UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Faculty of Natural and Agricultural Sciences

Phylogeography of the snapper kob *Otolithes ruber* (Bloch & Schneider 1801) from the South West Indian Ocean

By

Ilkser Erdem Kiper

Submitted in partial fulfilment of the requirements for the degree

MAGISTER SCIENTIAE

in the Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

January 2014

Phylogeography of the snapper kob *Otolithes ruber* (Bloch & Schneider 1801) from the South West Indian Ocean

By

Ilkser Erdem Kiper

Supervisor: Prof. P. Bloomer
Molecular Ecology and Evolution Programme
Department of Genetics
University of Pretoria
Pretoria

Co-supervisor: Dr. S.T. Fennessy
Oceanographic Research Institute
Durban

Department of Genetics

Magister Scientiae

Dissertation summary

Studying the distribution and demography of marine species is challenging due to the open nature of the oceans. This latter was historically believed to facilitate extensive gene flow among populations when clear barriers to gene flow were not apparent. Gene flow between populations could be mediated directly by adult migration across large distances for reproduction purposes, feeding and habitat preference or via larval dispersal with the aid of ocean currents, in order to recruit to optimal areas for development and to increase survival. Hence, gene flow among localized populations across large geographic distances most likely results in weak or no genetic differentiation. This recognized model has changed in recent years, as limited gene flow has been demonstrated for many species even in the absence of physical barriers to genetic mixing.

To understand fish population dynamics and to manage marine resources sustainably, assessing the degree of population connectivity has become an important focus. It has been reported globally that many marine species are exposed to intensive fishing activities. One such species from the South West Indian Ocean (SWIO) region is the widespread fish species, *Otolither ruber.* Based on our knowledge, there are no genetic investigations conducted for this species to date. Hence, I studied the evolutionary and ecological processes influencing genetic diversity and the population genetic structure of *O. ruber* throughout the region based on mitochondrial and nuclear data analyses.

Before attempting large scale data generation, critical consideration should be given to marker selection. I was able to find the most suitable mitochondrial gene from three evaluated genes. The cytochrome *b* gene gave consistently good amplification and showed the most variation within and among *O. ruber* sampling locations. The advantageous cross-species amplification approach yielded seven microsatellites, previously developed in other sciaenid fishes, with relatively high levels of polymorphism in *O. ruber*.

Two distinct *O. ruber* populations in the north and the south of the SWIO region were identified based on statistical analyses of cytochrome *b* sequences and microsatellite genotypes. Low, unidirectional, female gene flow (based on mtDNA) and overall asymmetrical gene flow (based on nuclear DNA markers) were inferred from north to south between these evolutionary units. The isolation with migration model demonstrated a recent past population divergence and a low level of ongoing gene flow. The observed genetic differentiation in this shallow-water demersal sciaenid is mostly likely determined by the

oceanography of the region, historical processes and life history traits, such as male-biased dispersal. The biology of the species in terms of reproductive and migratory behaviour should be further investigated to substantiate these findings.

The results obtained from this study contribute to improved biological knowledge of *O. ruber* and its genetic status in the SWIO region, and will be used for future studies undertaking a management assessment of *O. ruber*. The observed variation among these populations should be maintained by management interventions. The study will also add to comparative phylogeograpic studies of co-distributed species, particularly other sciaenids.

Declaration

I, ... declare that this dissertation, which I hereby submit for the degree ..... at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: ……. Date: …….

Acknowledgements

I would like to thank my supervisors Prof. Paulette Bloomer and Dr Sean Fennessy for giving me, as an international student, the opportunity to undertake this study. I am grateful for their encouragement, support, guidance and valuable input into this project. I wish to thank my committee members Prof Jaco Greeff and Dr Thierry Hoareau for their contributions; a special word of thanks to Dr Hoareau for his advice and discussions.

I would like to thank the South Western Indian Ocean Fisheries Project (SWIOFP) for funding this project and sample collections through Sean Fennessy (South Africa), Edward Kimane (Kenya), Baraka Kuguru (Tanzania), Rui Mutombene (Mozambique) and Faratiana Ratsifa (Madagascar).

Special thanks to my parents, Hanim and Kemal Kiper, to my siblings, Ilkadim and Ismier Kiper, for their endless support, believing in me and being there for me. I would like to thank my uncle, Yasar Kiper, for the financial support he provided through my studying years.

I would like to thank members of my research group, the Molecular Ecology and Evolution Programme (MEEP), for the friendly, helpful and supportive working environment. Finally, I would like to thank my friends for making my difficult times go super fast and smoothly and being patient with me.

Table of contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Literature review: Genetic diversity assessment of the snapper kob (*Otolithes ruber*) in the South Western Indian Ocean

1.1 Introduction

In the marine environment, patterns of population connectivity can be revealed by genetic signatures left from successful migration (Hellberg 2009). Widespread marine species often comprise single populations (Hauser & Carvalho 2008), as realised levels of gene flow lead to greater genetic connectivity over large geographical distances. One should, however, also consider the other forces that might be involved in shaping marine populations; this requires a framework within which effects of species' life history traits, environmental factors and evolutionary factors are characterised.

Knowledge of the biology and behaviour of species, including for example migratory and spawning behaviour and having different life stages (egg, larval, juvenile and adult stages) with specific requirements and adaptations, is essential to understanding diversity. For instance, longer duration of the eggs and larvae increases the potential for gene flow (White *et al.* 2009). Some characteristics decrease gene flow, resulting in more limited connectivity between populations, e.g. demersal eggs, large-sized larvae and a short planktonic larval stage (McCusker & Bentzen 2010). Since many benthic marine organisms show no or limited migratory ability (Hellberg 2009), gene flow between populations is expected to occur due to active and passive larval movements.

In some instances, the habitat preference of an organism can have an impact on population differentiation (Ayre *et al.* 2009; Hellberg 2009). Lewis & Fontoura (2005) suggested differential habitat occupation in *Paralonchurus brasiliensis*, a distribution pattern linked to sex where males are not present near the shoreline. Population differentiation might also occur due to sex-specific dispersal, for instance in species such as sharks that display female philopatry (Portnoy *et al.* 2010; Karl *et al.* 2011; Tillett *et al.* 2012). That is a system in which females stay in the home ranges or groups, while males disperse once they are sexually mature (Dudgeon *et al.* 2012). Limited movements of females can thus be responsible for structuring of populations and the level of genetic mixing will depend on male movements.

From an environmental/ecological perspective, habitat topography, salinity, temperature and food distribution can contribute to genetic diversity and species' responses to these variables

may differ for separate life stages (Ayre *et al.* 2009). Furthermore, ocean currents and counter-currents lead to formation of ocean fronts, eddies and gyres. These physical irregularities of the marine environment influence connectivity in ways that limit or enhance gene flow (Gonzalez & Zardoya 2007; Berntson & Moran 2009; Viñas *et al.* 2010). Additionally, contemporary population genetic differentiation and geographic distribution of marine fishes can be a result of historical events that occurred due to climatic changes and associated changes in the sea level (Fig. 1.1), e.g. the melting of the ice sheets, movement of the ocean floor, sedimentation (Martínez *et al.* 2006; Hoareau *et al.* 2012), climate-associated periodical extinctions and also changes in the direction of ocean currents (Gonzalez & Zardoya 2007). It has been demonstrated that an expansion-contraction model describes the effects of glacial cycles on species' evolutionary history (Provan & Bennett 2008), since the drop and rise in sea levels caused population size fluctuations; reductions are associated with genetic bottlenecks and increased genetic drift, while expansions may promote gene flow in the absence of barriers.

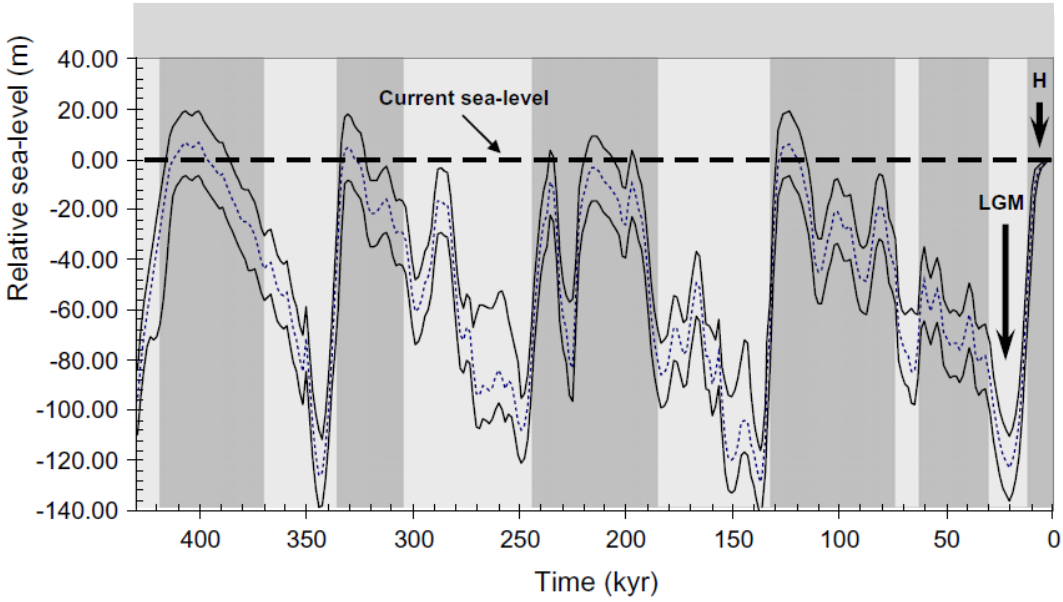


Fig. 1.1 Relative changes in sea level during the last 450 000 years (the late Pleistocene period). Rises and drops in sea levels are represented by dark and light shading, respectively. H: Holocene period, LGM: Last Glacial Period (modified from Hoareau *et al.* 2012).

For many marine species, geographic genetic structuring inferred from molecular markers has recently been utilized by fishery authorities to develop sustainable management plans, especially for overexploited species. Within this context, molecular data should be interpreted together with knowledge of ecology, oceanography, past distribution and demography to gain a better perspective on population genetic structure and history (Hauser & Ward 1998). Understanding population genetic differentiation is significant in terms of species conservation and management of the species (Ovenden *et al.* 2009). Without such information, the species may be managed as a single unit although it might be composed of separate populations. Subsequent overfishing results in depletion of part of the species' stocks with limited opportunity for regeneration from the populations that remain and the outcome is the loss of genetic diversity (Berntson & Moran 2009). Thus, exploring how populations cluster genetically is of primary interest for population studies.

1.2 Genetic structuring in marine fish populations

Limited geographic genetic structure is the expected pattern in marine species in the absence of obvious barriers to gene exchange (Palumbi 1994). However, the level of connectivity varies among species and may involve geographically restricted lineages (Boissin *et al.* 2011) or subdivision into smaller populations. Dispersal is the driving force behind connectivity and can contribute to population growth and gene flow (Lowe & Allendorf 2010). One can define dispersal as the movement of individuals between distinct populations, with long-term persistence in the new environment; this is compatible with the term "migration", allowing gene exchange to contribute to the gene pool of the new population (Lowe & Allendorf 2010). Studying such genetic connectivity can tell us what degree of gene flow affects evolutionary and ecological processes shaping natural populations.

Spatial and temporal changes can be exhibited by natural populations: Over space populations can subdivide into numerous subpopulations and connect with others, while over time populations vary in size, density and location (Hey & Machado 2003). Genetic differentiation among populations is illustrated in Fig.1.2, which shows the different levels of connectivity. The two extremes of observed patterns of differentiation are panmixia (Fig.1.2: Scenario D), with all the individuals in the population acting as a homogenous group with random mating, versus complete isolation (Fig.1.2: Scenario A), where subpopulations of the same species have no or extremely limited genetic connectivity, thus acting as independent,

separate populations. Populations in Fig.1.2 B represent modest connectivity, whereas those in Fig.1.2 C are linked through ongoing gene flow.

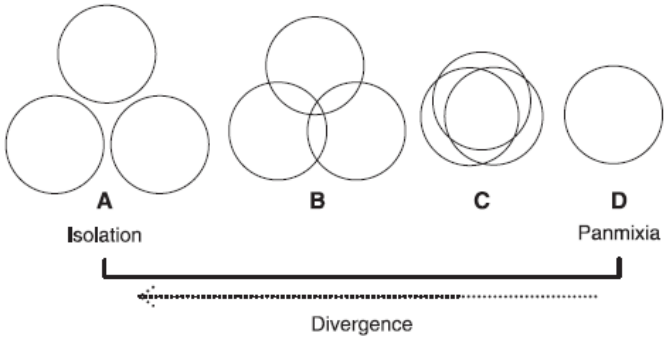


Fig. 1.2 Population (represented by circles) models based on the level of connectivity among subpopulations (Waples & Gaggiotti 2006): **A)** complete isolation of subpopulations, **B)** low levels of genetic connectivity, **C)** high levels of connectivity, **D)** panmixia (the null hypothesis for population structure in marine species).

There are four major spatial population genetic structure models: (i) the island population represents a model where a mainland population has migrants moving to one or more island population(s) (Hey & Machado 2003); (ii) in the stepping-stone model, gene exchange can take place only between adjacent populations and one can expect an inverse relationship between distance and genetic distance (White *et al.* 2010); (iii) the isolation by distance model (IBD), first introduced by Wright (1943), is the extreme case of the stepping-stone model, predicting genetic differentiation between populations that generally increases with geographic distances (Paquette & Lapointe 2009); while (iv) the metapopulation model explains the movement of populations, rather than individuals among populations. Over time, a population merges with another, becomes lost or represents a founder population (Hey & Machado 2003). Regional stability is strengthened as individuals disperse from habitat patches that have not suffered extinction to recolonize empty habitat patches (Hanski 1985, Fahrig & Paloheimo 1988, Adler & Nuernberger 1994). Consequently, subpopulations may experience extinction and recolonization, while the regional population (metapopulation) remains relatively stable.

The level of genetic variation within species displaying these different population models can be assessed using F-statistics, which has been the most commonly used method for

estimating levels of gene flow among populations (Berntson & Moran 2009; Lowe & Allendorf 2010). By this method the proportion of the observed genetic variation within each population is compared to that between populations; the value range from 0 indicating lack of differentiation to 1 indicating high levels of differentiation among populations. Figure 1.3 represents the different factors, such as population connectivity, size and history, affecting the level of genetic differentiation measured by F_{ST}. Both low levels of connectivity and small population size increases differentiation (Beerli & Felsenstein 2001; Lowe & Allendorf 2010). Effective population size is a measure of how many individuals in a population matter in evolutionary genetic terms (Berntson & Moran 2009). Consequences of low effective population size may involve the loss of adaptive variation and accumulation of detrimental alleles due to increased genetic drift and inbreeding (Hare *et al.* 2011; Dudgeon *et al.* 2012). On the contrary, a large effective population size counters the action of genetic drift. However, one should consider that not all mature adults spawn and not all spawning adults successfully contribute to the next generation due to external fertilization and the high rate of mortality of eggs and larvae (Hedgecock & Pudovkin 2011). From the perspective of population history, secondary contact (previously isolated and genetically differentiated populations coming into contact) and founder events (when a new population is established by a small number of individuals from a large population) promote population differentiation, while range expansion (colonization) reduces divergence (Hauser & Carvalho 2008).

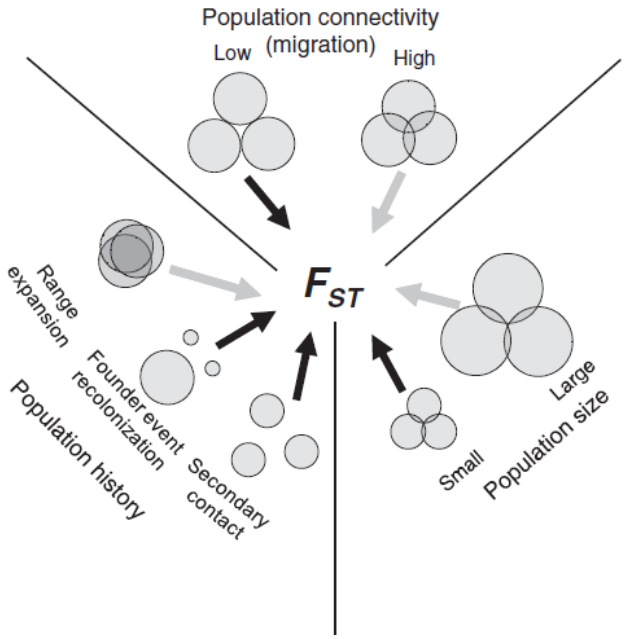


Fig. 1.3 Diagram showing the different contributors to genetic differentiation based on F_{ST} (proportion of the total variation in allele frequencies as a result of differences in average allelic frequencies among populations). Grey arrows indicate factors that reduce differentiation, while black arrows indicate factors that increase genetic differentiation (Hauser & Carvalho 2008).

Deducing historical and contemporary processes responsible for observed patterns of spatial genetic differentiation is the main aim of population studies (Beheregaray 2008; Ovenden *et al*. 2011). Even though IBD is one of the most frequently observed population models, its pattern may differ with species and habitat structure (Berntson & Moran 2009). Furthermore, F_{ST} cannot differentiate genetic differences that are due to gene flow–drift balance from those of completely isolated populations (Waples & Gaggiotti 2006). Genetic differentiation without significant IBD may best be explained through the recently developed isolation-with-migration (IM) model; it is a non-equilibrium model and represents a valuable tool to distinguish ancestral polymorphism from alleles shared as a consequence of ongoing gene flow between defined populations (Marko & Hart 2011).

The IM model is a coalescent-based method to assess gene exchange among populations (Fig. 1.4). Its likelihood approach allows the use of all the data, including the proportion of rare alleles (Hey & Nielsen 2007); it gives more precise and less biased estimates of population migration rate in terms of number of migrants (*Nm*) (Beerli & Felsenstein 2001).

The model estimates six demographic parameters, including the population mutation rate (θ) for each daughter population and the ancestral population, and mutation-scaled migration (m) to each population and the population splitting time (t). In the case of no strong population size increase or decrease after the splitting event, the sizes of the two daughter populations should approximately add up to the ancestral population size (Hey & Nielsen 2004).

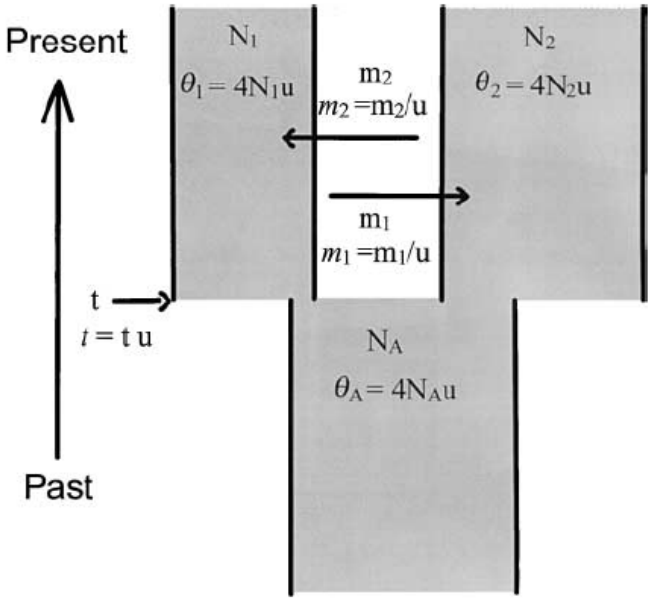


Fig. 1.4 The isolation-with-migration model is illustrated by two parameters: Demographic parameters and parameters scaled by the mutation rate *u*. N_1, N_2 and N_A are the effective population sizes for descendent populations 1 and 2 and the ancestral population, respectively; θ_1, θ_2 and θ_A are the population size estimates; m_1 and m_2 are the mutation-scaled migration rates while m_1 and m_2 are the migration rates per generation (m_1 represents gene movement from population 2 to population 1 and m_2 represents movement of genes from population 1 to population 2); t is the population splitting time (Hey & Nielsen 2004).

As with all coalescent-based analyses, an important requirement of the model is an accurately calibrated substitution rate. One of the widely used calibration points for substitution rates in marine organisms is the rise of the Isthmus of Panama around 2-3.5 million years ago (Schultz *et al.* 2008; Schluessel *et al.* 2010). Such a vicariant event has a benefit for dating, as divergence of many taxa can be used for the same calibration (Lessios

2008). Comparing the divergence between species pairs on both sides of this barrier, aids in determining the occurrence of simultaneous splitting events coincident with the appearance of the barrier (Lessios 2008). The IM model has been used to address the past and present geographical distribution of genetic lineages, dating lineage divergence and determining the pattern of the gene flow based on the calculated estimates (Kotlik *et al.* 2008; Papetti *et al.* 2012).

There are numerous molecular markers available for investigating the extent of connectivity among populations of marine species. A representative number of markers with high variation enable the correct assignment of individuals to populations even in the presence of extensive gene flow (Paetkau *et al.* 2004). The use of different types of molecular markers more accurately reflects population history since markers show differences in mutation rate, patterns of inheritance and the action of selection (Dudgeon *et al*. 2012).

1.3 Molecular markers for detecting genetic variation

All living organisms accumulate variations in their genomes due to interactions of evolutionary processes over time. Molecular markers can be used as representations of genetic variability for studies of population genetics, phylogeography and molecular evolution. These markers are required to be heritable and recognizable through molecular approaches (Liu & Cordes 2004). There are considerable differences in features between marker types, for instance their polymorphisms and inheritance patterns.

The marker of choice depends on the questions being asked in a study. In the case of population genetic studies, one can identify two main marker types to reveal isolation or population connectivity, namely frequency and sequence markers (Hellberg 2009). Microsatellites are nuclear DNA frequency markers, meaning that the recent genetic history of the study organism is deduced based on allele frequencies. Mitochondrial DNA markers are sequence markers from which the relationships between alleles can be inferred (Hellberg 2009). Although there are numerous existing molecular markers, mitochondrial DNA and microsatellites have become widely used (Boissin *et al.* 2011) due to their specific aspects discussed below.

1.3.1 Microsatellites

Microsatellites are tandemly repeated motifs of 1-6 base pairs (Kasapidis & Magoulas 2008) with conserved flanking regions. They are co-dominant markers and show mutation rates between 10^{-2} and 10^{-6} mutations per locus per generation (Selkoe & Toonen 2006). There have been several models of mutation proposed for microsatellites based on changes occurring in tandem repeats only. Most of the studies showed that changes arise following the stepwise mutational model (SMM) (Van Oppen *et al.* 2000; Marriage *et al.* 2009). In the SMM there is equal probability of gaining and losing a single repeat unit within the microsatellite region, thus producing new alleles (Barkley *et al.* 2009). It assumes that some mutations give rise to new alleles that could be identical in state to others present in the population (Barkley *et al.* 2009). Based on the infinite allele model (IAM), each mutation results in a new allele state in the population (van Oppen *et al.* 2000). As an extension of the SMM model, the two-phase model (TPM) was developed to consider larger mutation events. According to this model, most mutational events result in an increase or decrease of one repeat unit (as per the SMM), but some mutations give rise to a large number of repeats inserted or deleted simultaneously (Bhargava & Fuentes 2010). However, no consensus on mutation models has been found due to the fact that microsatellites display very complex mutation processes (Van Oppen *et al.* 2000).

One of the valuable characteristics of microsatellites is their high polymorphism, exhibiting high rates of mutation (Barkley *et al.* 2009). Most microsatellite loci are neutral markers, therefore, there is no pressure of natural selection acting on them (Kasapidis & Magoulas 2008). A requirement of only small amounts of DNA makes it possible to do non-destructive sampling, for instance by fin-clipping, as well as to analyse old samples with degraded DNA (Hansen *et al.* 2001). New statistical tools in the application of microsatellites are useful in detection of population structure, bottlenecks, estimation of migration rates and assessment of parentage (Hansen *et al.* 2001). Microsatellites have also become an important tool in the field of conservation biology, providing a better understanding of genetic differentiation (Schmidt *et al.* 2009). They have been used to define stock structure of fish species and to determine how individuals from a mixed population could be assigned to their population origins (O'Connell & Wright 1997).

