACTGGTGTTGAAGAAGGTCGGCTGATCTTTGACAATTTGAAGAAGTCTAT
PS_C36237_70_[C/G]	TGATGTTGATGAATTGCTGGGCATGCTAGATGAAAACGGTGACGGTCACGTCAGACGGGA[C/G]GTCCTGGAAAGGTACACAGAGGAAGCGCTCGCACACAGCGAGTAATACTGTCGAAGGTAC
PS_C34670_302_[A/C]	GTGGTTTGAGAGATTGTGAATATATTGGTCTGTGTGGGTAGGTTTGCGATAGACTGACGT[T/G]ATGAGCTGGTTATTGGAGTTGTCAATGGCGATGTCCAAGAAAGGGAGTTTGTTCTGTTGC
ILL_C21880_1003_[G/C]	TCAGCCTCAACACCCTCGCTGCTAAGCAGCAGAGCATCGTCGGAGTCCCGAAGGGAAATCT[C/G]AATCAGCTGCAAGAACTTGTTGATGCAGTGTCTCGGGGAGAGCTCCGTGTTCCTCTGTACAA

PS_C47330_170_[A/C]	GCGGCAGCATCTACATCACCACTCAGCGGTCAACCGTGCAGTTTCGTGTTGACAATAGCA[A/C]CAACGTTGCGCATACTTTAAATGTATCGTACACAGCCACCACGTCCTTGAATGAGGTGAA
PS_C36522_249_[G/C]	CAGGGTCTTCAATCCAGTCACCAGAGTCTACATCCAGTCTCAGCAGGATCATCCTGTCAC[C/G]AGCACTTTCTTCAAGTCGCAGAAGCAACTTCCATTCATCCAGTCACACCAGGGTCTTCCT
PS_C34604_423_[A/T]	GAAGAAGAGGAGCAAAGAAGATGATGACAAGAGTCAGAGTAAGAGTGAGAAGAGTTTGGA[A/T]AACAGGAATGACAGTACAGAGCGAGAAAGACAAGGGGAGAGGAAACGTGATAGGAGTAGG
PS_C34501_77_[T/G]	GTCGGTTTAGTAGGGGTTTGGTTATCTGTGGTGGCTTGGCGTTCTCCTATTTTAGGAGAA[A/C]AAAAGGTTGGTTTCAGTTTTTTATCTGGGTTTTTTTCGTGAATGTGATTTTTAGTTCCTC
PS_C28810_290_[A/C]	AGAACACCTGTCAGACATTGTGAGCTAAATCCAATCGAGTTGATCTGGGGAAACTGTAAA[A/C]ATTTCATAGCTCGCAATAACACCACATTCAAGCTTTCGGACATCAAGGAACTGGCATATG
PS_C28886_317_[T/C]	CTTCTTCTGCATCAGATCACGGATGTAAGTTCTAATGATAGCTTTGTAATCAGAGTATGG[A/G]ATAAGAAGTGGTGTCACAGATTTGTTGAGTGCTGCCTTGGCAGCAAGATCGGCCATTGTG
ILL_C31_1387_[T/C]	AACTTAAATGTGGAAAACAAACAAAAACCAAATTTCTTACCAACAGTGTGCCCACACTTCA[T/C]GTCTCCAAAGATTTGTTCGGTATTATTTTTGTAAAGGTATATTATCAGGTGATGGATACTACC
ILL_C844_440_[T/C]	CACCTTTCCTGTCTACCAAGACACCCCGCACCTCGGACTTTGGGAGCGAGTCTACGGCGG[T/C]CGCAAGGATGATTTCTTCATATTTGACAGATGTGGACAACAGACCTACCATATTGAACTCC
PS_C46857_366_[T/C]	TTCCACGCAGTAGGAATTTGTGAGGTGTACCAGATTGTGTTAAGTAGCTTTAGCAATAGC[A/G]TTTTGCCGTGTAAGGAGAACTGTTTTAGTAGTATGTAAGTTATGTCATCTCCTCCTGGAG
PS_C34605_382_[A/G]	AACACTGACTTTTGGCTAGTTTATGCCAAATTAGAAGTAAAATTAATATGCATTCAATCT[T/C]ATTTCAGTCACCTGTATCATTCTCTTGGTATGTCGGGCGCGGTGTGCGCGTGCGTGTATA
ILL_C853_1199_[A/G]	CGTTCCAGGCTGACGAGGACAGGAAGAACCAGGAGAGACTCCAGGACACCATCGACAAACT[A/G]AACGGAAAGATCAAGACCTACAAGCGCCAAGTGGAGGAGGCTGAGGAGATCGCTGCAGTTAA
ILL_C1002_85_[T/A]	TGATGACAAAGTAATACAATTGTGTGATCTATTAAAAAGGCATACAGTTGTTAACTGGCAT[A/T]GGGCTTTTGATGTTCAAACATGAATGACCATGTCACTACAGGAAAAAGTCTGACAATGGCCG
PS_C11659_399_[T/C]	CAGTGAGCGCAGTGGTCGTGGAGGGCGCCGTGAACCACGCGAAATCCGTGAAAATGATGG[A/G]GATAGGAGGCCACCCAGGCGCCAAGGGCCACGAGAATACAGAGACAACCGTGGAGGCGAT
ILL_C4791_1099_[T/G]	GCGACGACTGGGCGATGCTAGGCGCTGGTGCAAGGTACGTCTACACCTACGAGCTGAGACC[A/C]AGGTCAAGCAACCCCGGCTTCACCATCGGCGCTGACAACATCATCCCCAACGGCTACGAGTTC
ILL_C300_4982_[A/G]	TGTCTTGCTCCAGTCTCTGTTTTGTTTGTAAGCCTCTCTTGACTTCAGAGTCACTGTTTCT[T/C]TCACAGGTTCAGATGATTTTGATTTTGGCGCTAATACTGAATCAGGTTTGCTTGTTTTGTGC
ILL_C4593_326_[T/G]	GCCTGATCGTATTTTGGTCCACCAATTTTTAAAACGATCCTTGACTCTCTCTCTAAACCAA[A/C]GCCTAGCTCTCCAGGCGCTTTCTGCCTCCGGTAGAACCACCAGCAACGCCACCAGCATCAA
PS_C46533_541_[T/C]	CATGTGCTTCCAGAATGTTGGTCAAGGTCGAGGTCACGGCGTCACTGCTCAGCGCTACAT[T/C]GACCAAGTGATCAGACCCCACGTTGTCCCATTTTTCGCACAACATCGTCAACACATTTTC
PS_C15351_193_[A/G]	GGGACTGAATCTCTTCTCTCTCTTGCTGATCTTGCGAGCCTTCTCGCTCCTGAAGACCTG[T/C]CTGATGGGCATGATGTCTCCCTTGAGTTGGGTGGCTGTCTTGATCTGTTCAGCCGTGGCA
PS_C15018_147_[T/G]	CTAACACACATTATCCCTTGAGTCACATGTCCTTGAAGATGTAGAGTATGCAAACACAGC[A/C]AGACCAACCAGTTTAGAGTTGTGCAGTTACTGCTTCCTTTAATTTACTCTCATTGGCCCC
PS_C38608_168_[T/C]	CTCAGAGCTAAATGCTCGAAACAAGGTCAAAGCCACCAATACTTTTGCTGTGCCAGTCCT[T/C]TCCTATTCCTTTGGAGTAGTTGATTGGACAAAACAAGACATCCAGAGTCTTGACCGCCTG
ILL_C387_215_[A/G]	ATATAGAACTCCATCCATGTCACCTTGACAGCTACGCTATGACGGGACCCCAGTGTGGCCT[A/G]GGATGCTGTTACTGCTCATGTCCTTTTCCCAGAATTAGTGAGGCCAGGAAACAGACTCTCG
ILL_C5339_366_[T/C]	ACTCATCCTGACTCATCTGCAGAGCCTCAAGAATCTTTCCAGAGACAACTTTGCCTTCAAC[A/G]AAGCCTTGTGGGATAGCATTAGGGGCCAAGATGGAGTATCTCTGTTTGAATTCAGAGTATAC
PS_C34511_71_[T/G]	AGGCAAAGAGACAGAGCGAGGTAGGAAATATTTTAGTGCCGGAGAGAGAGAAGTGGCTCC[T/G]CAGAGCGTTAAGATCTGGACAGAGGTATTGAAACTTGTGAAGAGAAACGGACTAGATAGT
PS_C14838_228_[T/C]	GAGTACACGGGGTATCTGATGGGTGGCCATTATAATCATGCTGCTAGTGAATATCTGTGT[A/G]TTGATGGAGATGCAGAGAAGGATCGCAGTGGCCATGAAGACAAAAACGGCCAACTCTTGT
ILL_C20267_102_[T/C]	TGACGTTGGCTGAGTACCTTGCTGCATCCGGAGTCCCCCAGGCTTCAGCTAAAGACAACTT[T/C]GACTACTACGATACCAACAAGGATGGCGATGTTACGAAAGCTGACGTGAAGAAAATATTCTC
PS_C46674_204_[C/G]	GTAGCACATCGCGAATAATTTACAATGCCTTCTTGCAGAGATTTTTCCTAAAACTACAAT[C/G]AATTGGTTCAACTCTGTTCCTCTGCTAGTTTCTCGGCCTTCTGGACAACTTCCTCGACTC
ILL_C1878_506_[T/C]	CAAAGAAGAGTTCACGGAGGAAATGTTGAAAGATGTCTTCTCAGAGTTTGGTGAAGTAACATC[A/G]GTTGATTTCATCATGGACAAGGCAACCGGCAAACCAAAGGGCTTCTGTTTCATTTCATATGAT
PS_C36563_85_[T/G]	AGGAAGGGAATGCAAGAAACACACCATGCACAAAGTCACCCAGTATAAGGCTGGCAAAGC[A/C]TCCCTCTATGCCCAAGGTAAGAGGCGATACGACAGAAAACAGTCCGGATATGGTGGTCAG
ILL_C2735_326_[T/G]	TGTCTGTGGATGTCTTGTATCATCTTACGCCTGCTGTGTATCTGGATTGTGCCAGAGTCTC[T/G]AGTATGGGCATGTAATGTTAATGCATCGCTGACTTTCAACATTGCCATGTCTAATGTAATA

ILL_C250_199_[T/C]	TTGGTCTATCATGGTGACAGCCTCCGTGAGCTTTCTGCCCTCCTTCAGCCTTGGTAACACT[T/C]GGAGGTACAGAATCTGAAGATTGGGCTTGGCCATTTACTACACTTCACTCCTGTTGCCTTC
PS_C47845_238_[T/C]	CAGCAGAGGCCTGAAGCGTAAGAGGCTGGGACTGATCAAGAAACTGAGGAAGGCTAAGAA[A/G]GAAGCAGGACCTCTTGAAAAACCCGAGGTGGTGAAGACCCATCTGAGAGACATGATCATT
ILL_C1363_269_[T/C]	CTGCCTCTGGCACGCTGGAGACCCTTCTCACTTCAGACATATCTGTGGTGACAGTAATTCC[A/G]CTGCTCTTGAAGTACTCATCCGCTGCCTTGTCCACCACCAGCAGTGTCGTCTCGTCACCTCC
ILL_C18774_676_[T/C]	GTCAGTTTTCACAAATACAAAACTTGCTTACTATTCAATTTTGGTCCTGCTAAGAAGCAAC[T/C]ATTAAAATACTCTATACAAGCTGGAAATATTTTTAATGCAAAGACTTTGGATTTGGCCAGG
PS_C23647_375_[A/G]	TGGAGCTAAAGAGTCCTTCACCAGACAATACACCGGGCAATCAGTTGACATGACATTTAC[T/C]AGTGACGGAAGTGTTGTGCGCAGCGGTTTCTCCTTTGATTATGTCACTGAAATTGTCTAC
ILL_C2028_1228_[A/T]	CGTGAAGGCTGAGATGACTCCGAAGAGGCCAGGCAGGATGTAGGCGAAGCCAACAGCCCACA[A/T]CATCATGAATCCATGATGTCTGTGGTGGTGCTCGTGATGGTGCTCGTGTCGTCCAGGTCGGTG
PS_C11985_171_[A/G]	TCGTTGCTCTCGTCGTGGAAGGTGCAAGATAAAGATCAAATATCGTTCTCTCGAAGATCC[T/C]ATCGAAGACAGTCCTGATGATCTCAACTCTGGTCTGATCACTGACCGAAATGTCGACTAT
PS_C11911_576_[A/C]	TGAGAAATCAAAGAAAGCTGGTGCAAAAGCGAAGACAAGTGCCAAGTCGCCGAAGAAGGT[A/C]AAGAAAAGTGTAGAAACGACACCGAAGCTCTCCAGGACGAAAAAGGCTGAAATGAAGACT
ILL_C1254_187_[A/G]	CCAACAGGGACATACCACGGAGATTCTGATCTGCAGTTAGAGAGAATCAACGTTTACTACAA[T/C]GAAGCAACCGGCGGAAAATATGTCCCACGTGCAATCCTTGTCGATTTGGAGCCCGGGACAAT
ILL_C20427_267_[T/C]	AGTTAATCTTTGCTCTCGGGACAGAGCTTTTGATAATGTTTAGATGAAGGTAACGTCTTTG[A/G]TTGACCTTATATAACAGACCAACGTGGAAGTCATGAGAATTGTGAGACGTTGTTTGGAGAGTG
PS_C21989_160_[T/G]	AACACGTTCCATTGCCTTTACAAACTCCTGGAAAGATATCATGGTGTCTTTGTCCTCATC[A/C]ACCTCCACGATTGTTCGGTCGGCAATGCTGCCTAACTGTTCTTCTGAAATGTTGGCCCCC
PS_C12069_1181_[T/C]	CTACTACTACTACTACTACTACTACGAGGATGATGACTACTACTGCTACTGCAACAACAA[T/C]AACAACAACTATGACGACTACTGCTACCTATTCCAAATGTGTTTCATCCATATGTCTTTG
ILL_C327_1076_[C/G]	AGAAGCGCCCCTAAGATCGCTCCTTTCGACAAGGAGTTCTACATCGTCTTGAACGTCGCCGT[C/G]GGCGGTTTCAACTACTTCAACGACTACAACAACACCGCCTACCCAAAGCCATGGACAGATG
PS_C11847_438_[T/C]	GCTCCGTCTGAAGCCAACAAGCAGGTACTGTCAGCTGGACCGCCTGTCAAGTGAGGTTGG[T/C]TGGATGTACCAGGGAGTGGTGGCTAAGCTTGAAGAAAAGAGGAAGGTCAAATCTAATGAT
PS_C47375_253_[A/G]	AGGGAACACTATAATCTCACTCAGCAAATTAATAGAAGGAGGCCCAGCCATATTTGCCAC[A/G]CAAAACACAAATCACCACATAGACATATATGGAGAAATCATCAATATCCCTTTACTTATT
PS_C34420_787_[A/G]	AGGGCAGAGGAGAGGAAGAGGGAGAGGGAGACCAAAAGGAGATGCTGGAGAAATTCCACA[A/G]GTTCATTGCTCACATGAAGGAAAAGAAGGAGGCTGAGAAGATGGAACAACAAAACGATGA
ILL_C2915_875_[T/C]	GCACTGCACATGCGCACTAACCGCACGCTTACTTCTGTTGTCTTGCGCTGCTCTTAATGTA[T/C]ATAGCTATCCAGCATGCTTTGTAGTACTTCTTGTTGATCTGTTTTCTTCTGTTGTTTCCAG
ILL_C22491_727_[T/C]	AAAGCCTGACCACCTGTTCCTGCACGAAGGGCTGATGTGAAACCAAATGACTCGTTTACTGG[A/G]AGGTAAGCCTTCTGTTGGCTCATGCTGGTGCCAGTAACATCGTTGTCCTCGAACACATGACCA
ILL_C911_1343_[T/A]	CTGTCTACAGCCTTGATACCTGTCTGCATAGGCTCTTTCACAGATGTACGTGGAATGATACC[A/T]GGAGCCTTCACGCCCACTCGAGCTCTCTTACCAGGAATTGCTCCTTTTCCATCAATGGCATTT
PS_C12196_434_[G/C]	TTCATTGAGTTGCATTGTGGAAGATACTTATCTGTCTTAAAGCGTCAATAGTGGCATGAA[C/G]ATTTTGGTATTAAATTTGGTTATTAATAGCAAAAGGTTGTCCCGCGAGCCGGATTTGAAC
PS_C14297_369_[T/G]	TCGCTACCAAATGAGGATCCACAAGCGAATCATCGATCTGCACAGTCCCTCTGAGATCGT[T/G]AAGCAGATCACCTCCATCAGCATCGAGCCAGGTGTAGAGGTCGAGGTTACCATTGCTGAC
PS_C47340_198_[A/G]	GCCATAAGCTGCACCTTGATCTGACATCCTGAGAAGCTAACAACTGCAAGACCCTTGGGC[A/G]CCACAGTACAACAAGCACAGCCACAGCCACCCTTAACAAAACAACAGCTTGACTTCATAG
PS_C36706_579_[T/C]	TTGAGTTCCACGACGGCCGTTTCAATGTTTGCTTCCACTCCGACGGATCAGTGACAGGAC[A/G]GGGTTTCCAGGCCACGTATGCCATGAATCCAGTTCAAGAGGTTGACTCTGATGTGACTAC
PS_C46597_301_[C/G]	CTTCTCAGCCAGTTGGTTAGAGTCTAACCAGCTAATGACATCGTTACACTTGTCAGTGAT[C/G]GTCTTCTTGTCATCTTCACTGATCTTGTCCTTCAGTTTTTCATCCTCTACAGTTGACTTC
ILL_C2040_1251_[A/T]	AGCCCTGGTTTGACTGGTCTTAAGGGAGCCGCCGGAGCCCAAGGTCAAGCTGGATACAAAGG[A/T]GAACAGGGTCTGAAGGGTTCCCCAGGTACTCCTGGAGACAGAGGACAACCAGGACCCCCAG
PS_C35683_190_[A/G]	TTGTGTGATGACAGTGATTGTGGTGATGATAACGATGAAGAAAACGAAGGAGAGGACAAT[A/G]GCGATGCTTTTAGTCAGAGTAACCCAACCTACATCCCTCCAAAATGATTTATAGGGCACA
PS_C46532_529_[A/T]	ATACTCTCCAACCAGCCCAAGCTACAGTCCAACAAGCCCCTCATACTCTCCAAGCTCTCC[A/T]AACTTCAGCCCTCAGTCACCTTCTTACTCTCCAACCAGCCCATCCTATAGCCCAACAAGC
PS_C15379_78_[A/C]	CCAGGACAACCTAGCAATTCGCATGGTCGCGAACACTGTCTGGAACTAGTAAAGGAGCAT[A/C]GAATATTTGTCTGGAATGATATCAAATGTGACGTTCCCTTACATTTTATTTGTGAAAAGG
PS_C11871_171_[A/G]	GAAGGTGGCTGCCATCAAAGCCCATCTGAAGATGATTCAAGAAGAACAGGAGAAGCTGAG[A/G]CAAGAAGAAGAGGAGAGAATCAGGCTGGAGGAAGAGGCGGAGAAGGCGCGTGAGGAACAG

PS_C35823_1199_[A/G]	GTATGTAGGGGGTTTGCTTGTAATATTTGCTTATGTGGCTGCGTTAGTGCCTAATAGAAT[T/C]TTTTTGGGGTTTTGAGTTTTCTTTGGTTTCTTCTGTGGGGCATGCCTGTCTGTCGGACTT
PS_C39731_379_[A/C]	CGGGTTCCGACTTCCATGGCCACCGTCCTGCTGTCTATATCGACCAACACCTTTTATGGT[A/C]TCTCATTAGCGTCAGTGTTAGGCGCCTTAACCGGACGTTTGGTTCATCCCACAGCGCCAG
ILL_C6061_1289_[T/G]	AATGTATGAAGTAATTGGCTTATTCTTATGAATATTTTTAGCTTACATACAAAAGGCAAGGC[T/G]TGTTAAGATTGTATGTGGCTCATGGGTTTTGTACAAAGGGAGATAACTCAGTCATGCATAGC
ILL_C980_261_[A/G]	ACGATTGTATCGTATTTATGTAAGCAGAAAGCCACTAGACTCGTGTTAAGGACAAGTTAT[T/C]GTCTTTACAATGCAGGAGAAGTGCCAGCCCTCTCATGTGCAGCAAGCTGTGTCAATGACGG
PS_C 451_[A/T]	TTGCAACGACCGAATAACGTCATCGGTAGGGTAAAACTAACCTGTCTCACGACGGTCTAA[A/T]CCCAGCTCACGTTCCCTATTAGTGGGTGAACAATCCAACGCTTGGCGAATTCTGCTTCGC
PS_C10066_[A/T]	TGAAACATATAGCTATAAGATACGACTATATACATGACACTACCGTGTAACAACCTTTAC[A/T]CCACTTCAGCCTTCAAAGTTCTCGTTTGAATATTTGCTACTACCACCAAGATCTGCACCC
PS_C11654_[A/G]	AATGACACAACAAAGAAGCCATTCAATGATAAAGCAGAAACCGACAAATTGAGATATCAG[A/G]GAGAGATGGGAGACTACAAACCGCCAAAGGGTGGTAAGACAGGTGGGAAGGTAGAAGATG
PS_C11672_[C/G]	AATCATGTGGTTAAATAGGAAACTTTCCCCGATTAAGAAGTGATTATTCGGATACAAATG[C/G]ACGAGAAAGCACACGCGCCCCCGAGCTTGCAACTATCAGCCACAACAAGCAGGGTGTCTC
PS_C11679_[A/G]	ACTGCAGGAGGACCTTGATTCGGAGAAAGAGAGCAGGAATAAGGCAGAGAAACAGAAGAG[A/G]GACCTTAATGAGGAACTGGAGGCTCTGAAGACCGAGTTTGAGGCAGAATTGGACAGCACA
PS_11783_[T/G]	CACCTTGCAAGGGACAGTGTGTGGTTTTCCGATCTTGTTACCCCAGTATCCTCGTCGGAC[T/G]GGAATGACCGACAGTTTGGCCAGAATGATGGCACCACGGATAGCAGTGGCAACTTCCTTG
PS_11952_[A/G]	GTGGGCAGAGCAGAGGAGGGAGAGAGAGGATGCACAGGCCAAATACCGTCAGCGTCAAGA[A/G]GAGAAACAGAGAGAACGAGAACGAAAGATTCAGAAGATCCTTGAGAGAGAACAGCAGCTC
PS_11962_[A/G]	CACATCTCACATCACTCCTACGCTGACGACAGGCAACTTCTCAACAACTTCAATATTGAA[A/G]ATGCTAAAGATGAAATCAAACGGATGGAATCATGCACCACAGATATCAAATCCTGGATGC
PS_12081_[A/G]	GGAGAAAGTACCTGAGAAAGAATCAAAAGATTCAGAAGAAAAAACTGAAACTGCTGAGAA[A/G]GGAGAAGATGAAGAAAAGGAAAAGTCCAAACCAGAAGATGAGAGTGAGAAGCCTGAAGAC
PS_12272_[A/G]	AACTAGCTAATAAGCCGCAGGCTGATCCAAAAGCGAGGCCCGAAAGCCTCTTTCCTAGCA[A/G]TACAAGAACTGCTAGATTATCCGGTATTACCCCATGTTTCCATGAGCTATCCCCAACTTC
PS_12342_[T/G]	ATCAGTTTGACTGATTATATTGCTGTGAAAGAGAAGTATGCCAAATATTTGGCTCACTCC[T/G]CCGGCAGATACCAGGTGAAGAGATTCCGCAAGGCACAATGCCCCATTGTGGAGAGACTTG
PS_12499_[T/C]	GTCAGCGTTGCACAACAGTTACATTTCCAGCATTTCAAAATTGTTATACAAGGGAGGTAA[T/C]TCTCCACAGTACCTTTGTTTGATCTTTTGCATGACCTATTGGTCAGCTCCCTGTGATGTC
PS_12587_[A/C]	CATCACCAGAACAGGGGGCCGCGCACAAACAACAACCTGGAAGGTTTCCATAATAGACTA[A/C]ACCGCACATTGCCACATAATCATCCGAACATTTACAGATTTATCACTGTCATCCGGAAGA
PS_14242_[T/C]	CTACAAAGTTACAATAAACGATGTCAGAATAAAGAAATATGAAGATTTCCACAGAAACTA[T/C]GATGCCAAGGCTCACTACGACAGAATGATCAAGGCCGGAGTGTTCCAGTCAATACGATCG
PS_14379_[C/G]	GCCCTCTCCCGGATCACGTCAGCATCGTCGAGCCCAAGGATGAGCCCGCCCCAGCTCAAC[C/G]TTACAGTGAGCAGAAGGGAGCTAAGCCTTCAGCAGATGCTACTCCAGCTCCACCTATTGT
PS_14847_[T/C]	ACAGTAACATTGTTGTTGGCGAGATAATCCATGGTAACACGTGCAACATGTGGTCTCGCG[T/C]TGTCTTGCTGAAATATCTCCCTTCTGACGTCAATGGCGGGAATGAGGTGACGTTGGAGGA
PS_C15086_[T/C]	TCTACAGTTTTGCTGTCTATTTTCTACTCGGGCAAGTCTTTTGTCTTATTTTGTGCCACC[T/C]AGCAGTATTCATCAGTGACCCTTGAAACGGGATGCGAAAAGAAAGACGTTCATCTCCCAA
PS_C15108_[A/G]	TGCGACGTTTTTGTCAGTGGAGCAGGCAGGTCTCAAATTGCCAACGCCTGTCTTCGTAAT[A/G]GTTGTGATGAATGCTGAACACTCCCTTTAGGACATGCGCAGCGGGCAGTTTGAAGTTTAG
PS_C1517_[T/C]	CGTGCACAGTTCAGGAAGAATTTGCCCTCAAAAGCCATGGGCAAAAGAGTAAGAGTGATG[T/C]TGTACCCTTCCAGAATATGAGAAAATGTAAATGAAGGCATTAAAAGACGTCGAAACAATT
PS_C15389_[A/C]	ATTCTTTTCAGTTTTAGTTTTCTTCAGTTTGTAACTTACAATGGCTTCGGTTGAGGAACT[A/C]TGGAACAAACTTAAGAGTGCAGCTAACTGCAAATCTCTTCTGAAGAAACATCTTACAGAG
PS_C15513_[A/C]	TCGAACATAAGACGTCAAGATGCAAAACGAAGCAGGTGAATTCGTTGATATGTACATCCC[A/C]AGAAAATGCTCTGCAAGCAACCGCATCATTGCTGCTAAAGACCACGCCTCCATCCAGATC
PS_C16693_[C/G]	GGAGAAGAGACTTCAAGTGGTGCAACACCAGAAACAACAAGTCCAGTGACTGTTACCACT[C/G]TAGCAGCAGATGACAAACCAACAGTGACTGAAGACTTCTCTGGCATCACTGAAGATACGT
PS_C19057_[A/G]	CATTGCAACCAAGACTCATAGTCATGACGTCATCCTTCAAGCAATTCGGAAGTCACTTAA[A/G]GGAGAAGCAGCTATCATATCTATGAAGCTTGGGCCACATGCAGACTTCCAAGAATTACTT
PS_C23053_[T/G]	TCGTCTTTCTTGGCTGGTGTGGTGGCTTTAGGAGTCGTCTTTTGAGCAGGTGGGGTTGTA[T/G]TGTCTTTCTTCGGAGACTGGGCTGGTGGAGGCATCTGAGTTTTAAGTTCCTGCTTTTGAT
PS_C23079_[A/G]	CTGTTTCGATCCGGTCGGCGGCCAGTTGCTTTGTATGCGGCCGCATAAATAGGCTGAGTG[A/G]CTTCGTTTGCCGTTACTCGTGTGTTGTTGCATCTCGATAAAACTACCAATATGTCAGGAC

PS_C23169_[A/T]	GCCCGACTCATTATCAGAGGTGGGTACTCCTGCTCTTTAATAGTTGAATGCATTCTAGCT[A/T]GTTACTGGCATACATAAATCGATGCCACAAATCCAAAAGGATTCTGCCTATGTGTCTTTT
PS_C23442_[T/G]	GCCTGGTCCTGACGGTATAGTGAATGAGCACATAAAGAAGGCTGGTAGATCGCTGTTGCA[T/G]CATATTGCCTTACTATTCAATGCCATTATTACGTTTGAATACATTCCTAAATCATTTTCT
PS_C24026_[T/G]	TGGGTTTAGTTTGAATTTCTTAAGGGTGGATTTTTTGGTGTTGTTGTATTTATTTGTGAT[T/G]AGTTTGAATGTGGGGTTGTGTTGGGCCCCAGCCCAGTAGGTGAGTGTAGTTTGGTTGATT
PS_C24124_[T/G]	CTTCTTAGAGGGACGGGCGACGCTTAGTCGCACGAAGTAGAGCAATAACAGGTCTGTGAT[T/G]CCCTTAGATGTCCAGGGCCGCACGCGCGCTACACTGAAAGAGTCAACGGGGATGCCCACC
PS_C24216_[A/G]	AGGCTCATCTGCCTCGACATAGAGACTGTCAACAGGAGAGGTTCTAAAAGACCCAAGACA[A/G]AGTCTAAGACCTTGGTGGTGGACAGAATCAAGTAGTTTTAGGTTGCTTTTGCAGGCTCCA
PS_C24229_[A/G]	TCGGCGGGAGGCCTACCGGAAACGAAAGTCTTCGGGTTCCAGGGGAAGTATGGTTGCAAA[A/G]CTGAAACTTAAAGGAATTGACGGAAGGGCACCACCAGGAGTGGAGCCTGCGGCTTAATTT
PS_C25199_[T/G]	TCATGGACAGGGGACTTACACCTACGCTGCCACTGGTACTCGTTATGTCGGCACGTGGAT[T/G]GATGGGAAGAGGGATTGTCATGGGGAGCTTGTCCATGCTAACCACAAATATGTGGGCACA
PS_C25290_[A/C]	ATAGTCCTCTGTCTTAGCAGGCGGCCCTGTGGCGCAACGGTAGCGCGTCTGACTCCAGAT[A/C]AGAAGGTTGCGTGTTCAAATCACGTCGGGGTCAATTCAAAGTTTTGCTATTTTGAATAAA
PS_C25384_[T/C]	TTGTGGAATGTACGTCCACGCTGAAGACATCGACACACAATGTGGTGGTAAACACACCAC[T/C]GGTTAAGGCGCCTAACGCTGACGCTAAGCAGAGACCATAAAAGGTGTTGGTCGATATAGA
PS_C25642_[T/C]	TACTTTCGTTTCCGTCTGTTTGTAAGTGGCGTTGCAAGAGGTCAAATATTCTTGCAATAC[T/C]GCCTTACGACAACTCATGGTAGGCGTAACAGTGCCTGCGAGAACGATGAGACTGGGTTCA
PS_C26263_[C/G]	AAGGATCAACTTCTTGTACAGAAAATGATTTTGGAAGAGGCTAAAACATATCACCGCAAC[C/G]TCTCCATGTGCTGGATTGACTACAAAAAAGCATACGACTCTGTCCCTCATGAATGGATTC
PS_C27069_[T/G]	CAGACTCTTGACTTCAGTATCTAGTGTGACCACCACGAGCGGCAATAACTGCAGCACAGC[T/G]TCTATGCATGGAGTGCACCAATCGATTGATGAAGGCCACTGGCACCTGGAACCACACTCT
PS_C27339_[T/C]	GGAGATAGCAGAAAACGTGCTACCGATCCCATATGGTCAATGAAAACGTATAACATCGAC[T/C]TAGTCGACATCAAAGACGACGAACCTAACTTGTACTACTTGAAAGACGGACCGCCGAGAG
PS_C34745_[T/C]	CTCTGATGCCAGCGTTTATGAAGTGTATGCCCTTCCTAAGCTGTATGCCAAGCTCTTGTA[T/C]TGCGTGTCATGCGCCATCCACTCCAAGGTGGTGAGGAACAGGTCCCGTGAGGCCAGGAAG
PS_C35945_[A/T]	CCCGTTACATTGTCGGCGCAGAGACTCTCGACCAGTGAGCTATTACGCACTCTTTAAATG[A/T]ATGGCTGCTTCTAAGCCAACATCCTGGTTGTTTCAGAATCTCCACCTCCTTTCCCACTTA
PS_C36068_[A/C]	ATCACTCGCATACTATGGACCTAATATGGCATATCGGAGGTTACTGTGTGTACATCCCCA[A/C]GCGTCATGTACACAGGTAAAATTTCGTAAAACTTAGAAAAAATTCAAGGGTGGGACACAT
PS_C36540_[A/G]	TCGCGTCAGGTAACCTTTGAGCGACCTATGACCGTCCAATCAAAGGCGAGACGGCCGAAC[A/G]GATTTAACCAATCAGACAACGGCTACTATTTAGGGTTCGGCGGGACTTACAAAATGGATT
PS_C38360_[C/G]	GTCGTTGGTGAAGGCGCAGCAAAGGAACGACTTCAGGCCTTAGAATTTCATCAACATAAC[C/G]CTGTGCTGTAAGATTGCCCTGAATAACATGAAGATCGCTCCTGAAACGTGCGTTTATTGC
PS_C42647_[A/G]	AAGACGCCCTACAGCGAAAGCATTTGCCAAGTATGTTTTCATTAATCAAGAACGAAAGTC[A/G]GAGGTTCGAAGACGATCAGATACCGTCGTAGTTCTGACCGTAAACTATGCCAACTGGCAG
PS_C42692_[A/G]	TTCGTGCCCCGGTCCGTACCCATGTCCGCATCAGGTCTCCAAGGTTCACAGCCTCTAGCC[A/G]ATAGAACAATGTAGGTAAGGGAAGTCGGCAAATTCGATCCGTAACTTCGGGACAAGGATT
PS_C46532_[T/A]	ATACTCTCCAACCAGCCCAAGCTACAGTCCAACAAGCCCCTCATACTCTCCAAGCTCTCC[A/T]AACTTCAGCCCTCAGTCACCTTCTTACTCTCCAACCAGCCCATCCTATAGCCCAACAAGC
PS_C46579_[G/T]	GGCCACATAATATCTCAGCTCAGCTATTTTCTCGGCCTTTTCACAAGCATACTCGTAGTA[T/G]GCATTGCGAACGCCTTCCACACTGGCCATTCGGAGACGGAGGTTGTAACGGAACGAACGG
PS_C46624_[T/C]	GCACGTCTGTCCTCCTTGCGCTGGGCCAGTGCCACATACAGAGGTTTTGACACAATGATA[T/C]GACCATTCATCTCGGTCACAGCCTTGGTTGCCTCCTCTGGAGAGCTGAAGCACACAAAGC
PS_C46754_[T/C]	TTCTATTGTTTCTTCTAAATCCCCTTCATCTTCACTAGCTGTGAACTCTTCATCACTTCC[T/C]TTTGATTCTTTGGACTTTTGGGGAGTTTCTGTTTCTGTTTCATCTTGAATGCTTGCAGGC
PS_C47728_[T/G]	AGATCAGGAATGTGCAGCAAGATCCAAGACAATCTTCCGATGACGACAGTTATTTTACCC[T/G]ACCGATGACGTTATTCGGTCGTTGCAATGGTAATCCTGCTCAGTACGAGAGGAACCGCAG
PS_C48958_[A/G]	CTTGGGTAACCATACGTCCCGTACCCATCAATCAGTGTCTCCAAGGTCCAGCCTCTAGCC[A/G]ATAGAACAATGTAGGTAAGGGAAGTCGGCAAATTCGATCCGTAACTTCGGGATAAGGATT
PS_C49403_[C/T]	AACTTGATGAACAAGAACAAATACGTCGTTCACTATCGTAATCTGCAGCTGTACTTATCG[T/C]TGGGAATGAAACTCACGAAAATTCACAGGGCGCTCGAGTTCGATCAAACTCCTTGGATGG
PS_C459_[A/G]	TGCTGCACTCCCACCTATGCTACACCTTTCATGTCGCTTCACAATGCCTAACTGAAGTCA[A/G]GCTCAACAGGGTCTTCTTTCCCCGCCGAGGATTCCAAGCCCGTTCCCTTTGGCTGTGGTT
PS_C47062_[C/T]	AGAAGTTCAGTCTCCTGACAGCTGTTACTACAACGCAGGCGCCGATGACGCCGTAAGAGA[T/C]GTCGCTGTGTCGGAACTAGCTGCAACGGTGGCAACACTTAAGGAGAGAAAAGATGGATTT

PS_C15235_[T/C]	AGCAGTCTGAAACACTGACCAGAATACCCTGCCGCCTGTTAGTCGATCAAGACAAATGAC[T/C]CCTTAAACAGATGAAATCACTTTTAGACCACAGCATGATAAGAACTTACACTTGAATCTA
PS_C15450_[T/C]	ACAGACTTCTGGGTGTATCCTTCACCTTATTCCCCTGCGTCAGGAAGATCAGCGTTTCTT[T/C]CCGCGATGAGGTTTCTGTTTGCATATTGTATTGCGGCCTCGTCCGGCTCCACGGCAACCG
PS_C23852_[T/C]	AGTCCGACTCCTTACCCATTTAAAGTTTGAGAATAGGTTGAGGACGTTTCGTCCCCAAGG[T/C]CTCTAATCATTCGCTTTACCGGATAAAACTGTTTTTCCCGAGCGCCAGCTATCCTGAGGG
PS_C24289_[C/G]	CACTTTTGTACATAAAAATATTTTAAAATTATGCTTAAAAATTATGGGAAGGGGAGGGGG[C/G]CATTGATTTTGTACATGACATGCAGGGGGTCACTGTCACATGATTTCATCTTTACCACGA
PS_C24450_[A/G]	AAGACGCCCTACAGCGAAAGCATTTGCCAAGTATGTTTTCATTAATCAAGAACGAAAGTC[A/G]GAGGTTCGAAGACGATCAGATACCGTCGTAGTTCCGACCGTAAACTATGCCAACTGGCAG
PS_C25160_[A/C]	GCGTGCGTTTTGACTTTGTGCGTGTGTGGAGGTGGTGTGCGGATGTGGCAAGCGAAAAAA[A/C]CCGCATGTTTGGCACTTATTTGGGCCCGTTTTCGCTTAACCCCCTAACCCTAACCCTATC
PS_C25478_[T/G]	GCGAAGCAGAATTCGCCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTGAGCTGGG[T/G]TTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACCGATGACGTTATTCGGTCGTTGCAA
PS_C25608_[A/T]	AACCATGGAACGAATGGTAAACAATAGATTAACATGGTATCTTGAAACCAATAACCTTAT[A/T]ACAAATATTCAATGTGGTTTCAGGAAAAATCGTAGTACCATTGATCATTTGGTACGTTTA
PS_C26361_[T/C]	GTGAAATTGGGTTTAACGTCGTGCTTGAGAATATTTCGCTCATATGACGACGTGCATGCG[T/C]GTGTATGTGTATGTGCGGCTGCTTGTACTAGCCCAGACTGAAGCTGGTTTTATAGTGCTA
PS_C27036_[T/C]	TCAGACTCTTGACTTCAGTATCTAGTGTGACCACCACGAGCGGCAATAACTGCAGCACAG[T/C]GTCTATGCATGGAGTGCACCAATCGATTGATGAAGGCCATTGGCACCTGGAACCACACTC
PS_C36141_[T/C]	GTGGAGTTATATGCTGTTTTATGCCGTTTTGGGCAAATAATGAATTTGTGCCATTTTCAA[T/C]CCAATATGGCCGCCAAACGATGATGTTTTTGATATGCATGTAAGAACACGAAATGTTAAG
PS_C36180_[A/C]	AATTTCACAACAAATTAGGTTAAAATCACCACTGCCACAACAATGTGAGTGTCTGAGTTA[A/C]TCTCCTGAGGGTCAATTTGAAGTCAAATGATAAGAATATCATGAAATATGTTCACTGCTG
PS_C36256_[G/T]	ACGGAAGAATGGACCGAGGAGACAGGGGTGACCGACGACCCAAGCCATACTAGAACAGTT[T/G]ACTCATTTGCCGGCTAGAACAGCCATCTTCACTGAAGAGCATTTTAGCCATCATCTCATT
PS_C38489_[A/G]	CACTGGTCAGTTATGACCGTATATACCAAAGATTAAGTAGAGCCCTGCGTTTGCCTTATA[A/G]TCAGGTTTACAAAGTTGTAAACAAATCGCTAACAGAGCCAGGGGTCGTATCAACATCAAA
PS_C45957_[A/T]	ACTTTCTTCAAGTCGCAGAAGCAACTTCCATTCATCCAGTCACACCAGGGTCTTCCTTTC[A/T]ATCAGGTCAAGGTCGTCCATCCATCGTCTCACTAGGGTCTTCCATCCAGTCACCAGGGTC
PS_C46576_[T/G]	GTTCGTGGTATCACGTGATTTATAATCTTAATCCTTAGATAAGTGGTAAGGAAAGGGGGG[T/G]GTGTTTAGTTTTCTTGGGTTGTTAGTATTACCCATTGTGAGAAGTCGTCAACATTAATTG
PS_C46778_[C/T]	CCTACCTATATACAGACAGTCATTCATACCCAACATAGGATCAATCAACCACAGCATTTC[T/C]CTTCACACAGAATACGCAATTGTAGCTGTTGGTCAGTACAGGCGCGGGTTTAAACGTCGC
ILL_C45_3002_[G/T]	TCCAAATCCGTATCCCGAAGAACATGCAGTTTATGGGCAAGTTGAAGCATTGTCAAATGGA[T/G]GATTCTGAGCTTGAGGGTTATATTATCAACCGTGTTGTTGGGAAGACAGCTTTCTTCAACGT
ILL_C1384_793_[A/T]	ACAAAAACTCATCTCTAAATGCTAAAGTAAATAATTCCAGTCAGATTGGTCTAGGCTACAC[A/T]CAGAAACTTAGGGATGGTGTGAAGCTGACCGTTTCCTCTTTAATAGAAGGCAAGAACTTCAA
ILL_C5433_233_[A/G]	CTGATGGCTGGAGATTGTCAAGGTACTCTCTGAGGTATTCTCCATGCGTAACACAAAACAC[A/G]CAATGTCCTAGATGAAGATGACATTATTCATCACCCGGTCTCTTCTTGTACTCTTGAAGATT
ILL_C14033_777_[T/C]	CGGCCTGCATCGCAGAAATGTCACTATCCAGCTTCCTCTTCTGGCCTTGGAGTGATGAACT[T/C]GTAGCTGTCAGCTCATTGACTCTGTCAAGAGCCTCTGTCAGCTCAGCTTCTGCATTCTTGCGGC
ILL_C22347_319_[C/T]	ATTCTAATGATTAAATTACATTCTATTACAAACTCGAACTGATTGTGGAACTCTTCACATTT[T/C]CCATGCGTTGGTATGTTTGGTATAAATATCAGACGCAAGTTCTAATGGTTGAATAACGTTCT

Table S4.2: Diversity statistics, minor allele frequency (MAF), observed- (H_o) and expected (H_e) heterozygosity and F_{is} values, including Hardy-Weinberg equilibrium *P*-value per locus per population.

AS(F_1)					
Locus	**MAF**	**H_o**	**H_e**	**F_{is}**	***P*-value**
ILL_C2122_257_[A/G]	0.071	0.143	0.135	-0.077	1.000
PS_C15402_271_[T/A]	0.052	0.103	0.100	-0.055	1.000
ILL_C253_1545_[A/C]	0.086	0.172	0.160	-0.094	1.000
ILL_C428_225_[A/G]	0.345	0.483	0.460	-0.068	1.000
PS_C23094_578_[G/C]	0.196	0.250	0.321	0.208	0.248
ILL_C20682_843_[A/T]	0.040	0.080	0.078	-0.042	1.000
PS_C34490_403_[A/G]	0.393	0.571	0.486	-0.198	0.441
PS_C25208 _377_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C1783_492_[T/G]	0.034	0.069	0.068	-0.036	1.000
ILL_C929_734_[T/C]	0.096	0.192	0.177	-0.106	1.000
PS_C11665_287_[A/G]	0.328	0.517	0.448	-0.174	0.675
ILL_C2141_350_[T/A]	0.431	0.517	0.499	-0.055	1.000
PS_C23591_200_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15689_162_[A/T]	0.138	0.276	0.242	-0.160	1.000
ILL_C229_2772_[A/G]	0.086	0.172	0.160	-0.094	1.000
ILL_C22574_507_[A/C]	0.103	0.207	0.189	-0.115	1.000
PS_C15088_268_[A/G]	0.321	0.643	0.444	-0.474	0.028
ILL_C2406_641_[T/C]	0.207	0.276	0.334	0.159	0.557
PS_C12925_666_[A/G]	0.160	0.320	0.274	-0.190	1.000
PS_C11984_159_[T/G]	0.466	0.931	0.506	-0.871	0.000
ILL_C22449_261_[A/G]	0.190	0.379	0.313	-0.234	0.546
PS_C11970_157_[A/G]	0.196	0.393	0.321	-0.244	0.551
PS_C23051_368_[A/G]	0.250	0.500	0.386	-0.333	0.521
ILL_C6012_280_[T/C]	0.414	0.621	0.494	-0.279	0.243
ILL_C2903_1043_[T/A]	0.310	0.483	0.436	-0.128	0.691
PS_C34501_638_[A/G]	0.052	0.034	0.100	0.648	0.061
PS_C12352_527_[T/C]	0.393	0.571	0.486	-0.198	0.444
PS_C34725_229_[T/C]	0.190	0.310	0.313	-0.010	1.000
PS_C15230_93_[A/G]	0.426	0.481	0.498	0.015	1.000
PS_C23075_525_[G/C]	0.224	0.310	0.354	0.108	0.602
ILL_C3835_411_[A/G]	0.328	0.517	0.448	-0.174	0.664
PS_C24743_123_[T/G]	0.241	0.483	0.373	-0.318	0.153
ILL_C5634_234_[A/G]	0.448	0.345	0.503	0.303	0.140
PS_C23630_237_[T/C]	0.052	0.103	0.100	-0.055	1.000
ILL_C1813_300_[T/G]	0.339	0.536	0.456	-0.195	0.431
PS_C12218_188_[A/G]	0.121	0.241	0.216	-0.137	1.000
PS_C36136_185_[T/C]	0.276	0.483	0.407	-0.208	0.396
PS_C23070_1364_[A/G]	0.207	0.345	0.334	-0.051	1.000