Cross-species amplification of microsatellite markers is a time- and cost-effective technique for testing the utility of markers developed in a source species in other related species (Reid *et al.* 2012). Since the conserved flanking regions are the sites for primer design, loci could be successfully amplified in other species. However, amplification success and

polymorphism of loci might differ from one species to the other in a family or higher taxon, depending on the genetic divergence between the source and other species (Carreras-Carbonell *et al.* 2008; Reid *et al.* 2012). In this way, many studies benefit and contribute to the expanding use of microsatellite markers over a reasonable time.

Although microsatellites are currently the markers of choice in population genetics, there are some associated drawbacks. One of the factors that may influence the estimate of variability of microsatellite loci is the presence of null alleles, i.e. amplification failures for some alleles after Polymerase Chain Reaction (PCR) and separation (O'Connell & Wright 1997). Null alleles arise due to point mutations occurring within the primer binding sites, which explain the absence of amplification. It can be detected as a homozygote excess in a data set or from the complete failure of PCR amplification in the presence of two null alleles (null homozygote) (Selkoe & Toonen 2006). Inefficient PCR conditions also result in null allele formation (Selkoe & Toonen 2006). Another potential cause of missing alleles is large allele dropouts. This can be explained when the longer allele of a heterozygote does not amplify as efficiently as the shorter one. Their amplification appears too faint to be detected in the genotype scoring process (Selkoe & Toonen 2006). This occurs as PCR amplifies shorter sequences more efficiently than longer ones. The cases of large allele dropouts are more frequent when alleles are very different in size in heterozygotes and if dealing with low quality DNA. Moreover, homoplasy is arguably the most important drawback of microsatellites. It is a limiting factor, causing underestimation of population diversity, subdivision and genetic divergence, as alleles are identical in state (length) but not identical by descent (Barkley *et al.* 2009). These disadvantages should be considered in decision-making regarding the most informative microsatellite markers. Fast-evolving microsatellite markers are valuable to disentangle recent events, but what can more slowly evolving markers, such as mitochondrial genes, contribute to the understanding of intraspecific genetic variation?

1.3.2 Mitochondrial DNA (mtDNA)

Mitochondrial DNA is a circular double-stranded DNA molecule approximately 15-20 kb in size in sciaenid species such as the spinyhead croaker (*Collichthys lucida*) (Cui *et al.* 2009; Cheng *et al.* 2011). It contains two ribosomal genes (*12S* rDNA and *16S* rDNA), 13 protein coding genes (*ATPase 6* and *8*, *COI-III*, *Cyt b*, *ND1-6* and *4L*), 22 tRNA genes and a noncoding region, the Control Region (CR; Cui *et al.* 2009; Xu *et al.* 2011) (Fig. 1.5). The mitochondrial genome shows maternal inheritance in most vertebrate species and is haploid, thus, its effective population size is a quarter that of autosomal nuclear DNA (Keeney *et al.*

2005). In comparison to single copy nuclear DNA, mtDNA accumulates mutations more rapidly because it lacks an efficient repair mechanism during replication (Cheng *et al.* 2011). Mitochondrial DNA has been used for determination of stock structure of many fishes, including scianids such as red drum (*Sciaenops ocellatus*) (Gold *et al.* 1993) and the southern king fish (*Macrodon atricauda*) (Rodrigues *et al.* 2013).

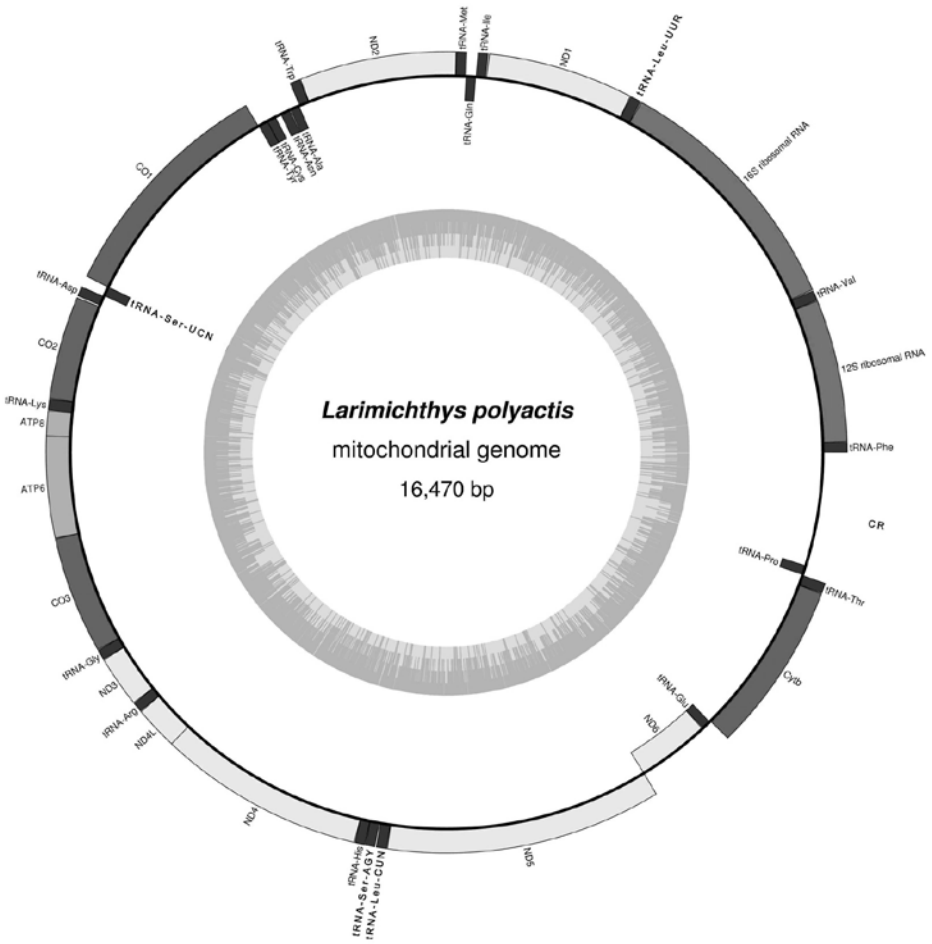


Fig. 1.5 A representative illustration of the gene organization of the mitochondrial genome of *Larimichthys polyactis* (Sciaenidae) with a total size of 16,470 bp (Cheng *et al.* 2011). Most of the genes are transcribed from the heavy strand (outer circle and gene labels); the nine genes transcribed from the light strand are shown in the inner circle; the CR is non-coding. GC content, which is the percentage of nitrogenous bases in the genome that are either G or C, is indicated by the inner circle.

Most eukaryotic cells contain many mitochondria, each having many copies of mtDNA (Ballard & Whitlock 2004). This allows relatively easy amplification of mtDNA regions, and the subsequent DNA sequencing is simple and effortless due to the haploid nature of the genome (Walther *et al.* 2011). Genetic variation analysis can be easy due to the non-recombinant maternal inheritance and evidence of neutral evolution of mtDNA (Walther *et al.* 2011). Some parts of the mtDNA molecule are highly conserved across species which makes universal primer development possible and, thus, the same mtDNA fragment can be analyzed in many species (Okumuş & Çiftci 2003). As with other markers, mtDNA also has some limitations. Its maternal inheritance only represents one aspect of history, which might be different compared to overall population or species' history (Okumuş & Çiftci 2003). The potential existence of nuclear copies of part of the mtDNA is another complicating feature (Zhang & Hewitt 2003).

Fish phylogenetics, phylogeography and studies dealing with molecular evolution have broadly benefited from mtDNA markers (Kawahara *et al.* 2008). The most studied mtDNA genes are cytochrome *b* (*Cyt b*), NADH dehydrogenase subunit 2 (*ND2*) and cytochrome *c* oxidase subunit I (*COI*). Due to the fact that these are protein coding genes, insertions and deletions are rare and such mutations usually involve addition or deletion of a codon(s) (Thacker 2003). In comparison to the control region, *ND2* has a slower rate of evolution which might provide more suitable levels of differentiation in investigating population structure (Bradman *et al.* 2011). The *COI* gene has been recognized as a good evolutionary marker in order to determine relationships within and among species in many marine fish (Hebert *et al.* 2003; Spies *et al.* 2006). It is also the suggested DNA barcode for species identification in animals. Cytochrome *b* was first used as a fish identification tool by Bartlett & Davidson (1991), which presented the recognition of four tuna species (*Thunnus* spp.). Even though *Cyt b* is a good representative gene for investigation of variation within and between species, it might become less informative with the increase in divergence levels (Unmack *et al.* 2011). The power of a genetic marker varies between species and the complexity of evolutionary histories.

Many studies of marine fishes use a framework of comparative analyses of both nuclear and mitochondrial markers to characterize population genetic structure and identify the evolutionary processes (Eytan & Hellberg 2010). The aim is the understanding of the effect of forces shaping the population genetic structure and the distribution of marine species over evolutionary and ecological time scales (Dudgeon *et al.* 2012). Microsatellites are best suited for investigating the contemporary changes, while mtDNA evolves at comparatively slower rates and, hence, is suited for questions of historical changes (Hickerson *et al.* 2010).

Although mtDNA dominates in phylogeographic studies, it should be kept in mind that this marker only contributes inference of a part of the overall population history and the resolution is not sufficient to explain current genetic structuring (Barber *et al.* 2012). Furthermore, contrasting nuclear and mtDNA data could be an indication of sex-biased gene flow (Okumuş & Çiftci 2003). The combined power of nuclear and mitochondrial markers has been highly informative within comparative phylogeographic frameworks, focused on specific regions (e.g. Carpenter *et al.* 2011; Sanciangco *et al.* 2013). One such a region is the South Western Indian Ocean (SWIO), a tropical and subtropical area, is of growing interest from a phylogeographic perspective (Wafar *et al.* 2011).

1.4 The South Western Indian Ocean (SWIO)

The region bordering Kenya and southwards within the Western Indian Ocean is called the SWIO (Fig. 1.6). This region is regarded as a biogeographical area of the Indo-West Pacific (IWP), with high levels of regional endemism and significant diversity of fish species as a consequence of the variety of habitats (Briggs & Bowen 2012) and the oceanographic conditions (Quartly & Srokosz 2004). The region is described as oligotrophic, meaning that the water is poor in nutrients, and includes both tropical and subtropical water masses. The complex oceanic features of the SWIO may have played a prominent role in shaping the evolutionary history of both widespread and endemic species that occur in the region; this encourages scientists to study evolutionary history of the inhabitants of the region.

1.4.1 The oceanography of the region

The major circulatory component is ocean currents (Smith & Heemstra 2003), which are involved in nutrient availability for marine animals and also assist their dispersal and distributions. In the tropical part of the SWIO region, the South Equatorial Current (SEC), which flows from east to west in the Indian Ocean, divides into two main circulation components when it reaches the eastern coast of Africa (Fig. 1.6A). The southern component includes the Mozambique Current (MC) and the North East Madagascar Current (NEMC), while the East African Coastal Current (EACC) comprises the northern component (Fratini *et al.* 2010). These currents are expected to separate marine populations off Tanzania and Kenya from populations off Mozambique and South Africa (Fratini *et al.* 2010). One also should consider the effect of seasonal monsoons that may change the direction of currents. For instance, the EACC becomes the Somali Current and combines with the Indian Monsoon Current during the southeast monsoon, contrasting with the northeast monsoon

period which causes a reduction in northward flow leading to the EACC becoming the Equatorial Counter Current (ECC) (Ragionieri *et al.* 2010) (Fig. 1.6B). The EACC, flowing from northern Kenya to the eastern Indian Ocean is more dynamic during the northeast monsoon period (December to April; Benny 2002, Visram *et al.* 2010). The SEC is a likely mechanism of larval transport from the islands to the east African mainland, whereas the EACC transports the larvae along the coast to the north (Visram *et al.* 2010).

Oceanographic features of the Mozambique Channel show variability, with mean circulation flowing southward along the continental slope, while the eastern part represents a northward flow (Matano *et al.* 2008). It was demonstrated that the western boundary current along the Mozambican shelf edge is continuous only for a short time and the channel is dominated by anticyclonic eddies with slow movements towards the south (Schouten *et al.* 2003) (Fig. 1.6A). These features are most likely formed in the north of the channel, where it narrows, and carry chlorophyll from the coastal regions into the centre of the channel (Lutjeharms *et al.* 2012). The volume of water transport increases in the south where the Agulhas Current (AC) takes action (Lutjeharms *et al.* 2012).

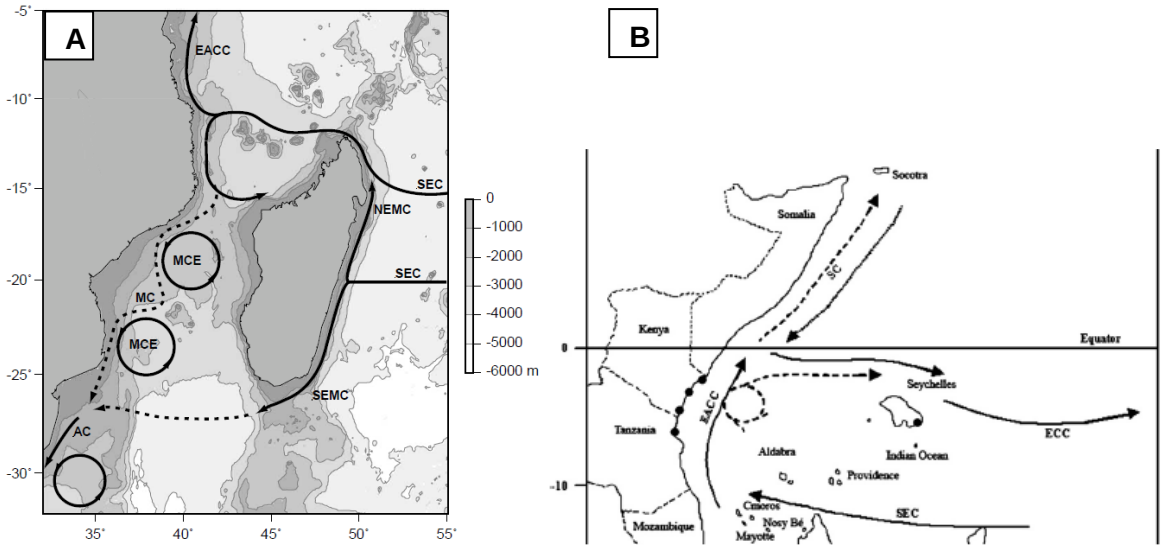


Fig. 1.6 Oceanography of the SWIO region: **A)** bathymetry of the region (Schouten *et al.* 2003). Shallow and deep regions are shaded according to the scale. Circulatory mechanisms represented are the main oceanic currents of the region and contributors to the Agulhas Current. SEC: South Equatorial Current, EACC: East African Coastal Current, NEMC and SEMC: North East and South East Madagascar Current, MC: Mozambique Current, AC: Agulhas Current, MCE: Mozambique Channel Eddies; **B)** circulatory mechanisms in the tropical part of the region. Dashed lines represent the southwest monsoon period (Ragionieri *et al.* 2010). SC: Somali Current, ECC: Equatorial Counter Current.

The central region of the channel shows regular variations as the movement of eddies from north to south is the dominant action. These eddies have diameters ranging between 300 and 350 km, with speeds of 3-6 km/day moving south (Schouten *et al.* 2003). They speed up to 8 km/day where they meet with the northern Agulhas Current flow (Schouten *et al.* 2003). It was shown that one eddy at 17°S reached the sea bed (~ 2000 m depth) with a speed of 10 cm/s (Schouten *et al.* 2003). High eddy variability, coupled with strong recirculation cells, influence the position of the mean flow in the south of the channel (Matano *et al.* 2008). Slow currents along the west coast of Madagascar may have no distinct influence on the main circulatory events or on the shelf regions (Lutjeharms 2006).

Eddies act as vehicles to transport zooplankton, therefore their activity has importance as productive components within the food chain; they might provide better conditions for larval survival due to containing phyto- and zooplankton in high concentrations (Durgadoo *et al.* 2010). They also play a protective role by retaining larvae in coastal areas, preventing their

drift into current systems (Hauser & Ward 1998). These characteristics of eddies might be responsible for existence of separate populations within species. In a general view, oceanic gyres also promote the retention of eggs and juveniles and this has led to the presence of isolated populations (Hauser & Carvalho 2008). Secondary effects of eddies are upwellings which leads to rising nutrient-rich water from the depth of the ocean to the surface and the extraction of shelf waters to mid-channel areas (Lutjeharms 2006).

Further south in the region, the main oceanographic feature is the AC. The AC is one of the strongest western boundary currents in the world, characterised as deep, fast-flowing and narrow (Smith & Heemstra 2003). It flows very close to the shelf edge in the north compared to the south (Fig. 1.6A). It has a surface speed of 1.5 m/s at the inshore edge and this periodically increases. Salinity levels range between 35.00 and 35.50 mL/L and temperatures during summer might be higher than 28 ºC (Lutjeharms 2006). The AC is likely formed from the combined action of other currents, which are suggested as its suppliers: MC, EMC and recirculation within a gyre in the southwest of the Indian Ocean (Quartly & Srokosz 2004) (Fig. 1.6). The MC originates from the NEMC which is the northern branch of SEC following its split when reaching eastern Madagascar (Swallow *et al.* 1988). The MC flows southward in the Mozambique Channel and joins up with the southern branch of SEC, the SEMC, that reaches the east coast of the African continent. This joins the AC, flowing southwards. The oceanographic forces of the region likely influence connectivity of marine populations. These impacts combined with unique biological features of species can aid in understanding population genetic patterns in marine species of the region (White *et al.* 2010).

1.4.2 Population genetics studies conducted in the SWIO region

Relatively few population genetics studies have been conducted on species of the region. However, available knowledge of co-distributed species from the SWIO can provide useful insights. Both high levels of genetic connectivity, as well as restrictions to gene flow have been reported for some of the SWIO species.

The panmictic nature of populations in the region has been shown by several studies which used different molecular tools. An allozyme study by Ridgway *et al.* (2001) showed high levels of connectivity inside the central and southern reef systems of northern KwaZulu-Natal for the coral species *Pocillopora verrucosa*. They found that this species functions as a single, genetically homogenous population due to long-distance dispersal that overcomes localized recruitment (Ridgway *et al.* 2001). Also, the effect of the main circulation feature in

the region, the AC, contributes to the observed pattern because, at times, this current flows close inshore as a result of the narrow continental shelf (Ridgway *et al.* 2001). Hence, they concluded that the absence of any significant differentiation was not unexpected. However, a more recent study by Ridgway *et al.* (2008) found a conflicting result based on microsatellite data compared to their allozyme study (see above). They found restriction of gene flow between South Africa and southern Mozambique with the use of more variable markers (Ridgway *et al.* 2008), illustrating the importance of marker choice and resolution in determining levels of connectivity.

Mitochondrial DNA variation showed a panmictic population of the mudcrab (*Scylla serrata*) in the SWIO (Fratini *et al*. 2010). These authors suggested that evidence of this homogenised population is related to the biology of the species as it is highly mobile, has a long planktonic larval stage (transported by ocean currents) and spawning occurs in the open waters off the continental shelf (Fratini *et al.* 2010). Their results further showed that there was no equilibrium between gene flow and genetic drift and that panmixia might be due to recent demographic or range extensions, with high levels of connectivity between neighbouring sites. A study where nuclear and mitochondrial markers were used, including samples from the Mozambique Channel and the eastern coast of Africa, showed high levels of genetic connectivity in a reef fish, the common bluestripe snapper (*Lutjanus kasmira*) (Muths *et al.* 2012). These authors explained the observed absence of genetic differentiation in this species based on its ecological and reproductive traits. The larvae have swimming potential which could be faster than average current speed, enhancing the dispersal capacity of the species (Muths *et al.* 2012). It was also mentioned that this species spawns throughout the year along the east coast of Africa and has a large effective population size (Muths *et al.* 2012). Another allozyme study, on the giant tiger prawn (*Penaeus monodon*) from South Africa, Mozambique and northwest Madagascar, also showed panmixia with high gene flow among all sampling localities (Forbes *et al.* 1999).

A few cases have demonstrated that the circulation mechanisms in the SWIO, especially in the Mozambique Channel, are involved in the disruption of connectivity. One study investigated mtDNA variation in the green turtle (*Chelonia mydas*) and found two populations in the north and south of the Mozambique Channel, respectively (Bourjea *et al.* 2007). These authors also showed that two separate genetic stocks, without ongoing gene flow, exist in the south Mozambique Channel (from the islands Europa and Juan de Nova). They proposed that the pattern of population structuring might be due to recent colonization from the Atlantic Ocean; future work is needed to clarify these results. The oceanography of the

Mozambique Channel may play a role by influencing the early stages of this species as they spend their first years in oceanic waters (Bourjea *et al.* 2007).

Muths *et al.* (2011) considered high connectivity among Pacific populations as more general feature of many marine fish, whereas this specifically related to the squirrel fish (*Myripristis berndti*) as per the study by Craig *et al.* (2007). This study found restricted population connectivity in the region based on microsatellites and mitochondrial *Cyt b* data. They illustrated limited gene flow between SWIO localities in the Mozambique Channel; Europa and northern localities being genetically distinct populations, and Juan de Nova showing intermediate mixing from either Europa or northern populations (Muths *et al.* 2011). The oceanography of the Mozambique Channel, the gyre action generated by eddy-eddy interaction and strong dynamic gradients were the proposed explanation for the existing independent gene pools (Muths *et al.* 2011). They also indicated small but significant differentiation among northern populations (Muths *et al.* 2011).

A study on the mangrove crab (*Neosarmatium meinerti*) from Kenya, Tanzania, Mozambique, South Africa (Durban) and Seychelles (Mahe Island) described the pattern of genetic variation in the *COI* gene (Ragionieri *et al.* 2010). They confirmed the presence of two evolutionary distinct groups; the East African mainland metapopulation and the Western Indian Ocean island group, which are believed to have appeared during the early Pleistocene by a fragmentation event (Ragionieri *et al.* 2010). A similar study by Visram *et al.* (2010) showed significant genetic structuring between two groups (Kenya and Seychelles versus Mauritius and Tanzania) for the blue-barred parrotfish (*Scarus ghobban*) based on the mitochondrial CR. In the study area, they illustrated asymmetrical gene flow with movement favoured from Mauritius to Tanzania, and Tanzania to Kenya (45.2 migrants per generation; more than twice that than from Kenya to Tanzania): These patterns are associated with the role of the SEC and the EACC, respectively (Visram *et al.* 2010). Ragionieri *et al.* (2010) indicated that gene flow was restricted between the African mainland and Seychelles and the reason for this observation was the absence of stepping stone populations. Genetic homogeneity was found among the East African populations from northern Kenya to South Africa, where stepping stone populations of mangroves are variably distributed (Ragionieri *et al.* 2010). However, the temporal difference in this species' spawning period between the subtropical and tropical regions was thought to be a likely factor for population differentiation.

Marine fish species inhabiting the SWIO can be good candidates to investigate the effects of the above-mentioned features in terms of population structuring, which can also contribute to

population genetics knowledge in the region. The region also has major gaps in terms of knowledge of species distributions (Obura 2012). A sciaenid, *Otolithes ruber,* is a good candidate since there is no genetic assessment conducted on it to date; it can potentially add valuable information due to its life history traits (different life stages differentially influenced by oceanographic features) and the distribution of its preferred habitat (i.e. turbid water).

1.5 The Sciaenidae family

The family Sciaenidae, also called drums, croakers or kobs, commonly inhabits warm coastal waters and estuaries worldwide (Smith & Heemstra 2003). They are important components of commercial and recreational fisheries (Phelan *et al.* 2008). This family includes 70 genera and about 270 species (Xu *et al.* 2011).