ILL_C618_116_[A/G]	0.214	0.357	0.343	-0.061	1.000
PS_C1652_228_[A/C]	0.138	0.138	0.242	0.420	0.067
PS_C25083_285_[A/G]	0.161	0.107	0.275	0.603	0.009
ILL_C394_1510_[A/G]	0.483	0.690	0.508	-0.381	0.071
PS_C24267_71_[A/G]	0.379	0.759	0.479	-0.611	0.001
PS_C34507_1191_[T/C]	0.241	0.276	0.373	0.247	0.301
PS_C36273_73_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C35977_153_[T/C]	0.071	0.000	0.135	1.000	0.001
PS_C16093_142_[T/C]	0.190	0.379	0.313	-0.234	0.552
PS_C16031_147_[A/G]	0.224	0.379	0.354	-0.091	1.000
ILL_C5106_273_[T/C]	0.069	0.138	0.131	-0.074	1.000
PS_C00512_245_[T/C]	0.052	0.103	0.100	-0.055	1.000
ILL_C140_2421_[A/G]	0.241	0.483	0.373	-0.318	0.158
PS_C36237_70_[C/G]	0.071	0.143	0.135	-0.077	1.000
PS_C34670_302_[A/C]	0.034	0.069	0.068	-0.036	1.000
ILL_C21880_1003_[G/C]	0.328	0.517	0.448	-0.174	0.671
PS_C47330_170_[A/C]	0.397	0.655	0.487	-0.369	0.116
PS_C36522_249_[G/C]	0.442	0.500	0.503	-0.013	1.000
PS_C34604_423_[A/T]	0.207	0.345	0.334	-0.051	1.000
PS_C34501_77_[T/G]	0.069	0.000	0.131	1.000	0.001
PS_C28810_290_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.293	0.379	0.422	0.085	0.656
ILL_C844_440_[T/C]	0.017	0.034	0.034	-0.018	1.000
PS_C46857_366_[T/C]	0.036	0.071	0.070	-0.037	1.000
PS_C34605_382_[A/G]	0.019	0.037	0.037	-0.019	1.000
ILL_C853_1199_[A/G]	0.052	0.103	0.100	-0.055	1.000
ILL_C1002_85_[T/A]	0.310	0.621	0.436	-0.450	0.028
PS_C11659_399_[T/C]	0.034	0.069	0.068	-0.036	1.000
ILL_C4791_1099_[T/G]	0.483	0.552	0.508	-0.105	0.715
ILL_C300_4982_[A/G]	0.036	0.071	0.070	-0.037	1.000
ILL_C4593_326_[T/G]	0.362	0.517	0.470	-0.120	0.702
PS_C46533_541_[T/C]	0.069	0.138	0.131	-0.074	1.000
PS_C15351_193_[A/G]	0.414	0.483	0.494	0.005	1.000
PS_C15018_147_[T/G]	0.089	0.107	0.166	0.341	0.176
PS_C38608_168_[T/C]	0.500	1.000	0.509	-1.000	0.000
ILL_C387_215_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C5339_366_[T/C]	0.052	0.103	0.100	-0.055	1.000
PS_C34511_71_[T/G]	0.038	0.077	0.075	-0.040	1.000
PS_C14838_228_[T/C]	0.018	0.036	0.036	-0.018	1.000
ILL_C20267_102_[T/C]	0.224	0.448	0.354	-0.289	0.290
PS_C46674_204_[C/G]	0.155	0.241	0.267	0.079	0.507
ILL_C1878_506_[T/C]	0.089	0.179	0.166	-0.098	1.000
PS_C36563_85_[T/G]	0.096	0.192	0.177	-0.106	1.000
ILL_C2735_326_[T/G]	0.017	0.034	0.034	-0.018	1.000
ILL_C250_199_[T/C]	0.071	0.071	0.135	0.462	0.105

PS_C47845_238_[T/C]	0.370	0.444	0.475	0.047	1.000
ILL_C1363_269_[T/C]	0.207	0.414	0.334	-0.261	0.304
ILL_C18774_676_[T/C]	0.052	0.103	0.100	-0.055	1.000
PS_C23647_375_[A/G]	0.068	0.136	0.130	-0.073	1.000
ILL_C2028_1228_[A/T]	0.071	0.143	0.135	-0.077	1.000
PS_C11985_171_[A/G]	0.212	0.192	0.340	0.424	0.052
PS_C11911_576_[A/C]	0.017	0.034	0.034	-0.018	1.000
ILL_C1254_187_[A/G]	0.286	0.429	0.416	-0.050	1.000
ILL_C20427_267_[T/C]	0.056	0.111	0.107	-0.059	1.000
PS_C21989_160_[T/G]	0.121	0.241	0.216	-0.137	1.000
PS_C12069_1181_[T/C]	0.052	0.103	0.100	-0.055	1.000
ILL_C327_1076_[C/G]	0.017	0.034	0.034	-0.018	1.000
PS_C11847_438_[T/C]	0.260	0.440	0.393	-0.143	1.000
PS_C47375_253_[A/G]	0.414	0.000	0.494	1.000	0.000
PS_C34420_787_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C2915_875_[T/C]	0.188	0.375	0.311	-0.231	0.552
ILL_C22491_727_[T/C]	0.276	0.483	0.407	-0.208	0.395
ILL_C911_1343_[T/A]	0.259	0.517	0.390	-0.349	0.135
PS_C12196_434_[G/C]	0.328	0.448	0.448	-0.018	1.000
PS_C14297_369_[T/G]	0.125	0.179	0.223	0.184	0.353
PS_C47340_198_[A/G]	0.100	0.040	0.184	0.778	0.006
PS_C36706_579_[T/C]	0.155	0.310	0.267	-0.184	1.000
PS_C46597_301_[C/G]	0.034	0.069	0.068	-0.036	1.000
ILL_C2040_1251_[A/T]	0.121	0.172	0.216	0.188	0.331
PS_C35683_190_[A/G]	0.310	0.414	0.436	0.033	1.000
PS_C46532_529_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C15379_78_[A/C]	0.328	0.448	0.448	-0.018	1.000
PS_C11871_171_[A/G]	0.310	0.414	0.436	0.033	1.000
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.052	0.103	0.100	-0.055	1.000
ILL_C6061_1289_[T/G]	0.466	0.517	0.506	-0.039	1.000
ILL_C980_261_[A/G]	0.089	0.179	0.166	-0.098	1.000
Mean	**0.197**	**0.301**	**0.281**	**-0.040**	

CR(2000)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.111	0.111	0.203	0.438	0.170
PS_C15402_271_[T/A]	0.028	0.056	0.056	-0.029	1.000
ILL_C253_1545_[A/C]	0.111	0.222	0.203	-0.125	1.000
ILL_C428_225_[A/G]	0.361	0.278	0.475	0.398	0.124
PS_C23094_578_[G/C]	0.056	0.111	0.108	-0.059	1.000
ILL_C20682_843_[A/T]	0.042	0.083	0.083	-0.043	1.000
PS_C34490_403_[A/G]	0.250	0.500	0.386	-0.333	0.526
PS_C25208 _377_[T/G]	0.026	0.053	0.053	-0.027	1.000
ILL_C1783_492_[T/G]	0.158	0.316	0.273	-0.187	1.000
ILL_C929_734_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C11665_287_[A/G]	0.139	0.167	0.246	0.303	0.268

ILL_C2141_350_[T/A]	0.474	0.316	0.512	0.367	0.160
PS_C23591_200_[T/C]	0.231	0.462	0.369	-0.300	1.000
PS_C15689_162_[A/T]	0.132	0.263	0.235	-0.152	1.000
ILL_C229_2772_[A/G]	0.056	0.111	0.108	-0.059	1.000
ILL_C22574_507_[A/C]	0.105	0.105	0.193	0.441	0.157
PS_C15088_268_[A/G]	0.306	0.389	0.437	0.084	1.000
ILL_C2406_641_[T/C]	0.056	0.111	0.108	-0.059	1.000
PS_C12925_666_[A/G]	0.182	0.364	0.312	-0.222	1.000
PS_C11984_159_[T/G]	0.444	0.889	0.508	-0.800	0.003
ILL_C22449_261_[A/G]	0.184	0.368	0.309	-0.226	1.000
PS_C11970_157_[A/G]	0.211	0.421	0.341	-0.267	0.537
PS_C23051_368_[A/G]	0.029	0.059	0.059	-0.030	1.000
ILL_C6012_280_[T/C]	0.289	0.263	0.422	0.360	0.122
ILL_C2903_1043_[T/A]	0.278	0.333	0.413	0.169	0.558
PS_C34501_638_[A/G]	0.026	0.053	0.053	-0.027	1.000
PS_C12352_527_[T/C]	0.353	0.353	0.471	0.227	0.333
PS_C34725_229_[T/C]	0.265	0.294	0.401	0.244	0.530
PS_C15230_93_[A/G]	0.474	0.211	0.512	0.578	0.017
PS_C23075_525_[G/C]	0.389	0.333	0.489	0.299	0.316
ILL_C3835_411_[A/G]	0.139	0.167	0.246	0.303	0.275
PS_C24743_123_[T/G]	0.289	0.579	0.422	-0.407	0.263
ILL_C5634_234_[A/G]	0.368	0.316	0.478	0.321	0.166
PS_C23630_237_[T/C]	0.026	0.053	0.053	-0.027	1.000
ILL_C1813_300_[T/G]	0.342	0.368	0.462	0.182	0.621
PS_C12218_188_[A/G]	0.194	0.389	0.322	-0.241	1.000
PS_C36136_185_[T/C]	0.421	0.316	0.501	0.352	0.159
PS_C23070_1364_[A/G]	0.158	0.316	0.273	-0.187	1.000
ILL_C618_116_[A/G]	0.194	0.389	0.322	-0.241	1.000
PS_C1652_228_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C25083_285_[A/G]	0.105	0.000	0.193	1.000	0.003
ILL_C394_1510_[A/G]	0.395	0.263	0.491	0.449	0.066
PS_C24267_71_[A/G]	0.421	0.842	0.501	-0.727	0.002
PS_C34507_1191_[T/C]	0.421	0.316	0.501	0.352	0.165
PS_C36273_73_[A/C]	0.083	0.167	0.157	-0.091	1.000
PS_C35977_153_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C16093_142_[T/C]	0.421	0.526	0.501	-0.080	1.000
PS_C16031_147_[A/G]	0.237	0.368	0.371	-0.019	1.000
ILL_C5106_273_[T/C]	0.250	0.250	0.387	0.333	0.212
PS_C00512_245_[T/C]	0.029	0.059	0.059	-0.030	1.000
ILL_C140_2421_[A/G]	0.250	0.278	0.386	0.259	0.261
PS_C36237_70_[C/G]	0.395	0.474	0.491	0.009	1.000
PS_C34670_302_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C21880_1003_[G/C]	0.194	0.278	0.322	0.113	0.507
PS_C47330_170_[A/C]	0.316	0.632	0.444	-0.462	0.117
PS_C36522_249_[G/C]	0.222	0.333	0.356	0.036	1.000
PS_C34604_423_[A/T]	0.158	0.316	0.273	-0.187	1.000

PS_C34501_77_[T/G]	0.316	0.000	0.444	1.000	0.000
PS_C28810_290_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.421	0.421	0.501	0.136	0.637
ILL_C844_440_[T/C]	0.028	0.056	0.056	-0.029	1.000
PS_C46857_366_[T/C]	0.088	0.176	0.166	-0.097	1.000
PS_C34605_382_[A/G]	0.026	0.053	0.053	-0.027	1.000
ILL_C853_1199_[A/G]	0.184	0.368	0.309	-0.226	1.000
ILL_C1002_85_[T/A]	0.263	0.526	0.398	-0.357	0.277
PS_C11659_399_[T/C]	0.056	0.111	0.108	-0.059	1.000
ILL_C4791_1099_[T/G]	0.342	0.368	0.462	0.182	0.612
ILL_C300_4982_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C4593_326_[T/G]	0.417	0.500	0.500	-0.029	1.000
PS_C46533_541_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15351_193_[A/G]	0.395	0.579	0.491	-0.212	0.632
PS_C15018_147_[T/G]	N/A	N/A	N/A	N/A	N/A
PS_C38608_168_[T/C]	0.500	1.000	0.514	-1.000	0.000
ILL_C387_215_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C5339_366_[T/C]	0.028	0.056	0.056	-0.029	1.000
PS_C34511_71_[T/G]	0.079	0.158	0.149	-0.086	1.000
PS_C14838_228_[T/C]	0.263	0.526	0.398	-0.357	0.264
ILL_C20267_102_[T/C]	0.132	0.263	0.235	-0.152	1.000
PS_C46674_204_[C/G]	0.263	0.421	0.398	-0.086	1.000
ILL_C1878_506_[T/C]	0.026	0.053	0.053	-0.027	1.000
PS_C36563_85_[T/G]	0.063	0.125	0.121	-0.067	1.000
ILL_C2735_326_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.342	0.368	0.462	0.182	0.612
ILL_C1363_269_[T/C]	0.406	0.313	0.498	0.352	0.289
ILL_C18774_676_[T/C]	0.182	0.364	0.312	-0.222	1.000
PS_C23647_375_[A/G]	0.286	0.571	0.423	-0.400	0.507
ILL_C2028_1228_[A/T]	0.139	0.167	0.246	0.303	0.268
PS_C11985_171_[A/G]	0.208	0.250	0.344	0.242	0.395
PS_C11911_576_[A/C]	0.132	0.263	0.235	-0.152	1.000
ILL_C1254_187_[A/G]	0.250	0.500	0.386	-0.333	0.528
ILL_C20427_267_[T/C]	0.029	0.059	0.059	-0.030	1.000
PS_C21989_160_[T/G]	0.053	0.105	0.102	-0.056	1.000
PS_C12069_1181_[T/C]	0.026	0.053	0.053	-0.027	1.000
ILL_C327_1076_[C/G]	0.079	0.053	0.149	0.638	0.082
PS_C11847_438_[T/C]	0.438	0.750	0.508	-0.524	0.118
PS_C47375_253_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C34420_787_[A/G]	0.028	0.056	0.056	-0.029	1.000
ILL_C2915_875_[T/C]	0.139	0.278	0.246	-0.161	1.000
ILL_C22491_727_[T/C]	0.333	0.333	0.457	0.250	0.308
ILL_C911_1343_[T/A]	0.237	0.474	0.371	-0.310	0.526
PS_C12196_434_[G/C]	0.316	0.632	0.444	-0.462	0.111

PS_C14297_369_[T/G]	0.316	0.421	0.444	0.026	1.000
PS_C47340_198_[A/G]	0.316	0.000	0.444	1.000	0.000
PS_C36706_579_[T/C]	0.079	0.158	0.149	-0.086	1.000
PS_C46597_301_[C/G]	0.118	0.235	0.214	-0.133	1.000
ILL_C2040_1251_[A/T]	0.211	0.316	0.341	0.050	1.000
PS_C35683_190_[A/G]	0.368	0.632	0.478	-0.357	0.319
PS_C46532_529_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C15379_78_[A/C]	0.250	0.375	0.387	0.000	1.000
PS_C11871_171_[A/G]	0.447	0.474	0.508	0.042	1.000
PS_C35823_1199_[A/G]	0.105	0.000	0.193	1.000	0.003
PS_C39731_379_[A/C]	0.289	0.474	0.422	-0.152	1.000
ILL_C6061_1289_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C980_261_[A/G]	0.194	0.278	0.322	0.113	0.501
Mean	**0.218**	**0.300**	**0.312**	**0.012**	

CR(2011)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.207	0.345	0.334	-0.051	1.000
PS_C15402_271_[T/A]	0.148	0.296	0.257	-0.174	1.000
ILL_C253_1545_[A/C]	0.069	0.138	0.131	-0.074	1.000
ILL_C428_225_[A/G]	0.317	0.433	0.440	-0.001	1.000
PS_C23094_578_[G/C]	0.080	0.160	0.150	-0.087	1.000
ILL_C20682_843_[A/T]	0.100	0.200	0.184	-0.111	1.000
PS_C34490_403_[A/G]	0.207	0.345	0.334	-0.051	1.000
PS_C25208 _377_[T/G]	0.071	0.143	0.135	-0.077	1.000
ILL_C1783_492_[T/G]	0.207	0.414	0.334	-0.261	0.298
ILL_C929_734_[T/C]	0.022	0.043	0.043	-0.022	1.000
PS_C11665_287_[A/G]	0.138	0.276	0.242	-0.160	1.000
ILL_C2141_350_[T/A]	0.400	0.467	0.488	0.028	1.000
PS_C23591_200_[T/C]	0.031	0.063	0.063	-0.032	1.000
PS_C15689_162_[A/T]	0.111	0.222	0.201	-0.125	1.000
ILL_C229_2772_[A/G]	0.121	0.172	0.216	0.188	0.318
ILL_C22574_507_[A/C]	0.083	0.167	0.155	-0.091	1.000
PS_C15088_268_[A/G]	0.375	0.679	0.477	-0.448	0.043
ILL_C2406_641_[T/C]	0.121	0.172	0.216	0.188	0.346
PS_C12925_666_[A/G]	0.136	0.273	0.241	-0.158	1.000
PS_C11984_159_[T/G]	0.483	0.966	0.508	-0.933	0.000
ILL_C22449_261_[A/G]	0.283	0.367	0.413	0.097	0.655
PS_C11970_157_[A/G]	0.217	0.233	0.345	0.313	0.104
PS_C23051_368_[A/G]	0.114	0.227	0.206	-0.128	1.000
ILL_C6012_280_[T/C]	0.283	0.367	0.413	0.097	0.655
ILL_C2903_1043_[T/A]	0.352	0.704	0.465	-0.543	0.009
PS_C34501_638_[A/G]	0.017	0.033	0.033	-0.017	1.000
PS_C12352_527_[T/C]	0.500	0.571	0.509	-0.143	0.698
PS_C34725_229_[T/C]	0.224	0.379	0.354	-0.091	1.000
PS_C15230_93_[A/G]	0.481	0.423	0.509	0.153	0.448
PS_C23075_525_[G/C]	0.345	0.414	0.460	0.084	0.682

ILL_C3835_411_[A/G]	0.217	0.300	0.345	0.116	0.582
PS_C24743_123_[T/G]	0.328	0.655	0.448	-0.487	0.009
ILL_C5634_234_[A/G]	0.367	0.533	0.472	-0.148	0.695
PS_C23630_237_[T/C]	0.050	0.100	0.097	-0.053	1.000
ILL_C1813_300_[T/G]	0.267	0.467	0.398	-0.193	0.637
PS_C12218_188_[A/G]	0.196	0.321	0.321	-0.018	1.000
PS_C36136_185_[T/C]	0.467	0.467	0.506	0.063	0.730
PS_C23070_1364_[A/G]	0.217	0.233	0.345	0.313	0.111
ILL_C618_116_[A/G]	0.231	0.385	0.362	-0.083	1.000
PS_C1652_228_[A/C]	0.033	0.000	0.066	1.000	0.015
PS_C25083_285_[A/G]	0.483	0.233	0.508	0.533	0.002
ILL_C394_1510_[A/G]	0.233	0.400	0.364	-0.118	1.000
PS_C24267_71_[A/G]	0.117	0.233	0.210	-0.132	1.000
PS_C34507_1191_[T/C]	0.339	0.393	0.456	0.124	0.665
PS_C36273_73_[A/C]	0.050	0.100	0.097	-0.053	1.000
PS_C35977_153_[T/C]	0.034	0.000	0.068	1.000	0.016
PS_C16093_142_[T/C]	0.467	0.600	0.506	-0.205	0.458
PS_C16031_147_[A/G]	0.200	0.333	0.325	-0.042	1.000
ILL_C5106_273_[T/C]	0.185	0.296	0.307	0.018	1.000
PS_C00512_245_[T/C]	0.023	0.045	0.045	-0.023	1.000
ILL_C140_2421_[A/G]	0.250	0.367	0.381	0.022	1.000
PS_C36237_70_[C/G]	0.383	0.500	0.481	-0.058	1.000
PS_C34670_302_[A/C]	0.033	0.067	0.066	-0.034	1.000
ILL_C21880_1003_[G/C]	0.375	0.393	0.477	0.162	0.438
PS_C47330_170_[A/C]	0.241	0.483	0.373	-0.318	0.147
PS_C36522_249_[G/C]	0.420	0.600	0.497	-0.232	0.408
PS_C34604_423_[A/T]	0.150	0.233	0.259	0.085	0.506
PS_C34501_77_[T/G]	0.483	0.100	0.508	0.800	0.000
PS_C28810_290_[A/C]	0.017	0.033	0.033	-0.017	1.000
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.300	0.533	0.427	-0.270	0.202
ILL_C844_440_[T/C]	0.034	0.069	0.068	-0.036	1.000
PS_C46857_366_[T/C]	0.179	0.286	0.299	0.026	1.000
PS_C34605_382_[A/G]	0.034	0.069	0.068	-0.036	1.000
ILL_C853_1199_[A/G]	0.103	0.207	0.189	-0.115	1.000
ILL_C1002_85_[T/A]	0.450	0.900	0.503	-0.818	0.000
PS_C11659_399_[T/C]	0.067	0.133	0.127	-0.071	1.000
ILL_C4791_1099_[T/G]	0.500	0.586	0.509	-0.172	0.469
ILL_C300_4982_[A/G]	0.054	0.107	0.103	-0.057	1.000
ILL_C4593_326_[T/G]	0.500	0.667	0.508	-0.333	0.137
PS_C46533_541_[T/C]	0.067	0.133	0.127	-0.071	1.000
PS_C15351_193_[A/G]	0.483	0.500	0.508	-0.001	1.000
PS_C15018_147_[T/G]	0.067	0.133	0.127	-0.071	1.000
PS_C38608_168_[T/C]	0.482	0.964	0.508	-0.931	0.000
ILL_C387_215_[A/G]	0.074	0.000	0.140	1.000	0.001
ILL_C5339_366_[T/C]	0.050	0.100	0.097	-0.053	1.000

PS_C34511_71_[T/G]	0.093	0.185	0.171	-0.102	1.000
PS_C14838_228_[T/C]	0.125	0.250	0.223	-0.143	1.000
ILL_C20267_102_[T/C]	0.276	0.483	0.407	-0.208	0.390
PS_C46674_204_[C/G]	0.200	0.267	0.325	0.167	0.303
ILL_C1878_506_[T/C]	0.150	0.233	0.259	0.085	0.497
PS_C36563_85_[T/G]	0.143	0.190	0.251	0.222	0.339
ILL_C2735_326_[T/G]	0.017	0.033	0.033	-0.017	1.000
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.328	0.448	0.448	-0.018	1.000
ILL_C1363_269_[T/C]	0.379	0.276	0.479	0.414	0.044
ILL_C18774_676_[T/C]	0.212	0.423	0.340	-0.268	0.555
PS_C23647_375_[A/G]	0.154	0.308	0.265	-0.182	1.000
ILL_C2028_1228_[A/T]	0.150	0.300	0.259	-0.176	1.000
PS_C11985_171_[A/G]	0.375	0.583	0.479	-0.244	0.387
PS_C11911_576_[A/C]	0.138	0.207	0.242	0.130	0.426
ILL_C1254_187_[A/G]	0.232	0.393	0.363	-0.102	1.000
ILL_C20427_267_[T/C]	0.154	0.308	0.265	-0.182	1.000
PS_C21989_160_[T/G]	0.231	0.385	0.362	-0.083	1.000
PS_C12069_1181_[T/C]	0.117	0.233	0.210	-0.132	1.000
ILL_C327_1076_[C/G]	N/A	N/A	N/A	N/A	N/A
PS_C11847_438_[T/C]	0.386	0.591	0.485	-0.246	0.394
PS_C47375_253_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C34420_787_[A/G]	0.017	0.034	0.034	-0.018	1.000
ILL_C2915_875_[T/C]	0.225	0.250	0.358	0.283	0.209
ILL_C22491_727_[T/C]	0.317	0.367	0.440	0.153	0.414
ILL_C911_1343_[T/A]	0.300	0.467	0.427	-0.111	0.682
PS_C12196_434_[G/C]	0.414	0.621	0.494	-0.279	0.261
PS_C14297_369_[T/G]	0.379	0.621	0.479	-0.318	0.131
PS_C47340_198_[A/G]	0.400	0.000	0.488	1.000	0.000
PS_C36706_579_[T/C]	0.155	0.310	0.267	-0.184	1.000
PS_C46597_301_[C/G]	0.155	0.310	0.267	-0.184	1.000
ILL_C2040_1251_[A/T]	0.138	0.207	0.242	0.130	0.417
PS_C35683_190_[A/G]	0.267	0.400	0.398	-0.023	1.000
PS_C46532_529_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C15379_78_[A/C]	0.375	0.393	0.477	0.162	0.422
PS_C11871_171_[A/G]	0.333	0.400	0.452	0.100	0.695
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.125	0.250	0.223	-0.143	1.000
ILL_C6061_1289_[T/G]	0.033	0.067	0.066	-0.034	1.000
ILL_C980_261_[A/G]	0.143	0.286	0.249	-0.167	1.000
Mean	**0.219**	**0.316**	**0.305**	**-0.028**	

GB(2003)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C15402_271_[T/A]	0.065	0.130	0.125	-0.070	1.000
ILL_C253_1545_[A/C]	0.045	0.091	0.089	-0.048	1.000