Sciaenids are benthic or demersal fishes, meaning they live on or near the bottom of the tropical and subtropical regions of the Indian, Atlantic and Pacific oceans (Fennessy 2000; Phelan *et al.* 2008). Sciaenids are zoobenthos and their diet mostly consists of small penaeid prawns, brachyurans and small fish (Pillai 1983; Pauly & Neale 1985; Fennessy 1995). Estuaries and inshore waters are used for nurseries (Fennessy 2000), providing food, turbidity and shelter for the species' survival at early life stages. The Tugela Bank off central KwaZulu-Natal, South Africa, is suggested for this purpose for some sciaenids such as squaretail kob (*Argyrosomus thorpei*), snapper kob (*Otolithes ruber*), blackmouth croaker (*Atrobucca nibe*) and *Johnius fuscolineatus,* based on length frequencies, sizes at maturity and feeding (Fennessy 2000). In terms of movement, recapture studies for red drum (*Sciaenops ocellatus*), silver kob (*Argyrosomus inodorus*) and spotted seatrout (*Cynoscion nebulosus*) revealed movements greater than 100 km, 240 km and over 500 km, respectively (Griffiths 1997). Also, seasonal migrations of striped weakfish (*Cynoscion guatucupa*) and blackspotted croaker (*Protonibea diacanthus*) adults were reported (De Miranda & Haimovici 2007; Phelan *et al.* 2008). Another sciaenid, the king fish (*Macrodon ancylodon*) has limited dispersal ability (Santos *et al.* 2006). Most members are generally characterized as slow-growing and long-lived individuals (Brash & Fennessy 2005) that consequently make them more susceptible to overexploitation (Sumaila *et al.* 2010).

1.5.1 Otolithes ruber

Otolithes ruber is one of the species recognized in the family and the model organism for this study. It is an elongated kob which has bronze-red colour above and silvery to white below, with grey or yellowish fins (van der Elst 1981) (Fig. 1.7). This species is widely distributed in the Indo-West Pacific (IWP), including the SWIO (Brash & Fennessy 2005), although there is contradictory information as to whether the species occurs off Australia. This species is an important demersal species of the SWIO region. It is known as snapper kob or long tooth kob in South Africa, and tiger-toothed croaker elsewhere (van der Elst 1981). It is also called the Malindi herring in Kenya, corvina in Mozambique and telonify au trident in Madagascar (Brash & Fennessy 2005).

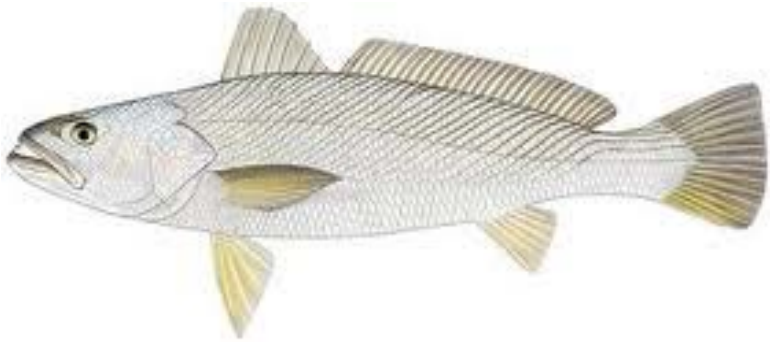


Fig. 1.7 Representative illustration of *Otolithes ruber* (http://www.efishalbum.com).

Otolithes ruber is a shoaling fish of coastal waters, large estuarine systems and sheltered bays (van der Elst 1981). It is characterised as a sluggish carnivore, inhabiting turbid water (van der Elst 1981; Brash & Fennessy 2005), with adults reported to reach 70-80 cm in total length (van der Elst 1981; Smith & Heemstra 2003). 50%-sexual maturity is attained at 22 – 24 cm and spring and summer are spawning seasons (Fennessy 2000; Brash & Fennessy 2005). There are very few studies on the biology of this fish and its migratory and breeding behaviour is not known.

In many fisheries across its distribution range, *O. ruber* is an excellent protein source (Heemstra & Heemstra 2004). It is caught by trawlers, gill-nets, beach seines, line and hook fishing in the SWIO countries, including South Africa, Mozambique, Tanzania and Kenya (Brash & Fennessy 2005). In addition, it is incidentally caught as by-catch on penaeid prawn trawlers (Fennessy *et al.* 2008); large-sized individuals are used for sale, whereas smaller

ones are discarded at sea due to their non-commercial value and during this process many individuals are killed (Schultz 1992; Olbers & Fennessy 2007). This is important, as these smaller sexually immature individuals will then not have had the opportunity to spawn. The rising concern for this species is overfishing in the SWIO region. Two previous studies showed that *O. ruber* is overexploited in South Africa and in Mozambique (Olbers & Fennessy 2007). Overfishing can give rise to changes, such as reduced genetic diversity and enhanced inbreeding, and productivity of the species will be negatively affected in the long term (Berntson & Moran 2009; Belgrano & Fowler 2013 and references therein).

The aim of the present study was to investigate the historical, evolutionary and contemporary processes influencing the genetic structure of *O. ruber* throughout the SWIO region based on the data of mitochondrial and nuclear markers. From the obtained results, management plans should be illuminated in terms of the existence of a single or multiple populations that require separate management interventions. Also the gathered information from this study will further contribute to the knowledge of *O. ruber* in this region.

1.6 Research questions

The main question addressed in this study was: What is the geographic population genetic structure within *O. ruber*?

Two components of the main question were: (i) is the species panmictic or highly structured?, and (ii) what are the underlying processes associated with the observed pattern of differentiation?

1.7 Alternative potential scenarios

Two scenarios were envisaged:

- High level of connectivity: A homogenised, panmictic population structure throughout the sampling locations due to possible passive larval dispersal which is enhanced by current systems and/or active adult migration.

- Limited connectivity: Structured populations exist in the south (South Africa, Mozambique) and north (Kenya, Tanzania) of the SWIO range, based on the influence of the South Equatorial Current on dispersal and/or the requirement of this species in turbid habitats. In this scenario, populations from Madagascar could fall into either the northern or southern sub-population.

1.8 Dissertation outline

All the chapters were written following the referencing style of the journal *Molecular Ecology*.

Chapter 2: Selection of molecular markers for the phylogeographic study of the snapper kob (*Otolithes ruber*). This is the technical chapter explaining all the laboratory work for optimization of molecular markers and the steps involved in selection of the most informative markers from small pilot studies.

Chapter 3: Historical isolation and contemporary isolation-with-migration in the South West Indian Ocean sciaenid, *Otolithes ruber*. This is the main research chapter where the formulated questions have been addressed. This chapter will be submitted to *Molecular Ecology* with co-authors Sean Fennessy and Paulette Bloomer.

Chapter 4: Dissertation conclusion. This is a short synthesis chapter bringing the main results and discussions of the chapters together. It also deals with the future aspects, recommendations for management and projects that should be conducted in the near future.

1.9 References

Adler F, Nuernberger B (1994) Persistence in patchy irregular landscapes. *Theoretical Population Biology,* **45**, 41-75.

Ayre DJ, Minchinton TE, Perrin C (2009) Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Molecular Ecology,* **18**, 1887-1903.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology,* **13**, 729-744.

Barber BR, Xu J, Pérez-Losada M, Jara CG, Crandall KA (2012) Conflicting evolutionary patterns due to mitochondrial introgression and multilocus phylogeography of the Patagonian freshwater crab *Aegla neuquensis*. *PLoS ONE,* **7**, 1-22.

Barkley NA, Krueger RR, Federici CT, Roose ML (2009) What phylogeny and gene genealogy analyses reveal about homoplasy in citrus microsatellite alleles. *Plant Systematics and Evolution,* **282**, 71-86.

Bartlett SE, Davidson WS (1991) Identification of *Thunnus* tuna species by the polymerase chain reaction and direct sequence analysis of their mitochondrial cytochrome *b* genes. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**, 309-317.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Evolution,* **98**, 4563–4568.

Beheregaray LB (2008) Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Molecular Ecology,* **17**, 3754–3774.

Belgrano A, Fowler CW (2013) How fisheries affect evolution. *Science*, **342**, 1176.

Benny PN (2002) Variability of western Indian Ocean currents. *West Indian Ocean Journal of Marine Science,* **1**, 81-90.

Berntson EA, Moran P (2009) The utility and limitations of genetic data for stock identification and management of North Pacific rockfish (*Sebastes spp*.). *Reviews in Fish Biology and Fisheries,* **19**, 233–247.

Bhargava A, Fuentes FF (2010) Mutational dynamics of microsatellites. *Molecular Biotechnology*, **44**, 250-266.

Boissin E, Hoareau TB, Berrebi P (2011) Effects of current and historical habitat fragmentation on the genetic structure of the sand goby *Pomatoschistus minutes* (Osteichthys, Gobiidae). *Biological Journal of the Linnean Society*, **102**, 175-198.

Bourjea J, Lapègue S, Gagnevin L, Broderick D, Mortimer JA, Ciccione S, Roos D, Taquet C, Grizel H (2007) Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. *Molecular Ecology*, **16**, 175-186.

Bradman H, Grewe P, Appleton B (2011) Direct comparison of mitochondrial markers for the analysis of swordfish population structure. *Fisheries Research,* **109**, 95-99.

Brash JM, Fennessy ST (2005) A preliminary investigation of age and growth of *Otolithes ruber* from KwaZulu-Natal, South Africa. *Western Indian Ocean Journal of Marine Science,* **4**, 21-28.

Carpenter KE, Barber PH, Crandall ED, Ablan-Lagman Ma. CA, Ambariyanto, NgurahMahardika G, Manjaji-Matsumoto BM, Juinio-Meñez MA, Santos MD, Starger CJ, Abdul Hamid A. Toha AHA (2011) Comparative phylogeography of the Coral Triangle and implications for marine management. *Journal of Marine Biology* 1-14.

Carreras-Carbonell J, Macpherson E, Pascual M (2008) Utility of pairwise mtDNA genetic distance for predicting cross-species microsatellite amplification and polymorphism success in fishes. *Conservation Genetics,* **9**, 181-190.

Cheng Y, Wang R, Xu T (2011) The mitochondrial genome of the spinyhead croaker *Collichthys lucida*: Genome organization and phylogenetic consideration. *Marine Genomics,* **4**, 17-23.

Craig MT, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series*, **334**, 245-254.

Cui Z, Liu Y, Li CP, You F, Chu KH (2009) The complete mitochondrial genome of the large yellow croaker, *Larimichthys crocea* (Perciformes, Sciaenidae): Unusual features of its control region and the phylogenetic position of the Sciaenidae. *Gene,* **432**, 33-43.

De Miranda LV, Haimovici M (2007) Changes in the population structure, growth and mortality of striped weakfish *Cynoscion guatucupa* (Sciaenidae, Teleostei) of southern Brazil between 1976 and 2002. *Hydrobiologia,* **589**, 69-78.

Dudgeon CL, Blower DC, Broderick D, Giles JL, Holmes BJ, Kashiwagi T, Krück NC, Morgan JAT, Tillett BJ, Ovenden JR (2012) A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology,* **80**, 1789–1843.

Durgadoo JV, Ansorge IJ, Lutjeharms JRE (2010) Oceanographic observations of eddies impacting the Prince Edward Islands, South Africa. *Antarctic Science,* **22**, 211-219.

Eytan RI, Hellberg ME (2010) Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution*, **64**, 3380-3397.

Fahrig L, Paloheimo J (1988) Effect of Spatial Arrangement of Habitat Patches on Local Population Size. *Ecology*, **69**, 468-75.

Fennessy ST (1995) Abundances of non-commercial crustaceans caught by Tugela Bank prawn trawlers. *Lammergeyer,* **43**, 1-5.

Fennessy ST (2000) Aspects of the biology of four species of Sciaenidae from the east coast of South Africa. *Estuarine, Coastal and Shelf Science,* **50**, 259-269.

Fennessy ST, Vincent X, Budeba Y, Mueni EM, Gove DZ (2008) An update on initiatives to reduce prawn trawl bycatch in the Western Indian Ocean. *Western Indian Ocean Journal of Marine Science*, **7**, 217-222.

Forbes AT, Demetriades NT, Benzie JAH, Ballment E (1999) Allozyme frequencies indicate little geographic variation among stocks of giant tiger prawn *Penaeus monodon* in the south-west Indian Ocean. *South African Journal of Marine Science*, **21**, 271-277.

Fratini S, Ragionieri L, Cannicci S (2010) Stock structure and demographic history of the Indo-West Pacific mud crab *Scylla serrata*. *Estuarine, Coastal and Shelf Science,* **86**, 51-61.

Gold JR, Richardson LR, Furman C, King TL (1993) Mitochondrial DNA differentiation and population structure in red drum (*Sciaenops ocellatus*) from the Gulf of Mexico and Atlantic Ocean. *Marine Biology,* **116**, 175-185.

Gonzalez EG, Zardoya R (2007) Relative role of life history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *BMC Evolutionary Biology,* **7**, 197-209.

Griffiths MH (1997) The life history and stock separation of silver kob, *Argyrosomus inodorus*, in South African waters. *Fishery Bulletin,* **95**, 47-67.

Hansen MM, Kenchington E, Nielsen EE (2001) Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries,* **2**, 93-112.

Hanski I (1985) Single-species spatial dynamics may contribute to long-term rarity and commonness. *Ecology*, **66**, 335-43.

Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, Ruegg K, Palstra F (2011) Understanding and estimating effective population size for practical application in marine species management. *Conservation Biology,* **25**, 438–449.

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. *Fish and Fisheries,* **9**, 333-362.

Hauser L, Ward RD (1998) Population identification in pelagic fish: The limits of molecular markers. In: *Advances in molecular ecology* (ed. Carvalho GR), pp. 191-224. IOS Press, Amsterdam.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identification through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321.

Hedgecock D, Pudovkin AI (2011) Sweepstakes reproductive success in highly fecund Marine Fish and Shellfish : A review and commentary. *Bulletin of Marine Science*, **87**, 971–1002.

Heemstra P, Heemstra E (2004) *Coastal fishes of Southern Africa* NISC. SAIAB.

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. In: *Annual Review of Ecology, Evolution, and Systematics*, **40**, 291-310.

Hey J, Machado CA (2003) The study of structured populations - New hope for a difficult and divided science. *Nature Reviews Genetics,* **4**, 535-543.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics,* **167**, 747-760.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *PNAS,* **104**, 2785-2790.

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetic and Evolution,* **54**, 291-301.

Hoareau TB, Boissin E, Berrebi, P (2012) Evolutionary history of a widespread Indo-Pacific goby: The role of Pleistocene sea-level changes on demographic contraction/expansion dynamics. *Molecular Phylogenetics and Evolution*, **62**, 566-572.

Karl S, Castro A, Lopez J, Charvet P, Burgess G (2011) Phylogeography and conservation of the bull shark (*Carcharhinus leucas*) inferred from mitochondrial and microsatellite DNA. *Conservation Genetics,* **12,** 371–382.

Kasapidis P, Magoulas A (2008) Development and application of microsatellite markers to address the population structure of the horse mackerel *Trachurus trachurus*. *Fisheries Research,* **89**, 132-135.

Kawahara R, Miya M, Mabuchi K, Lavoué S, Inoue JG, Satoh TP, Kawaguchi A, Nishida M (2008) Interrelationships of the 11 gasterosteiform families (sticklebacks, pipefishes, and their relatives): A new perspective based on whole mitogenome sequences from 75 higher teleosts. *Molecular Phylogenetics and Evolution,* **46**, 224-236.

Keeney DB, Heupel MR, Hueter RE, Heist EJ (2005) Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea. *Molecular Ecology,* **14**, 1911-1923.

Kotlik P, Markova S, Choleva L, Bogutskaya NG, Ekmekci FG, Ivanova PP (2008) Divergence with gene flow between Ponto-Caspian refugia in an anadromous cyprinid *Rutilus frisii* revealed by multiple gene phylogeography. *Molecular Ecology*, **17**, 1076-1088.

Lessios HA (2008) The Great American schism: Divergence of marine organisms after the rise of the central American Isthmus. *Annual Reviews in Ecologoy Evolution and Systematics*, **39**, 63–91.

Lewis DD, Fontoura NF (2005) Maturity and growth of *Paralonchurus brasiliensis* females in southern Brazil (Teleostei, Perciformes, Sciaenidae). *Journal of Applied Ichthyology,* **21**, 94-100.

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

Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Molecular Ecology,* **19**, 3038-3051.

Lutjeharms RE (2006) The Coastal Oceans of South-Eastern Africa (15,W). In: *The Sea*, volume 14B (eds Robinson AR, Brink KH), pp. 783-834. Harvard University Press, Cambridge, MA.

Lutjeharms JRE, Biastoch A, Van der Werf PM, Ridderinkhof H, De Ruijter WPM (2012) On the discontinuous nature of the Mozambique Current. *South African Journal of Science,* **108**, 1-5.

Marko PB, Hart MW (2011) The complex analytical landscape of gene flow inference. *Trends in Ecology and Evolution,* **26**, 448-456.

Marriage TN, Hudman S, Mort ME, Orive ME, Shaw RG, Kelly JK (2009) Direct estimation of the mutation rate at dinucleotide microsatellite loci in *Arabidopsis thaliana* (Brassicaceae). *Heredity*, **103**, 310-317.

Martínez P, Gonzalez EG, Castilho R, Zardoya R (2006) Genetic diversity and historical demography of Atlantic bigeye tuna (*Thunnus obesus*). *Molecular Phylogenetics and Evolution,* **39**, 404-416.

Matano RP, Beier EJ, Strub PT (2008) The seasonal variability of the circulation in the South Indian Ocean: Model and observations. *Journal of Marine Systems,* **74**, 315-328.

McCusker MR, Bentzen P (2010) Historical influences dominate the population genetic structure of a sedentary marine fish, Atlantic wolfish (*Anarhichas lupus*), across the North Atlantic Ocean. *Molecular Ecology*, **19**, 4228–4241.

Muths D, Gouws G, Mwale M, Tessier E, Bourjea J (2012) Genetic connectivity of the reef fish *Lutjanus kasmira* at the scale of the western Indian Ocean. *Canadian Journal of Fisheries and Aquatic Sciences,* **69**, 1-12.

Muths D, Tessier E, Gouws G, Craig M, Mwale M, Mwaluma J, Mwandya A, Bourjea J (2011) Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. *Marine Ecology Progress Series,* **443**, 167-180.

Obura D (2012) Evolutionary mechanisms and diversity in a western Indian Ocean centre of diversity. Proceedings of the 12th International Coral Reef Symposium, Cairns, Australia.

O'Connell M, Wright JM (1997) Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries,* **7**, 331-363.

Okumuş I, Çiftci Y (2003) Fish population genetics and molecular markers: II- molecular markers and their applications in fisheries and aquaculture. *Turkish Journal of Fisheries and Aquatic Sciences,* **3**, 51-79.

Olbers JM, Fennessy ST (2007) A retrospective assessment of the stock status of *Otolithes ruber* (Pisces: Sciaenidae) as bycatch on prawn trawlers from KwaZulu-Natal, South Africa. *African Journal of Marine Science,* **29**, 247-252.

Ovenden JR, Kashiwagi T, Broderick D, Giles J, Salini J (2009) The extent of population genetic subdivision differs among four co-distributed shark species in the Indo-Australian archipelago. *BMC Evolutionary Biology,* **9**.

Ovenden JR, Morgan J, Street R, Tobin A, Simpfendorfer CA, Macbeth W, Welch D (2011) Negligible evidence for regional genetic population structure for two shark species (*Rhizoprionodon acutus*, Rüppell, 1837 and *Sphyrna lewini*, Griffith & Smith, 1834) with contrasting biology. *Marine Biology,* **158,** 1497–1509.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: A simulation-based exploration of accuracy and power. *Molecular Ecology*, **13**, 55–65.

Palumbi SR (1994) Genetic divergence, reproductive isolation and marine speciation. *Annual Reviews in Ecology and Systematics*, **25**, 547–572.

Papetti C, Pujolar JS, Mezzavilla M, La Mesa M, Rock J, Zane L, Patarnello T (2012) Population genetic structure and gene flow patterns between populations of the Antartic icefish *Chionodraco rastrospinosus*. *Journal of Biogeography*, **39**, 1361-1372.

Paquette SR, Lapointe FJ (2009) A statistical procedure to assess the significance level of barriers to gene flow. *Journal of Genetics and Genomics,* **36**, 685-693.

Pauly D, Neal R (1985) Shrimp vs. fish in Southeast Asian fisheries: The biological, technological and social problems. In: *Recursos Pesqueros Potenciales de Mexico: La Pesca Acompanante del Camaron,* pp. 487-510. Progr. Univ. de Alimentos, Inst. Cienc. del Mary Limnol., Inst. Nal. de Pesca. UNAM, Mexico DF.

Phelan MJ, Gribble NA, Garrett RN (2008) Fishery biology and management of *Protonibea diacanthus* (Sciaenidae) aggregations in far Northern Cape York Peninsula waters. *Continental Shelf Research,* **28**, 2143-2151.

Pillai PKM (1983) On the biometry, food and feeding and spawning habits of *Otolithes ruber* (Schneider) from Porto Novo. *Indian Journal of Fisheries,* **30**, 69-73.

Portnoy DS, McDowell JR, Heist EJ, Musick JA, Graves JE (2010) World phylogeography and male-mediated gene flow in the sandbar shark, *Carcharhinus plumbeus*. *Molecular Ecology,* **19**, 1994–2010.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, **23**, 564-571.

Quartly GD, Srokosz MA (2004) Eddies in the southern Mozambique Channel. *Deep-Sea Research Part II: Topical Studies in Oceanography,* **51**, 69-83.

Ragionieri L, Cannicci S, Schubart CD, Fratini S (2010) Gene flow and demographic history of the mangrove crab *Neosarmatium meinerti*: A case study from the western Indian Ocean. *Estuarine, Coastal and Shelf Science,* **86**, 179-188.

Reid K, Hoareau TB, Bloomer P (2012) High-throughput microsatellite marker development in two sparid species and verification of their transferability in the family Sparidae. *Molecular Ecology Resources,* **12**, 740-752.

Ridgway T, Hoegh-Guldberg O, Ayre D (2001) Panmixia in *Pocillopora verrucosa* from South Africa. *Marine Biology,* **139**, 175-181.

Ridgway T, Riginos C, Davis J, Hoegh-Guldberg O (2008) Genetic connectivity patterns of *Pocillopora verrucosa* in southern African Marine Protected Areas. *Marine Ecology Progress Series,* **354**, 161-168.

Rodrigues R, Santos S, Haimovici M, Saint-Paul U, Sampaio I, Schneider H (2013) Mitochondrial DNA reveals population structuring in *Macrodon atricauda* (Perciformes: Sciaenidae): A study covering the whole geographic distribution of the species in the southwestern Atlantic. *Mitochondrial DNA*, Early Online: 1–7.

Sanciangco JC, Carpenter KE, Etnoyer PJ, Moretzsohn F (2013) Habitat availability and heterogeneity and the Indo-Pacific warm pool as predictors of marine species richness in the tropical Indo-Pacific. *PLoS ONE*, **8**, e56245.

Santos S, Hrbek T, Farias IP, Schneider H, Sampaio I (2006) Population genetic structuring of the king weakfish, *Macrodon ancylodon* (Sciaenidae), in Atlantic coastal waters of South America: Deep genetic divergence without morphological change. *Molecular Ecology,* **15**, 4361-4373.

Schmidt JV, Schmidt CL, Ozer F*,* Ernst RE, Feldheim KA, Ashley MW, Levine M (2009) Low genetic differentiation across three major ocean populations of the whale shark, *Rhincodon typus*. *PLoS ONE,* **4**.

Schouten MW, De Ruijter WPM, Van Leeuwen PJ, Ridderinkhof H (2003) Eddies and variability in the Mozambique Channel. *Deep-Sea Research Part II,* **50**, 1987-2003.

Schluessel V, Broderick D, Collin SP, Ovenden JR (2010) Evidence for extensive population structure in the white-spotted eagle ray within the Indo-Pacific inferred from mitochondrial gene sequences. *Journal of Zoology,* **281**, 46-55.

Schultz N (1992) Preliminary investigations on the population dynamics of Otolithes ruber (Sciaenidae) on Sofala Bank, Mozambique. *Revista de Investigacao Psequeria,* **21**, 41-49.

Schultz JK, Feldheim KA, Gruber Ashley MV, McGovern TM, Bowen BW (2008) Global phylogeography and seascape genetics of the lemon sharks (genus *Negaprion*). Molecular Ecology, **17**, 5336–5348.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters,* **9**, 615-629.

Smith MM, Heemstra PC (2003) *Smiths' Sea Fishes* Struik Publishers.

Spies IB, Gaichas S, Stevenson DE, Orr JW, Canino MF (2006) DNA-based identification of Alaska skates (Amblyraja, Bathyraja and Raja: Rajidae) using cytochrome c oxidase subunit I (*COI*) variation. *Journal of Fish Biology,* **69**, 283-292.

Sumaila UR, Khan A, Teh L, Watson R, Tyedmers P, Pauly D (2010) Subsidies to high seas bottom trawl fleets and the sustainability of deep-sea demersal fish stocks. *Marine Policy,* **34**, 495-497.

Swallow J, Fieux M, Schott F (1988) The boundary currents east and north of Madagascar 1. Geostrophic currents and transports. *Journal of Geophysical Research,* **93**, 4951–4962.

Thacker CE (2003) Molecular phylogeny of the gobioid fishes (Teleostei: Perciformes: Gobioidei). *Molecular Phylogenetics and Evolution,* **26**, 354-368.

Tillett BJ, Meekan MG, Broderick D, Field IC, Cliff G, Ovenden J (2012) Pleistocene isolation, secondary introgression and restricted contemporary gene flow in the pigeye shark, *Carcharhinus amboinensis* across northern Australia. *Conservation Genetics,* **13**, 99–115.

Unmack PJ, Hammer MP, Adams M, Dowling TE (2011) A phylogenetic analysis of pygmy perches (Teleostei: Percichthyidae) with an assessment of the major historical influences on aquatic biogeography in Southern Australia. *Systematic Biology,* **60**, 797-812.

van der Elst R (1981) *A guide to the common sea fishes of southern Africa*, 1st edn. C. Struik Publishers, Cape Town.

van Oppen MJH, Rico C, Turner GF, Hewitt GM (2000) Extensive homoplasy, nonstepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi cichlids. *Molecular Biology and Evolution,* **17**, 489-498.

Viñas J, Gordoa A, Fernández-Cebrián R, Pla C, Vahdet Ü, Araguas RM (2010) Facts and uncertainties about the genetic population structure of Atlantic bluefin tuna (*Thunnus thynnus*) in the Mediterranean: Implications for fishery management. *Reviews in Fish Biology and Fisheries*, **21**, 1-15.

Visram S, Yang MC, Pillay RM, Said S, Henriksson O, Grahn M, Chen CA (2010) Genetic connectivity and historical demography of the blue barred parrotfish (*Scarus ghobban*) in the western Indian Ocean. *Marine Biology,* **157**, 1475-1487.

Wafar M, Venkataraman K, Ingole B, Khan SA, LokaBharathi P (2011) State of knowledge of coastal and marine biodiversity of Indian Ocean countries. *PloS ONE,* **6**, e14613.

Walther E, Schöfl G, Mrotzek G, Haryanti, Sugama K, Saluz HP (2011) Paralogous mitochondrial control region in the giant tiger shrimp, *Penaeus monodon* (F.) affects population genetics inference: A cautionary tale. *Molecular Phylogenetics and Evolution,* **58**, 404-408.

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. *Molecular Ecology,* **15**, 1419-1439.

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences,* **277**, 1685-1694.

White TA, Stefanni S, Stamford J, Hoelzel AR (2009) Unexpected panmixia in a long-lived, deep-sea fish with well-defined spawning habitat and relatively low fecundity. *Molecular Ecology,* **18**, 2563-2573.

Wright S (1943) Isolation by Distance. *Genetics,* **28**, 114-138.

Xu Tj, Cheng Yz, Sun Yn, Shi G, Wang Rx (2011) The complete mitochondrial genome of bighead croaker, *Collichthys niveatus* (Perciformes, Sciaenidae): Structure of control region and phylogenetic considerations. *Molecular Biology Reports*, **38**, 4673-4685.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. *Molecular Ecology*, **12**, 563–584.

Chapter 2

Selection of molecular markers for the phylogeographic study of the snapper kob (*Otolithes ruber*)

Abstract

The assessment of genetic variation, which is detectable by molecular markers at a given locus, will be helpful to facilitate targeted management and conservation plans with respect to changing environmental conditions. Molecular tools have successfully been used to address a large range of biological questions at the individual, family, population and species levels. Cautious consideration is important when choosing appropriate genetic markers; this requires a preliminary analysis for accurate estimates of population parameters before large scale data generation. The aim of this chapter was to select the most informative markers based on a pilot study for the proposed phylogeographic study of an overexploited, demersal, marine fish *Otolithes ruber*. I tested three protein coding mitochondrial genes (*COI*, *ND2* and *Cyt b*) and 38 microsatellites, developed for different members of the family Sciaenidae. Cytochrome *b* was the most informative based on the observed sequence variation within and among the samples from the selected sampling localities. The degree of microsatellite amplification and the maintenance of polymorphism were compared between *O. ruber* and different source species. In total 22 of the 38 markers were successfully amplified, and only seven of them were polymorphic, with no marker-associated problems. Thus *Cyt b* and the seven microsatellites were used to generate data to answer questions from the perspective of both historical and ecological time scales.

Keywords: Molecular markers, mitochondrial DNA, microsatellites, cross-species amplification, marker optimization

2.1 Introduction

Investigating the genetic variation and relationships at the individual level through to genetic structuring of populations and species distributions, evolutionary relationships can be inferred by comparing neutral variation within selected regions of the genome (Buschiazzo & Gemmell 2006; Dudgeon *et al.* 2012). Hence, the level of this variation is an important measure for conservation and management programs to maintain natural genetic diversity (Lakra *et al.* 2009). This variation can be studied by a genetic marker that is required to be heritable and recognizable through molecular approaches (Liu & Cordes 2004). Various molecular markers have been used for questions of interest in different disciplines; in the case of population genetics and phylogeography, microsatellites (frequency markers) and mitochondrial DNA sequences (sequence markers) are the most commonly applied tools due to their specific features (Hellberg 2009).

Mitochondrial DNA is present in multiple copies in most vertebrate cells (Ballard & Whitlock 2004), thus allowing easy amplification from a small amount of DNA material. It is a maternally-inherited, haploid genome with rare or no recombination, mainly presenting neutral evolution and there are no introns (Okumuş & Çiftci 2003). Furthermore, analysis of genetic variation across selected mtDNA regions (all representing a single linked locus) are relatively easy (Walther *et al.* 2011). The effective population size is a quarter that of autosomal nuclear DNA (Keeney *et al.* 2005), leading to mtDNA being more sensitive at tracing population features such as bottlenecks (Okumuş & Çiftci 2003). The most studied mtDNA genes in phylogenetics and phylogeography are cytochrome *b* (*Cyt b*), NADH dehydrogenase (*ND2*) and cytochrome *c* oxidase subunit I (*COI*). Because these genes encode for proteins, insertions and deletions (indels) are rare and the nucleotide sequences are relatively conserved. These genes are ideal for sequence-based analyses, since the obtained sequence data are easy to align and alignments can be validated against a reading frame (Baharum & Nurdalila 2012). These genes have been recognized as good evolutionary markers to determine relationships within and among species of marine fish (Spies *et al.* 2006; Ward *et al.* 2009; Gaither *et al.* 2010; Bradmana *et al.* 2011), and determination of stock structure of many fishes, e.g. red drum, *Sciaenops ocellatus* (a sciaenid) (Gold *et al.* 1993). Specifically, *COI* (Hebert *et al.* 2003) and *Cyt b* (Borsa *et al.* 2007) have been used for species identifications, resolving taxonomic uncertainty and allowing the recognition of cryptic species by DNA barcoding (Hubert *et al*. 2012).

Microsatellites are short nuclear DNA sequences with tandemly repeated motifs of 1-6 base pairs and conserved regions flanking the repeat region (Kasapidis & Magoulas 2008). They

are abundant, are distributed throughout most eukaryotic genomes (Schlötterer 2004). They are co-dominant markers and are highly polymorphic due to exhibiting of high rates of mutation (Barkley *et al.* 2009), between 10^{-2} and 10^{-6} mutations per locus per generation (Selkoe & Toonen 2006). Most microsatellite loci are neutral markers, therefore, there is no pressure of natural selection acting on them (Kasapidis & Magoulas 2008). Microsatellites are extensively used in numerous different fish species, defining stock structure and determining how individuals from each population can be assigned to their population origins (O'Connell & Wright 1997). Greater confidence regarding differentiation (none or weak genetic divergence) can be obtained by using several variable microsatellite markers (Dudgeon *et al.* 2012). However, there are some associated disadvantages that should be considered in deciding on the most informative microsatellite markers (see Chapter 1).

Development of microsatellites is highly costly since identifying potential markers requires sequence information. Once the sequence information is available it becomes very useful for other studies working on the same species or on closely related species (Guichoux *et al.* 2011). The latter involves the approach of cross-species amplification, which can be a time and cost effective approach for microsatellite genotyping without the need for *de novo* cloning and sequencing. The conserved flanking regions, which are the sites for primer design, make locus amplification possible in multiple target species (Guichoux *et al.* 2011). It has been successfully performed in vertebrates, invertebrates, fungi and plants (Barbará *et al.* 2007). However, amplification success and polymorphism of loci might differ from one species to the other in a family, depending on the genetic divergence between the source and target species (Carreras-Carbonell *et al.* 2008; Reid *et al.* 2012). The simplicity of cross-species amplification, associated low cost and high reproducibility are features favouring microsatellites as popular markers (Bhargava & Fuentes 2010).

Although mtDNA variation offers an independent view of population and species history, it is not enough to explain current genetic structuring, since marine fish species are typically characterized by large effective population sizes and high potential for dispersal during several life stages (Zhang & Hewitt 2003; Gonzalez & Zardoya 2007). Comparative analyses based on nuclear and mitochondrial data have been applied to characterize population genetic structure and identify the underlying evolutionary processes in marine fishes (Eytan & Hellberg 2010). Microsatellites enable the distinction between high gene flow and highly structured populations with restricted gene flow over ecological time scales (Zhang & Hewitt 2003; Waples & Gaggiotti 2006; Hauser & Carvalho 2008). Hence a framework including both classes of markers gives a relatively real history of the study species based on past and recent timescales.

The model organism of the present study is a demersal marine fish species, the snapper kob (*Otolithes ruber*). This species is under pressure of overfishing due to its popularity as a protein source in many countries throughout its distribution (Indo-west Pacific, including the South Western Indian Ocean region). Thus far, almost no molecular studies have been conducted on this species based on mtDNA; a few broad phylogenetic studies of sciaenids have applied *COI* and *16S* rDNA genes (Lakra *et al.* 2009), and *COI*, *16S* and *Cyt b* genes (Xu *et al.* 2011). There are no studies on this species that have used microsatellites yet. I aimed to identify the most informative molecular markers in order to investigate the phylogeography of this species (see Chapter 3). I tested three mtDNA genes, *COI*, *ND2* and *Cyt b,* and 38 nuclear microsatellite loci that were available in the literature, developed in various relatives. The selection of a mtDNA gene for in-depth analyses was based on haplotype variation, and microsatellite loci were chosen based on allelic diversity for each marker.

2.2 Material and Methods

2.2.1 Sampling, DNA extraction and DNA quality

Fin tissue samples of *O. ruber* were obtained from seven locations throughout the SWIO region: South Africa (40), south Mozambique (33), central Mozambique (30), Tanzania (30), Kenya (30), west Madagascar (30), north-west Madagascar (30) (see Fig. 3.1 in chapter 3). DNA was extracted with the DNeasy Blood and Tissue extraction kit (Qiagen), according to the manufacturer's instructions.

The quality of extracted DNA was confirmed by mtDNA *16S* rDNA amplification. Polymerase Chain Reaction (PCR) was carried out in 10 µL volumes containing 1 x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNPTs (Promega), 0.5 pmol primers (Oligos), 0.04 U *Taq* polymerase (Southern Cross Biotechnologies) and approximately 30 ng genomic DNA. PCR conditions were: 94 °C for 4 min followed by 40 cycles of 94 °C for 30 s, 60 °C for 1 min, 72 °C for 1 min and a final extension step of 72 °C for 7 min. The PCR products were electrophoresed on 2% agarose gels and visualised with GelRed™ stain (Biotium) under UV light applied in Molecular Imager, Gel Doc™ XR+ Imagining System (Bio Rad).

2.2.2 Mitochondrial DNA optimization steps

Three mtDNA genes were tested: *COI*, *ND2* and *Cyt b*.

STEP 1: *PCR amplification and sequencing tests*

The PCR reactions were performed in 25 µL volumes containing 1 x PCR buffer, $MgCl_2$ (refer to Appendix 2A for the concentrations for the amplification of each gene region), 0.2 mM dNTPs (Promega), forward and reverse primers (Applied Biosystems) (see Appendix 2A for the concentrations), 0.04 U Supertherm *Taq* polymerase (Southern Cross Biotechnologies) and approximately 90 ng of genomic DNA. Cycling conditions including denaturation, elongation and final elongation steps for the amplifications of each gene are indicated in Appendix 2A. The PCR products were electrophoresed on 2% agarose gels and visualised with GelRed™ stain. The tests were done using a few samples (2 to 3) for each gene.

Test sequencing was performed for the amplified products of each gene in the same way. The amplified products were precipitated using 2.8 volumes of absolute ethanol, 0.4 volumes of Sabax water and 0.08 volumes of 3 M sodium acetate (NaAc), followed by a washing step using 90 µL 70% ethanol. The products were eluted in 15 µL of Sabax water and confirmed on a 2% agarose gel. Cycle sequencing reactions were prepared with the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (Applied Biosystems). The manufacturer's recommended conditions were used for the cycling reaction and it contained approximately 30 ng purified DNA, 1 µL BigDye reaction mix, 5 x sequencing buffer and 3.2 pmol primer. The products were precipitated as described above. Sequences were separated on an ABI 3130 sequencer (Applied Biosystems). Samples were sequenced in forward and reverse directions. Trace data analyses of sequences were done using CLC Bio®. An online BLAST search (Altschul *et al*. 1997) was performed for each individual sequence of each gene against the available sequences in the database.

STEP 2: *Small scale PCR amplification and sequencing*

For this step *COI* was eliminated due to difficulties in sequencing. *ND2* (~ 785 bp) and *Cyt b* (624 bp) were sequenced for a representative number of samples (15) to examine the genetic diversity within and among sampling localities. Trace data analyses of sequences were done using CLC Bio® and aligned in BioEdit Sequence Alignment Editor (Hall 1999), using ClustalW (Thompson *et al.* 1994). Mega version 5 (Tamura *et al.* 2011) was used to

assess the number of variable sites, transition and transversions, and the presence/absence of any sequence ambiguities (e.g. unexpected stop codons in the amino acid alignments).

2.2.3 Microsatellite optimization steps

Cross-species amplification was performed using available loci developed for species of the Sciaenidae family. The focus was on tri- and tetra-nucleotide loci. In total 38 markers described as polymorphic in closely related species to *O. ruber,* were tested (Appendix 2B).

STEP 1: *Amplification test*

For amplification optimization and the selection of polymorphic loci, a panel of eight samples was used. Amplification reactions were performed in 10 µL volumes containing approximately 40 ng of template DNA, 1 x PCR buffer, $MgCl_2$ (see Appendix 2B for the concentrations), 0.2 mM dNTPs (Promega), 0.2 pmol forward and reverse primers (Whitehead Scientific), 0.05 U Supertherm *Taq* polymerase (Southern Cross Biotechnologies). The PCR cycling conditions were: An initial denaturation at 94°C for 5 min, followed by 30 cycles: Denaturation at 94 °C for 30 s, annealing at the specific temperature (see Appendix 2B for each locus) for 1 min, and elongation at 72 °C for 1 min and a final elongation at 72 °C for 7 min. The PCR products were electrophoresed on 2% agarose gels and visualised with GelRed™ stain.

STEP 2: *Polymorphism tests*

The PCR products were electrophoresed on 3% agarose gels for one hour to one and half hours at 100 V to allow separation of different alleles and visualised with GelRed. The optimized loci also were re-amplified using fluorescently-labelled ChromaTide(R) Alexa Fluor™ 488-5-dUTPs (Invitrogen) to confirm the presence of more than two different alleles at each locus. These fragments were analysed on an ABI 3500xl genetic analyzer (Applied Biosystems) with the GeneScan Liz™500 Size Standard (Applied Biosystems).

STEP 3: *Small scale genotyping*

Forward primers of loci that were found to be polymorphic were labelled fluorescently (G5 dye set) (Applied Biosystems) for multiplexing (co-amplification of multiple loci in a PCR mix) reactions (refer to Table 2.2). Loci with non-overlapping size ranges were labelled with the same dye, while loci with overlapping size ranges were labelled with different dyes. These

loci were genotyped in a representative number of individuals (72) using the Quantitect Multiplex PCR kit (Qiagen) in a single multiplex combination with the conditions recommended by the manufacturer.

STEP 4: *Sequencing of homozygotes and summary statistics*

To confirm allele scoring and the repeat motif variation, an assortment of alleles were amplified using unlabeled primers following the protocol described above. The amplified products were precipitated using the NaAc and ethanol protocol and cycle sequenced as mentioned in the mtDNA section. The amplified alleles were sequenced in the forward direction only. Sequences were analysed using CLC Bio^{R}.

Scoring of genotypes was done using GENEMARKER version 1.5 software (SoftGenetics, State Collage, Pennsylvania, USA). Basic summary statistics for each locus were calculated in GENETIX version 4.05 software (Belkhir *et al.* 1996-2004). These included number of alleles, observed heterozygosity (H_O), expected heterozygosity (H_E) and the inbreeding coefficient F_{IS}. Deviations from Hardy-Weinberg Equilibrium (HWE) were tested with GENEPOP version 4.0.10 (Rousset 2008) using 1000 dememorizations and 100 batches, with 1000 iterations per batch. MICROCHECKER version 2.2.3 (van Oosterhout *et al.* 2004) was used to check for the presence of null alleles and large allele dropouts. Linkage disequilibrium (LD) between markers was tested in GENEPOP, using 1000 dememorizations and 100 batches, with 1000 iterations per batch.

2.3 Results and Discussion

The protection of biodiversity and sustainable exploitation require investigation of genetic diversity of threatened species. For this purpose, selection of molecular markers is important as markers should supply the most informative data for the question of interest. Many studies have benefited from universal molecular markers that have an advantage of easy usage in non-model organisms. In this study, I tested available mitochondrial and nuclear DNA markers in *O. ruber*.

2.3.1 DNA extraction and quality

The success of *16S* amplification (Fig. 2.1) showed that good quality DNA was recovered from the extraction procedure. There was no consistent amplification failure of any samples.

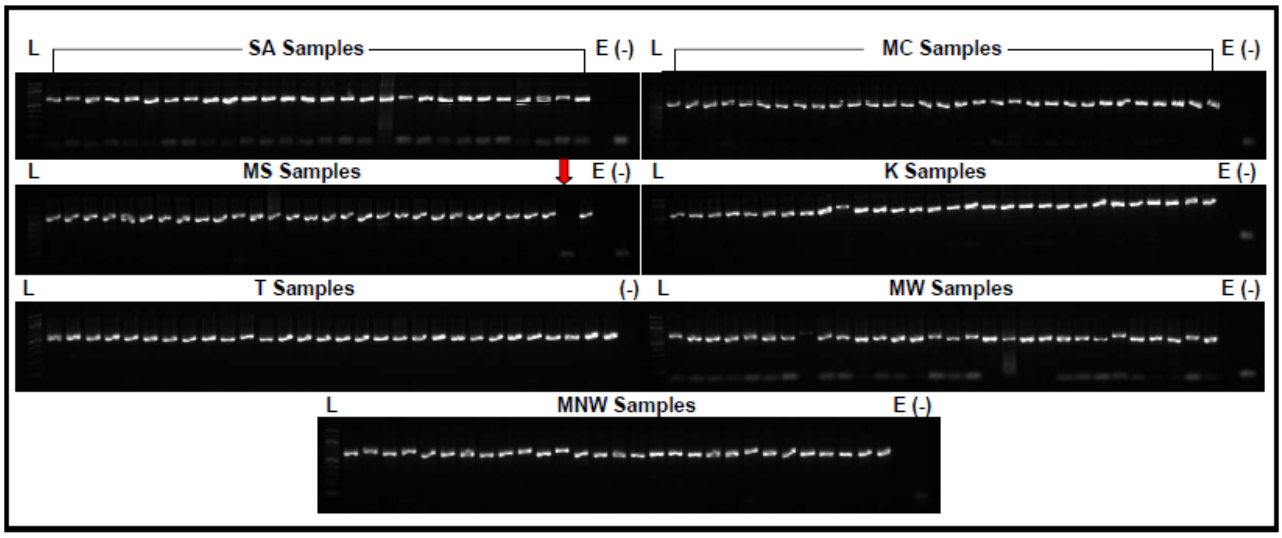


Fig. 2.1 Successful *16S* amplifications in *Otolithes ruber* visualised on 2% agarose gels. PCR products of samples from SA: South Africa, MC: Central Mozambique, MS: South Mozambique, K: Kenyan, T: Tanzanian, MW: West Madagascar, MNW: Northwest Madagascar. L: 100 bp DNA marker, (-): Negative control, E: Empty well. The red arrow represents amplification failure (the amplification was repeated and the product was obtained). In some cases primer dimers, seen as the faint bands at the bottom of the gel, are present, e.g. in SA and MW samples.

2.3.2 Mitochondrial DNA amplification and sequencing

The difficulty associated with the PCR optimization of the *COI* gene using the first set of primers, FF2d and FR1d, was the presence of a second unspecific amplified product. The desired PCR product with the correct size (~ 750-800 bp) was obtained after several amplification trials. However, sequencing trials were not successful as the generated sequences had too much background noise (Fig.2.2A) to allow accurate base calling. From a literature search, a new pair of primers was used due to the fact that the primers were tested in *O. ruber* together with other sciaenids (Lakra *et al.* 2009). Although the sequencing was improved in comparison to the previous set, generated sequences still had background noise (Fig.2.2B), which made it impossible to assemble the forward and reverse sequences into consensus sequences.

The same amplification problem was dealt with in the case of the *ND2* optimizations using the ND2cv-F and -R primers. The best amplified products, the clearest and brightest bands with the correct size, were sequenced. The sequence data could not be used; although good sequence was obtained with the reverse primer, the forward reactions failed (see the sequence trace data with ND2cv-F primer in Fig. 2.3). Aligning the forward primer with all the available sciaenid *ND2* sequences showed that the primer sequence does not have enough similarity to the target site. Hence, the forward primer was redesigned (ND2sF), which improved sequencing. As can be seen in Fig. 2.3, the forward sequence was as good as the reverse sequence. Before further sequencing, the obtained *ND2* sequences were BLAST searched and the best hits (96% coverage, 79% maximum identity) were the *ND2* gene region of fish species in the sciaenid family (e.g. silver croaker, *Pennahia argentata*).