ILL_C428_225_[A/G]	0.229	0.458	0.361	-0.297	0.271
PS_C23094_578_[G/C]	0.125	0.250	0.223	-0.143	1.000
ILL_C20682_843_[A/T]	0.022	0.043	0.043	-0.022	1.000
PS_C34490_403_[A/G]	0.271	0.375	0.403	0.051	1.000
PS_C25208 _377_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C1783_492_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C929_734_[T/C]	0.022	0.043	0.043	-0.022	1.000
PS_C11665_287_[A/G]	0.229	0.375	0.361	-0.061	1.000
ILL_C2141_350_[T/A]	0.500	0.500	0.511	0.000	1.000
PS_C23591_200_[T/C]	0.313	0.625	0.439	-0.455	0.057
PS_C15689_162_[A/T]	0.083	0.167	0.156	-0.091	1.000
ILL_C229_2772_[A/G]	0.167	0.167	0.284	0.400	0.086
ILL_C22574_507_[A/C]	0.104	0.208	0.191	-0.116	1.000
PS_C15088_268_[A/G]	0.409	0.636	0.495	-0.316	0.211
ILL_C2406_641_[T/C]	0.146	0.292	0.254	-0.171	1.000
PS_C12925_666_[A/G]	0.174	0.348	0.294	-0.211	1.000
PS_C11984_159_[T/G]	0.417	0.833	0.496	-0.714	0.002
ILL_C22449_261_[A/G]	0.313	0.542	0.439	-0.261	0.347
PS_C11970_157_[A/G]	0.229	0.375	0.361	-0.061	1.000
PS_C23051_368_[A/G]	0.375	0.750	0.484	-0.600	0.031
ILL_C6012_280_[T/C]	0.458	0.583	0.507	-0.175	0.677
ILL_C2903_1043_[T/A]	0.375	0.583	0.479	-0.244	0.381
PS_C34501_638_[A/G]	0.125	0.000	0.223	1.000	0.000
PS_C12352_527_[T/C]	0.354	0.458	0.467	-0.002	1.000
PS_C34725_229_[T/C]	0.271	0.292	0.403	0.262	0.301
PS_C15230_93_[A/G]	0.458	0.417	0.507	0.161	0.440
PS_C23075_525_[G/C]	0.261	0.522	0.394	-0.353	0.276
ILL_C3835_411_[A/G]	0.333	0.500	0.454	-0.125	1.000
PS_C24743_123_[T/G]	0.250	0.500	0.383	-0.333	0.279
ILL_C5634_234_[A/G]	0.396	0.375	0.488	0.216	0.386
PS_C23630_237_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C1813_300_[T/G]	0.167	0.333	0.284	-0.200	1.000
PS_C12218_188_[A/G]	0.125	0.250	0.223	-0.143	1.000
PS_C36136_185_[T/C]	0.417	0.667	0.496	-0.371	0.107
PS_C23070_1364_[A/G]	0.250	0.333	0.383	0.111	0.600
ILL_C618_116_[A/G]	0.313	0.625	0.439	-0.455	0.048
PS_C1652_228_[A/C]	0.125	0.167	0.223	0.238	0.303
PS_C25083_285_[A/G]	0.021	0.042	0.042	-0.021	1.000
ILL_C394_1510_[A/G]	0.354	0.375	0.467	0.180	0.382
PS_C24267_71_[A/G]	0.417	0.833	0.496	-0.714	0.000
PS_C34507_1191_[T/C]	0.250	0.333	0.383	0.111	0.594
PS_C36273_73_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C35977_153_[T/C]	0.167	0.000	0.284	1.000	0.000
PS_C16093_142_[T/C]	0.292	0.250	0.422	0.395	0.059
PS_C16031_147_[A/G]	0.271	0.542	0.403	-0.371	0.131
ILL_C5106_273_[T/C]	0.063	0.125	0.120	-0.067	1.000

PS_C00512_245_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C140_2421_[A/G]	0.167	0.333	0.284	-0.200	1.000
PS_C36237_70_[C/G]	0.043	0.087	0.085	-0.045	1.000
PS_C34670_302_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C21880_1003_[G/C]	0.208	0.250	0.337	0.242	0.247
PS_C47330_170_[A/C]	0.313	0.625	0.439	-0.455	0.047
PS_C36522_249_[G/C]	0.295	0.500	0.426	-0.201	0.616
PS_C34604_423_[A/T]	0.250	0.417	0.383	-0.111	1.000
PS_C34501_77_[T/G]	0.063	0.042	0.120	0.644	0.068
PS_C28810_290_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.458	0.500	0.507	-0.007	1.000
ILL_C844_440_[T/C]	0.042	0.083	0.082	-0.043	1.000
PS_C46857_366_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C34605_382_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C853_1199_[A/G]	0.146	0.292	0.254	-0.171	1.000
ILL_C1002_85_[T/A]	0.250	0.417	0.383	-0.111	1.000
PS_C11659_399_[T/C]	0.042	0.083	0.082	-0.043	1.000
ILL_C4791_1099_[T/G]	0.229	0.292	0.361	0.174	0.550
ILL_C300_4982_[A/G]	0.043	0.087	0.085	-0.045	1.000
ILL_C4593_326_[T/G]	0.229	0.292	0.361	0.174	0.564
PS_C46533_541_[T/C]	0.021	0.042	0.042	-0.021	1.000
PS_C15351_193_[A/G]	0.500	0.583	0.511	-0.167	0.689
PS_C15018_147_[T/G]	N/A	N/A	N/A	N/A	N/A
PS_C38608_168_[T/C]	0.500	1.000	0.511	-1.000	0.000
ILL_C387_215_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C5339_366_[T/C]	0.063	0.125	0.120	-0.067	1.000
PS_C34511_71_[T/G]	0.095	0.190	0.177	-0.105	1.000
PS_C14838_228_[T/C]	0.021	0.042	0.042	-0.021	1.000
ILL_C20267_102_[T/C]	0.250	0.417	0.383	-0.111	1.000
PS_C46674_204_[C/G]	0.146	0.292	0.254	-0.171	1.000
ILL_C1878_506_[T/C]	0.042	0.083	0.082	-0.043	1.000
PS_C36563_85_[T/G]	0.136	0.273	0.241	-0.158	1.000
ILL_C2735_326_[T/G]	0.042	0.083	0.082	-0.043	1.000
ILL_C250_199_[T/C]	0.042	0.083	0.082	-0.043	1.000
PS_C47845_238_[T/C]	0.313	0.292	0.439	0.321	0.151
ILL_C1363_269_[T/C]	0.333	0.333	0.454	0.250	0.357
ILL_C18774_676_[T/C]	0.043	0.087	0.085	-0.045	1.000
PS_C23647_375_[A/G]	0.065	0.130	0.125	-0.070	1.000
ILL_C2028_1228_[A/T]	0.042	0.083	0.082	-0.043	1.000
PS_C11985_171_[A/G]	0.130	0.174	0.232	0.233	0.319
PS_C11911_576_[A/C]	0.104	0.125	0.191	0.330	0.215
ILL_C1254_187_[A/G]	0.375	0.500	0.479	-0.067	1.000
ILL_C20427_267_[T/C]	0.063	0.125	0.120	-0.067	1.000
PS_C21989_160_[T/G]	0.125	0.250	0.223	-0.143	1.000
PS_C12069_1181_[T/C]	0.021	0.042	0.042	-0.021	1.000

ILL_C327_1076_[C/G]	0.042	0.000	0.082	1.000	0.024
PS_C11847_438_[T/C]	0.341	0.500	0.460	-0.113	1.000
PS_C47375_253_[A/G]	0.500	0.000	0.511	1.000	0.000
PS_C34420_787_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C2915_875_[T/C]	0.174	0.348	0.294	-0.211	1.000
ILL_C22491_727_[T/C]	0.478	0.609	0.510	-0.220	0.413
ILL_C911_1343_[T/A]	0.354	0.458	0.467	-0.002	1.000
PS_C12196_434_[G/C]	0.354	0.625	0.467	-0.366	0.181
PS_C14297_369_[T/G]	0.458	0.583	0.507	-0.175	0.666
PS_C47340_198_[A/G]	0.067	0.000	0.129	1.000	0.034
PS_C36706_579_[T/C]	0.083	0.167	0.156	-0.091	1.000
PS_C46597_301_[C/G]	0.125	0.250	0.223	-0.143	1.000
ILL_C2040_1251_[A/T]	0.146	0.208	0.254	0.164	0.390
PS_C35683_190_[A/G]	0.438	0.375	0.503	0.238	0.254
PS_C46532_529_[A/T]	0.042	0.083	0.082	-0.043	1.000
PS_C15379_78_[A/C]	0.292	0.417	0.422	-0.008	1.000
PS_C11871_171_[A/G]	0.458	0.667	0.507	-0.343	0.217
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C6061_1289_[T/G]	0.326	0.478	0.449	-0.088	1.000
ILL_C980_261_[A/G]	0.146	0.292	0.254	-0.171	1.000
Mean	**0.218**	**0.320**	**0.305**	**-0.032**	

RP(2003)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C15402_271_[T/A]	0.032	0.064	0.062	-0.033	1.000
ILL_C253_1545_[A/C]	0.122	0.204	0.217	0.050	0.540
ILL_C428_225_[A/G]	0.304	0.451	0.427	-0.066	0.740
PS_C23094_578_[G/C]	0.098	0.118	0.179	0.335	0.050
ILL_C20682_843_[A/T]	0.063	0.125	0.119	-0.067	1.000
PS_C34490_403_[A/G]	0.337	0.388	0.451	0.132	0.348
PS_C25208 _377_[T/G]	0.010	0.020	0.020	-0.010	1.000
ILL_C1783_492_[T/G]	0.137	0.235	0.239	0.006	1.000
ILL_C929_734_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C11665_287_[A/G]	0.167	0.292	0.281	-0.050	1.000
ILL_C2141_350_[T/A]	0.418	0.510	0.492	-0.048	1.000
PS_C23591_200_[T/C]	0.281	0.521	0.409	-0.288	0.074
PS_C15689_162_[A/T]	0.214	0.429	0.340	-0.273	0.096
ILL_C229_2772_[A/G]	0.100	0.200	0.182	-0.111	1.000
ILL_C22574_507_[A/C]	0.098	0.078	0.179	0.557	0.002
PS_C15088_268_[A/G]	0.489	0.383	0.505	0.234	0.123
ILL_C2406_641_[T/C]	0.049	0.098	0.094	-0.052	1.000
PS_C12925_666_[A/G]	0.100	0.200	0.182	-0.111	1.000
PS_C11984_159_[T/G]	0.490	0.860	0.505	-0.721	0.000
ILL_C22449_261_[A/G]	0.216	0.275	0.342	0.189	0.206
PS_C11970_157_[A/G]	0.240	0.440	0.368	-0.206	0.240

PS_C23051_368_[A/G]	0.071	0.143	0.135	-0.077	1.000
ILL_C6012_280_[T/C]	0.353	0.471	0.461	-0.030	1.000
ILL_C2903_1043_[T/A]	0.311	0.356	0.433	0.171	0.301
PS_C34501_638_[A/G]	0.020	0.040	0.040	-0.020	1.000
PS_C12352_527_[T/C]	0.378	0.347	0.475	0.262	0.075
PS_C34725_229_[T/C]	0.230	0.300	0.358	0.153	0.249
PS_C15230_93_[A/G]	0.460	0.520	0.502	-0.047	1.000
PS_C23075_525_[G/C]	0.265	0.373	0.393	0.043	0.724
ILL_C3835_411_[A/G]	0.225	0.216	0.353	0.382	0.013
PS_C24743_123_[T/G]	0.108	0.216	0.194	-0.121	1.000
ILL_C5634_234_[A/G]	0.353	0.510	0.461	-0.116	0.542
PS_C23630_237_[T/C]	0.108	0.176	0.194	0.083	0.447
ILL_C1813_300_[T/G]	0.196	0.353	0.318	-0.120	0.655
PS_C12218_188_[A/G]	0.157	0.314	0.267	-0.186	0.329
PS_C36136_185_[T/C]	0.373	0.314	0.472	0.329	0.018
PS_C23070_1364_[A/G]	0.265	0.412	0.393	-0.058	1.000
ILL_C618_116_[A/G]	0.069	0.098	0.129	0.233	0.199
PS_C1652_228_[A/C]	0.137	0.157	0.239	0.338	0.038
PS_C25083_285_[A/G]	0.069	0.098	0.129	0.233	0.209
ILL_C394_1510_[A/G]	0.420	0.440	0.492	0.097	0.552
PS_C24267_71_[A/G]	0.500	1.000	0.505	-1.000	0.000
PS_C34507_1191_[T/C]	0.316	0.388	0.437	0.104	0.527
PS_C36273_73_[A/C]	0.051	0.061	0.098	0.368	0.098
PS_C35977_153_[T/C]	0.060	0.000	0.114	1.000	0.000
PS_C16093_142_[T/C]	0.500	0.490	0.505	0.020	1.000
PS_C16031_147_[A/G]	0.284	0.490	0.411	-0.205	0.301
ILL_C5106_273_[T/C]	0.245	0.367	0.374	0.007	1.000
PS_C00512_245_[T/C]	0.020	0.039	0.039	-0.020	1.000
ILL_C140_2421_[A/G]	0.255	0.314	0.384	0.174	0.246
PS_C36237_70_[C/G]	0.398	0.388	0.484	0.191	0.229
PS_C34670_302_[A/C]	0.020	0.039	0.039	-0.020	1.000
ILL_C21880_1003_[G/C]	0.363	0.294	0.467	0.364	0.017
PS_C47330_170_[A/C]	0.315	0.587	0.436	-0.360	0.022
PS_C36522_249_[G/C]	0.327	0.449	0.444	-0.021	1.000
PS_C34604_423_[A/T]	0.225	0.373	0.353	-0.067	1.000
PS_C34501_77_[T/G]	0.461	0.255	0.502	0.487	0.001
PS_C28810_290_[A/C]	0.049	0.020	0.094	0.790	0.000
PS_C28886_317_[T/C]	0.010	0.020	0.020	-0.010	1.000
ILL_C31_1387_[T/C]	0.390	0.460	0.481	0.033	0.774
ILL_C844_440_[T/C]	0.029	0.059	0.058	-0.030	1.000
PS_C46857_366_[T/C]	0.130	0.140	0.228	0.381	0.019
PS_C34605_382_[A/G]	0.011	0.022	0.022	-0.011	1.000
ILL_C853_1199_[A/G]	0.147	0.255	0.253	-0.016	1.000
ILL_C1002_85_[T/A]	0.167	0.294	0.281	-0.059	1.000
PS_C11659_399_[T/C]	0.010	0.020	0.020	-0.010	1.000
ILL_C4791_1099_[T/G]	0.441	0.412	0.498	0.165	0.256

ILL_C300_4982_[A/G]	0.031	0.061	0.060	-0.032	1.000
ILL_C4593_326_[T/G]	0.353	0.510	0.461	-0.116	0.553
PS_C46533_541_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15351_193_[A/G]	0.480	0.569	0.504	-0.139	0.426
PS_C15018_147_[T/G]	0.020	0.039	0.039	-0.020	1.000
PS_C38608_168_[T/C]	0.500	0.959	0.505	-0.918	0.000
ILL_C387_215_[A/G]	0.031	0.021	0.061	0.656	0.031
ILL_C5339_366_[T/C]	0.030	0.060	0.059	-0.031	1.000
PS_C34511_71_[T/G]	0.049	0.098	0.094	-0.052	1.000
PS_C14838_228_[T/C]	0.198	0.229	0.321	0.278	0.072
ILL_C20267_102_[T/C]	0.388	0.367	0.480	0.226	0.135
PS_C46674_204_[C/G]	0.270	0.300	0.398	0.239	0.141
ILL_C1878_506_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C36563_85_[T/G]	0.125	0.250	0.221	-0.143	1.000
ILL_C2735_326_[T/G]	0.074	0.106	0.139	0.228	0.208
ILL_C250_199_[T/C]	0.020	0.039	0.039	-0.020	1.000
PS_C47845_238_[T/C]	0.410	0.460	0.489	0.049	0.773
ILL_C1363_269_[T/C]	0.296	0.306	0.421	0.265	0.084
ILL_C18774_676_[T/C]	0.075	0.150	0.141	-0.081	1.000
PS_C23647_375_[A/G]	0.130	0.261	0.229	-0.150	1.000
ILL_C2028_1228_[A/T]	0.100	0.160	0.182	0.111	0.376
PS_C11985_171_[A/G]	0.250	0.239	0.379	0.362	0.027
PS_C11911_576_[A/C]	0.157	0.314	0.267	-0.186	0.324
ILL_C1254_187_[A/G]	0.235	0.388	0.363	-0.079	1.000
ILL_C20427_267_[T/C]	0.050	0.100	0.096	-0.053	1.000
PS_C21989_160_[T/G]	0.130	0.220	0.228	0.027	1.000
PS_C12069_1181_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C327_1076_[C/G]	0.010	0.020	0.020	-0.010	1.000
PS_C11847_438_[T/C]	0.436	0.404	0.497	0.178	0.226
PS_C47375_253_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C34420_787_[A/G]	0.010	0.020	0.020	-0.010	1.000
ILL_C2915_875_[T/C]	0.149	0.298	0.256	-0.175	0.564
ILL_C22491_727_[T/C]	0.422	0.569	0.493	-0.166	0.381
ILL_C911_1343_[T/A]	0.363	0.490	0.467	-0.060	0.766
PS_C12196_434_[G/C]	0.433	0.422	0.497	0.140	0.384
PS_C14297_369_[T/G]	0.441	0.529	0.498	-0.074	0.790
PS_C47340_198_[A/G]	0.440	0.000	0.498	1.000	0.000
PS_C36706_579_[T/C]	0.157	0.314	0.267	-0.186	0.326
PS_C46597_301_[C/G]	0.088	0.176	0.162	-0.097	1.000
ILL_C2040_1251_[A/T]	0.127	0.255	0.225	-0.146	1.000
PS_C35683_190_[A/G]	0.392	0.431	0.481	0.095	0.549
PS_C46532_529_[A/T]	0.010	0.020	0.020	-0.010	1.000
PS_C15379_78_[A/C]	0.447	0.511	0.500	-0.033	1.000
PS_C11871_171_[A/G]	0.480	0.569	0.504	-0.139	0.399
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.150	0.220	0.258	0.137	0.256

ILL_C6061_1289_[T/G]	0.010	0.021	0.021	-0.011	1.000
ILL_C980_261_[A/G]	0.098	0.196	0.179	-0.109	1.000
Mean	**0.214**	**0.281**	**0.290**	**0.037**	
	RP(2011)				
Locus	**MAF**	**H_o**	**H_e**	**F_{is}**	***P*-value**
ILL_C2122_257_[A/G]	0.103	0.207	0.189	-0.115	1.000
PS_C15402_271_[T/A]	0.086	0.172	0.160	-0.094	1.000
ILL_C253_1545_[A/C]	0.065	0.129	0.123	-0.069	1.000
ILL_C428_225_[A/G]	0.167	0.333	0.282	-0.200	0.567
PS_C23094_578_[G/C]	0.081	0.161	0.151	-0.088	1.000
ILL_C20682_843_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C34490_403_[A/G]	0.375	0.536	0.477	-0.143	0.691
PS_C25208 _377_[T/G]	0.048	0.097	0.094	-0.051	1.000
ILL_C1783_492_[T/G]	0.145	0.290	0.252	-0.170	1.000
ILL_C929_734_[T/C]	0.037	0.074	0.073	-0.038	1.000
PS_C11665_287_[A/G]	0.167	0.333	0.282	-0.200	0.567
ILL_C2141_350_[T/A]	0.400	0.667	0.488	-0.389	0.062
PS_C23591_200_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15689_162_[A/T]	0.113	0.161	0.204	0.195	0.314
ILL_C229_2772_[A/G]	0.050	0.100	0.097	-0.053	1.000
ILL_C22574_507_[A/C]	0.067	0.133	0.127	-0.071	1.000
PS_C15088_268_[A/G]	0.403	0.677	0.489	-0.408	0.058
ILL_C2406_641_[T/C]	0.109	0.219	0.198	-0.123	1.000
PS_C12925_666_[A/G]	0.204	0.333	0.331	-0.027	1.000
PS_C11984_159_[T/G]	0.484	0.903	0.508	-0.808	0.000
ILL_C22449_261_[A/G]	0.323	0.452	0.444	-0.033	1.000
PS_C11970_157_[A/G]	0.167	0.267	0.282	0.040	1.000
PS_C23051_368_[A/G]	0.114	0.227	0.206	-0.128	1.000
ILL_C6012_280_[T/C]	0.150	0.300	0.259	-0.176	1.000
ILL_C2903_1043_[T/A]	0.411	0.607	0.493	-0.254	0.270
PS_C34501_638_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C12352_527_[T/C]	0.317	0.500	0.440	-0.155	0.676
PS_C34725_229_[T/C]	0.317	0.300	0.440	0.307	0.090
PS_C15230_93_[A/G]	0.440	0.640	0.503	-0.299	0.230
PS_C23075_525_[G/C]	0.258	0.452	0.389	-0.179	0.639
ILL_C3835_411_[A/G]	0.200	0.267	0.325	0.167	0.311
PS_C24743_123_[T/G]	0.290	0.516	0.419	-0.253	0.372
ILL_C5634_234_[A/G]	0.467	0.600	0.506	-0.205	0.470
PS_C23630_237_[T/C]	0.078	0.156	0.146	-0.085	1.000
ILL_C1813_300_[T/G]	0.167	0.333	0.282	-0.200	0.553
PS_C12218_188_[A/G]	0.177	0.290	0.297	0.005	1.000
PS_C36136_185_[T/C]	0.317	0.500	0.440	-0.155	0.672
PS_C23070_1364_[A/G]	0.203	0.344	0.329	-0.062	1.000
ILL_C618_116_[A/G]	0.052	0.103	0.100	-0.055	1.000
PS_C1652_228_[A/C]	0.031	0.063	0.062	-0.032	1.000
PS_C25083_285_[A/G]	0.367	0.267	0.472	0.426	0.027

ILL_C394_1510_[A/G]	0.383	0.567	0.481	-0.199	0.448
PS_C24267_71_[A/G]	0.281	0.563	0.411	-0.391	0.068
PS_C34507_1191_[T/C]	0.250	0.367	0.381	0.022	1.000
PS_C36273_73_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C35977_153_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C16093_142_[T/C]	0.452	0.387	0.503	0.218	0.267
PS_C16031_147_[A/G]	0.323	0.452	0.444	-0.033	1.000
ILL_C5106_273_[T/C]	0.241	0.345	0.373	0.058	0.651
PS_C00512_245_[T/C]	0.017	0.033	0.033	-0.017	1.000
ILL_C140_2421_[A/G]	0.183	0.367	0.305	-0.224	0.545
PS_C36237_70_[C/G]	0.333	0.400	0.452	0.100	0.676
PS_C34670_302_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C21880_1003_[G/C]	0.250	0.433	0.381	-0.156	0.635
PS_C47330_170_[A/C]	0.300	0.533	0.427	-0.270	0.228
PS_C36522_249_[G/C]	0.293	0.379	0.422	0.085	0.664
PS_C34604_423_[A/T]	0.183	0.367	0.305	-0.224	0.551
PS_C34501_77_[T/G]	0.344	0.063	0.458	0.861	0.000
PS_C28810_290_[A/C]	0.031	0.063	0.062	-0.032	1.000
PS_C28886_317_[T/C]	0.031	0.063	0.062	-0.032	1.000
ILL_C31_1387_[T/C]	0.339	0.226	0.455	0.496	0.012
ILL_C844_440_[T/C]	0.048	0.097	0.094	-0.051	1.000
PS_C46857_366_[T/C]	0.129	0.258	0.228	-0.148	1.000
PS_C34605_382_[A/G]	0.032	0.065	0.063	-0.033	1.000
ILL_C853_1199_[A/G]	0.177	0.355	0.297	-0.216	0.551
ILL_C1002_85_[T/A]	0.422	0.844	0.496	-0.730	0.000
PS_C11659_399_[T/C]	0.065	0.129	0.123	-0.069	1.000
ILL_C4791_1099_[T/G]	0.387	0.645	0.482	-0.360	0.074
ILL_C300_4982_[A/G]	0.048	0.097	0.094	-0.051	1.000
ILL_C4593_326_[T/G]	0.450	0.567	0.503	-0.145	0.709
PS_C46533_541_[T/C]	0.141	0.281	0.246	-0.164	1.000
PS_C15351_193_[A/G]	0.435	0.548	0.500	-0.115	0.726
PS_C15018_147_[T/G]	0.016	0.032	0.032	-0.016	1.000
PS_C38608_168_[T/C]	0.484	0.968	0.508	-0.938	0.000
ILL_C387_215_[A/G]	0.020	0.040	0.040	-0.020	1.000
ILL_C5339_366_[T/C]	0.161	0.258	0.275	0.046	1.000
PS_C34511_71_[T/G]	0.018	0.036	0.036	-0.018	1.000
PS_C14838_228_[T/C]	0.067	0.133	0.127	-0.071	1.000
ILL_C20267_102_[T/C]	0.226	0.387	0.355	-0.107	1.000
PS_C46674_204_[C/G]	0.283	0.433	0.413	-0.067	1.000
ILL_C1878_506_[T/C]	0.094	0.188	0.173	-0.103	1.000
PS_C36563_85_[T/G]	0.107	0.214	0.195	-0.120	1.000
ILL_C2735_326_[T/G]	0.050	0.100	0.097	-0.053	1.000
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.355	0.452	0.465	0.014	1.000
ILL_C1363_269_[T/C]	0.304	0.250	0.431	0.409	0.067
ILL_C18774_676_[T/C]	0.086	0.172	0.160	-0.094	1.000