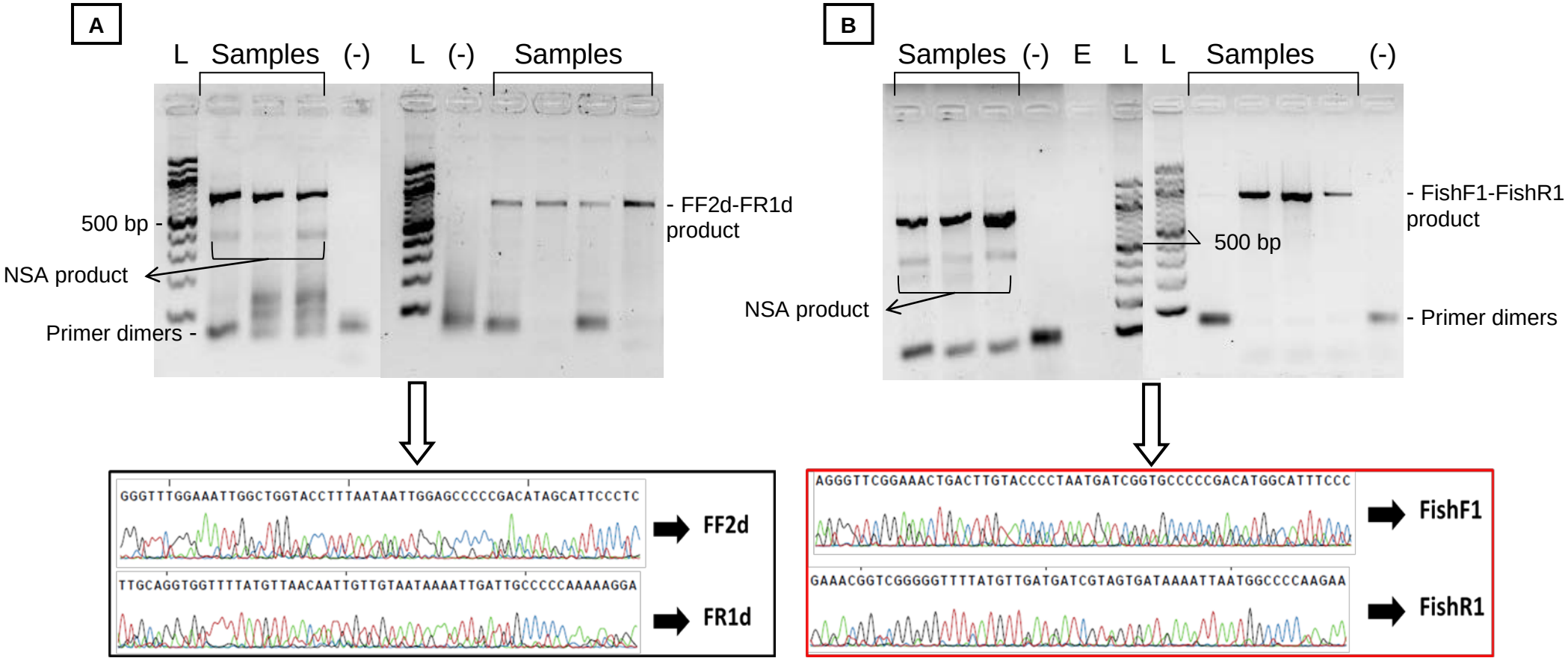


Fig. 2.2 Illustration of the non-optimal and optimal PCR, and sequencing results for the *COI* gene region. **A)** 2% agarose gel representation of the first set of *COI* primers (FF2d, FR1d) PCR products, L: 100bp DNA marker, (-): Negative control, NSA: Non-specific amplification. Related forward and reverse sequence trace data are represented below the gel picture. **B)** Second set of *COI* primers' (FishF1, FishR1) PCR products, E: Empty well.

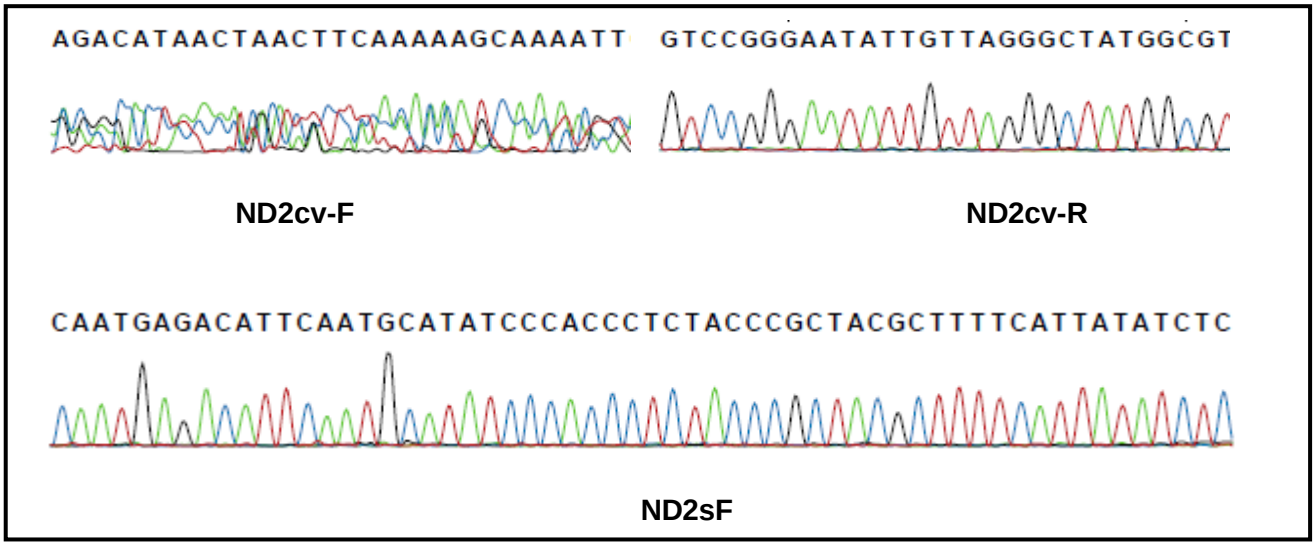


Fig. 2.3 Sequence trace data representation of *ND2* products with primer ND2cv-F (forward sequence), primer ND2cv-R (reverse sequence) and the sciaenid-specific ND2sF (new forward sequence). Further *ND2* amplification was performed using the ND2sF and ND2cv-R primer set.

Compared to *COI* and *ND2,* optimization of the *Cyt b* gene fragment amplification and sequencing was more successful. There was no difficulty in obtaining sequence data in both directions and generating consensus sequences. The best hits from a BLAST search were against an *Argyrosomus regius* (Sciaenidae) *Cyt b* gene and a *Miichthys miiuy* (Sciaenidae) partial mitochondrial *Cyt b* gene.

Due to sequencing difficulties, *COI* was eliminated as a candidate mtDNA marker for the study. *ND2* and *Cyt b* were sequenced in a representative number of samples from Kenya, Mozambique and South Africa in order to compare the polymorphism. Both genes had a small number of variable sites, *ND2* = 4 and *Cyt b* = 7; all the identified substitutions were transition mutations. No nuclear copies were observed as obtained sequences had the expected size and there were no stop codons within the sequenced gene regions after translating them into amino acids. Alignments of the sequences showed that *Cyt b* had more variation than *ND2* (Appendix 2C). It also represented variation both within and between the locations. Hence, *Cyt b* was chosen as mtDNA marker for the study.

2.3.3 Cross-species amplification of microsatellites

From the 38 tested microsatellite loci, 22 were amplified successfully (Appendix 2B). Polymorphism of these ampified loci were tested using two methods: Separation of alleles on 3% agarose gels and genotyping using labelled dNTPs.

2.3.3.1 Polymorphism

Analyzing the PCR products on 3% gels electrophoresed for a longer time period allowed us to observe the presence of different alleles per locus. Polymorphism of the loci with tetra-nucleotide repeats was more clearly distinguished compared to loci with tri-nucleotide repeats. Fig. 2.4 shows examples of polymorphic loci identified using this technique.

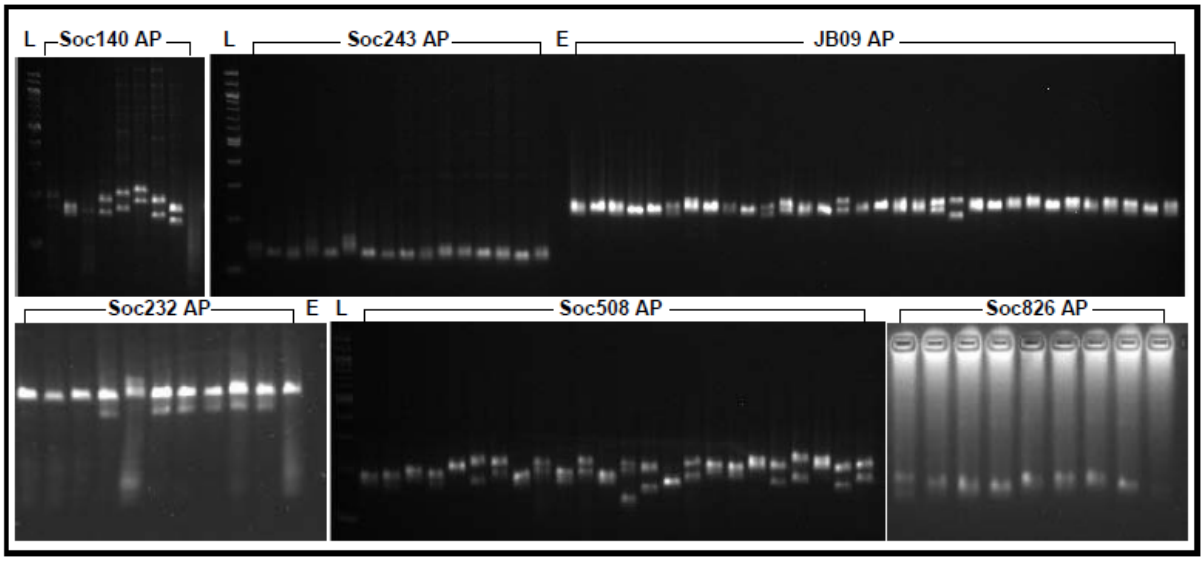


Fig. 2.4 Representation of microsatellite polymorphism using 3% gel electrophoresis. Soc140, Soc508 and Soc826 are loci with tetra-nucleotide repeats; Soc243, Soc232 and JB09 are loci with tri-nucleotide repeats. Two bands in a lane indicate the presence of two alleles at that locus in the specific individual. Variation in PCR product size (different alleles) and the thickness of PCR bands (presence of two alleles close in size) are also signs of polymorphism. L: 100 bp DNA marker, E: Empty well.

Although this is a useful method, one should keep in mind that alleles differing by only a few nucleotide repeat units cannot be visualized with 3% gels. Therefore, all the amplified loci were re-amplified using labelled dNTPs for genotyping purposes. This is also a necessity to

design multiplex combinations by allowing the determination of the size ranges of alleles per locus. Fig. 2.5 illustrates the different alleles identified at two loci, Soc140 and Soc243, as an example of genotyping with labelled dNTPs. Polymorphism tests revealed that only seven of the 22 amplified loci were polymorphic. The remainder of the tested markers were monomorphic or in some cases products could not be scored reliably due to the presence of too much background noise.

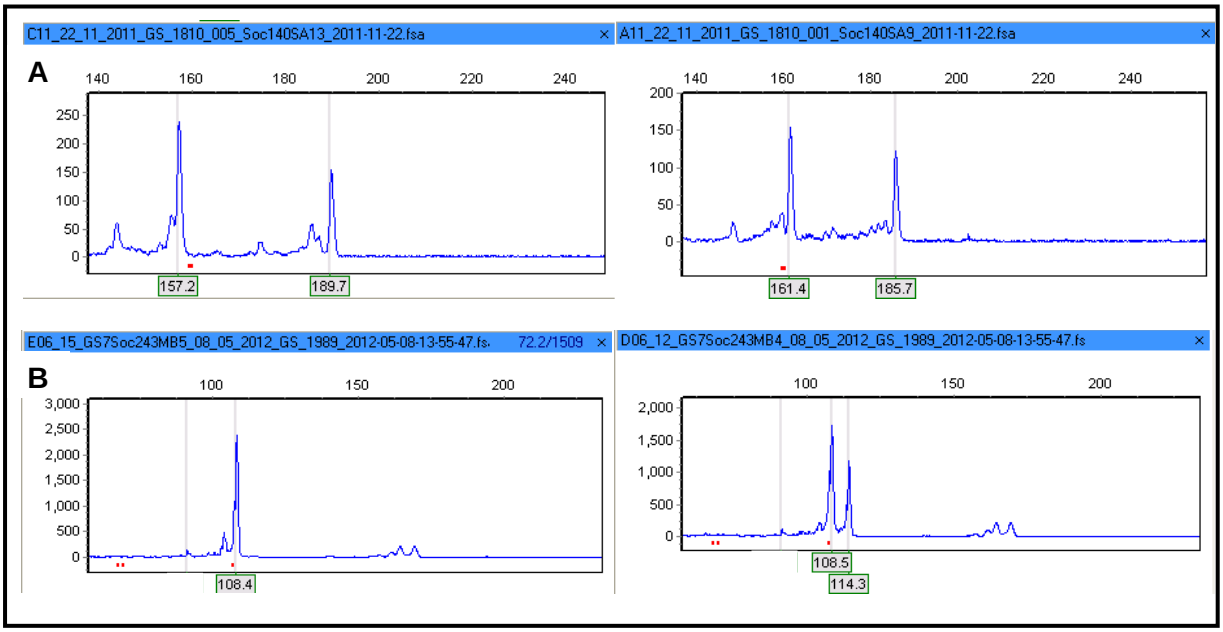


Fig. 2.5 Electropherogram results with labelled dNTPs. **A)** Representation of different alleles identified in two heterozygous individuals for the locus Soc140, tetra-nucleotide. **B)** Representation of a homozygous and a heterozygous individual for the locus Soc243, tri-nucleotide.

2.3.3.2 Small scale genotyping

Determination of allele sizes allowed us to make a single multiplex combination including all seven loci with no overlapping alleles (Fig. 2.6). These loci were genotyped in a total of 72 individuals from Kenya, Tanzania, South Africa and Madagascar. Allele scoring of each locus was not problematic as there was no issue with stutters, no background noise and each locus had different characteristics such as unique peak shapes and peak patterns.

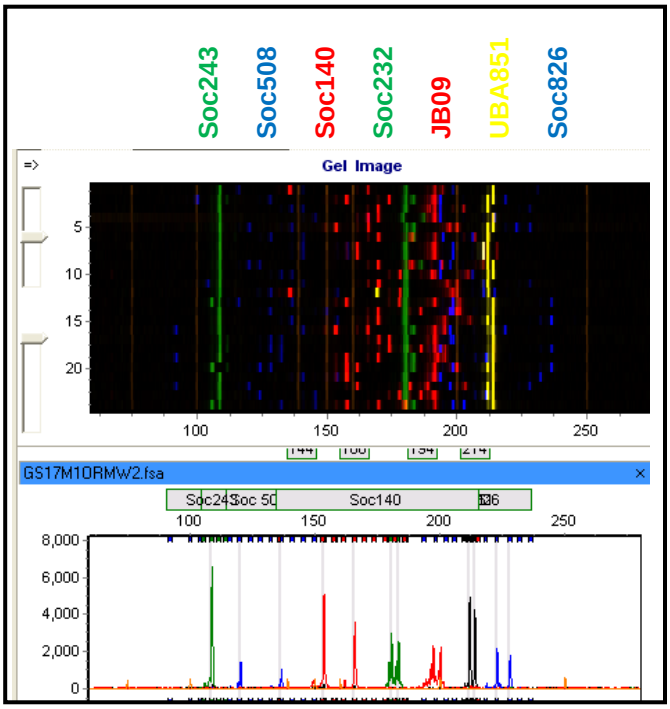


Fig. 2.6 Electropherogram of a multiplex combination of the seven microsatellite loci with the corresponding raw gel image represented at the top. Two loci were labelled with green (VIC), two loci with blue (6-FAM), two loci with red (PET) and one locus with yellow (NED). The specific fluorescent dyes for each locus are given in Table 2.1.

2.3.3.3 Homozygote sequencing and summary statistics

The last step of optimization was sequencing a selective number of different alleles per locus. Trace data of sequences were good enough to confirm the allele size (based on the number of repeats) and repeat motifs for all loci, except Soc243, which had failed sequences, and Soc232 (Appendix 2D). Soc232 alleles fit with a tri-nucleotide pattern with labelled dNTPs, as well as labelled primers. However, literature information characterises the locus as a tetra-nucleotide (AGAC) (Turner *et al.* 1998). Sequencing showed that Soc232 had an interrupted repeat region (Appendix 2D), which might explain the contrasting genotyping result in the different species. Additionally, JB09 sequences had background noise, but the repeat regions were representative enough to confirm the repeat unit and the allele size.

Summary statistics for the seven polymorphic loci are represented in Table 2.1, based on the genotypic information of 72 individuals from four locations. In total, 70 alleles were identified. Expected heterozygosity ranged from 0.213 to 0.919 (for Soc243 and Soc508, respectively). Soc140 and Soc508 were the most polymorphic, while Soc243 was the least

polymorphic marker. No evidence of null alleles, large allele dropouts or linkage disequilibrium was detected. Testing all marker pairs for linkage disequilibrium showed no specific pattern as the cases were reported for different pairs of loci and for different locations. No significant deviations from Hardy-Weinberg Equilibrium (HWE) were detected. F_{IS} (inbreeding coefficient) values were not significant. Summary statistics were calculated of each marker for each location and are indicated in Appendix 2E.

Table 2.1 Information for the seven polymorphic loci tested using 72 *Otolithes ruber* individuals, including locus name, fluorescent dye, primer sequence, repeat type according to the literature, allele sizes, Na (number of alleles), observed heterozygosity (H_O), expected heterozygosity (H_E) and F_{IS} (inbreeding coefficient) (Weir & Cockerham 1984).

Locus name	Dye	Primer sequence	Repeat motif	Allele range (bp)	Na	H_O	H_E	F_{IS}
Soc 140[1]	PET	F:GGT GCA AAC ACA GCC ATA CAG T R:GCA AAA TCG AAG ACC GAG TTT AG	$(CTGT)_8$	134-238	18	0.895	0.898	- 0.015
Soc232[1]	VIC	F:AGG GCA CAG TTG CAT CTC TG R:CCC ATC CTC AAG GCA GAA C	$(AGAC)_4$	180-186	3	0.260	0.233	- 0.113
Soc243[1]	VIC	F:GAC GGG GAT GCC ATC TGC R:AAT GCG AAA AAG ACG AAA CAG T	$(CCT)_9$	106-112	3	0.125	0.213	0.418
Soc508[2]	6-FAM	F:GCA GCA CAT TTC AGC ACA C R:TAA TGC CCC TGT TAT CTA TCT A	$(GATA)_{18}$	92-168	18	0.967	0.919	- 0.005
Soc826[3]	6-FAM	F:GGC AGG ATT TAG GCA ATT CA R:ACA CAC TCC TGT GTG CAA CC	$(GTGA)_{11}$	194-246	13	0.736	0.810	0.107
JB09[4]	PET	F:AAC GTG GAC AGG TTT AAT CTA T R:GCA GAT AAA AGC ACA AAC ACT T	$(AGC)_5$	180-210	11	0.701	0.712	- 0.005
UBA851[5]	NED	F:AGA AAG CAG GTA TAC TCC AAA C R:GGA CAG AGG AGA AAG AGA AGT	$(AG)_{14}$	170-214	4	0.573	0.500	- 0.196

1: Turner *et al.* (1998)
2: Karlsson *et al.* (2008)
3: Renshaw *et al.* (2012)
4: Ma *et al.* (2011)
5: Archangi *et al.* (2009)

2.3.4 Choice of the mitochondrial marker

I initially used *COI*, *ND2* and *Cyt b* mtDNA genes, which are generally preferred in population studies of animal species. These are useful for the applications of evolutionary genetics due to the fact that they are protein coding genes; they are functionally important and have conserved sequences for primer design, but are variable enough to show differences within and between species. Hence, a low incidence of insertion and deletions, and low nucleotide divergence makes sequence alignments easy (Townsend *et al.* 2008). In our case, *COI* was not useful for the study as obtaining good sequence data was not possible. The specific *COI* primer set (FishF1, FishR1) was not degenerate (Lakra *et al.* 2009) and the gene was previously successfully sequenced in *O. ruber* samples. We could not find a reason for the non-specific amplification and background noise in the *COI* sequences in our samples. However, possible contamination of the tissue sample by bacteria (amplification of the bacterial *COI*) might be a reason.

Cytochrome *b* was the more informative mtDNA marker compared to *ND2*. Although both genes represented low sequence variation, the observed substitutions in *Cyt b* were present in individuals within and among the selected sampling locations. Also, troubleshooting with *Cyt b* was negligible in terms of amplification and sequencing compared to *ND2*. These reasons justified *Cyt b* as the mitochondrial region for the present study. It has been shown by other phylogenetic and population genetic studies that *Cyt b* is a commonly used marker for fish species (Gaither *et al.* 2010, 2011; Reece *et al.* 2010; Shikano *et al.* 2010). Also, the large accumulation of *Cyt b* data in sequence databases compared to *ND2* is evidence of the easy access to this gene and to primer design in different regions on the mtDNA. By generating data on a large scale, *Cyt b* haplotype variation provides estimates of genetic differentiation over longer time scales compared with microsatellites, providing resolution of recent historical effects (refer to Chapter 3).

2.3.5 Microsatellite amplification success

Amplification of existing microsatellite markers from related species is time and cost effective. However, it is not clear how closely related the species should be to ensure success of cross-species amplification. In several animal taxa such as birds and fish, genetic distance was used as a measure of cross-species amplification and polymorphism success rates (Primmer *et al.* 2005; Reid *et al.* 2012). In the present study, cross-species amplification showed amplification success of 58% (22 loci), but only 31% of loci were

polymorphic (7 loci). Most of the amplified loci were from distantly related species (Fig. 2.7): Red drum (*Sciaenops ocellatus*) (10/19 amplified, with five polymorphic loci), Belanger's croaker (*Johnius belengerii*) (3/5 amplified, with one polymorphic locus) and mulloway (*Argyrosomus japonicus*) (9/12 amplified, one polymorphic). This is contrary to the negative relationship between genetic distance and the amplification and polymorphism of loci that was identified by other studies (Primmer *et al.* 2005; Carreras-Carbonell *et al.* 2008; Reid *et al.* 2012).

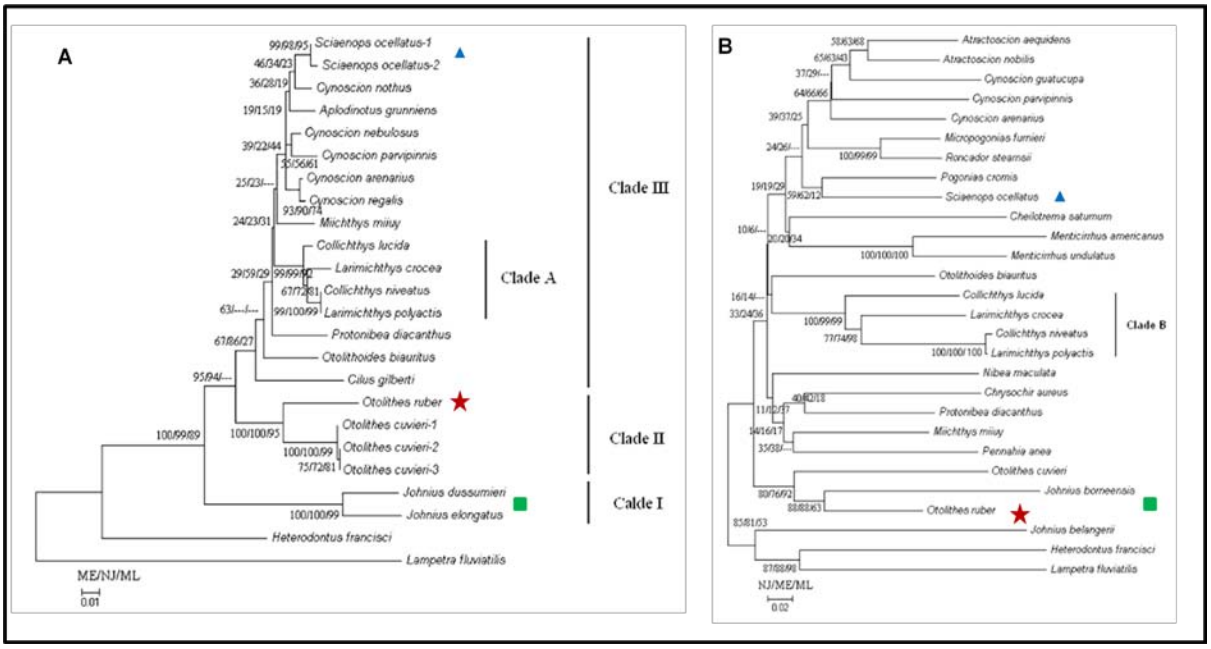


Fig. 2.7 Representation of phylogenetic relationships of the Sciaenidae based on the partial *16S* rDNA gene **(A)** and *COI* gene **(B)** sequences. Numbers above branches indicate bootstrap percentages for Minimum Evolution (ME), Neighbour Joining (NJ), and Maximum Likelihood (ML) analyses (Xu *et al.* 2011). In each tree a red star shows the position of *O. ruber*, a blue triangle *Sciaenops* and a green square *Johnius*.

2.3.6 Polymorphism in O. ruber

Compared to the source species, all loci showed high levels of polymorphism in *O. ruber,* except UBA851. I detected four alleles (in 72 individuals) with low values of H_O and H_E (0.573 and 0.500, respectively), whereas in Australian mulloway (*Argyrosomus japonicus*), nine alleles (in 29 individuals) were recovered with $H_O = 0.77$ and $H_E = 0.72$ (Archangi *et al.* 2009). They also tested this locus to investigate its efficacy in the South African conspecific, where they found 10 alleles (in 20 individuals) with high values of $H_O = 0.95$ and $H_E = 0.90$ (Archangi *et al.* 2009). The most polymorphic microsatellite loci in *O. ruber* were Soc140 and Soc 508 with 18 alleles each and high levels of heterozygosity. The levels of their polymorphism were higher than the source species: Soc140 had 4 alleles (7 to 11 individuals) with $H_O = 0.72$ and $H_E = 0.67$ in red drum (*Sciaenops ocellatus*). Soc 140 also cross-species amplified in spotted sea trout (*Cynoscion nebulosus*), black drum (*Pogonias cromis*), Atlantic croaker (*Micropogonias undulates*) recovered one, four and six alleles respectively (Turner *et al.* 1998). Soc508 had 9 alleles (in 24 individuals) with $H_O = 0.917$ and $H_E = 0.868$ (Karlsson *et al.* 2008). Soc826 had 13 alleles in *O. ruber* ($H_O = 0.736$ and $H_E = 0.810$) whilst it was less polymorphic in the source species with six alleles (in 20 individuals; $H_O = 0.650$ and $H_E = 0.749$). This locus was also tested in another sciaenid, the sea trout (*Cynoscion nebulosus*), which also showed high polymorphism, 18 alleles (in 20 individuals) with $H_O = 0.800$ and $H_E = 0.946$ (Renshaw *et al.* 2012). JB09 had 11 alleles in *O. ruber*, twice the number of alleles recovered in Belanger's croaker (*J. belengerii*) (Ma *et al.* 2011), with $H_O = 0.701$ and $H_E = 0.712$. Soc232 was not found to be polymorphic in red drum (*Sciaenops ocellatus*) (Turner *et al.* 1998), but had three alleles and low values of $H_O = 0.260$ and $H_E = 0.233$ in *O. ruber*. One should therefore be cautious when selecting potential loci from source species as loci could be more polymorphic in target species.

2.3.7 Conclusion

By conducting this pilot study, I was able to find both informative mtDNA (*Cyt b*) and microsatellite markers (seven loci) for a phylogeographic study of *O. ruber,* which enabled me to address questions from historical and contemporary perspectives. Also, from microsatellite markers, the small data set represented a contradictory result to the negative relationship between evolutionary distance to the model species and cross-species amplification of loci; *S. ocellatus*, distantly-related, loci were mostly amplified and showed high polymorphism in *O. ruber.* The illustrated steps should be taken into consideration before using molecular markers in large scale data generation. A pilot study is essential to determine the power of markers for accurate estimates, deciding the number of markers

needed, and for the identification and elimination of the markers with detectable drawbacks (the presence of null alleles associated with microsatellites can be detected through deviations from HWE). If there are insufficient microsatellite loci in the literature for a species of interest, it is worth testing loci with no polymorphism in the source species, since my results showed that these loci might still be polymorphic in the target species. Cross-species amplification approach and accumulation of mtDNA sequences benefit studies; the selected markers in this study will be useful for those working on this species or close relatives.

2.4 References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology,* **215**, 403-410.

Archangi B, Chand V, Mather PB (2009) Isolation and characterization of 15 polymorphic microsatellite DNA loci from *Argyrosomus japonicus* (mulloway), a new aquaculture species in Australia. *Molecular Ecology Resources,* **9**, 412-414.

Avise JC, Nelson WS, Sibley CG (1994) DNA sequence support for a close phylogenetic relationship between some storks and New World vultures. *Proceedings of the National Academy of Sciences of the USA*, **91**, 5173-5177.

Baharum SN, Nurdalila AA (2012) Application of *16S* rDNA and cytochrome *b* ribosomal markers in studies of lineage and fish populations structure of aquatic species. *Molecular Biology Reports,* **39**, 5225–5232.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology,* **13**, 729–744.

Barbará T, Palma-Silva C, Paggi GM, Bered F, Fay MF, Lexer C (2007) Cross-species transfer of nuclear microsatellite markers: Potential and limitations. *Molecular Ecology,* **16**, 3759-3767.

Barkley NA, Krueger RR, Federici CT, Roose ML (2009) What phylogeny and gene genealogy analyses reveal about homoplasy in citrus microsatellite alleles. *Plant Systematics and Evolution,* **282**, 71-86.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05 Logciel sous Windows pour la Genetique des Populations, Laboratorie Genome, Populations, Interactions, CNRS UMR 5171, Universite de Montpellier II, Montpellier (France).

Bhargava A, Fuentes FF (2010) Mutational dynamics of microsatellites. *Molecular Biotechnology,* **44**, 250–266.

Blandon IR, Morales C, Vega RR, Deborah Overath R, Stunz GW, Ward R (2011) Development of primers for 29 microsatellite loci in spotted seatrout (*Cynoscion nebulosus*). *Conservation Genetics Resources*, **3**, 123-126.

Borsa P, Lemer S, Aurelle D (2007) Patterns of lineage diversification in rabbitfishes. *Molecular Phylogenetics and Evolution*, **44**, 427–435.

Bradmana H, Greweb P, Appletona B (2011) Direct comparison of mitochondrial markers for the analysis of swordfish population structure. *Fisheries Research,* **109**, 95-99.

Buschiazzo E, Gemmell NJ (2006) The rise, fall and renaissance of microsatellites in eukaryotic genomes. *BioEssays,* **28**, 1040–1050.

Carreras-Carbonell J, Macpherson E, Pascual M (2008) Utility of pairwise mtDNA genetic distance for predicting cross-species microsatellite amplification and polymorphism success in fishes. *Conservation Genetic,* **9**, 181-190.

Dudgeon CL, Blower DC, Broderick D, Giles JL, Holmes BJ, Kashiwagi T, N. C. Krück NC, Morgan JAT, Tillett BJ, Ovenden JR (2012) A review of the application of molecular genetics for fisheries management and conservation of sharks and rays. *Journal of Fish Biology,* **80**, 1789–1843.

Eytan RI, Hellberg ME (2010) Nuclear and mitochondrial sequence data reveal and conceal different demographic histories and population genetic processes in Caribbean reef fishes. *Evolution,* **64***,* 3380–3397.

Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW (2010) Genetic evaluation of marine biogeographical barriers: Perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). *Journal of Biogeography,* **37**, 133-147.

Gaither MR, Jones SA, Kelley C*,* Newman SJ, Sorenson L, Bowen BW (2011) High connectivity in the deepwater snapper *Pristipomoides filamentosus* (Lutjanidae) across the Indo-Pacific with isolation of the Hawaiian Archipelago. *PLoS ONE,* **6**, 1-13.

Gold JR, Richardson LR, Furman C, King TL (1993) Mitochondrial DNA differentiation and population structure in red drum (*Sciaenops ocellatus*) from the Gulf of Mexico and Atlantic Ocean. *Marine Biology,* **116**, 175-185.

Gonzalez EG, Zardoya R (2007) Relative role of life history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *BMC Evolutionary Biology,* **7**.

Guichoux E, Lagache L, Wagner S, Chaumeil P, Leger L, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. *Molecular Ecology Resources,* **11**, 591-611.

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Symposium Series,* **41**, 95-98.

Hansen MM, Kenchington E, Nielsen EE (2001) Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries,* **2**, 93-112.

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. *Fish and Fisheries,* **9**, 333-362.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identification through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences,* **270**, 313–321.

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics*, Vol, 291-310.

Hubert N, Meyer CP, Bruggemann HJ, Guérin F, Komeno RJL, Espiau B, Causse R, Williams JT, Planes S (2012) Cryptic diversity in Indo-Pacific coral-reef fishes revealed by DNA-barcoding provides new support to the Centre-of-Overlap hypothesis. *PLoS ONE,* **7**, e28987.

Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN (2007) Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes***, 7**, 544-548.

Karlsson S, Renshaw MA, Rexroad CE, Gold JR (2008) PCR primers for 100 microsatellites in red drum (*Sciaenops ocellatus*). *Molecular Ecology Resources,* **8**, 393-398.

Kasapidis P, Magoulas A (2008) Development and application of microsatellite markers to address the population structure of the horse mackerel *Trachurus trachurus*. *Fisheries Research,* **89**, 132-135.

Keeney DB, Heupel MR, Hueter RE, Heist EJ (2005) Microsatellite and mitochondrial DNA analyses of the genetic structure of blacktip shark (*Carcharhinus limbatus*) nurseries in the northwestern Atlantic, Gulf of Mexico, and Caribbean Sea. *Molecular Ecology,* **14**, 1911-1923.

Lakra WS, Goswami M, Gopalakrishnan A (2009) Molecular identification and phylogenetic relationships of seven Indian Sciaenids (Pisces: Perciformes, Sciaenidae) based on *16S* rRNA and cytochrome c oxidase subunit I mitochondrial genes. *Molecular Biology Reports,* **36**, 831-839.

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

Ma C, Ma H, Ma L, Shen A, Gao L (2011) Development and characterization of 12 microsatellite markers for *Johnius belengerii*. *Conservation Genetic Resources,* **3**, 717-719.

Mirimin L, Ruiz Guajardo JC, Vervalle J, Bester-Van der Merwe A, Kerwath S, Macey B, Bloomer P, Roodt-Wilding R (2013) Isolation and validation of microsatellite markers from a depleted South African sciaenid species, the dusky kob (*Argyrosomus japonicus*), by means of the FIASCO/454 approach. *Conservation Genetics Resources*, **5**, 841-844.

O'Connell M, Wright JM (1997) Microsatellite DNA in fishes. *Reviews in Fish Biology and Fisheries,* **7**, 331-363.

Okumuş I, Çiftci Y (2003) Fish population genetics and molecular markers: II- molecular markers and their applications in fisheries and aquaculture. *Turkish Journal of Fisheries and Aquatic Sciences,* **3**, 51-79.

O'Malley KG, Abbey CA, Ross K, Gold JR (2003) Microsatellite DNA markers for kinship analysis and genetic mapping in red drum, *Sciaenops ocellatus* (Sciaenidae, Teleostei). *Molecular Ecology Notes*, **3**, 155-158.

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski G, (1991) The simple fools guide to PCR. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu, Hawaii.

Primmer CR, Painter JN, Koskinen MT, Palo JU, Merila J (2005) Factors affecting avian cross-species microsatellite amplification. *Journal of Avian Biology,* **36**, 348-360.

Reece JS, Bowen BW, Joshi K, Goz V, Larson A (2010) Phylogeography of two moray eels indicates high dispersal throughout the Indo-Pacific. *Journal of Heredity,* **101**, 391-402.

Reid K, Hoareau TB, Bloomer P (2012) High-throughput microsatellite marker development in two sparid species and verification of their transferability in the family Sparidae. *Molecular Ecology Resources,* **12**, 740-752.

Renshaw MA, Hollenbeck CM, Gold JR (2012) Isolation of microsatellite markers from red drum, *Sciaenops ocellatus*, and characterization in red drum and spotted seatrout, *Cynoscion nebulosus*. *Molecular Ecology Resources,* **12**, 570-572.

Rousset F (2008) Genepop'007: A complete reimplementation of Genepop software for Windows and Linux. *Molecular Ecology Resources,* **8**, 103-106.

Saillant E, Gizdziel K, O'Malley KG, Turner TF, Pruett CL, Gold JR (2004) Microsatellite markers for Red Drum, *Sciaenops ocellatus*. *Gulf of Mexico Science*, **1**, 100-107.

Schlötterer C (2004) The evolution of molecular markers — just a matter of fashion? *Nature reviews Genetics,* **5,** 63-69.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters,* **9**, 615-629.

Shikano T, Shimada Y, Herczeg G, Merila J (2010) History vs. habitat type: Explaining the genetic structure of European nine-spined stickleback (*Pungitius pungitius*) populations. *Molecular Ecology,* **19**, 1147-1161.

Spies IB, Gaichas S, Stevenson DE, Orr JW, Canino MF (2006) DNA-based identification of Alaska skates (Amblyraja, Bathyraja and Raja: Rajidae) using cytochrome *c* oxidase subunit I (*COI*) variation. *Journal of Fish Biology,* **69**, 283-292.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* (submitted).

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research,* **22**, 4673-4680.

Townsend TM, E. AR, Kelley ST, Wien JJ, Reeder TW (2008) Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: An example from squamate reptiles. *Molecular Phylogenetics and Evolution,* **47**, 129-142.

Turner TF, Richardson LR, Gold JR (1998) Polymorphic microsatellite DNA markers in red drum (*Sciaenops ocellatus*). *Molecular Ecology,* **7**, 1771-1788.

Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes,* **4**, 535-538.

Walther E, Schöfl G, Mrotzek G, Haryanti, Sugama K, Saluz HP (2011) Paralogous mitochondrial control region in the giant tiger shrimp, *Penaeus monodon* (F.) affects population genetics inference: A cautionary tale. *Molecular Phylogenetics and Evolution,* **58**, 404-408.

Waples RS, Gaggiotti O (2006) What is a population? An emperical evaluation of some genetic methods for identifying the number of gene pools and their connectivity. *Molecular Ecology,* **15**, 1419-1439.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **360**, 1847-1857.

Ward RD, Hanner R, Hebert PDN (2009) The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology,* **74**, 329-356.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution,* **38**, 1358-1370.

Xu T, Cheng Y, Sun Y, Shi G, Wang R (2011) The complete mitochondrial genome of bighead croaker, *Collichthys niveatus* (Perciformes, Sciaenidae): Structure of control region and phylogenetic considerations. *Molecular Biology Reports,* **38**, 4673-4685.

Zhang DX, Hewitt GM (2003) Nuclear DNA analysis in genetic analysis of populations: Practice, problems and prospects. . *Molecular Ecology,* **12**, 563-584.

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PD (2009) DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources,* **9**, 237-242.

2.5 Appendices

Appendix 2A Primer information of the three mtDNA genes, including primer names, sequences in the 5' to 3' direction, $MgCl_2$ and primer concentrations in PCR reactions, and optimized PCR cycling conditions.

Gene	Primer name	5'→ 3' sequence	[$MgCl_2$] [*primer*]	Cycling conditions
COI	FF2d[1]	TTCTCCACCAACCACAARGAYATYGG	1.5 Mm *0.2 pmol* (each)	94 °C 2 min; 94 °C 30 s, 56 °C 40 s, 72 °C 1 min (x35); 72 °C 10 min
	FR1d[1]	CACCTCAGGGTGTCCGAARAAYCARAA		
	FishF1[2]	TCAACCAACCACAAAGACATTGGCAC	0.8 mM *0.1 pmol*	94 °C 2 min; 94 °C 40 s, 56 °C 40 s, 72 °C 1:30 (x35); 72 °C 10 min
	FishR1[2]	TAGACTTCTGGGTGGCCAAAGAATCA		
ND2	ND2cv-F[3]	TGTCCGACGACGGCCACGGGCCCATACCCC	1 mM *0.2 pmol*	94 °C 3 min; 94 °C 45 s, 60 °C 1:30, 72 °C 1:30 (x30); 72 °C 10 min
	ND2cv-R[3]	GTCACTCCTACGGAAGGCTTTGAAGGC		
	ND2sF[3]	CGGACTTGGAACTACAATCACATTTGC	1.2 mM *0.2 pmol*	94 °C 3 min; 94 °C 45 s, 48 °C 1:20, 72 °C 1:30 (x35); 72 °C 10 min
Cyt b	L14724[4]	TGACTTGAAAAACCACCGTTG	1.5 mM *0.25 pmol*	94 °C 2 min; 94 °C 30 s, 51 °C 30 s, 72 °C 45 s (x35); 72 °C 5 min
	H15499[5]	GGTTGTTTGAGCCTGATTC		

1: Ivanova *et al.* (2007)

2: Zemlak *et al.* (2009); Ward *et al.* (2005)

3: This study

4: Palumbi *et al.* (1991)

5: Avise *et al.* (1994)

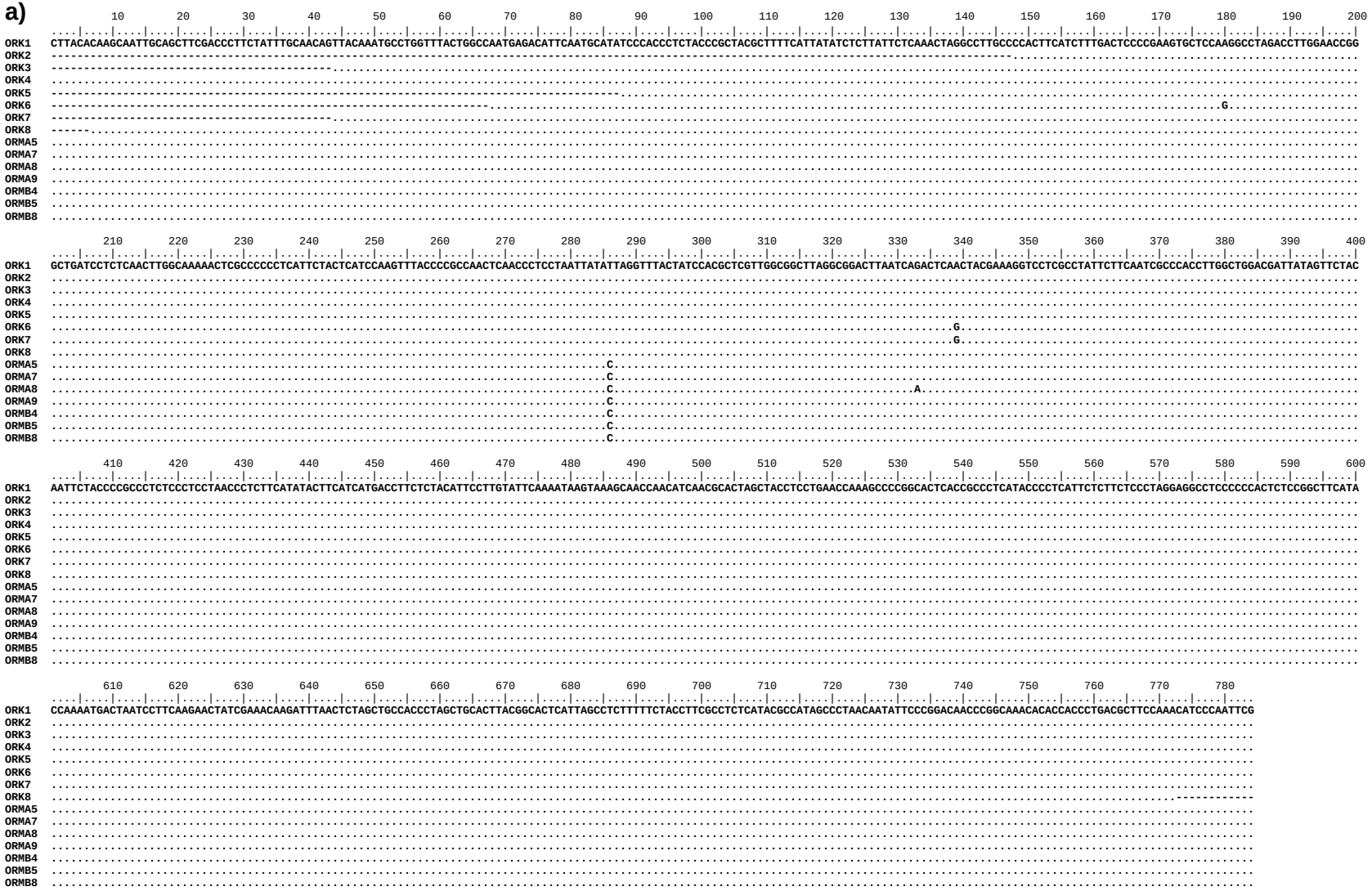
Appendix 2B Microsatellite primer information for loci tested for cross-amplification in *Otolithes ruber*. A: Amplification, N: No, Y: Yes, U: Non-specific, [$MgCl_2$]: Optimal $MgCl_2$ concentration, TA: Optimal annealing temperature, ?: Information is not known. The seven polymorphic markers are indicated in bold.

Name	Repeat type	Species	A	[$MgCl_2$] (mM)	TA (°C)	Polymorphic
Cne02D[1]	$(TCTA)_{31}$	*Cynoscion nebulosus*	N	-	-	-
Cne15D[1]	$(GGA)_{15}$	*Cynoscion nebulosus*	N	-	-	-
UBA03[2]	$(AC)_{18}$	*Argyrosomus japonicus*	Y	1.5	63	N
UBA05[2]	$(CT)_{16}$	*Argyrosomus japonicus*	Y	1.5	51	N
UBA06[2]	$(CA)_{13}$	*Argyrosomus japonicus*	Y	1.5	59	N
UBA42[2]	$(TGC)_{21}$	*Argyrosomus japonicus*	N	-	-	-
UBA851[2]	$(AG)_{14}$	*Argyrosomus japonicus*	Y	1.5	61	Y
UBA853[2]	$(GA)_{24}$	*Argyrosomus japonicus*	Y	1.5	52	N
Ajap06[3]	$(GGAT)_{?}$	*Argyrosomus japonicus*	Y	1.5	57	N
Ajap12[3]	$(ATCT)_{?}$	*Argyrosomus japonicus*	U	-	-	-
Ajap14[3]	$(ATCT)_{?}$	*Argyrosomus japonicus*	U	-	-	-
Ajap24[3]	$(AGAT)_{?}$	*Argyrosomus japonicus*	Y	2	45	N
Ajap37[3]	$(AGC)_{?}$	*Argyrosomus japonicus*	Y	2	45	N
Ajap40[3]	?	*Argyrosomus japonicus*	Y	2	45	N
JB01[4]	$(AC)_6$	*Johnius belengerii*	N	-	-	-
JB02[4]	$(AAG)_4$	*Johnius belengerii*	Y	1.5	53	N
JB07[4]	$(TG)_8$	*Johnius belengerii*	Y	0.75	53	N
JB09[4]	$(AGC)_5$	*Johnius belengerii*	Y	1.5	58	Y
JB11[4]	$(GCT)_6$	*Johnius belengerii*	U	-	-	-
Soc019[5]	$(GATA)_{16}$	*Sciaenops ocellatus*	N	-	-	-
Soc60[5]	$(AGG)_8$	*Sciaenops ocellatus*	Y	1.5	48	N
Soc86[5]	$(TGTC)_9$	*Sciaenops ocellatus*	Y	0.75	52	N
Soc129[5]	$(TATC)_{11}$	*Sciaenops ocellatus*	N	-	-	-
Soc133[5]	$(TGC)_8$	*Sciaenops ocellatus*	N	-	-	-
Soc137[6]	$(TGTC)_8$	*Sciaenops ocellatus*	N	-	-	-
Soc138[6]	$(TGTC)_6$	*Sciaenops ocellatus*	N	-	-	-
Soc140[5]	$(CTGT)_8$	*Sciaenops ocellatus*	Y	1.5	52	Y
Soc177[5]	$(TAGA)_{10}$	*Sciaenops ocellatus*	N	-	-	-
Soc201[5]	$(CCT)_6$	*Sciaenops ocellatus*	Y	1.5	62	N
Soc204[5]	$(CTG)_{12}$	*Sciaenops ocellatus*	Y	0.75	55	N
Soc206[6]	(GCAC)	*Sciaenops ocellatus*	U	-	-	-
Soc232[5]	$(AGAC)_4$	*Sciaenops ocellatus*	Y	0.75	52	Y
Soc243[5]	$(CCT)_9$	*Sciaenops ocellatus*	Y	1.5	61	Y
Soc247[5]	$(TAT)_7$	*Sciaenops ocellatus*	U	-	-	-
Soc445[7]	$(TCC)_{10}$	*Sciaenops ocellatus*	Y	1.5	54	N
Soc507[8]	$(GATA)_{24}$	*Sciaenops ocellatus*	U	-	-	-
Soc508[8]	$(GATA)_{18}$	*Sciaenops ocellatus*	Y	1.5	53	Y
Soc826[9]	$(GTGA)_{11}$	*Sciaenops ocellatus*	Y	1.5	61	Y