PS_C23647_375_[A/G]	0.167	0.250	0.284	0.100	0.511
ILL_C2028_1228_[A/T]	0.067	0.133	0.127	-0.071	1.000
PS_C11985_171_[A/G]	0.346	0.538	0.462	-0.190	0.659
PS_C11911_576_[A/C]	0.083	0.100	0.155	0.345	0.161
ILL_C1254_187_[A/G]	0.350	0.433	0.463	0.048	1.000
ILL_C20427_267_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C21989_160_[T/G]	0.145	0.226	0.252	0.090	0.503
PS_C12069_1181_[T/C]	0.033	0.067	0.066	-0.034	1.000
ILL_C327_1076_[C/G]	N/A	N/A	N/A	N/A	N/A
PS_C11847_438_[T/C]	0.340	0.680	0.458	-0.515	0.019
PS_C47375_253_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C34420_787_[A/G]	0.017	0.033	0.033	-0.017	1.000
ILL_C2915_875_[T/C]	0.286	0.476	0.418	-0.167	0.631
ILL_C22491_727_[T/C]	0.283	0.500	0.413	-0.231	0.371
ILL_C911_1343_[T/A]	0.333	0.467	0.452	-0.050	1.000
PS_C12196_434_[G/C]	0.500	0.438	0.508	0.125	0.498
PS_C14297_369_[T/G]	0.339	0.548	0.455	-0.224	0.420
PS_C47340_198_[A/G]	0.344	0.000	0.458	1.000	0.000
PS_C36706_579_[T/C]	0.129	0.258	0.228	-0.148	1.000
PS_C46597_301_[C/G]	0.100	0.200	0.183	-0.111	1.000
ILL_C2040_1251_[A/T]	0.031	0.063	0.062	-0.032	1.000
PS_C35683_190_[A/G]	0.350	0.500	0.463	-0.099	0.711
PS_C46532_529_[A/T]	0.016	0.032	0.032	-0.016	1.000
PS_C15379_78_[A/C]	0.483	0.552	0.508	-0.105	0.727
PS_C11871_171_[A/G]	0.367	0.467	0.472	-0.005	1.000
PS_C35823_1199_[A/G]	0.031	0.000	0.062	1.000	0.015
PS_C39731_379_[A/C]	0.078	0.156	0.146	-0.085	1.000
ILL_C6061_1289_[T/G]	0.017	0.034	0.034	-0.018	1.000
ILL_C980_261_[A/G]	0.150	0.300	0.259	-0.176	1.000
Mean	**0.209**	**0.314**	**0.294**	**-0.063**	

SD(2004)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.011	0.022	0.022	-0.011	1.000
PS_C15402_271_[T/A]	0.077	0.154	0.144	-0.083	1.000
ILL_C253_1545_[A/C]	0.113	0.175	0.202	0.124	0.398
ILL_C428_225_[A/G]	0.228	0.370	0.356	-0.049	1.000
PS_C23094_578_[G/C]	0.076	0.152	0.142	-0.082	1.000
ILL_C20682_843_[A/T]	0.061	0.073	0.116	0.361	0.117
PS_C34490_403_[A/G]	0.370	0.348	0.471	0.254	0.104
PS_C25208 _377_[T/G]	0.011	0.022	0.022	-0.011	1.000
ILL_C1783_492_[T/G]	0.076	0.152	0.142	-0.082	1.000
ILL_C929_734_[T/C]	0.054	0.054	0.104	0.471	0.083
PS_C11665_287_[A/G]	0.081	0.116	0.151	0.222	0.232
ILL_C2141_350_[T/A]	0.344	0.467	0.457	-0.033	1.000
PS_C23591_200_[T/C]	0.095	0.189	0.174	-0.104	1.000
PS_C15689_162_[A/T]	0.189	0.378	0.310	-0.233	0.318

ILL_C229_2772_[A/G]	0.057	0.114	0.108	-0.060	1.000
ILL_C22574_507_[A/C]	0.185	0.326	0.305	-0.082	1.000
PS_C15088_268_[A/G]	0.478	0.511	0.505	-0.024	1.000
ILL_C2406_641_[T/C]	0.211	0.378	0.337	-0.134	0.651
PS_C12925_666_[A/G]	0.158	0.263	0.269	0.010	1.000
PS_C11984_159_[T/G]	0.488	0.881	0.506	-0.763	0.000
ILL_C22449_261_[A/G]	0.261	0.522	0.390	-0.353	0.016
PS_C11970_157_[A/G]	0.141	0.239	0.245	0.015	1.000
PS_C23051_368_[A/G]	0.278	0.481	0.409	-0.200	0.623
ILL_C6012_280_[T/C]	0.477	0.455	0.505	0.089	0.562
ILL_C2903_1043_[T/A]	0.242	0.424	0.373	-0.155	0.641
PS_C34501_638_[A/G]	0.043	0.000	0.084	1.000	0.000
PS_C12352_527_[T/C]	0.364	0.500	0.468	-0.080	0.748
PS_C34725_229_[T/C]	0.344	0.378	0.457	0.163	0.331
PS_C15230_93_[A/G]	0.409	0.455	0.489	0.060	0.765
PS_C23075_525_[G/C]	0.130	0.261	0.229	-0.150	1.000
ILL_C3835_411_[A/G]	0.402	0.587	0.486	-0.221	0.211
PS_C24743_123_[T/G]	0.304	0.522	0.428	-0.232	0.173
ILL_C5634_234_[A/G]	0.389	0.378	0.481	0.205	0.225
PS_C23630_237_[T/C]	0.116	0.140	0.208	0.321	0.088
ILL_C1813_300_[T/G]	0.152	0.217	0.261	0.158	0.255
PS_C12218_188_[A/G]	0.174	0.304	0.290	-0.059	1.000
PS_C36136_185_[T/C]	0.435	0.435	0.497	0.115	0.551
PS_C23070_1364_[A/G]	0.109	0.174	0.196	0.102	0.424
ILL_C618_116_[A/G]	0.196	0.348	0.318	-0.105	1.000
PS_C1652_228_[A/C]	0.083	0.083	0.154	0.455	0.020
PS_C25083_285_[A/G]	0.033	0.065	0.064	-0.034	1.000
ILL_C394_1510_[A/G]	0.435	0.435	0.497	0.115	0.542
PS_C24267_71_[A/G]	0.490	0.979	0.505	-0.959	0.000
PS_C34507_1191_[T/C]	0.337	0.349	0.452	0.220	0.177
PS_C36273_73_[A/C]	0.054	0.109	0.104	-0.057	1.000
PS_C35977_153_[T/C]	0.045	0.000	0.088	1.000	0.001
PS_C16093_142_[T/C]	0.359	0.500	0.465	-0.087	0.743
PS_C16031_147_[A/G]	0.261	0.435	0.390	-0.127	0.703
ILL_C5106_273_[T/C]	0.190	0.238	0.312	0.228	0.149
PS_C00512_245_[T/C]	0.023	0.045	0.045	-0.023	1.000
ILL_C140_2421_[A/G]	0.163	0.239	0.276	0.124	0.313
PS_C36237_70_[C/G]	0.054	0.109	0.104	-0.057	1.000
PS_C34670_302_[A/C]	0.011	0.022	0.022	-0.011	1.000
ILL_C21880_1003_[G/C]	0.239	0.478	0.368	-0.314	0.049
PS_C47330_170_[A/C]	0.333	0.556	0.451	-0.250	0.242
PS_C36522_249_[G/C]	0.400	0.400	0.485	0.167	0.341
PS_C34604_423_[A/T]	0.217	0.304	0.344	0.106	0.430
PS_C34501_77_[T/G]	0.208	0.375	0.333	-0.137	0.660
PS_C28810_290_[A/C]	0.198	0.146	0.321	0.541	0.001
PS_C28886_317_[T/C]	0.044	0.089	0.086	-0.047	1.000

ILL_C31_1387_[T/C]	0.478	0.565	0.505	-0.133	0.560
ILL_C844_440_[T/C]	0.044	0.089	0.086	-0.047	1.000
PS_C46857_366_[T/C]	0.078	0.111	0.145	0.225	0.225
PS_C34605_382_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C853_1199_[A/G]	0.152	0.304	0.261	-0.179	0.571
ILL_C1002_85_[T/A]	0.292	0.500	0.418	-0.210	0.295
PS_C11659_399_[T/C]	0.098	0.109	0.178	0.384	0.041
ILL_C4791_1099_[T/G]	0.489	0.674	0.505	-0.348	0.032
ILL_C300_4982_[A/G]	0.012	0.024	0.024	-0.012	1.000
ILL_C4593_326_[T/G]	0.196	0.348	0.318	-0.105	1.000
PS_C46533_541_[T/C]	0.021	0.042	0.041	-0.021	1.000
PS_C15351_193_[A/G]	0.337	0.457	0.452	-0.022	1.000
PS_C15018_147_[T/G]	0.033	0.067	0.065	-0.034	1.000
PS_C38608_168_[T/C]	0.500	1.000	0.506	-1.000	0.000
ILL_C387_215_[A/G]	0.098	0.000	0.178	1.000	0.000
ILL_C5339_366_[T/C]	0.076	0.152	0.142	-0.082	1.000
PS_C34511_71_[T/G]	0.044	0.089	0.086	-0.047	1.000
PS_C14838_228_[T/C]	0.136	0.273	0.238	-0.158	1.000
ILL_C20267_102_[T/C]	0.228	0.413	0.356	-0.172	0.407
PS_C46674_204_[C/G]	0.286	0.476	0.413	-0.167	0.447
ILL_C1878_506_[T/C]	0.043	0.043	0.084	0.477	0.060
PS_C36563_85_[T/G]	0.107	0.119	0.194	0.378	0.046
ILL_C2735_326_[T/G]	0.061	0.122	0.116	-0.065	1.000
ILL_C250_199_[T/C]	0.054	0.109	0.104	-0.057	1.000
PS_C47845_238_[T/C]	0.318	0.409	0.439	0.057	0.734
ILL_C1363_269_[T/C]	0.391	0.478	0.482	-0.004	1.000
ILL_C18774_676_[T/C]	0.081	0.097	0.151	0.347	0.168
PS_C23647_375_[A/G]	0.024	0.049	0.048	-0.025	1.000
ILL_C2028_1228_[A/T]	0.033	0.065	0.064	-0.034	1.000
PS_C11985_171_[A/G]	0.133	0.133	0.234	0.423	0.016
PS_C11911_576_[A/C]	0.044	0.089	0.086	-0.047	1.000
ILL_C1254_187_[A/G]	0.276	0.447	0.405	-0.119	0.691
ILL_C20427_267_[T/C]	0.047	0.093	0.090	-0.049	1.000
PS_C21989_160_[T/G]	0.196	0.391	0.318	-0.243	0.173
PS_C12069_1181_[T/C]	0.054	0.109	0.104	-0.057	1.000
ILL_C327_1076_[C/G]	0.054	0.022	0.104	0.789	0.002
PS_C11847_438_[T/C]	0.477	0.364	0.505	0.271	0.080
PS_C47375_253_[A/G]	0.457	0.000	0.502	1.000	0.000
PS_C34420_787_[A/G]	0.011	0.022	0.022	-0.011	1.000
ILL_C2915_875_[T/C]	0.209	0.279	0.335	0.157	0.343
ILL_C22491_727_[T/C]	0.304	0.348	0.428	0.179	0.288
ILL_C911_1343_[T/A]	0.250	0.370	0.379	0.014	1.000
PS_C12196_434_[G/C]	0.371	0.514	0.474	-0.101	0.728
PS_C14297_369_[T/G]	0.402	0.500	0.486	-0.040	1.000
PS_C47340_198_[A/G]	0.029	0.000	0.056	1.000	0.015
PS_C36706_579_[T/C]	0.174	0.348	0.290	-0.211	0.305

PS_C46597_301_[C/G]	0.178	0.311	0.296	-0.064	1.000
ILL_C2040_1251_[A/T]	0.105	0.116	0.190	0.380	0.045
PS_C35683_190_[A/G]	0.341	0.409	0.455	0.090	0.509
PS_C46532_529_[A/T]	0.043	0.087	0.084	-0.045	1.000
PS_C15379_78_[A/C]	0.397	0.590	0.485	-0.231	0.202
PS_C11871_171_[A/G]	0.500	0.478	0.505	0.043	0.773
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.056	0.111	0.106	-0.059	1.000
ILL_C6061_1289_[T/G]	0.386	0.455	0.480	0.041	0.761
ILL_C980_261_[A/G]	0.152	0.304	0.261	-0.179	0.581
Mean	**0.201**	**0.281**	**0.279**	**0.036**	

SD(2010)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.065	0.065	0.123	0.466	0.106
PS_C15402_271_[T/A]	0.058	0.115	0.111	-0.061	1.000
ILL_C253_1545_[A/C]	0.065	0.129	0.123	-0.069	1.000
ILL_C428_225_[A/G]	0.242	0.355	0.373	0.033	1.000
PS_C23094_578_[G/C]	0.083	0.100	0.155	0.345	0.157
ILL_C20682_843_[A/T]	0.053	0.105	0.102	-0.056	1.000
PS_C34490_403_[A/G]	0.444	0.519	0.503	-0.050	1.000
PS_C25208 _377_[T/G]	0.016	0.032	0.032	-0.016	1.000
ILL_C1783_492_[T/G]	0.017	0.033	0.033	-0.017	1.000
ILL_C929_734_[T/C]	0.063	0.125	0.120	-0.067	1.000
PS_C11665_287_[A/G]	0.143	0.143	0.249	0.417	0.059
ILL_C2141_350_[T/A]	0.500	0.586	0.509	-0.172	0.483
PS_C23591_200_[T/C]	0.045	0.091	0.091	-0.048	1.000
PS_C15689_162_[A/T]	0.145	0.226	0.252	0.090	0.495
ILL_C229_2772_[A/G]	0.117	0.167	0.210	0.191	0.326
ILL_C22574_507_[A/C]	0.113	0.226	0.204	-0.127	1.000
PS_C15088_268_[A/G]	0.500	0.667	0.508	-0.333	0.147
ILL_C2406_641_[T/C]	0.133	0.267	0.235	-0.154	1.000
PS_C12925_666_[A/G]	0.100	0.200	0.185	-0.111	1.000
PS_C11984_159_[T/G]	0.400	0.800	0.488	-0.667	0.001
ILL_C22449_261_[A/G]	0.226	0.258	0.355	0.262	0.141
PS_C11970_157_[A/G]	0.258	0.323	0.389	0.158	0.363
PS_C23051_368_[A/G]	0.111	0.222	0.203	-0.125	1.000
ILL_C6012_280_[T/C]	0.484	0.452	0.508	0.096	0.717
ILL_C2903_1043_[T/A]	0.323	0.452	0.444	-0.033	1.000
PS_C34501_638_[A/G]	0.161	0.000	0.275	1.000	0.000
PS_C12352_527_[T/C]	0.407	0.593	0.492	-0.227	0.429
PS_C34725_229_[T/C]	0.417	0.367	0.494	0.246	0.252
PS_C15230_93_[A/G]	0.383	0.500	0.481	-0.058	1.000
PS_C23075_525_[G/C]	0.217	0.367	0.345	-0.080	1.000
ILL_C3835_411_[A/G]	0.371	0.355	0.474	0.240	0.252
PS_C24743_123_[T/G]	0.233	0.400	0.364	-0.118	1.000
ILL_C5634_234_[A/G]	0.371	0.226	0.474	0.516	0.005

PS_C23630_237_[T/C]	0.065	0.129	0.123	-0.069	1.000
ILL_C1813_300_[T/G]	0.258	0.452	0.389	-0.179	0.638
PS_C12218_188_[A/G]	0.100	0.200	0.183	-0.111	1.000
PS_C36136_185_[T/C]	0.383	0.433	0.481	0.083	0.707
PS_C23070_1364_[A/G]	0.177	0.226	0.297	0.226	0.209
ILL_C618_116_[A/G]	0.268	0.393	0.399	-0.002	1.000
PS_C1652_228_[A/C]	0.113	0.161	0.204	0.195	0.321
PS_C25083_285_[A/G]	0.117	0.100	0.210	0.515	0.029
ILL_C394_1510_[A/G]	0.339	0.419	0.455	0.064	0.697
PS_C24267_71_[A/G]	0.468	0.935	0.506	-0.879	0.000
PS_C34507_1191_[T/C]	0.379	0.345	0.479	0.268	0.230
PS_C36273_73_[A/C]	0.048	0.097	0.094	-0.051	1.000
PS_C35977_153_[T/C]	0.161	0.000	0.275	1.000	0.000
PS_C16093_142_[T/C]	0.435	0.548	0.500	-0.115	0.718
PS_C16031_147_[A/G]	0.323	0.452	0.444	-0.033	1.000
ILL_C5106_273_[T/C]	0.117	0.233	0.210	-0.132	1.000
PS_C00512_245_[T/C]	0.020	0.040	0.040	-0.020	1.000
ILL_C140_2421_[A/G]	0.167	0.267	0.282	0.040	1.000
PS_C36237_70_[C/G]	0.200	0.333	0.325	-0.042	1.000
PS_C34670_302_[A/C]	0.016	0.032	0.032	-0.016	1.000
ILL_C21880_1003_[G/C]	0.339	0.419	0.455	0.064	0.693
PS_C47330_170_[A/C]	0.328	0.517	0.448	-0.174	0.679
PS_C36522_249_[G/C]	0.350	0.500	0.463	-0.099	0.708
PS_C34604_423_[A/T]	0.323	0.387	0.444	0.114	0.673
PS_C34501_77_[T/G]	0.048	0.032	0.094	0.650	0.052
PS_C28810_290_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.367	0.667	0.472	-0.435	0.042
ILL_C844_440_[T/C]	0.016	0.032	0.032	-0.016	1.000
PS_C46857_366_[T/C]	0.121	0.241	0.216	-0.137	1.000
PS_C34605_382_[A/G]	0.016	0.032	0.032	-0.016	1.000
ILL_C853_1199_[A/G]	0.145	0.290	0.252	-0.170	1.000
ILL_C1002_85_[T/A]	0.339	0.613	0.455	-0.368	0.109
PS_C11659_399_[T/C]	0.017	0.033	0.033	-0.017	1.000
ILL_C4791_1099_[T/G]	0.419	0.516	0.495	-0.060	1.000
ILL_C300_4982_[A/G]	0.048	0.097	0.094	-0.051	1.000
ILL_C4593_326_[T/G]	0.250	0.500	0.381	-0.333	0.150
PS_C46533_541_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15351_193_[A/G]	0.500	0.613	0.508	-0.226	0.291
PS_C15018_147_[T/G]	0.033	0.067	0.066	-0.034	1.000
PS_C38608_168_[T/C]	0.500	1.000	0.508	-1.000	0.000
ILL_C387_215_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C5339_366_[T/C]	0.017	0.034	0.034	-0.018	1.000
PS_C34511_71_[T/G]	0.020	0.040	0.040	-0.020	1.000
PS_C14838_228_[T/C]	0.194	0.387	0.317	-0.240	0.566
ILL_C20267_102_[T/C]	0.183	0.300	0.305	-0.002	1.000

PS_C46674_204_[C/G]	0.217	0.300	0.345	0.116	0.591
ILL_C1878_506_[T/C]	0.081	0.161	0.151	-0.088	1.000
PS_C36563_85_[T/G]	0.068	0.136	0.130	-0.073	1.000
ILL_C2735_326_[T/G]	0.107	0.214	0.195	-0.120	1.000
ILL_C250_199_[T/C]	0.097	0.129	0.178	0.262	0.230
PS_C47845_238_[T/C]	0.400	0.400	0.488	0.167	0.451
ILL_C1363_269_[T/C]	0.226	0.194	0.355	0.446	0.021
ILL_C18774_676_[T/C]	0.130	0.259	0.230	-0.149	1.000
PS_C23647_375_[A/G]	0.212	0.423	0.340	-0.268	0.550
ILL_C2028_1228_[A/T]	0.054	0.107	0.103	-0.057	1.000
PS_C11985_171_[A/G]	0.160	0.160	0.274	0.405	0.084
PS_C11911_576_[A/C]	0.048	0.097	0.094	-0.051	1.000
ILL_C1254_187_[A/G]	0.208	0.333	0.337	-0.011	1.000
ILL_C20427_267_[T/C]	0.100	0.200	0.183	-0.111	1.000
PS_C21989_160_[T/G]	0.117	0.233	0.210	-0.132	1.000
PS_C12069_1181_[T/C]	0.065	0.129	0.123	-0.069	1.000
ILL_C327_1076_[C/G]	0.016	0.032	0.032	-0.016	1.000
PS_C11847_438_[T/C]	0.327	0.423	0.449	0.039	1.000
PS_C47375_253_[A/G]	0.323	0.000	0.444	1.000	0.000
PS_C34420_787_[A/G]	0.017	0.033	0.033	-0.017	1.000
ILL_C2915_875_[T/C]	0.313	0.375	0.439	0.127	0.636
ILL_C22491_727_[T/C]	0.362	0.379	0.470	0.179	0.418
ILL_C911_1343_[T/A]	0.323	0.387	0.444	0.114	0.675
PS_C12196_434_[G/C]	0.435	0.613	0.500	-0.247	0.274
PS_C14297_369_[T/G]	0.355	0.387	0.465	0.155	0.435
PS_C47340_198_[A/G]	0.034	0.000	0.068	1.000	0.019
PS_C36706_579_[T/C]	0.133	0.267	0.235	-0.154	1.000
PS_C46597_301_[C/G]	0.113	0.161	0.204	0.195	0.317
ILL_C2040_1251_[A/T]	0.117	0.233	0.210	-0.132	1.000
PS_C35683_190_[A/G]	0.387	0.452	0.482	0.048	1.000
PS_C46532_529_[A/T]	0.033	0.067	0.066	-0.034	1.000
PS_C15379_78_[A/C]	0.446	0.393	0.503	0.205	0.282
PS_C11871_171_[A/G]	0.468	0.548	0.506	-0.101	0.726
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.194	0.323	0.317	-0.033	1.000
ILL_C6061_1289_[T/G]	0.333	0.400	0.452	0.100	0.696
ILL_C980_261_[A/G]	0.100	0.067	0.183	0.630	0.013
Mean	**0.208**	**0.285**	**0.290**	**0.027**	

WC(F_1)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.056	0.111	0.108	-0.059	1.000
PS_C15402_271_[T/A]	0.079	0.158	0.149	-0.086	1.000
ILL_C253_1545_[A/C]	0.028	0.056	0.056	-0.029	1.000
ILL_C428_225_[A/G]	0.211	0.316	0.341	0.050	1.000
PS_C23094_578_[G/C]	0.118	0.235	0.214	-0.133	1.000

ILL_C20682_843_[A/T]	0.029	0.059	0.059	-0.030	1.000
PS_C34490_403_[A/G]	0.342	0.474	0.462	-0.052	1.000
PS_C25208 _377_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C1783_492_[T/G]	0.158	0.316	0.273	-0.187	1.000
ILL_C929_734_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C11665_287_[A/G]	0.184	0.263	0.309	0.124	0.489
ILL_C2141_350_[T/A]	0.316	0.526	0.444	-0.218	0.614
PS_C23591_200_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15689_162_[A/T]	0.026	0.053	0.053	-0.027	1.000
ILL_C229_2772_[A/G]	0.053	0.105	0.102	-0.056	1.000
ILL_C22574_507_[A/C]	0.026	0.053	0.053	-0.027	1.000
PS_C15088_268_[A/G]	0.368	0.526	0.478	-0.131	1.000
ILL_C2406_641_[T/C]	0.132	0.263	0.235	-0.152	1.000
PS_C12925_666_[A/G]	0.147	0.294	0.258	-0.172	1.000
PS_C11984_159_[T/G]	0.500	1.000	0.514	-1.000	0.000
ILL_C22449_261_[A/G]	0.263	0.526	0.398	-0.357	0.264
PS_C11970_157_[A/G]	0.222	0.333	0.356	0.036	1.000
PS_C23051_368_[A/G]	0.059	0.118	0.114	-0.063	1.000
ILL_C6012_280_[T/C]	0.263	0.316	0.398	0.186	0.553
ILL_C2903_1043_[T/A]	0.105	0.211	0.193	-0.118	1.000
PS_C34501_638_[A/G]	0.079	0.053	0.149	0.638	0.087
PS_C12352_527_[T/C]	0.472	0.389	0.513	0.220	0.371
PS_C34725_229_[T/C]	0.263	0.316	0.398	0.186	0.548
PS_C15230_93_[A/G]	0.500	0.667	0.514	-0.333	0.338
PS_C23075_525_[G/C]	0.263	0.421	0.398	-0.086	1.000
ILL_C3835_411_[A/G]	0.158	0.211	0.273	0.208	0.367
PS_C24743_123_[T/G]	0.105	0.211	0.193	-0.118	1.000
ILL_C5634_234_[A/G]	0.421	0.632	0.501	-0.295	0.361
PS_C23630_237_[T/C]	0.028	0.056	0.056	-0.029	1.000
ILL_C1813_300_[T/G]	0.211	0.316	0.341	0.050	1.000
PS_C12218_188_[A/G]	0.184	0.368	0.309	-0.226	1.000
PS_C36136_185_[T/C]	0.289	0.474	0.422	-0.152	1.000
PS_C23070_1364_[A/G]	0.342	0.368	0.462	0.182	0.603
ILL_C618_116_[A/G]	0.184	0.263	0.309	0.124	0.487
PS_C1652_228_[A/C]	0.167	0.238	0.285	0.143	0.446
PS_C25083_285_[A/G]	0.269	0.538	0.409	-0.368	0.500
ILL_C394_1510_[A/G]	0.500	0.684	0.514	-0.368	0.198
PS_C24267_71_[A/G]	0.452	0.905	0.508	-0.826	0.000
PS_C34507_1191_[T/C]	0.237	0.263	0.371	0.272	0.241
PS_C36273_73_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C35977_153_[T/C]	0.053	0.000	0.102	1.000	0.026
PS_C16093_142_[T/C]	0.368	0.316	0.478	0.321	0.167
PS_C16031_147_[A/G]	0.237	0.368	0.371	-0.019	1.000
ILL_C5106_273_[T/C]	0.395	0.474	0.491	0.009	1.000
PS_C00512_245_[T/C]	0.056	0.111	0.108	-0.059	1.000
ILL_C140_2421_[A/G]	0.105	0.211	0.193	-0.118	1.000

PS_C36237_70_[C/G]	0.237	0.474	0.371	-0.310	0.521
PS_C34670_302_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C21880_1003_[G/C]	0.474	0.421	0.512	0.156	0.637
PS_C47330_170_[A/C]	0.222	0.444	0.356	-0.286	0.522
PS_C36522_249_[G/C]	0.353	0.588	0.471	-0.288	0.595
PS_C34604_423_[A/T]	0.237	0.474	0.371	-0.310	0.514
PS_C34501_77_[T/G]	0.429	0.095	0.502	0.806	0.001
PS_C28810_290_[A/C]	0.048	0.095	0.093	-0.050	1.000
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.447	0.263	0.508	0.468	0.062
ILL_C844_440_[T/C]	0.056	0.111	0.108	-0.059	1.000
PS_C46857_366_[T/C]	0.158	0.211	0.273	0.208	0.367
PS_C34605_382_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C853_1199_[A/G]	0.079	0.158	0.149	-0.086	1.000
ILL_C1002_85_[T/A]	0.381	0.762	0.483	-0.615	0.014
PS_C11659_399_[T/C]	0.026	0.053	0.053	-0.027	1.000
ILL_C4791_1099_[T/G]	0.289	0.474	0.422	-0.152	1.000
ILL_C300_4982_[A/G]	0.026	0.053	0.053	-0.027	1.000
ILL_C4593_326_[T/G]	0.417	0.611	0.500	-0.257	0.620
PS_C46533_541_[T/C]	0.048	0.095	0.093	-0.050	1.000
PS_C15351_193_[A/G]	0.447	0.579	0.508	-0.171	0.655
PS_C15018_147_[T/G]	0.026	0.053	0.053	-0.027	1.000
PS_C38608_168_[T/C]	0.500	1.000	0.514	-1.000	0.000
ILL_C387_215_[A/G]	0.038	0.077	0.077	-0.040	1.000
ILL_C5339_366_[T/C]	0.053	0.000	0.102	1.000	0.028
PS_C34511_71_[T/G]	0.111	0.222	0.203	-0.125	1.000
PS_C14838_228_[T/C]	0.184	0.368	0.309	-0.226	1.000
ILL_C20267_102_[T/C]	0.237	0.474	0.371	-0.310	0.531
PS_C46674_204_[C/G]	0.105	0.211	0.193	-0.118	1.000
ILL_C1878_506_[T/C]	0.147	0.294	0.258	-0.172	1.000
PS_C36563_85_[T/G]	0.167	0.333	0.287	-0.200	1.000
ILL_C2735_326_[T/G]	0.079	0.158	0.149	-0.086	1.000
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.306	0.389	0.437	0.084	1.000
ILL_C1363_269_[T/C]	0.361	0.278	0.475	0.398	0.120
ILL_C18774_676_[T/C]	0.200	0.400	0.331	-0.250	1.000
PS_C23647_375_[A/G]	0.214	0.429	0.349	-0.273	1.000
ILL_C2028_1228_[A/T]	0.105	0.211	0.193	-0.118	1.000
PS_C11985_171_[A/G]	0.194	0.167	0.322	0.468	0.082
PS_C11911_576_[A/C]	0.105	0.211	0.193	-0.118	1.000
ILL_C1254_187_[A/G]	0.211	0.316	0.341	0.050	1.000
ILL_C20427_267_[T/C]	0.105	0.211	0.193	-0.118	1.000
PS_C21989_160_[T/G]	0.211	0.211	0.341	0.367	0.146
PS_C12069_1181_[T/C]	0.053	0.105	0.102	-0.056	1.000
ILL_C327_1076_[C/G]	0.026	0.053	0.053	-0.027	1.000
PS_C11847_438_[T/C]	0.206	0.412	0.337	-0.259	1.000

PS_C47375_253_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C34420_787_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C2915_875_[T/C]	0.139	0.167	0.246	0.303	0.272
ILL_C22491_727_[T/C]	0.395	0.579	0.491	-0.212	0.633
ILL_C911_1343_[T/A]	0.237	0.368	0.371	-0.019	1.000
PS_C12196_434_[G/C]	0.421	0.526	0.501	-0.080	1.000
PS_C14297_369_[T/G]	0.447	0.474	0.508	0.042	1.000
PS_C47340_198_[A/G]	0.421	0.000	0.501	1.000	0.000
PS_C36706_579_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C46597_301_[C/G]	N/A	N/A	N/A	N/A	N/A
ILL_C2040_1251_[A/T]	0.368	0.526	0.478	-0.131	1.000
PS_C35683_190_[A/G]	0.368	0.316	0.478	0.321	0.168
PS_C46532_529_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C15379_78_[A/C]	0.289	0.474	0.422	-0.152	1.000
PS_C11871_171_[A/G]	0.500	0.474	0.514	0.053	1.000
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.132	0.263	0.235	-0.152	1.000
ILL_C6061_1289_[T/G]	0.053	0.105	0.102	-0.056	1.000
ILL_C980_261_[A/G]	0.132	0.263	0.235	-0.152	1.000
Mean	**0.218**	**0.316**	**0.307**	**-0.035**	

WC(F_2)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.021	0.042	0.041	-0.021	1.000
PS_C15402_271_[T/A]	0.031	0.021	0.061	0.656	0.031
ILL_C253_1545_[A/C]	0.010	0.021	0.021	-0.011	1.000
ILL_C428_225_[A/G]	0.327	0.490	0.444	-0.114	0.524
PS_C23094_578_[G/C]	0.076	0.152	0.142	-0.082	1.000
ILL_C20682_843_[A/T]	0.042	0.083	0.081	-0.043	1.000
PS_C34490_403_[A/G]	0.102	0.204	0.185	-0.114	1.000
PS_C25208 _377_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C1783_492_[T/G]	0.071	0.143	0.134	-0.077	1.000
ILL_C929_734_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C11665_287_[A/G]	0.061	0.122	0.116	-0.065	1.000
ILL_C2141_350_[T/A]	0.286	0.408	0.412	0.000	1.000
PS_C23591_200_[T/C]	0.034	0.069	0.068	-0.036	1.000
PS_C15689_162_[A/T]	0.010	0.020	0.020	-0.010	1.000
ILL_C229_2772_[A/G]	0.010	0.020	0.020	-0.010	1.000
ILL_C22574_507_[A/C]	0.031	0.061	0.060	-0.032	1.000
PS_C15088_268_[A/G]	0.286	0.531	0.412	-0.300	0.065
ILL_C2406_641_[T/C]	0.133	0.224	0.232	0.024	1.000
PS_C12925_666_[A/G]	0.136	0.273	0.238	-0.158	1.000
PS_C11984_159_[T/G]	0.388	0.776	0.480	-0.633	0.000
ILL_C22449_261_[A/G]	0.378	0.510	0.475	-0.086	0.767
PS_C11970_157_[A/G]	0.143	0.286	0.247	-0.167	0.559
PS_C23051_368_[A/G]	0.243	0.486	0.373	-0.321	0.150

ILL_C6012_280_[T/C]	0.316	0.469	0.437	-0.085	0.735
ILL_C2903_1043_[T/A]	0.286	0.449	0.412	-0.100	0.727
PS_C34501_638_[A/G]	0.429	0.000	0.495	1.000	0.000
PS_C12352_527_[T/C]	0.479	0.458	0.504	0.082	0.553
PS_C34725_229_[T/C]	0.185	0.326	0.305	-0.082	1.000
PS_C15230_93_[A/G]	0.245	0.362	0.374	0.021	1.000
PS_C23075_525_[G/C]	0.347	0.531	0.458	-0.171	0.334
ILL_C3835_411_[A/G]	0.265	0.449	0.394	-0.152	0.453
PS_C24743_123_[T/G]	0.245	0.490	0.374	-0.324	0.045
ILL_C5634_234_[A/G]	0.194	0.347	0.316	-0.110	0.669
PS_C23630_237_[T/C]	0.010	0.020	0.020	-0.010	1.000
ILL_C1813_300_[T/G]	0.439	0.510	0.498	-0.036	1.000
PS_C12218_188_[A/G]	0.020	0.041	0.040	-0.021	1.000
PS_C36136_185_[T/C]	0.398	0.469	0.484	0.020	1.000
PS_C23070_1364_[A/G]	0.031	0.061	0.060	-0.032	1.000
ILL_C618_116_[A/G]	0.102	0.204	0.185	-0.114	1.000
PS_C1652_228_[A/C]	0.480	0.184	0.504	0.632	0.000
PS_C25083_285_[A/G]	0.032	0.064	0.062	-0.033	1.000
ILL_C394_1510_[A/G]	0.316	0.510	0.437	-0.180	0.332
PS_C24267_71_[A/G]	0.490	0.980	0.505	-0.960	0.000
PS_C34507_1191_[T/C]	0.041	0.082	0.079	-0.043	1.000
PS_C36273_73_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C35977_153_[T/C]	0.306	0.000	0.429	1.000	0.000
PS_C16093_142_[T/C]	0.276	0.388	0.403	0.029	1.000
PS_C16031_147_[A/G]	0.235	0.469	0.363	-0.307	0.050
ILL_C5106_273_[T/C]	0.235	0.347	0.363	0.034	0.714
PS_C00512_245_[T/C]	0.012	0.024	0.024	-0.012	1.000
ILL_C140_2421_[A/G]	0.480	0.592	0.504	-0.186	0.263
PS_C36237_70_[C/G]	0.053	0.106	0.102	-0.056	1.000
PS_C34670_302_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C21880_1003_[G/C]	0.173	0.224	0.290	0.217	0.146
PS_C47330_170_[A/C]	0.480	0.878	0.504	-0.758	0.000
PS_C36522_249_[G/C]	0.344	0.521	0.456	-0.154	0.363
PS_C34604_423_[A/T]	0.347	0.490	0.458	-0.081	0.752
PS_C34501_77_[T/G]	0.327	0.000	0.444	1.000	0.000
PS_C28810_290_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.429	0.694	0.495	-0.417	0.011
ILL_C844_440_[T/C]	0.041	0.082	0.079	-0.043	1.000
PS_C46857_366_[T/C]	0.052	0.104	0.100	-0.055	1.000
PS_C34605_382_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C853_1199_[A/G]	0.102	0.204	0.185	-0.114	1.000
ILL_C1002_85_[T/A]	0.255	0.469	0.384	-0.235	0.134
PS_C11659_399_[T/C]	0.204	0.408	0.328	-0.256	0.172
ILL_C4791_1099_[T/G]	0.480	0.469	0.504	0.060	0.774
ILL_C300_4982_[A/G]	N/A	N/A	N/A	N/A	N/A

ILL_C4593_326_[T/G]	0.479	0.667	0.504	-0.336	0.038
PS_C46533_541_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15351_193_[A/G]	0.479	0.404	0.504	0.190	0.235
PS_C15018_147_[T/G]	N/A	N/A	N/A	N/A	N/A
PS_C38608_168_[T/C]	0.500	1.000	0.505	-1.000	0.000
ILL_C387_215_[A/G]	0.031	0.020	0.060	0.656	0.025
ILL_C5339_366_[T/C]	0.020	0.041	0.040	-0.021	1.000
PS_C34511_71_[T/G]	0.085	0.170	0.157	-0.093	1.000
PS_C14838_228_[T/C]	0.102	0.204	0.185	-0.114	1.000
ILL_C20267_102_[T/C]	0.296	0.551	0.421	-0.322	0.040
PS_C46674_204_[C/G]	0.020	0.041	0.040	-0.021	1.000
ILL_C1878_506_[T/C]	0.010	0.020	0.020	-0.010	1.000
PS_C36563_85_[T/G]	0.050	0.100	0.096	-0.053	1.000
ILL_C2735_326_[T/G]	0.010	0.020	0.020	-0.010	1.000
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.149	0.298	0.256	-0.175	0.572
ILL_C1363_269_[T/C]	0.319	0.426	0.439	0.021	1.000
ILL_C18774_676_[T/C]	0.043	0.085	0.082	-0.044	1.000
PS_C23647_375_[A/G]	0.091	0.182	0.167	-0.100	1.000
ILL_C2028_1228_[A/T]	0.020	0.041	0.040	-0.021	1.000
PS_C11985_171_[A/G]	0.103	0.051	0.186	0.721	0.001
PS_C11911_576_[A/C]	0.020	0.041	0.040	-0.021	1.000
ILL_C1254_187_[A/G]	0.357	0.429	0.464	0.067	0.760
ILL_C20427_267_[T/C]	0.031	0.061	0.060	-0.032	1.000
PS_C21989_160_[T/G]	0.286	0.449	0.412	-0.100	0.721
PS_C12069_1181_[T/C]	0.092	0.184	0.169	-0.101	1.000
ILL_C327_1076_[C/G]	0.010	0.020	0.020	-0.010	1.000
PS_C11847_438_[T/C]	0.341	0.634	0.455	-0.410	0.015
PS_C47375_253_[A/G]	0.378	0.020	0.475	0.957	0.000
PS_C34420_787_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C2915_875_[T/C]	0.125	0.250	0.221	-0.143	1.000
ILL_C22491_727_[T/C]	0.316	0.469	0.437	-0.085	0.754
ILL_C911_1343_[T/A]	0.194	0.388	0.316	-0.241	0.177
PS_C12196_434_[G/C]	0.500	0.592	0.505	-0.184	0.271
PS_C14297_369_[T/G]	0.375	0.625	0.474	-0.333	0.045
PS_C47340_198_[A/G]	0.333	0.000	0.449	1.000	0.000
PS_C36706_579_[T/C]	0.122	0.245	0.217	-0.140	1.000
PS_C46597_301_[C/G]	0.235	0.429	0.363	-0.193	0.249
ILL_C2040_1251_[A/T]	0.041	0.082	0.079	-0.043	1.000
PS_C35683_190_[A/G]	0.276	0.388	0.403	0.029	1.000
PS_C46532_529_[A/T]	0.102	0.204	0.185	-0.114	1.000
PS_C15379_78_[A/C]	0.388	0.286	0.480	0.398	0.006
PS_C11871_171_[A/G]	0.316	0.510	0.437	-0.180	0.318
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	N/A	N/A	N/A	N/A	N/A
ILL_C6061_1289_[T/G]	0.418	0.633	0.492	-0.300	0.087

ILL_C980_261_[A/G]	0.444	0.578	0.499	-0.170	0.365
Mean	**0.211**	**0.297**	**0.284**	**-0.036**	
	WS(2004)				
Locus	**MAF**	**H_o**	**H_e**	**F_{is}**	***P*-value**
ILL_C2122_257_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C15402_271_[T/A]	0.023	0.045	0.045	-0.023	1.000
ILL_C253_1545_[A/C]	0.100	0.100	0.185	0.444	0.156
ILL_C428_225_[A/G]	0.172	0.344	0.289	-0.208	0.548
PS_C23094_578_[G/C]	0.094	0.188	0.173	-0.103	1.000
ILL_C20682_843_[A/T]	0.040	0.080	0.078	-0.042	1.000
PS_C34490_403_[A/G]	0.333	0.400	0.452	0.100	0.687
PS_C25208 _377_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C1783_492_[T/G]	0.094	0.188	0.173	-0.103	1.000
ILL_C929_734_[T/C]	0.050	0.100	0.097	-0.053	1.000
PS_C11665_287_[A/G]	0.081	0.161	0.151	-0.088	1.000
ILL_C2141_350_[T/A]	0.452	0.516	0.503	-0.042	1.000
PS_C23591_200_[T/C]	0.045	0.091	0.089	-0.048	1.000
PS_C15689_162_[A/T]	0.076	0.091	0.142	0.351	0.146
ILL_C229_2772_[A/G]	0.133	0.267	0.235	-0.154	1.000
ILL_C22574_507_[A/C]	0.242	0.364	0.373	0.010	1.000
PS_C15088_268_[A/G]	0.333	0.467	0.452	-0.050	1.000
ILL_C2406_641_[T/C]	0.065	0.129	0.123	-0.069	1.000
PS_C12925_666_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C11984_159_[T/G]	0.500	1.000	0.510	-1.000	0.000
ILL_C22449_261_[A/G]	0.188	0.375	0.310	-0.231	0.567
PS_C11970_157_[A/G]	0.250	0.438	0.381	-0.167	0.649
PS_C23051_368_[A/G]	0.250	0.500	0.387	-0.333	0.513
ILL_C6012_280_[T/C]	0.355	0.452	0.465	0.014	1.000
ILL_C2903_1043_[T/A]	0.250	0.375	0.387	0.000	1.000
PS_C34501_638_[A/G]	0.063	0.000	0.119	1.000	0.001
PS_C12352_527_[T/C]	0.333	0.467	0.452	-0.050	1.000
PS_C34725_229_[T/C]	0.274	0.226	0.405	0.433	0.021
PS_C15230_93_[A/G]	0.422	0.469	0.496	0.039	1.000
PS_C23075_525_[G/C]	0.439	0.515	0.500	-0.046	1.000
ILL_C3835_411_[A/G]	0.281	0.375	0.411	0.072	0.668
PS_C24743_123_[T/G]	0.197	0.333	0.321	-0.054	1.000
ILL_C5634_234_[A/G]	0.258	0.455	0.388	-0.188	0.642
PS_C23630_237_[T/C]	0.173	0.115	0.292	0.597	0.013
ILL_C1813_300_[T/G]	0.106	0.212	0.193	-0.119	1.000
PS_C12218_188_[A/G]	0.290	0.387	0.419	0.061	0.687
PS_C36136_185_[T/C]	0.313	0.250	0.437	0.418	0.034
PS_C23070_1364_[A/G]	0.197	0.212	0.321	0.329	0.079
ILL_C618_116_[A/G]	0.152	0.242	0.261	0.057	0.560
PS_C1652_228_[A/C]	0.167	0.152	0.282	0.455	0.024
PS_C25083_285_[A/G]	0.016	0.031	0.031	-0.016	1.000
ILL_C394_1510_[A/G]	0.469	0.500	0.506	-0.004	1.000

PS_C24267_71_[A/G]	0.500	1.000	0.508	-1.000	0.000
PS_C34507_1191_[T/C]	0.431	0.448	0.499	0.086	0.699
PS_C36273_73_[A/C]	0.094	0.125	0.173	0.264	0.235
PS_C35977_153_[T/C]	0.091	0.000	0.168	1.000	0.000
PS_C16093_142_[T/C]	0.422	0.344	0.496	0.295	0.152
PS_C16031_147_[A/G]	0.375	0.625	0.476	-0.333	0.138
ILL_C5106_273_[T/C]	0.222	0.296	0.352	0.143	0.571
PS_C00512_245_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C140_2421_[A/G]	0.125	0.188	0.222	0.143	0.388
PS_C36237_70_[C/G]	0.226	0.194	0.355	0.446	0.022
PS_C34670_302_[A/C]	0.015	0.030	0.030	-0.015	1.000
ILL_C21880_1003_[G/C]	0.281	0.438	0.411	-0.082	1.000
PS_C47330_170_[A/C]	0.391	0.609	0.487	-0.278	0.369
PS_C36522_249_[G/C]	0.333	0.424	0.451	0.045	1.000
PS_C34604_423_[A/T]	0.203	0.219	0.329	0.324	0.082
PS_C34501_77_[T/G]	0.470	0.636	0.506	-0.277	0.180
PS_C28810_290_[A/C]	0.364	0.121	0.470	0.738	0.000
PS_C28886_317_[T/C]	0.030	0.061	0.060	-0.031	1.000
ILL_C31_1387_[T/C]	0.371	0.419	0.474	0.101	0.701
ILL_C844_440_[T/C]	0.032	0.065	0.063	-0.033	1.000
PS_C46857_366_[T/C]	0.111	0.222	0.201	-0.125	1.000
PS_C34605_382_[A/G]	0.063	0.125	0.121	-0.067	1.000
ILL_C853_1199_[A/G]	0.031	0.063	0.062	-0.032	1.000
ILL_C1002_85_[T/A]	0.424	0.727	0.496	-0.489	0.011
PS_C11659_399_[T/C]	0.032	0.065	0.063	-0.033	1.000
ILL_C4791_1099_[T/G]	0.273	0.364	0.403	0.083	0.666
ILL_C300_4982_[A/G]	0.107	0.214	0.195	-0.120	1.000
ILL_C4593_326_[T/G]	0.167	0.273	0.282	0.018	1.000
PS_C46533_541_[T/C]	0.015	0.030	0.030	-0.015	1.000
PS_C15351_193_[A/G]	0.422	0.406	0.496	0.167	0.454
PS_C15018_147_[T/G]	0.063	0.125	0.119	-0.067	1.000
PS_C38608_168_[T/C]	0.483	0.967	0.508	-0.935	0.000
ILL_C387_215_[A/G]	0.117	0.033	0.210	0.838	0.002
ILL_C5339_366_[T/C]	0.048	0.097	0.094	-0.051	1.000
PS_C34511_71_[T/G]	0.016	0.031	0.031	-0.016	1.000
PS_C14838_228_[T/C]	0.136	0.273	0.239	-0.158	1.000
ILL_C20267_102_[T/C]	0.141	0.219	0.246	0.095	0.469
PS_C46674_204_[C/G]	0.304	0.464	0.431	-0.098	1.000
ILL_C1878_506_[T/C]	0.015	0.030	0.030	-0.015	1.000
PS_C36563_85_[T/G]	0.224	0.310	0.354	0.108	0.588
ILL_C2735_326_[T/G]	0.017	0.034	0.034	-0.018	1.000
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.367	0.400	0.472	0.139	0.453
ILL_C1363_269_[T/C]	0.359	0.406	0.468	0.118	0.475
ILL_C18774_676_[T/C]	0.053	0.105	0.102	-0.056	1.000
PS_C23647_375_[A/G]	0.129	0.194	0.228	0.139	0.399