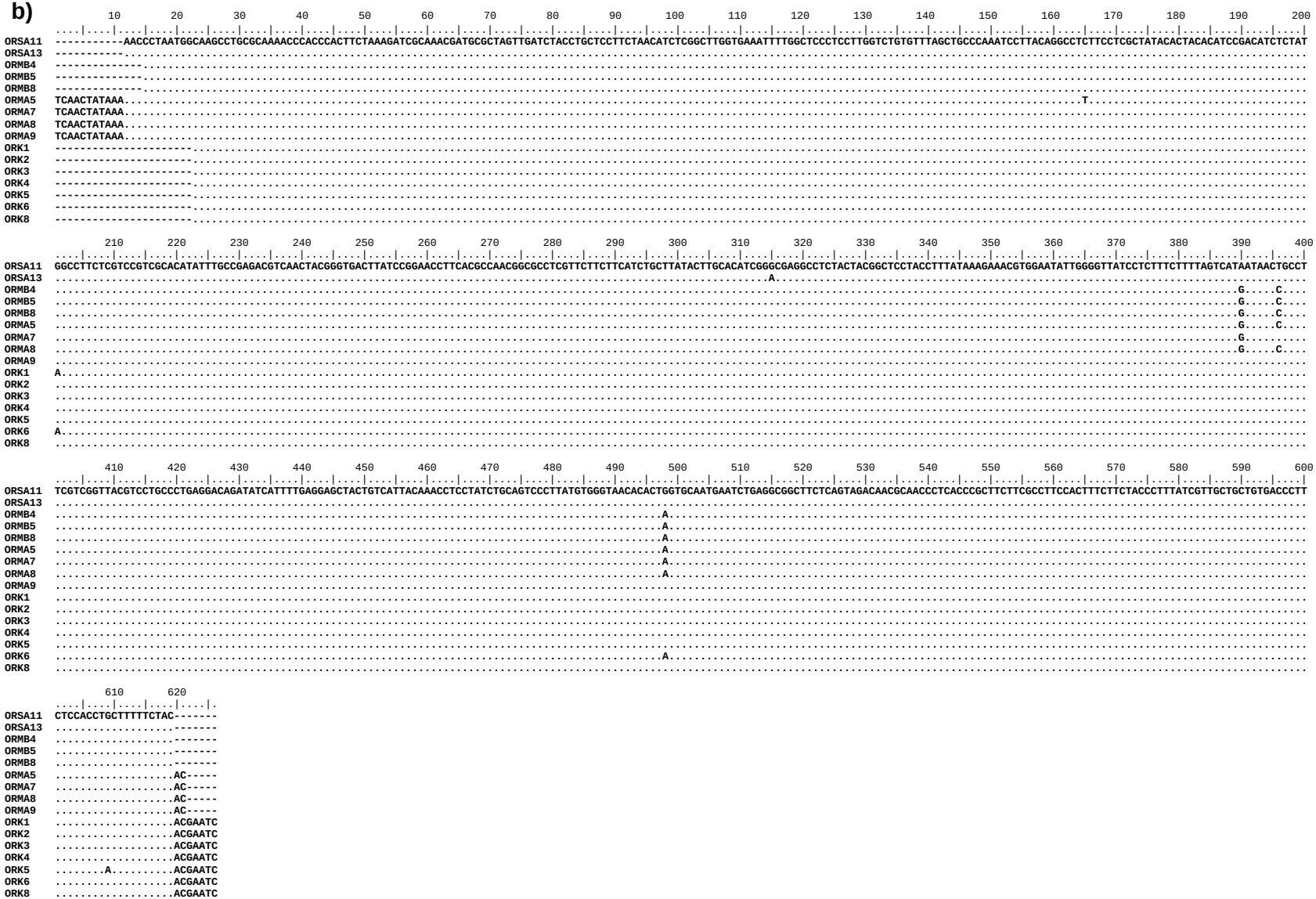
1: Blandon *et al.* (2011)
2: Archangi *et al.* (2009)
3: Mirimin *et al.* (2013)
4: Ma *et al.* (2011)
5: Turner *et al.* (1998)
6: Saillant *et al.* (2004)
7: O' Malley *et al.* (2003)
8: Karlsson *et al.* (2008)
9: Renshaw *et al.* (2012)

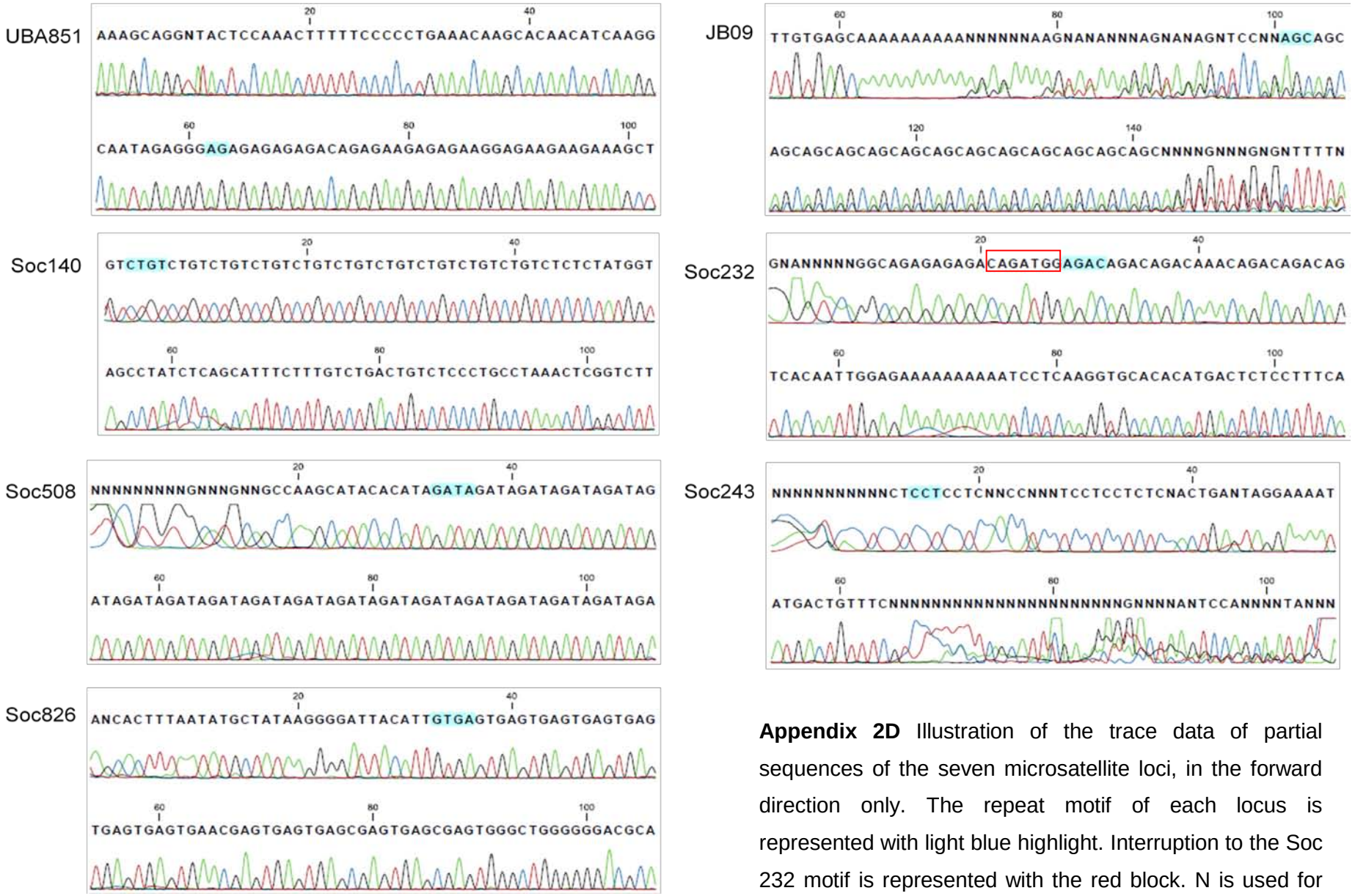
Appendix 2C Sequence alignment of *ND2* **(a)** and *Cyt b* **(b)** gene segments for selected *Otolithes ruber* samples. The following symbols are used: Identical = .; Missing = -.

a)



b)





Appendix 2D Illustration of the trace data of partial sequences of the seven microsatellite loci, in the forward direction only. The repeat motif of each locus is represented with light blue highlight. Interruption to the Soc 232 motif is represented with the red block. N is used for uncertain base calls.

Appendix 2E Microsatellite diversity in *Otolithes ruber* from for four locations. N: Number of samples, H_E: Expected heterozygosity, H_O: Observed heterozygosity, F_{IS}: Fixation index, N_A: Number of alleles recovered in *O. ruber* from each locality. "-": Due to the low concentration of the primer Soc826 in the multiplex combination SA samples were not amplified for the trial.

	Soc508	**Soc826**	**Soc243**	**Soc232**	**UBA851**	**Soc140**	**JB09**
K (N)	24	24	24	24	24	24	22
H_e	0.9140	0.7367	0.5186	0.0816	0.4965	0.8883	0.6786
H_o	1.0000	0.6250	0.2917	0.0833	0.5833	0.9583	0.6818
N_A	12	10	3	2	2	12	8
F_{IS}	-0.096	0.154	0.443	-0.022	-0.179	-0.081	-0.005
T (N)	12	12	12	12	12	12	12
H_e	0.9275	0.8080	0.2899	0.2899	0.2899	0.8696	0.7138
H_o	1.0000	0.5833	0.1667	0.3333	0.3333	0.8333	0.6667
N_A	11	8	2	2	2	9	6
F_{IS}	-0.082	0.287	0.436	-0.158	-0.158	0.043	0.069
SA (N)	23	-	24	24	24	22	22
H_e	0.9169	-	0.0417	0.1941	0.5922	0.9249	0.7526
H_o	0.8696	-	0.0417	0.2083	0.7917	0.9545	0.8182
N_A	14	-	2	3	3	13	6
F_{IS}	0.053	-	-0.000	-0.075	-0.347	-0.033	-0.089
MW (N)	12	12	12	12	12	12	12
H_e	0.9167	0.8841	0.0000	0.3659	0.6123	0.9130	0.7013
H_o	1.0000	1.0000	0.0000	0.4167	0.5833	0.8333	0.6364
N_A	13	9	1	3	4	12	8
F_{IS}	-0.095	-0.138	-	-0.146	0.049	0.091	0.097

Chapter 3

Historical isolation and contemporary isolation-with-migration in the South West Indian Ocean sciaenid, *Otolithes ruber*

Abstract

The South Western Indian Ocean (SWIO) region has reported environmental degradation and reductions in natural resources and biodiversity; it also has complex oceanography which might affect the distribution of its inhabitants, also influenced by historical processes. *Otolithes ruber* is a demersal, sciaenid fish inhabiting this region. It is an excellent protein source and is caught in several fishery sectors in East Africa and Madagascar, including as incidental by-catch on trawlers. There is increasing concern regarding overexploitation of this species and yet not much is known about its biology. The aim of the study was to investigate the geographic genetic population structure of *O. ruber* from the SWIO and to identify the underlying historical and evolutionary processes affecting its populations. Genetic variation was examined in 223 specimens from seven localities based on the mitochondrial cytochrome *b* gene and seven microsatellites. Both marker types were able to demonstrate the existence of two populations in the northern and southern regions of the SWIO, respectively. Madagascar needs further investigation; it might also represent a separate population based on some of the analyses. The coalescent model showed that there is a signature of historical divergence among these subpopulations but they are connected by contemporary gene flow, with more migration from the northern to the southern subpopulation. The analyses also showed that male-biased migration might be a life history trait of the species. The oceanographic features of the region, as well as past and contemporary demographic processes, likely play major roles in determining the degree of connectivity between populations of *O. ruber*.

Keywords: Historical demography, gene flow, isolation with migration, cytochrome *b*, microsatellites, *Otolithes ruber*, South West Indian Ocean.

3.1 Introduction

The contemporary geographic distribution of individuals in terms of patterns of gene genealogy has been investigated in many marine species. The high dispersal potential of larvae and the open nature of the marine environment mostly suggest high levels of gene flow, irrespective of possible juvenile and adult behaviour (Huijbers 2013). Hence, homogenized gene pools with little to no genetic differentiation commonly exist (Gonzalez & Zardoya 2007; Von der Heyden *et al.* 2007; Was *et al.* 2010), also due to large effective population sizes restricting genetic drift (Waples 1998). In contrast, life-history traits such as small spatial scale of recruitment, philopatric behaviour or natal homing, local adaptations and local larval retention act as drivers of isolation by distance and give rise to significant levels of genetic differentiation among some fish populations (Kelly & Palumbi 2010). The duration of egg and larval stages is another important factor, since longer periods spent in the pelagic environment enhance opportunities for gene flow (White *et al.* 2009). For marine organisms with sedentary adult stages, larval dispersal is the source of population connectivity (Nicastro *et al.* 2008; Galarza *et al.* 2009).

Ocean fronts, currents and eddies are physical irregularities of the marine environment which influence connectivity (Gonzalez & Zardoya 2007; Viñas *et al.* 2010; White *et al.* 2010), while bathymetry could also affect connectivity of demersal species' populations (Knutsen *et al*. 2009). Eddies act as vehicles for transporting plankton and may, thus, create conditions enhancing larval survival since they contain high concentrations of phyto- and zooplankton (Durgadoo *et al.* 2010). Eddies are also responsible for drawing off shelf waters into deeper oceanic regions (Lutjeharms 2006); they therefore promote the retention of eggs and larvae, protecting them from being transported by ocean currents and may lead to the isolation of populations (Hauser & Carvalho 2008). However, retaining of eggs and larvae in eddies also can provide a mechanism for transport, leading connectivity. In addition to contemporary processes, the genetic differentiation and geographic distribution of marine fishes are also impacted by historical events linked to climatic and sea level changes (Martínez *et al.* 2006; Lessios 2008; Gaither *et al*. 2011), geographical fragmentations, and climate-associated periodical extinctions (Gonzalez & Zardoya 2007).

The model organism of this study was *Otolithes ruber* (Bloch & Schneider 1801), a demersal sciaenid fish species distributed throughout the tropical Indo-West Pacific (IWP), including the South West Indian Ocean (SWIO) region (Fig. 3.1) (Brash & Fennessy 2005). In the SWIO region, large-scale population studies have been done mostly on reef fishes (Craig *et*

al. 2007; DiBattista *et al.* 2013) and wide-ranging pelagic species, e.g. tuna (Durand *et al.* 2005; Theisen *et al.* 2008). Hence, there is a gap in genetic studies for a species such as *O. ruber,* which has the potential to provide different perspectives on connectivity because of its habitat/life history. *Otolithes ruber* is a relatively inactive carnivore, inhabiting turbid water over sandy and muddy substrates (van der Elst 1981; Brash & Fennessy 2005). In the SWIO region the species has a coastal distribution, with a narrow depth range (15-45 m; Olbers & Fennessy 2007). Some important spawning and nursery areas have been identified on the Thukela Bank off South Africa (Fennessy *et al*. 1994; Fennessy 2000) and the Sofala Bank off Mozambique (Schultz 1992). Even though fecundity is high, natural mortality of early life stages is also high (Navaluna 1982; Olbers & Fennessy 2007). Crucial information pertinent to population structuring, such as larval duration, dispersal and adult movement patterns, is lacking. *Otolithes ruber* is an excellent protein source, targeted by trawlers, gill-nets, beach seines, and line and hook fishing off South Africa, Mozambique, Tanzania and Kenya (Brash & Fennessy 2005). In addition, it is also caught as by-catch on penaeid prawn trawlers (Heemstra & Heemstra 2004; Olbers & Fennessy 2007). A rising concern for this species is overfishing in the SWIO region (Schultz 1992; Olbers & Fennessy 2007).

The SWIO region, which borders Kenya and the area southwards within the Western Indian Ocean region (Fig. 3.1), is regarded as a biogeographic subregion of the IWP showing high levels of fish species diversity and endemism, believed to be due to the variety of habitats (Briggs & Bowen 2012; Hoareau *et al.* 2013) and the oceanographic conditions of the region (Quartly & Srokosz 2004). In the tropical part of the SWIO region, the South Equatorial Current (SEC) which flows from east to west in the Indian Ocean divides into two main circulation components when it reaches the eastern coast of Africa (Fig. 3.1); the East African Coastal Current (EACC) comprises the northern component (Fratini *et al.* 2010). These currents are expected to separate populations off Tanzania and Kenya from the populations off Mozambique and South Africa (Fratini *et al.* 2010). The SEC is likely responsible for larval transport from the Indian Ocean islands (Mascarene Plateau and further east in the IWP) to the coast of the East African mainland, whereas the EACC transports larvae along the coast in a northerly direction (Visram *et al.* 2010). In the Mozambique Channel, oceanographic features show variability, with mean circulation southwards along the continental slope and northwards in the eastern part (Matano *et al.* 2008). Further south in this region, the Agulhas Current (AC) is the main feature (Quartly & Srokosz 2004). It flows very close to the shelf edge in the northern part of its flow route. The oceanographic features are expected to play a major role in determining fish population functioning and connectivity within the region.

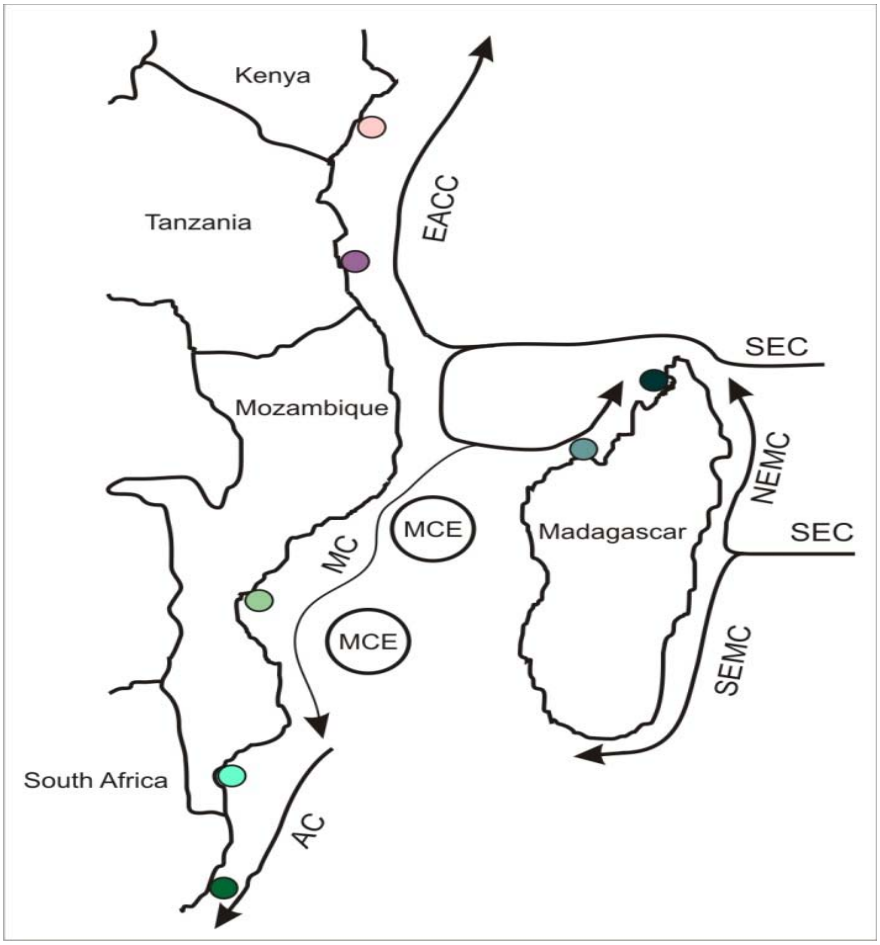


Fig. 3.1 Oceanographic features of the South West Indian Ocean region and the sampling localities of *Otolithes ruber* in the present study. SEC: South Equatorial Current, NEMC: North East Madagascar Current, SEMC: South East Madagascar Current, EACC: East African Coastal Current, MC: Mozambique Current, AC: Agulhas Current, MCE: Mozambique Channel Eddies. The map was modified from Ragionieri *et al.* (2010). The sampling locations are indicated by coloured circles: South Africa (SA) (green), south Mozambique (MS) (light blue), central Mozambique (MC) (faded green), west Madagascar (MW) (ocean green), northwest Madagascar (MNW) (dark green), Tanzania (T) (purple) and Kenya (K) (faded pink).

Broad-scale phylogeographic patterns have been investigated for several taxa in the IWP (e.g. for fish species: Bay *et al*. 2004; Fitzpatrick *et al*. 2011; Gaither *et al*. 2011). Some authors have suggested limited gene flow between the WIO and the rest of the Indo-Pacific (Ridgway & Sampayo 2005). Subdivision has been reported between the West and East Indian Oceans (Williams *et al*. 2002) and studies have revealed varying degrees of connectivity within the WIO (Visram *et al*. 2010; Muths *et al*. 2011, 2012). To date, genetic variation within *O. ruber* throughout the IWP has not been investigated, although a recent

DNA barcoding study of fish species from South African and Australian waters (Zemlak *et al.* 2009) suggests that many widespread species likely harbour undiscovered biodiversity. A few samples of *O. ruber* from the two regions were compared using the standard cytochrome *c* oxidase 1 (*CO1*) barcoding gene and revealed species-level rather than population-level divergence between the two regions (Zemlak *et al*. 2009). There are no published studies on connectivity among sciaenid populations from the SWIO, but some evidence from other regions is informative.

Population substructuring was reported throughout the Australian range of the estuarine and coastal mulloway (*Argyrosomus japonicus*), ranging from Queensland to Western Australia (Silberschneider & Gray 2008). On a broader scale, a phylogenetic investigation based on both mtDNA and nuclear DNA sequences showed that *A. japonicus* from Australia is distinct from the fragmented populations of the species from the northwestern Pacific and the WIO off South Africa (Klopper, Whitfield and Bloomer, unpublished data). King fish (*Macrodon ancylodon*) from Brazil has two distinct evolutionary lineages (tropical and subtropical) that appear to be related to temperature variation and oceanic currents (Santos *et al.* 2006). Moreover, two recent studies also indicated the presence of two evolutionary units for the meagre (*A. regius*) from the Atlantic and Mediterranean (Haffray *et al.* 2012), and the southern king fish (*M. atricauda*) (Rodrigues *et al.* 2013) from south Brazil and Argentina based on microsatellite and mitochondrial control region data, respectively.

No migratory behaviour has been recorded in *O. ruber* and barcoding suggests subdivision of at least Western and Eastern Indian Ocean populations on a broad scale (Zemlak *et al*. 2009). The ocean current systems of the SWIO can be predicted to have a significant influence on connectivity patterns in this species, especially with respect to passive dispersal of eggs and larvae. The system of eddies that dominate the Mozambique Channel (Schouten *et al.* 2003) likely facilitates both movements from, and retention at spawning and nursery areas. Given the conservation concern for the species, the need of sustainable use of the resource and absence of any genetic assessment conducted to date, the aim of the present study was to investigate the geographic genetic structure of *O. ruber* among localities within the SWIO region, and the underlying processes of the observed patterns. I evaluated genetic differentiation, historical demography, levels and patterns of current nuclear gene flow and historical female gene flow using nuclear microsatellite markers and the mitochondrial cytochrome *b* gene (*Cyt b*), respectively. Mitochondrial DNA is mainly suitable for detecting historical events, thus informing over an evolutionary rather than an ecological time scale (Hickerson *et al*. 2010), while highly variable microsatellites can detect

contemporary population genetic structure and have been demonstrated to be more efficient and informative in marine fishes (Selkoe & Toonen 2006). Hence, I discussed the variation from the perspective of two time scales within the context of the current geographic distribution of the species. A coalescent approach, isolation with migration, was applied, enabling analysis of the data in a more natural framework (Hey & Nielsen 2007), since the model lacks the assumption of mutation and genetic drift equilibrium.