ILL_C2028_1228_[A/T]	0.167	0.333	0.282	-0.200	0.563
PS_C11985_171_[A/G]	0.233	0.267	0.364	0.255	0.166
PS_C11911_576_[A/C]	0.156	0.250	0.268	0.052	0.566
ILL_C1254_187_[A/G]	0.442	0.577	0.503	-0.169	0.689
ILL_C20427_267_[T/C]	0.018	0.036	0.036	-0.018	1.000
PS_C21989_160_[T/G]	0.182	0.303	0.302	-0.019	1.000
PS_C12069_1181_[T/C]	0.030	0.061	0.060	-0.031	1.000
ILL_C327_1076_[C/G]	0.047	0.031	0.091	0.650	0.052
PS_C11847_438_[T/C]	0.233	0.400	0.364	-0.118	1.000
PS_C47375_253_[A/G]	0.015	0.030	0.030	-0.015	1.000
PS_C34420_787_[A/G]	0.016	0.032	0.032	-0.016	1.000
ILL_C2915_875_[T/C]	0.232	0.321	0.363	0.098	0.601
ILL_C22491_727_[T/C]	0.435	0.484	0.500	0.016	1.000
ILL_C911_1343_[T/A]	0.234	0.281	0.365	0.216	0.295
PS_C12196_434_[G/C]	0.375	0.650	0.481	-0.387	0.159
PS_C14297_369_[T/G]	0.422	0.469	0.496	0.039	1.000
PS_C47340_198_[A/G]	0.182	0.000	0.302	1.000	0.000
PS_C36706_579_[T/C]	0.167	0.273	0.282	0.018	1.000
PS_C46597_301_[C/G]	0.097	0.129	0.178	0.262	0.233
ILL_C2040_1251_[A/T]	0.125	0.250	0.223	-0.143	1.000
PS_C35683_190_[A/G]	0.297	0.406	0.424	0.027	1.000
PS_C46532_529_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C15379_78_[A/C]	0.389	0.704	0.484	-0.481	0.032
PS_C11871_171_[A/G]	0.485	0.545	0.507	-0.092	0.741
PS_C35823_1199_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C39731_379_[A/C]	0.197	0.333	0.321	-0.054	1.000
ILL_C6061_1289_[T/G]	0.121	0.241	0.216	-0.137	1.000
ILL_C980_261_[A/G]	0.097	0.194	0.178	-0.107	1.000
Mean	**0.209**	**0.290**	**0.293**	**0.026**	

WS(2011)

Locus	MAF	H_o	H_e	F_{is}	*P*-value
ILL_C2122_257_[A/G]	0.056	0.111	0.111	-0.059	1.000
PS_C15402_271_[T/A]	0.125	0.250	0.233	-0.143	1.000
ILL_C253_1545_[A/C]	0.056	0.111	0.111	-0.059	1.000
ILL_C428_225_[A/G]	0.222	0.444	0.366	-0.286	1.000
PS_C23094_578_[G/C]	0.056	0.111	0.111	-0.059	1.000
ILL_C20682_843_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C34490_403_[A/G]	0.222	0.444	0.366	-0.286	1.000
PS_C25208 _377_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C1783_492_[T/G]	0.125	0.250	0.233	-0.143	1.000
ILL_C929_734_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C11665_287_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C2141_350_[T/A]	0.375	0.500	0.500	-0.067	1.000
PS_C23591_200_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C15689_162_[A/T]	0.125	0.250	0.233	-0.143	1.000
ILL_C229_2772_[A/G]	N/A	N/A	N/A	N/A	N/A

ILL_C22574_507_[A/C]	0.167	0.333	0.294	-0.200	1.000
PS_C15088_268_[A/G]	0.500	0.556	0.529	-0.111	1.000
ILL_C2406_641_[T/C]	0.063	0.125	0.125	-0.067	1.000
PS_C12925_666_[A/G]	0.100	0.200	0.200	-0.111	1.000
PS_C11984_159_[T/G]	0.500	1.000	0.529	-1.000	0.010
ILL_C22449_261_[A/G]	0.389	0.333	0.503	0.299	0.491
PS_C11970_157_[A/G]	0.250	0.250	0.400	0.333	0.391
PS_C23051_368_[A/G]	0.071	0.143	0.143	-0.077	1.000
ILL_C6012_280_[T/C]	0.188	0.125	0.325	0.590	0.205
ILL_C2903_1043_[T/A]	0.500	0.333	0.529	0.333	0.510
PS_C34501_638_[A/G]	0.222	0.000	0.366	1.000	0.013
PS_C12352_527_[T/C]	0.389	0.778	0.503	-0.636	0.171
PS_C34725_229_[T/C]	0.167	0.111	0.294	0.600	0.171
PS_C15230_93_[A/G]	0.500	0.429	0.538	0.143	1.000
PS_C23075_525_[G/C]	0.222	0.000	0.366	1.000	0.011
ILL_C3835_411_[A/G]	0.111	0.222	0.209	-0.125	1.000
PS_C24743_123_[T/G]	0.222	0.444	0.366	-0.286	1.000
ILL_C5634_234_[A/G]	0.500	0.778	0.529	-0.556	0.215
PS_C23630_237_[T/C]	0.167	0.333	0.294	-0.200	1.000
ILL_C1813_300_[T/G]	0.111	0.222	0.209	-0.125	1.000
PS_C12218_188_[A/G]	0.125	0.250	0.233	-0.143	1.000
PS_C36136_185_[T/C]	0.278	0.556	0.425	-0.385	1.000
PS_C23070_1364_[A/G]	0.333	0.000	0.471	1.000	0.005
ILL_C618_116_[A/G]	0.278	0.556	0.425	-0.385	1.000
PS_C1652_228_[A/C]	0.222	0.000	0.366	1.000	0.007
PS_C25083_285_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C394_1510_[A/G]	0.222	0.444	0.366	-0.286	1.000
PS_C24267_71_[A/G]	0.389	0.778	0.503	-0.636	0.178
PS_C34507_1191_[T/C]	0.444	0.222	0.523	0.550	0.167
PS_C36273_73_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C35977_153_[T/C]	0.222	0.000	0.366	1.000	0.012
PS_C16093_142_[T/C]	0.444	0.222	0.523	0.550	0.171
PS_C16031_147_[A/G]	0.222	0.444	0.366	-0.286	1.000
ILL_C5106_273_[T/C]	0.125	0.250	0.233	-0.143	1.000
PS_C00512_245_[T/C]	0.071	0.143	0.143	-0.077	1.000
ILL_C140_2421_[A/G]	0.056	0.111	0.111	-0.059	1.000
PS_C36237_70_[C/G]	0.333	0.444	0.471	0.000	1.000
PS_C34670_302_[A/C]	0.056	0.111	0.111	-0.059	1.000
ILL_C21880_1003_[G/C]	0.125	0.250	0.233	-0.143	1.000
PS_C47330_170_[A/C]	0.429	0.571	0.527	-0.167	1.000
PS_C36522_249_[G/C]	0.500	0.500	0.533	0.000	1.000
PS_C34604_423_[A/T]	0.222	0.222	0.366	0.357	0.335
PS_C34501_77_[T/G]	0.222	0.000	0.366	1.000	0.015
PS_C28810_290_[A/C]	N/A	N/A	N/A	N/A	N/A
PS_C28886_317_[T/C]	N/A	N/A	N/A	N/A	N/A
ILL_C31_1387_[T/C]	0.222	0.444	0.366	-0.286	1.000

ILL_C844_440_[T/C]	0.056	0.111	0.111	-0.059	1.000
PS_C46857_366_[T/C]	0.222	0.444	0.366	-0.286	1.000
PS_C34605_382_[A/G]	N/A	N/A	N/A	N/A	N/A
ILL_C853_1199_[A/G]	0.056	0.111	0.111	-0.059	1.000
ILL_C1002_85_[T/A]	0.333	0.667	0.471	-0.500	0.457
PS_C11659_399_[T/C]	0.167	0.333	0.294	-0.200	1.000
ILL_C4791_1099_[T/G]	0.333	0.444	0.471	0.000	1.000
ILL_C300_4982_[A/G]	0.063	0.125	0.125	-0.067	1.000
ILL_C4593_326_[T/G]	0.389	0.778	0.503	-0.636	0.174
PS_C46533_541_[T/C]	0.111	0.222	0.209	-0.125	1.000
PS_C15351_193_[A/G]	0.389	0.333	0.503	0.299	0.495
PS_C15018_147_[T/G]	N/A	N/A	N/A	N/A	N/A
PS_C38608_168_[T/C]	0.500	1.000	0.533	-1.000	0.023
ILL_C387_215_[A/G]	0.250	0.000	0.400	1.000	0.016
ILL_C5339_366_[T/C]	0.167	0.333	0.294	-0.200	1.000
PS_C34511_71_[T/G]	0.071	0.143	0.143	-0.077	1.000
PS_C14838_228_[T/C]	0.188	0.375	0.325	-0.231	1.000
ILL_C20267_102_[T/C]	0.222	0.444	0.366	-0.286	1.000
PS_C46674_204_[C/G]	0.389	0.333	0.503	0.299	0.502
ILL_C1878_506_[T/C]	0.056	0.111	0.111	-0.059	1.000
PS_C36563_85_[T/G]	0.417	0.833	0.530	-0.714	0.402
ILL_C2735_326_[T/G]	N/A	N/A	N/A	N/A	N/A
ILL_C250_199_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C47845_238_[T/C]	0.438	0.375	0.525	0.238	0.520
ILL_C1363_269_[T/C]	0.444	0.444	0.523	0.100	1.000
ILL_C18774_676_[T/C]	0.188	0.375	0.325	-0.231	1.000
PS_C23647_375_[A/G]	0.167	0.333	0.303	-0.200	1.000
ILL_C2028_1228_[A/T]	0.063	0.125	0.125	-0.067	1.000
PS_C11985_171_[A/G]	0.417	0.167	0.530	0.657	0.157
PS_C11911_576_[A/C]	0.167	0.333	0.294	-0.200	1.000
ILL_C1254_187_[A/G]	0.250	0.500	0.400	-0.333	1.000
ILL_C20427_267_[T/C]	N/A	N/A	N/A	N/A	N/A
PS_C21989_160_[T/G]	0.222	0.444	0.366	-0.286	1.000
PS_C12069_1181_[T/C]	0.056	0.111	0.111	-0.059	1.000
ILL_C327_1076_[C/G]	N/A	N/A	N/A	N/A	N/A
PS_C11847_438_[T/C]	0.278	0.556	0.425	-0.385	1.000
PS_C47375_253_[A/G]	N/A	N/A	N/A	N/A	N/A
PS_C34420_787_[A/G]	0.056	0.111	0.111	-0.059	1.000
ILL_C2915_875_[T/C]	0.250	0.500	0.409	-0.333	1.000
ILL_C22491_727_[T/C]	0.125	0.250	0.233	-0.143	1.000
ILL_C911_1343_[T/A]	0.278	0.556	0.425	-0.385	1.000
PS_C12196_434_[G/C]	0.444	0.444	0.523	0.100	1.000
PS_C14297_369_[T/G]	0.438	0.125	0.525	0.746	0.056
PS_C47340_198_[A/G]	0.222	0.000	0.366	1.000	0.010
PS_C36706_579_[T/C]	0.167	0.333	0.294	-0.200	1.000
PS_C46597_301_[C/G]	0.188	0.375	0.325	-0.231	1.000

ILL_C2040_1251_[A/T]	0.188	0.375	0.325	-0.231	1.000
PS_C35683_190_[A/G]	0.389	0.333	0.503	0.299	0.501
PS_C46532_529_[A/T]	N/A	N/A	N/A	N/A	N/A
PS_C15379_78_[A/C]	0.444	0.444	0.523	0.100	1.000
PS_C11871_171_[A/G]	0.278	0.556	0.425	-0.385	1.000
PS_C35823_1199_[A/G]	0.111	0.000	0.209	1.000	0.061
PS_C39731_379_[A/C]	0.167	0.333	0.294	-0.200	1.000
ILL_C6061_1289_[T/G]	0.111	0.222	0.209	-0.125	1.000
ILL_C980_261_[A/G]	0.250	0.500	0.400	-0.333	1.000
Mean	**0.239**	**0.327**	**0.346**	**-0.011**	

Table S4.3: Summary of outlier results, numerical values highlighted denotes loci putatively under selection as indicated by the respective tests. Where the locus name has been highlighted it shows a locus that has been identified as a locus under selection by both test methods across any of the population cohorts (Bold: locus under directional selection; Underlined: locus under balancing selection).

	Across all populations		Across wild populations (2000-4)		Across wild populations (2010/11)		Across F_1 cultured populations	
	Lositan	BayeScan	Lositan	BayeScan	Lositan	BayeScan	Lositan	BayeScan
Locus	*P*-value (Simulated F_{st} < sample F_{st})	log_{10}(PO)	*P*-value (Simulated F_{st} < sample F_{st})	log_{10}(PO)	*P*-value (Simulated F_{st} < sample F_{st})	log_{10}(PO)	*P*-value (Simulated F_{st} < sample F_{st})	log_{10}(PO)
ILL_C2122_257_[A/G]	ns	ns	ns	ns	ns	ns	0.046	ns
PS_C15402_271_[T/A]	0.001	ns	ns	ns	0.001	ns	ns	ns
ILL_C253_1545_[A/C]	0.000	ns	0.002	ns	0.000	ns	ns	ns
ILL_C428_225_[A/G]	0.004	ns	ns	ns	ns	ns	ns	ns
PS_C23094_578_[G/C]	0.000	ns	0.000	ns	0.006	ns	ns	ns
ILL_C20682_843_[A/T]	0.004	ns	0.000	ns	ns	ns	0.031	ns
PS_C34490_403_[A/G]	ns	ns	0.004	ns	ns	ns	ns	ns
PS_C25208 _377_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C1783_492_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C929_734_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C11665_287_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C2141_350_[T/A]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C23591_200_[T/C]	**0.967**	**2.104**	ns	ns	0.000	ns	ns	ns
PS_C15689_162_[A/T]	0.041	ns	ns	ns	0.000	ns	ns	ns
ILL_C229_2772_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C22574_507_[A/C]	ns	ns	ns	ns	0.010	ns	ns	ns
PS_C15088_268_[A/G]	0.008	ns	ns	ns	ns	ns	ns	ns

ILL_C2406_641_[T/C]	0.017	ns	ns	ns	0.002	ns	ns	ns
PS_C12925_666_[A/G]	0.000	ns	ns	ns	0.006	ns	0.010	ns
PS_C11984_159_[T/G]	0.000	ns	0.006	ns	0.007	ns	ns	ns
ILL_C22449_261_[A/G]	0.005	ns	0.047	ns	ns	ns	ns	ns
PS_C11970_157_[A/G]	0.000	ns	ns	ns	0.008	ns	0.021	ns
PS_C23051_368_[A/G]	ns	ns	ns	ns	0.000	ns	ns	ns
ILL_C6012_280_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C2903_1043_[T/A]	ns	ns	0.008	ns	ns	ns	ns	ns
PS_C34501_638_[A/G]	**1.000**	**1000.000**	**1.000**	**1000.000**	**1.000**	**1000.000**	**1.000**	**1000.000**
PS_C12352_527_[T/C]	0.002	-0.710	0.005	ns	ns	ns	ns	ns
PS_C34725_229_[T/C]	0.007	-0.959	0.005	ns	ns	ns	ns	ns
PS_C15230_93_[A/G]	0.008	-1.017	ns	ns	0.049	ns	ns	ns
PS_C23075_525_[G/C]	**1.000**	**1000.000**	**0.985**	**1000.000**	**0.995**	ns	ns	ns
ILL_C3835_411_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C24743_123_[T/G]	0.011	ns	ns	ns	0.038	ns	ns	ns
ILL_C5634_234_[A/G]	ns	ns	0.032	ns	ns	ns	ns	ns
PS_C23630_237_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C1813_300_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C12218_188_[A/G]	0.030	ns	ns	ns	ns	ns	ns	ns
PS_C36136_185_[T/C]	ns	ns	ns	ns	ns	ns	0.043	ns
PS_C23070_1364_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C618_116_[A/G]	ns	ns	ns	ns	ns	ns	0.047	ns
PS_C1652_228_[A/C]	**1.000**	**1000.000**	**1.000**	**1000.000**	**1.000**	**1000.000**	**0.995**	**1.273**
PS_C25083_285_[A/G]	**0.997**	**1000.000**	ns	ns	**0.994**	ns	ns	ns
ILL_C394_1510_[A/G]	ns	ns	0.044	ns	**0.966**	ns	0.048	ns
PS_C24267_71_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C34507_1191_[T/C]	ns	ns	ns	ns	ns	ns	0.001	ns
PS_C36273_73_[A/C]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C35977_153_[T/C]	**1.000**	**1000.000**	**1.000**	**1000.000**	**1.000**	**1000.000**	**1.000**	**1000.000**
PS_C16093_142_[T/C]	ns	ns	ns	ns	0.029	ns	ns	ns

PS_C16031_147_[A/G]	0.000	ns	ns	ns	ns	ns	0.002	ns
ILL_C5106_273_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C00512_245_[T/C]	0.000	ns	0.042	ns	0.000	ns	0.038	ns
ILL_C140_2421_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C36237_70_[C/G]	ns	**1.560**	ns	**1.433**	ns	ns	ns	ns
PS_C34670_302_[A/C]	0.000	ns	0.042	ns	0.000	ns	ns	ns
ILL_C21880_1003_[G/C]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C47330_170_[A/C]	ns	ns	0.000	ns	ns	ns	ns	ns
PS_C36522_249_[G/C]	0.003	ns	0.033	ns	ns	ns	ns	ns
PS_C34604_423_[A/T]	0.002	ns	0.003	ns	ns	ns	0.042	ns
PS_C34501_77_[T/G]	ns	**3.097**	ns	ns	ns	ns	ns	ns
PS_C28810_290_[A/C]	**0.992**	**1000.000**	**0.977**	**3.097**	ns	ns	ns	ns
PS_C28886_317_[T/C]	ns	ns	ns	ns	**1.000**	ns	ns	ns
ILL_C31_1387_[T/C]	0.047	ns	ns	ns	ns	ns	ns	ns
ILL_C844_440_[T/C]	0.000	ns	0.000	ns	0.000	ns	ns	ns
PS_C46857_366_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C34605_382_[A/G]	0.008	ns	ns	ns	ns	ns	ns	ns
ILL_C853_1199_[A/G]	0.011	ns	ns	ns	ns	ns	ns	ns
ILL_C1002_85_[T/A]	0.015	ns	ns	ns	ns	ns	ns	ns
PS_C11659_399_[T/C]	ns	ns	ns	ns	ns	ns	0.050	ns
ILL_C4791_1099_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C300_4982_[A/G]	0.005	ns	ns	ns	0.000	ns	ns	ns
ILL_C4593_326_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C46533_541_[T/C]	ns	ns	0.026	ns	ns	ns	ns	ns
PS_C15351_193_[A/G]	0.004	ns	ns	ns	0.042	ns	ns	ns
PS_C15018_147_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C38608_168_[T/C]	0.000	1.042	0.001	ns	0.000	ns	0.032	ns
ILL_C387_215_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C5339_366_[T/C]	ns	ns	0.003	ns	ns	ns	0.042	ns
PS_C34511_71_[T/G]	0.000	ns	0.003	ns	ns	ns	ns	ns

PS_C14838_228_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C20267_102_[T/C]	0.010	ns	ns	ns	0.010	ns	0.002	ns
PS_C46674_204_[C/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C1878_506_[T/C]	ns	ns	0.005	ns	ns	ns	ns	ns
PS_C36563_85_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C2735_326_[T/G]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C250_199_[T/C]	ns	ns	ns	ns	**0.955**	ns	ns	ns
PS_C47845_238_[T/C]	0.043	ns	ns	ns	ns	ns	ns	ns
ILL_C1363_269_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C18774_676_[T/C]	0.044	ns	ns	ns	0.035	ns	ns	ns
PS_C23647_375_[A/G]	0.047	ns	ns	ns	0.000	ns	ns	ns
ILL_C2028_1228_[A/T]	0.007	ns	ns	ns	ns	ns	0.039	ns
PS_C11985_171_[A/G]	ns	ns	0.030	ns	ns	ns	0.006	ns
PS_C11911_576_[A/C]	0.022	ns	ns	ns	ns	ns	ns	ns
ILL_C1254_187_[A/G]	0.006	ns	ns	ns	ns	ns	ns	ns
ILL_C20427_267_[T/C]	ns	ns	0.000	ns	ns	ns	ns	ns
PS_C21989_160_[T/G]	0.019	ns	ns	ns	0.047	ns	ns	ns
PS_C12069_1181_[T/C]	0.002	ns	0.025	ns	ns	ns	0.042	ns
ILL_C327_1076_[C/G]	0.016	ns	ns	ns	ns	ns	ns	ns
PS_C11847_438_[T/C]	0.016	ns	ns	ns	0.000	ns	0.038	ns
PS_C47375_253_[A/G]	**1.000**	**1000.000**	**1.000**	**1000.000**	**1.000**	**0.555**	**0.999**	**1.660**
PS_C34420_787_[A/G]	0.000	ns	0.000	ns	0.000	ns	ns	ns
ILL_C2915_875_[T/C]	0.001	ns	0.004	ns	0.001	ns	ns	ns
ILL_C22491_727_[T/C]	ns	ns	ns	ns	ns	ns	ns	ns
ILL_C911_1343_[T/A]	0.000	ns	0.047	ns	0.000	ns	0.006	ns
PS_C12196_434_[G/C]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C14297_369_[T/G]	ns	ns	0.025	ns	ns	ns	ns	ns
PS_C47340_198_[A/G]	ns	**3.398**	**0.950**	**1.598**	ns	ns	ns	ns
PS_C36706_579_[T/C]	0.004	ns	ns	ns	0.001	ns	ns	ns
PS_C46597_301_[C/G]	ns	ns	0.009	ns	0.005	ns	ns	ns

ILL_C2040_1251_[A/T]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C35683_190_[A/G]	0.035	ns	0.026	ns	ns	ns	ns	ns
PS_C46532_529_[A/T]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C15379_78_[A/C]	ns	ns	ns	ns	0.029	ns	ns	ns
PS_C11871_171_[A/G]	ns	ns	0.005	ns	ns	ns	ns	ns
PS_C35823_1199_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
PS_C39731_379_[A/C]	ns	ns	ns	ns	0.036	ns	ns	ns
ILL_C6061_1289_[T/G]	**0.996**	**1000.000**	**0.977**	**3.699**	**0.970**	ns	**0.963**	ns
ILL_C980_261_[A/G]	ns	ns	ns	ns	ns	ns	ns	ns
Directional selection	9	12	8	9	10	4	5	4
Balancing selection	48	1	33	0	33	0	21	0
Total	57	13	41	9	43	4	26	4
	SD(2004) vs SD(2010)	**WS(2004) vs WS(2011)**	**CR(2003) vs CR(2011)**	**RP(2003) vs RP(2011)**	**WC(F_1) vs WC(F_2)**	**AS(F_1) vs SD(2004)**	**AS(F_1) vs SD(2010)**	**WC(F_1) vs RP(2003)**
	Lositan	**Lositan**	**Lositan**	**Lositan**	**Lositan**	**Lositan**	**Lositan**	**Lositan**
Locus	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**
ILL_C2122_257_[A/G]	0.719	0.000	ns	ns	ns	ns	ns	**1.000**
PS_C15402_271_[T/A]	0.407	ns	**0.955**	ns	ns	ns	0.000	ns
ILL_C253_1545_[A/C]	0.602	ns	ns	ns	ns	ns	ns	ns
ILL_C428_225_[A/G]	0.353	ns	ns	ns	ns	ns	ns	ns
PS_C23094_578_[G/C]	0.382	ns	ns	ns	ns	ns	ns	ns
ILL_C20682_843_[A/T]	0.091	0.000	ns	**1.000**	ns	ns	0.000	ns
PS_C34490_403_[A/G]	0.581	ns	ns	ns	ns	ns	ns	ns
PS_C25208 _377_[T/G]	0.510	ns	ns	ns	ns	ns	ns	ns
ILL_C1783_492_[T/G]	0.835	0.000	ns	ns	ns	ns	ns	ns
ILL_C929_734_[T/C]	0.139	0.000	ns	ns	ns	ns	ns	ns
PS_C11665_287_[A/G]	0.663	**1.000**	ns	ns	ns	**0.979**	ns	ns
ILL_C2141_350_[T/A]	0.902	ns	ns	ns	ns	ns	ns	ns

PS_C23591_200_[T/C]	0.542	0.000	**0.990**	**1.000**	ns	**1.000**	ns	**1.000**
PS_C15689_162_[A/T]	0.465	ns	ns	ns	ns	ns	ns	**0.996**
ILL_C229_2772_[A/G]	0.689	**1.000**	ns	ns	ns	ns	ns	ns
ILL_C22574_507_[A/C]	0.710	ns	ns	ns	ns	ns	ns	ns
PS_C15088_268_[A/G]	0.345	ns	ns	ns	ns	ns	ns	ns
ILL_C2406_641_[T/C]	0.693	0.000	ns	ns	ns	ns	ns	ns
PS_C12925_666_[A/G]	0.573	ns	0.030	ns	ns	ns	ns	ns
PS_C11984_159_[T/G]	0.647	ns	ns	ns	ns	ns	ns	ns
ILL_C22449_261_[A/G]	0.440	ns	ns	ns	ns	ns	ns	ns
PS_C11970_157_[A/G]	0.856	0.000	ns	ns	ns	ns	ns	ns
PS_C23051_368_[A/G]	**0.951**	ns	ns	ns	ns	0.000	ns	0.030
ILL_C6012_280_[T/C]	0.318	ns	ns	ns	ns	ns	ns	ns
ILL_C2903_1043_[T/A]	0.622	ns	ns	ns	ns	ns	ns	**0.987**
PS_C34501_638_[A/G]	0.946	ns	0.000	ns	**0.955**	ns	ns	ns
PS_C12352_527_[T/C]	0.389	ns	ns	ns	ns	ns	0.037	ns
PS_C34725_229_[T/C]	0.543	ns	ns	ns	ns	ns	ns	ns
PS_C15230_93_[A/G]	0.351	ns	ns	ns	ns	ns	ns	ns
PS_C23075_525_[G/C]	0.738	ns	ns	ns	ns	ns	ns	ns
ILL_C3835_411_[A/G]	0.379	ns	ns	ns	ns	ns	ns	ns
PS_C24743_123_[T/G]	0.609	ns	ns	ns	ns	ns	ns	ns
ILL_C5634_234_[A/G]	0.343	ns	ns	ns	ns	ns	ns	ns
PS_C23630_237_[T/C]	0.635	0.016	ns	ns	ns	ns	ns	ns
ILL_C1813_300_[T/G]	0.814	0.000	ns	ns	ns	ns	ns	ns
PS_C12218_188_[A/G]	0.745	ns	ns	ns	ns	ns	ns	ns
PS_C36136_185_[T/C]	0.491	ns	ns	ns	ns	ns	ns	ns
PS_C23070_1364_[A/G]	0.688	ns	ns	ns	**0.971**	ns	ns	ns
ILL_C618_116_[A/G]	0.613	ns	ns	ns	ns	ns	ns	ns
PS_C1652_228_[A/C]	0.454	ns	0.000	ns	**0.952**	ns	ns	ns
PS_C25083_285_[A/G]	0.874	ns	**1.000**	**0.996**	ns	ns	ns	**0.994**
ILL_C394_1510_[A/G]	**0.974**	ns	ns	ns	ns	ns	ns	ns

PS_C24267_71_[A/G]	0.365	ns	**0.997**	ns	ns	ns	ns	ns
PS_C34507_1191_[T/C]	0.423	ns	ns	ns	ns	ns	ns	ns
PS_C36273_73_[A/C]	0.261	**1.000**	ns	ns	ns	ns	**1.000**	**1.000**
PS_C35977_153_[T/C]	0.941	ns	0.000	**1.000**	ns	ns	ns	ns
PS_C16093_142_[T/C]	0.613	ns	ns	ns	ns	ns	**0.958**	ns
PS_C16031_147_[A/G]	0.547	ns	ns	ns	ns	ns	ns	ns
ILL_C5106_273_[T/C]	0.707	ns	ns	ns	ns	ns	ns	ns
PS_C00512_245_[T/C]	0.500	0.000	0.000	ns	ns	ns	ns	ns
ILL_C140_2421_[A/G]	0.296	ns	ns	ns	**0.969**	ns	ns	**0.961**
PS_C36237_70_[C/G]	**0.978**	ns	ns	ns	ns	ns	ns	ns
PS_C34670_302_[A/C]	0.513	0.000	0.000	ns	ns	ns	ns	ns
ILL_C21880_1003_[G/C]	0.745	ns	ns	ns	ns	ns	ns	ns
PS_C47330_170_[A/C]	0.270	ns	ns	ns	ns	ns	ns	ns
PS_C36522_249_[G/C]	0.444	ns	ns	ns	0.000	ns	ns	ns
PS_C34604_423_[A/T]	0.779	ns	ns	ns	ns	ns	ns	ns
PS_C34501_77_[T/G]	**0.990**	ns	ns	ns	ns	ns	ns	ns
PS_C28810_290_[A/C]	**0.999**	**0.998**	ns	ns	ns	**1.000**	ns	0.000
PS_C28886_317_[T/C]	0.500	ns	ns	ns	ns	ns	ns	ns
ILL_C31_1387_[T/C]	0.770	ns	ns	ns	ns	ns	ns	ns
ILL_C844_440_[T/C]	0.570	0.000	ns	ns	ns	ns	ns	ns
PS_C46857_366_[T/C]	0.529	ns	ns	ns	ns	ns	ns	ns
PS_C34605_382_[A/G]	0.500	0.000	ns	ns	ns	ns	ns	ns
ILL_C853_1199_[A/G]	0.339	0.000	ns	ns	ns	ns	ns	ns
ILL_C1002_85_[T/A]	0.457	ns	ns	**0.972**	ns	ns	ns	**0.987**
PS_C11659_399_[T/C]	0.900	ns	ns	ns	ns	ns	ns	ns
ILL_C4791_1099_[T/G]	0.565	ns	ns	ns	ns	ns	ns	ns
ILL_C300_4982_[A/G]	0.613	ns	**1.000**	ns	ns	ns	ns	0.000
ILL_C4593_326_[T/G]	0.527	ns	ns	ns	ns	ns	ns	ns
PS_C46533_541_[T/C]	0.804	ns	**1.000**	**0.990**	ns	ns	**1.000**	**1.000**
PS_C15351_193_[A/G]	0.920	ns	ns	ns	ns	ns	ns	ns

PS_C15018_147_[T/G]	0.405	0.000	**1.000**	ns	ns	ns	ns	0.000
PS_C38608_168_[T/C]	0.288	0.050	ns	ns	ns	ns	ns	ns
ILL_C387_215_[A/G]	**0.986**	ns	**1.000**	ns	0.003	**1.000**	ns	0.000
ILL_C5339_366_[T/C]	0.831	ns	ns	ns	ns	ns	ns	ns
PS_C34511_71_[T/G]	0.527	ns	ns	ns	ns	ns	ns	ns
PS_C14838_228_[T/C]	0.541	ns	ns	ns	ns	ns	**0.975**	ns
ILL_C20267_102_[T/C]	0.502	ns	ns	ns	ns	ns	ns	ns
PS_C46674_204_[C/G]	0.602	ns	ns	ns	ns	ns	ns	**0.975**
ILL_C1878_506_[T/C]	0.599	0.000	**0.962**	**1.000**	ns	ns	ns	**0.997**
PS_C36563_85_[T/G]	0.491	ns	ns	ns	ns	ns	ns	ns
ILL_C2735_326_[T/G]	0.569	ns	ns	ns	ns	ns	ns	ns
ILL_C250_199_[T/C]	0.613	ns	ns	ns	ns	ns	ns	ns
PS_C47845_238_[T/C]	0.632	ns	ns	ns	ns	ns	ns	ns
ILL_C1363_269_[T/C]	0.940	ns	ns	ns	ns	ns	ns	ns
ILL_C18774_676_[T/C]	0.542	ns	0.000	ns	ns	ns	ns	ns
PS_C23647_375_[A/G]	**0.998**	0.000	ns	ns	ns	ns	ns	ns
ILL_C2028_1228_[A/T]	0.544	ns	ns	ns	ns	ns	ns	ns
PS_C11985_171_[A/G]	0.346	ns	ns	ns	ns	ns	ns	ns
PS_C11911_576_[A/C]	0.411	0.002	ns	ns	ns	ns	ns	ns
ILL_C1254_187_[A/G]	0.558	ns	ns	ns	ns	ns	ns	ns
ILL_C20427_267_[T/C]	0.696	ns	**0.957**	ns	ns	ns	ns	ns
PS_C21989_160_[T/G]	0.737	ns	**0.979**	ns	ns	ns	ns	ns
PS_C12069_1181_[T/C]	0.266	0.000	ns	ns	ns	ns	ns	**1.000**
ILL_C327_1076_[C/G]	0.625	0.000	ns	ns	ns	ns	ns	ns
PS_C11847_438_[T/C]	0.888	ns	ns	ns	ns	ns	ns	**0.983**
PS_C47375_253_[A/G]	**0.970**	ns	ns	ns	**0.997**	ns	ns	ns
PS_C34420_787_[A/G]	0.510	0.000	0.000	ns	ns	ns	ns	ns
ILL_C2915_875_[T/C]	0.747	0.000	ns	ns	0.026	ns	ns	ns
ILL_C22491_727_[T/C]	0.515	**0.953**	ns	ns	ns	ns	ns	ns
ILL_C911_1343_[T/A]	0.608	ns	ns	ns	ns	ns	ns	ns

PS_C12196_434_[G/C]	0.519	ns	ns	ns	ns	ns	ns	ns
PS_C14297_369_[T/G]	0.440	ns	ns	ns	ns	**0.982**	**0.960**	ns
PS_C47340_198_[A/G]	0.368	ns	ns	ns	ns	ns	ns	ns
PS_C36706_579_[T/C]	0.512	ns	ns	ns	**1.000**	ns	ns	**0.998**
PS_C46597_301_[C/G]	0.673	ns	ns	ns	**0.971**	ns	ns	**1.000**
ILL_C2040_1251_[A/T]	0.349	ns	ns	ns	ns	ns	ns	**0.995**
PS_C35683_190_[A/G]	0.429	ns	ns	ns	ns	ns	ns	ns
PS_C46532_529_[A/T]	0.441	ns	ns	ns	ns	ns	ns	ns
PS_C15379_78_[A/C]	0.442	ns	ns	ns	ns	ns	ns	ns
PS_C11871_171_[A/G]	0.393	ns	ns	ns	ns	ns	ns	ns
PS_C35823_1199_[A/G]	0.500	**1.000**	**0.965**	ns	ns	ns	ns	ns
PS_C39731_379_[A/C]	**0.971**	ns	ns	ns	**0.991**	ns	ns	ns
ILL_C6061_1289_[T/G]	0.492	0.000	0.000	ns	**0.973**	ns	ns	ns
ILL_C980_261_[A/G]	0.609	ns	ns	ns	ns	ns	ns	ns
Directional selection	9	6	12	7	9	5	5	16
Balancing selection	0	23	9	0	3	1	3	5
Total	9	29	21	7	12	6	8	21

	WC(F_1) vs RP(2011)	**WC(F_2) vs RP(2003)**	**WC(F_2) vs RP(2011)**
	Lositan	**Lositan**	**Lositan**
Locus	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**	***P*-value (Simulated F_{st} < sample F_{st})**
ILL_C2122_257_[A/G]	ns	ns	ns
PS_C15402_271_[T/A]	ns	ns	ns
ILL_C253_1545_[A/C]	ns	ns	ns
ILL_C428_225_[A/G]	ns	ns	ns
PS_C23094_578_[G/C]	ns	ns	ns
ILL_C20682_843_[A/T]	ns	ns	ns
PS_C34490_403_[A/G]	ns	ns	ns

PS_C25208 _377_[T/G]	**1.000**	ns	ns
ILL_C1783_492_[T/G]	ns	ns	ns
ILL_C929_734_[T/C]	**0.993**	ns	ns
PS_C11665_287_[A/G]	ns	ns	ns
ILL_C2141_350_[T/A]	ns	ns	ns
PS_C23591_200_[T/C]	ns	ns	ns
PS_C15689_162_[A/T]	ns	ns	ns
ILL_C229_2772_[A/G]	0.000	ns	ns
ILL_C22574_507_[A/C]	ns	ns	ns
PS_C15088_268_[A/G]	ns	ns	ns
ILL_C2406_641_[T/C]	ns	ns	ns
PS_C12925_666_[A/G]	ns	ns	ns
PS_C11984_159_[T/G]	ns	ns	ns
ILL_C22449_261_[A/G]	ns	ns	ns
PS_C11970_157_[A/G]	ns	ns	ns
PS_C23051_368_[A/G]	ns	ns	ns
ILL_C6012_280_[T/C]	ns	ns	ns
ILL_C2903_1043_[T/A]	**0.999**	ns	ns
PS_C34501_638_[A/G]	**1.000**	**0.996**	**0.999**
PS_C12352_527_[T/C]	ns	ns	ns
PS_C34725_229_[T/C]	ns	ns	ns
PS_C15230_93_[A/G]	ns	ns	ns
PS_C23075_525_[G/C]	ns	ns	ns
ILL_C3835_411_[A/G]	ns	ns	ns
PS_C24743_123_[T/G]	ns	ns	ns
ILL_C5634_234_[A/G]	ns	**0.964**	ns
PS_C23630_237_[T/C]	ns	ns	ns
ILL_C1813_300_[T/G]	ns	ns	ns
PS_C12218_188_[A/G]	ns	ns	ns
PS_C36136_185_[T/C]	ns	ns	ns

PS_C23070_1364_[A/G]	ns	ns	ns
ILL_C618_116_[A/G]	ns	ns	ns
PS_C1652_228_[A/C]	ns	**0.965**	**1.000**
PS_C25083_285_[A/G]	ns	ns	**0.972**
ILL_C394_1510_[A/G]	ns	ns	ns
PS_C24267_71_[A/G]	ns	ns	ns
PS_C34507_1191_[T/C]	ns	ns	ns
PS_C36273_73_[A/C]	ns	ns	ns
PS_C35977_153_[T/C]	**1.000**	ns	**0.990**
PS_C16093_142_[T/C]	ns	ns	ns
PS_C16031_147_[A/G]	ns	ns	ns
ILL_C5106_273_[T/C]	ns	ns	ns
PS_C00512_245_[T/C]	ns	ns	ns
ILL_C140_2421_[A/G]	ns	ns	ns
PS_C36237_70_[C/G]	ns	**0.971**	ns
PS_C34670_302_[A/C]	ns	ns	ns
ILL_C21880_1003_[G/C]	**0.964**	ns	ns
PS_C47330_170_[A/C]	ns	ns	ns
PS_C36522_249_[G/C]	ns	ns	ns
PS_C34604_423_[A/T]	ns	ns	ns
PS_C34501_77_[T/G]	ns	ns	ns
PS_C28810_290_[A/C]	0.000	ns	ns
PS_C28886_317_[T/C]	**0.980**	ns	ns
ILL_C31_1387_[T/C]	ns	ns	ns
ILL_C844_440_[T/C]	0.000	ns	ns
PS_C46857_366_[T/C]	ns	ns	ns
PS_C34605_382_[A/G]	**0.983**	ns	ns
ILL_C853_1199_[A/G]	ns	ns	ns
ILL_C1002_85_[T/A]	ns	ns	ns
PS_C11659_399_[T/C]	ns	ns	ns

ILL_C4791_1099_[T/G]	ns	ns	ns
ILL_C300_4982_[A/G]	ns	ns	ns
ILL_C4593_326_[T/G]	ns	ns	ns
PS_C46533_541_[T/C]	ns	ns	ns
PS_C15351_193_[A/G]	ns	ns	ns
PS_C15018_147_[T/G]	ns	ns	ns
PS_C38608_168_[T/C]	ns	ns	ns
ILL_C387_215_[A/G]	0.000	ns	ns
ILL_C5339_366_[T/C]	ns	ns	ns
PS_C34511_71_[T/G]	ns	ns	ns
PS_C14838_228_[T/C]	ns	ns	ns
ILL_C20267_102_[T/C]	ns	ns	ns
PS_C46674_204_[C/G]	ns	ns	**0.953**
ILL_C1878_506_[T/C]	ns	ns	ns
PS_C36563_85_[T/G]	ns	ns	ns
ILL_C2735_326_[T/G]	ns	ns	ns
ILL_C250_199_[T/C]	ns	ns	ns
PS_C47845_238_[T/C]	ns	ns	ns
ILL_C1363_269_[T/C]	ns	ns	ns
ILL_C18774_676_[T/C]	ns	ns	ns
PS_C23647_375_[A/G]	ns	ns	ns
ILL_C2028_1228_[A/T]	ns	ns	ns
PS_C11985_171_[A/G]	ns	ns	ns
PS_C11911_576_[A/C]	ns	ns	ns
ILL_C1254_187_[A/G]	ns	ns	ns
ILL_C20427_267_[T/C]	**0.963**	ns	ns
PS_C21989_160_[T/G]	ns	ns	ns
PS_C12069_1181_[T/C]	0.000	**1.000**	ns
ILL_C327_1076_[C/G]	ns	ns	ns
PS_C11847_438_[T/C]	ns	ns	ns

PS_C47375_253_[A/G]	ns	**0.999**	**0.999**
PS_C34420_787_[A/G]	ns	ns	ns
ILL_C2915_875_[T/C]	ns	ns	ns
ILL_C22491_727_[T/C]	ns	ns	ns
ILL_C911_1343_[T/A]	ns	ns	ns
PS_C12196_434_[G/C]	ns	ns	ns
PS_C14297_369_[T/G]	ns	ns	ns
PS_C47340_198_[A/G]	ns	ns	ns
PS_C36706_579_[T/C]	**0.989**	ns	ns
PS_C46597_301_[C/G]	**0.979**	ns	ns
ILL_C2040_1251_[A/T]	**1.000**	ns	ns
PS_C35683_190_[A/G]	ns	ns	ns
PS_C46532_529_[A/T]	ns	ns	ns
PS_C15379_78_[A/C]	ns	ns	ns
PS_C11871_171_[A/G]	ns	ns	ns
PS_C35823_1199_[A/G]	**0.980**	ns	ns
PS_C39731_379_[A/C]	ns	ns	**1.000**
ILL_C6061_1289_[T/G]	ns	**0.999**	**0.995**
ILL_C980_261_[A/G]	ns	**0.960**	ns
Directional selection	13	8	8
Balancing selection	5	0	0
Total	18	8	8

Appendix D

Scientific Contributions during Doctoral Candidature (2010-2012)

1. Published papers with indirect relevance to the work presented in this dissertation

2. Published papers, to date, directly emanating from the work presented in this dissertation

3. Local conference presentations

4. International conference presentations

1. Published papers with indirect relevance to the work presented in this dissertation:

Vervalle J, Hepple J, Jansen S, Du Plessis J, Wang P, **Rhode C**, Roodt-Wilding R (2012) Integrated linkage map of *Haliotis midae* Linnaeus based on microsatellites and SNPs. J Shellfish Res *in press.*

Slabbert R, Hepple J, **Rhode C**, Bester-Van der Merwe AE, Roodt-Wilding R (2012) Microsatellite marker development in the abalone *Haliotis midae* using pyrosequencing (454): characterisation and in silico analyses. Genet Mol Res 11: 2769-2779.

Rhode C, Roodt-Wilding R (2011) Bioinformatic survey of *Haliotis midae* microsatellites reveals a non-random distribution of repeat motifs. Biol Bull 221: 147-54.

2. Published papers, to date, directly emanating from the work presented in this dissertation:

Rhode C, Hepple J, Jansen S, Davis T, Vervalle J, Bester-van der Merwe AE, Roodt-Wilding R (2012) A population genetic analysis of abalone domestication events in South Africa: Implications for the management of the abalone resource. Aquaculture 356-357: 235-242.

3. Local conference presentations:

Rhode C*, Bester-van der Merwe AE, Roodt-Wilding R. Oral presentation: Molecular signatures of selection in the genome of the South African abalone, *Haliotis midae*. Joint

Conference of the South African Genetic Society and the South African Bioinformatics Society. September 2012, Stellenbosch, Western Cape, South Africa.

Rhode C*, Hepple J, Jansen S, Davis T, Vervalle J, Bester-van der Merwe AE, Roodt-Wilding R. Poster presentation: Population genetics of abalone (*Haliotis midae*) domestication events in South Africa. Joint Conference of the South African Genetic Society and the South African Bioinformatics Society. September 2012, Stellenbosch, Western Cape, South Africa.

Rhode C*, Bester-van der Merwe A, Roodt-Wilding R. Oral presentation: From classical breeding to genomics: An integrative approach to abalone domestication in South Africa. Congress of the South African Society for Animal Science. July 2012, East London, Eastern Cape, South Africa.

Rhode C*, Bester-van der Merwe AE, Roodt-Wilding R. Oral (invited) presentation: Genetic variability of the South African abalone: Implications for aquaculture and conservation. Critical Thinkers' Platform on Aquaculture and Emerging Technologies in South Africa hosted jointly by the Department of Agriculture Forestry and Fisheries (DAFF), Department of Science and Technology (DST) and the National Science and Technology Forum (NSTF). October 2011, Port Elizabeth, Eastern Cape, South Africa.

Van der Merwe AE*, **Rhode C**, Roodt-Wilding R. Oral presentation: Preserving South African abalone for the future: Genetics to genomics and integrative resource management. The Academy of Science of South Africa's Annual Young Scientists' Conference. October 2010, Pretoria, Guateng, South Africa.

Rhode C*, Roodt-Wilding R. Oral Presentation: Development of type I molecular markers for the South African abalone (*Haliotis midae*), using an *in silico* mining approach. Biannual Conference of the South African Genetics Society. April 2010, Bloemfontein, Free State, South Africa

[**Presenting author*]

4. International conference presentations

Roodt-Wilding R*, Blaauw S, Du Plessis J, **Rhode C**, Bester-van der Merwe. Oral and poster presentation: SNP marker technologies and applications for genetic improvement in the South African abalone (*Haliotis midae*). Conference for the International Society for Animal Genetics. July 2012, Cairns, Australia.

Rhode C*, Bester-van der Merwe A, Roodt-Wilding R. Oral presentation: Population genetic analysis of cultured abalone (*Haliotis midae*) in South Africa: Considerations for sustainable genetic improvement. International Symposium on Genetics in Aquaculture. June 2012, Auburn, Alabama, USA.

Bester-van der Merwe A, Blaauw S, Du Plessis J., **Rhode C***, Rouvay Roodt-Wilding. Oral presentation: Development and utilization of *in silico* SNPs in the South African aquaculture species, *Haliotis midae*. International Symposium on Genetics in Aquaculture. June 2012, Auburn, Alabama, USA.

Roodt-Wilding R*, **Rhode C**, Bester-van der Merwe AE Oral presentation: Sustainable genetic improvement of the South African abalone (*Haliotis midae*). International Abalone Symposium. May 2012, Hobart, Tasmania, Australia.

Roodt-Wilding R*, Jansen S, Blaauw S, Du Plessis J, Vervalle J, **Rhode C**, Bester-van der Merwe AE. Oral presentation: SNP technologies in cultivated South African abalone (*Haliotis midae*). International Abalone Symposium. May 2012, Hobart, Tasmania, Australia.

Rhode C*, Roodt-Wilding R. Poster presentation: *In silico* strategies for type I molecular marker development in South African Abalone (*Haliotis midae*). Annual International Plant and Animal Genome Conference. January 2010, San Diego, CA, USA.

[**Presenting author*]