3.2 Materials and Methods

3.2.1 Sampling information and DNA extractions

Fin tissue samples of *O. ruber* were obtained, as per sampling protocols developed during SWIOFP planning workshops, from seven locations throughout the SWIO (Fig. 3.1): South Africa (SA, Thukela Bank: 40), south Mozambique (MS, Maputo Bay/Machangulo, Inhaca: 33), central Mozambique (MC, Savane, Sofala Bank: 30), Tanzania (T, off Kisiju: 30), Kenya (K, Ungwana Bay, Malindi: 30), west Madagascar (MW, Mahajanga: 30), north-west Madagascar (MNW, Nosy Be: 30). DNA was extracted with the DNeasy Blood and Tissue extraction kit (Qiagen), according to the manufacturer's instructions. The quality of extracted DNA was confirmed with mtDNA 16*S* rDNA amplification. Polymerase Chain Reactions (PCR) were carried out in 10 µL volumes containing 1 x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Promega), 0.5 pmol primers (Whitehead Scientific) 16Sar-5' and 16Sbr-3' (Palumbi 1996; Ivanova *et al.* 2007), 0.04 U *Taq* polymerase (Southern Cross Biotechnologies) and approximately 30 ng genomic DNA. PCR conditions were: 94 °C for 4 min followed by 94 °C for 30 s, 60 °C for 1 min, 72 °C for 1 min repeated for 40 cycles, and a final extension step of 72 °C for 7 min. The PCR products were electrophoresed in 2% agarose gels (Separations) and visualised with GelRed™ stain (Biotium) under UV light applied in Molecular Imager, Gel Doc™ XR⁺ Imagining System (Bio Rad).

3.2.2 Mitochondrial DNA amplification and sequencing

A 624 base pair (bp) cytochrome *b* fragment was amplified in 20 randomly chosen individuals per site using the primers L14724 (Palumbi *et al.* 1991) and H15499 (Avise *et al.* 1994), which are widely used universal primers for vertebrates (Whitehead Scientific). PCR reactions were performed in 25 µL volumes containing 1 x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Promega), 0.25 pmol of each primer (Applied Biosystems), 0.04 U Supertherm *Taq* polymerase (Southern Cross Biotechnologies) and approximately 90 ng of genomic

DNA. Cycling conditions were: 94 °C for 2 min followed by 94 °C for 30 s, 51 °C for 30 s and 72 °C for 45 s, repeated for 35 cycles, followed by a final elongation step of 72 °C for 5 min. The PCR products were visualised as explained above.

The amplified products were precipitated using 2.8 volumes of absolute ethanol, 0.4 volumes of Sabax water and 0.08 volumes of 3 M sodium acetate (NaAc), followed by a washing step using 70% ethanol. The products were eluted in 15 µL of Sabax water and the presence of precipitated product confirmed on a 2% agarose gel by electrophoresis. Cycle sequencing reactions were performed by using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (Applied Biosystems). The manufacturer's recommended conditions were used for the cycling reaction and it contained approximately 30 ng purified DNA, 1 µL BigDye reaction mix, 5 x sequencing buffer and 3.2 pmol forward or reverse primer. The products were precipitated as described above. Sequences were separated on an ABI 3130 sequencer (Applied Biosystems) and samples were sequenced in both forward and reverse directions. Trace data analyses of sequences were done using CLC Bio^{R}. The forward and reverse sequences of every sample were aligned in BioEdit Sequence Alignment Editor (Hall 1999) to create a consensus sequence. Consensus sequences for all samples were aligned using ClustalW (Thompson *et al.* 1994).

3.2.3 Mitochondrial DNA data analyses

3.2.3.1 Statistical analysis from mtDNA variations

Mega 5 (Tamura *et al.* 2011) was used to investigate the segregating sites in the alignment to identify unique haplotypes. The level of DNA polymorphism was estimated by determining the haplotype diversity (*h*; the probability that two randomly chosen haplotypes are different) and nucleotide diversity (π; the proportion of homologous nucleotides that are different between any two copies of a gene) using DNA Sequence Polymorphism (DnaSP) version 5.1 (Librado & Rozas 2009). Population growth is a demographic event which can be traced from DNA sequence data with the assumption that selection did not act on the gene. Statistical tests, such as neutrality tests, can be used to investigate this and also provide additional information about historical changes in population size. Two neutrality tests, Tajima's D (Tajima 1989) and Fu's Fs (Fu 1997), were performed in ARLEQUIN version 3.11 (Excoffier *et al.* 2005). From a demographic perspective, negative values of Tajima's D indicate a population expansion after a recent bottleneck event, while positive values reflect admixture of two distinct populations (Tajima 1989). Fu's Fs test is based on the information from allele distributions; the value is negative when there is an excess of low frequency

variants (Fu 1997). The demographic inferences are made using these tests in the absence of selection. A mismatch distribution analysis (an analysis of the frequency distribution of pairwise differences among all individuals) was performed using ARLEQUIN, with a statistical evaluation [sum of squared deviation and Hardending's (1994) raggedness index] of the observed data relative to a null model of sudden expansion (Harpending 1994; Schneider & Excoffier 1999).

3.2.3.2 Evaluating genetic differentiation

Population differentiation based on pairwise F_{ST} values was calculated in ARLEQUIN (Excoffier *et al.* 2005). Significance levels for multiple tests were corrected using Bonferroni corrections (Rice 1989). A haplotype network, to visualise the frequency and spatial distribution of unique maternal alleles, was constructed using TCS (Clement *et al.* 2000). TCS assesses the genealogical relationship between sequences based on statistical parsimony and only connects haplotypes with a 95% confidence limit.

An Analysis of Molecular Variance (AMOVA) (Excoffier *et al.* 1992) was performed using ARLEQUIN to estimate variance of haplotypic partition between populations (Excoffier *et al.* 1992). Three variance components and associated F-statistics were evaluated: Among groups (F_{CT}), among populations within groups (F_{SC}) and within populations (F_{ST}). Significance of these components and the associated summary statistics were determined through 1000 random permutations. Different groupings were tested: North (K, T) versus south (SA, MS, MC, MNW, MW); and north (K, T) versus south (SA, MS, MC) versus Madagascar (MNW, MW). Different grouping arrangements were decided based on the haplotype network and the oceanographic knowledge of the region.

Isolation by distance (IBD) was tested using GENEPOP (Rousset 2008). The geographic distance between sampling locations was determined according to coordinates obtained from Google Earth, using the shortest distance between two points along a water path. The genetic distance, $F_{ST}/(1-F_{ST})$ (Rousset 1997), was calculated using GENEPOP. The Isolde option was used to determine IBD with 1000 random permutations. This option applies a Mantel test (Mantel 1967) to examine the correlation between genetic distance ($F_{ST}/(1-F_{ST})$) and geographic distance (distances between locations in kilometers).

3.2.3.3 Historical demography and pattern of female gene flow

The program MIGRATE-n version 3.3.1 (Beerli & Felsenstein 2001; Beerli & Palczewski, 2010) was used for estimating the theta parameter (θ = Nefμ, where Nef is the effective female population size and μ is the mutation rate per site per generation) for each population (south population = SA, MS, MC, MW, MNW, and north population = K and T) and gene flow between populations (M = m/μ, which is a measure of how much immigration introduces new variants into a population compared to mutation; m is the migration rate) (Visram *et al.* 2010). A Bayesian search strategy was selected, as explained in Beerli (2006), with five replicates of one long chain sampling across 2.5 million genealogies. Theta and M values were generated from the F_{ST} calculation with two possible options: θ variable, M symmetric (F_{ST}-theta) vs. M variable, θ the same (F_{ST}-migration). A full migration matrix model was selected. For analyzing a scenario of three populations (south, Madagascar and north) only the F_{ST}-theta option is available.

Population parameters were also estimated using the Isolation with Migration (IMa) coalescent framework (Hey & Nielsen, 2007) for the southern population (SA, MS, MC, MW, MNW) and the northern population (T and K). The estimated parameters included: Θ_1 (effective population size, Ne, estimate of population 1), Θ_2 (Ne estimate of population 2), Θ_A (Ne estimate of ancestral population), m_1 (m, migration rate, estimate from population 2 to 1), m_2 (m, migration rate, estimate from population 1 to 2) and t (t, time since population split estimate). The Hasegawa-Kishino-Yano (HKY) model (Hasegawa *et al.* 1985) was used with an inheritance scalar of 0.25. The HKY model takes into account multiple mutations at each site, differences in nucleotide frequencies and the presence of transition/transversion bias (Faulks *et al.* 2010). Initially, broad prior probabilities were used for each analysis. Thereafter, for every parameter, upper boundaries reduced (when a well-defined probability plot was obtained) or increased (when the estimated probability was not approaching zero) in repeated runs to determine the optimized probabilities. Convergence (good mixing) of the analysis was assessed by monitoring effective sample size (ESS) values (the lowest value being 50) and by examining the parameter trend plots. After optimization, the final/fixed priors for the parameters were: *q1* 20, *q2* 10, *qa* 10 (priors of *Θ*), *m1* 10, *m2* 10 and *t* 8, applying a 10 chain geometric heat mode (g1 0.94, g2 0.9), with 45 chain swaps per step, 10 million burn-ins and 100 000 genealogies. This final run was repeated three times to ensure consistency of the estimated values. Conversions of the estimated parameters to demographic quantities (Ne, m and t) were performed using a mutation rate (μ) scale. Using the five most accurate divergence rate estimates for the *Cyt b* gene in fish (Lessios 2008),

an average (3.92%/million years) was applied as the mutation rate calculation for the present study; therefore, $\mu = 624 \times 1.96\%/10^6$ (number of bp of the gene region x the divergence rate within the lineage). Generation time was not included in μ calculations for t and m, since the scale should be in years. However, a generation time of three years (biological basis) was included for Ne (in the coalescent models Ne is proportional to the inverse of the coalescent rate per generation; Hey & Nielsen (2007).

3.2.4 Microsatellite cross-species amplification

In total, 38 markers described as polymorphic in closely related species were tested for reliable amplification and polymorphism in *O. ruber* (see Chapter 2). Of these, seven were found to be polymorphic in the species. These seven loci were genotyped in all individuals using the Quantitect Multiplex PCR kit (Qiagen) in a single multiplex combination with the recommended conditions of the manufacturer. These fragments were analysed on an ABI 3500xl genetic analyzer (Applied Biosystems) with the GeneScan Liz^{TM}500 Size Standard (Applied Biosystems). All loci were used for further analyses; Soc232 was not included in the IMa analysis due to requirements of the programme (loci with the repeat number ≤ four are excluded). However, with or without this locus the same pattern was observed.

3.2.4.1 Microsatellite data analyses and summary statistics

Scoring of genotypes was done using GENEMARKER version 1.5 software (SoftGenetics, State College, Pennsylvania, USA). Summary statistics for each location were calculated in GENETIX version 4.05 (Belkhir *et al.* 1996-2004), including observed heterozygosity (H_O), expected heterozygosity (H_E) and the inbreeding coefficient F_{IS}. Allelic richness (AR), allelic diversity/mean number of alleles per locus, was calculated in FSTAT version 2.9.3.2 (Goudet 2002) based on a minimum sample size of 24 diploid individuals per locality. Deviations from Hardy-Weinberg Equilibrium (HWE) were calculated with GENEPOP using 1000 dememorization steps, 100 batches and 1000 iterations per batch. MICROCHECKER version 2.2.3 (van Oosterhout *et al.* 2004) was used to test for the presence of null alleles and large allele dropouts. Linkage Disequilibrium (LD) between markers was tested in GENEPOP using 1000 dememorization steps, 100 batches and 1000 iterations per batch.

3.2.4.2 Evaluation of genetic differentiation

Pairwise F_{ST} values was calculated in GENETIX (Belkhir *et al.* 1996-2004) to investigate the population differentiation and Bonferroni corrections (Rice 1989) were used to correct the significance levels for multiple tests. A Factorial Correspondence Analysis (FCA) was performed in GENETIX (Belkhir *et al.* 1996-2004). This analysis illustrates the relationship among individuals as a three-dimensional visual clustering, expressing the differences as factors. It is based on the pairwise genetic distance between individuals and sampling location was treated as a prior.

To estimate the most likely number of populations in the sample, the program STRUCTURE version 2.3.2 (Pritchard *et al.* 2000) was used with sampling localities as prior information. It was performed using K = 1 to 5 with burn-in length of 200 000 and a MCMC of 200 000 steps for 10 iterations. The log probability of the data (LnK) and the standard deviation across the 10 iterations was plotted for these K-values to determine the highest likelihood value. In addition, ΔK indicates the rate of change of the log probability of the data for tested K values and the largest change indicates the most likely number of populations (Evanno *et al.* 2005). An admixture model with correlated allele frequencies was used. For assignment of individuals to the inferred clusters, chains of 1×10^6 were run three times to ensure convergence. Individuals assigned with a probability (q) of 0.8 were regarded as belonging to a single cluster, while values of <0.8 were inferred as admixture.

AMOVA was performed using ARLEQUIN with 1000 permutations. It was used to evaluate the level of genetic structuring of defined populations. The same groupings as defined for the mtDNA analysis were evaluated for the microsatellite data. Isolation by distance and migration were evaluated as explained in the mtDNA section.

3.2.4.3 Historical demography and nuclear gene flow pattern

As explained in the mtDNA section, IMa was used to investigate the demographic population parameters in the observed southern and northern populations of *O. ruber*. The Stepwise Mutation Model (SMM) (Kimura & Ohta 1978), which is the recommended model for microsatellite data, was used with an inheritance scalar of 1 for nuclear DNA. Following optimization the final/fixed priors were used for the parameters: *q1* 2, *q2* 1, *qa* 10, *m1* 40, *m2* 50 and *t* 4, applying a 10 chain geometric heat mode (g1 0.94, g2 0.9), with 45 chain swaps per step, 10 million burn-ins and 10 000 genealogies. This final run was repeated three times

to ensure consistency of the estimated parameter values. To convert the estimated parameters to demographic quantities (Ne, m and t), a mutation rate of $1x10^{-4}$ was used. This value is considered as the average microsatellite mutation rate over many species (Yue *et al.* 2007; Fontaine *et al.* 2010; McCusker & Bentzen 2010).

3.3 Results

3.3.1 Mitochondrial DNA data analysis and summary statistics

The 624 bp fragment of the cytochrome *b* gene showed 27 polymorphic sites (15 singletons; all transitions). Polymorphism levels were high as indicated by the number of haplotypes per locality and the haplotype diversity ranging from 0.195 to 0.805; nucleotide diversity values were low with a range from 0.0005 to 0.0031 (Table 3.1). Tanzania showed the lowest genetic diversity, whereas the northern location in Mozambique (MC) showed the highest values. With the exception of MS (negative but not significant), MC and SA, Tajima's D and Fu's Fs values (Table 3.1) were negative and significant for all other locations indicating an excess of low frequency variants. A mismatch distribution analysis was used to examine the historical demography of *O. ruber*. The mismatch distribution shows a unimodal distribution (Fig. 3.2) and a low and non-significant value for the sum of squared deviations (SSD = 0.006, p = 0.517), indicating that the null hypothesis of an expansion could not be rejected. The demographic expansion is confirmed by the low Harpending's (1994) raggedness index (0.026), indicative of a smooth distribution.

Table 3.1 Mitochondrial DNA cytochrome *b* diversity indices for seven South West Indian Ocean localities of *Otolithes ruber*. N: Number of sequences, N_{hap}: Number of haplotypes, N_{uni}: Number of unique haplotypes, *h*: Haplotype diversity, π: Nucleotide diversity, neutrality tests (Tajima's D and Fu's Fs): Significant values are in bold *: p < 0.05, ***: p < 0.001. The locality designations follow Fig. 3.1.

Locations	N	N_{hap}	N_{uni}	*h*	π	Tajima's D	Fu's Fs
Kenya	20	9	8	0.753	0.0018	**-1.67***	**-6.02*****
Tanzania	20	3	2	0.195	0.0005	**-1.72***	**-1.14***
Madagascar northwest	20	8	3	0.747	0.0020	**-1.52***	**-4.07*****
Madagascar west	20	8	4	0.732	0.0017	**-1.82***	**-4.86*****
Mozambique central	20	5	0	0.805	0.0031	1.08	0.56
Mozambique south	20	6	3	0.768	0.0018	-1.05	-1.89
South Africa	20	5	1	0.758	0.0025	0.37	0.05
Overall	140	29	21	0.836	0.0034	**-1.69***	**-20.61***

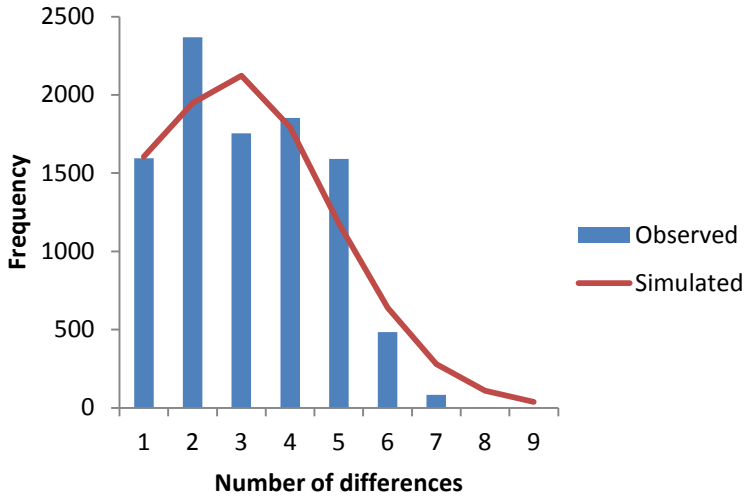


Fig. 3.2 Mismatch distribution of pairwise cytochrome *b* sequence differences observed among 140 *Otolithes ruber* from the South West Indian Ocean. The solid line represents the expected curve under a population expansion model. The sum of squared deviations (SSD = 0.006 p = 0.517) and a low Harpending's (1994) raggedness index (r = 0.026, p = 0.026) indicate a good fit to the model.

3.3.2 Genetic differentiation based on mtDNA

A haplotype network, constructed using 140 sequences (i.e. *O. ruber* individuals), showed a star-like shape with the most common haplotypes at the center connecting with lower-frequency haplotypes. In total, 29 maternal haplotypes were observed and several localities were characterised by unique haplotypes. Two mtDNA haplogroups were mainly represented in the northern and southern parts of the region, respectively (Fig. 3.3). Star-shaped haplotype networks for both haplogroups suggested population expansion events.

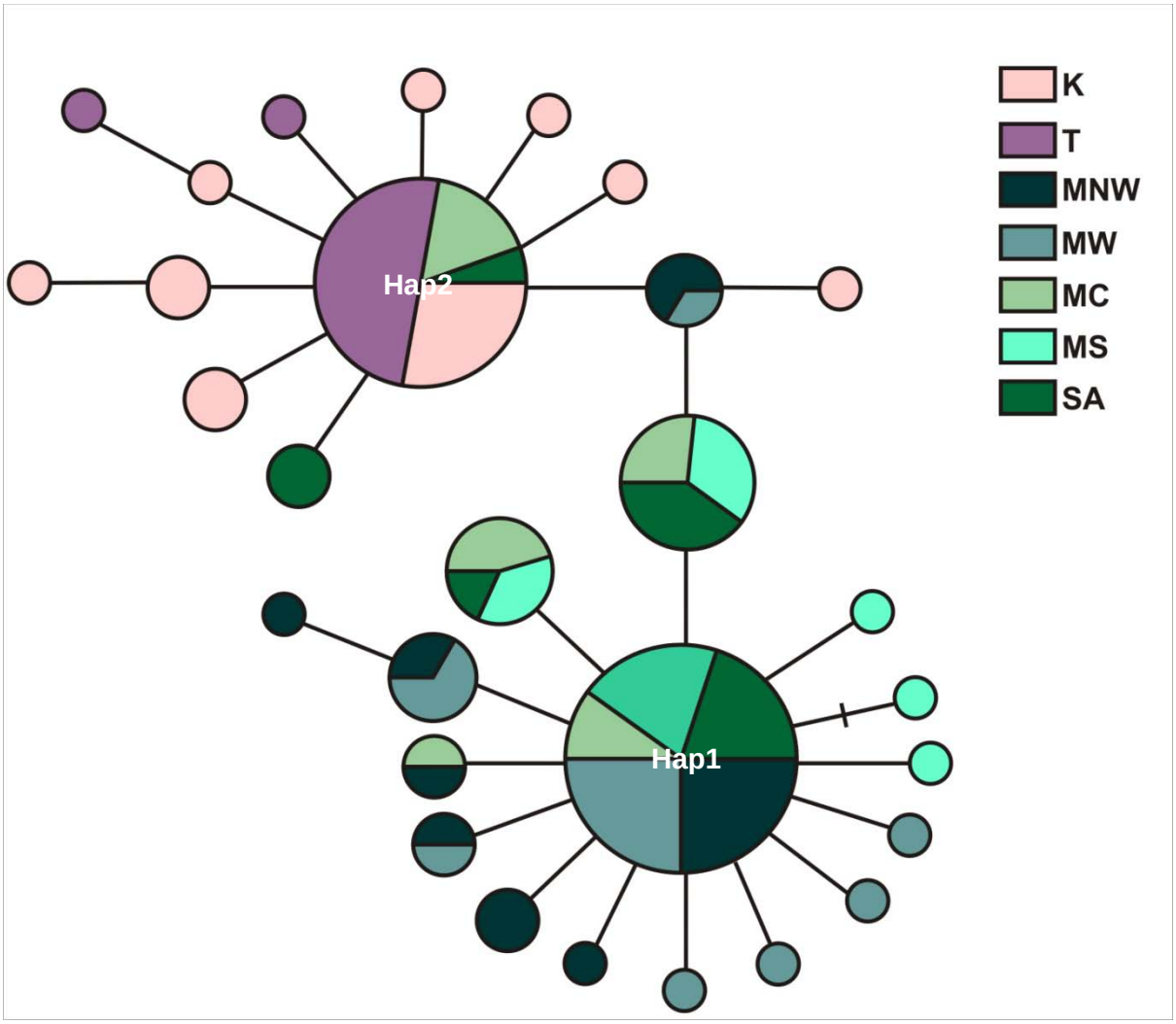


Fig. 3.3 Haplotype network showing the evolutionary relationships between mtDNA cytochrome *b* haplotypes of 140 *Otolithes ruber* from the South West Indian Ocean. Each unique maternal haplotype is indicated by a circle and the size of each circle is proportional to the number of individuals sharing the haplotype. The colours designate the seven sampling localities (as indicated in the legend and Fig. 3.1). Connecting lines between

haplotypes indicate one mutational step. TCS assigned the haplotype 1 (Hap1) as the ancestral haplotype.

The observed pattern of distinct haplogroups from the haplotype network was confirmed with AMOVA. Based on the haplotype network and oceanographic knowledge of the region, two grouping scenarios were evaluated (Table 3.2; Fig. 3.4). For the scenario of two population groupings (K, T versus SA, MS, MC, MNW, MW), the largest percentage of the overall variance could be attributed to differentiation among the two defined groups (Table 3.2). There was low but significant differentiation among populations within these two groups and a high level of variation within populations. When considering Madagascar (MNW, MW) as a third group, variation between the three groups and variation within populations were both significant, but the variance among populations within groups was not significant.

Table 3.2 Analysis of Molecular Variance based on 140 cytochrome *b* sequences of *Otolithes ruber* and two groupings of the samples (see Fig. 3.4). GV%: Percentage contribution to the overall genetic variance, df: Degrees of freedom. Significant values are shown in bold; *: p<0.05, ***: p<0.001.

Source of variation	**df**	**GV%**	**Fixation index**	**p-value**
(K,T**) (**MNW, MW, MC, MS, SA**)**	1	60.34	F_{CT}: 0.6034	**0.0489***
Among populations within groups	5	3.98	F_{SC}: 0.1005	**0.0000*****
Within populations	133	35.68	F_{ST}: 0.6432	**0.0000*****
(K, T**) (**MNW, MW**) (**MC, MS, SA**)**	2	52.61	F_{CT}: 0.5261	**0.0176***
Among populations within groups	4	1.75	F_{SC}: 0.0370	0.0763
Within populations	133	45.63	F_{ST}: 0.5437	**0.0000*****

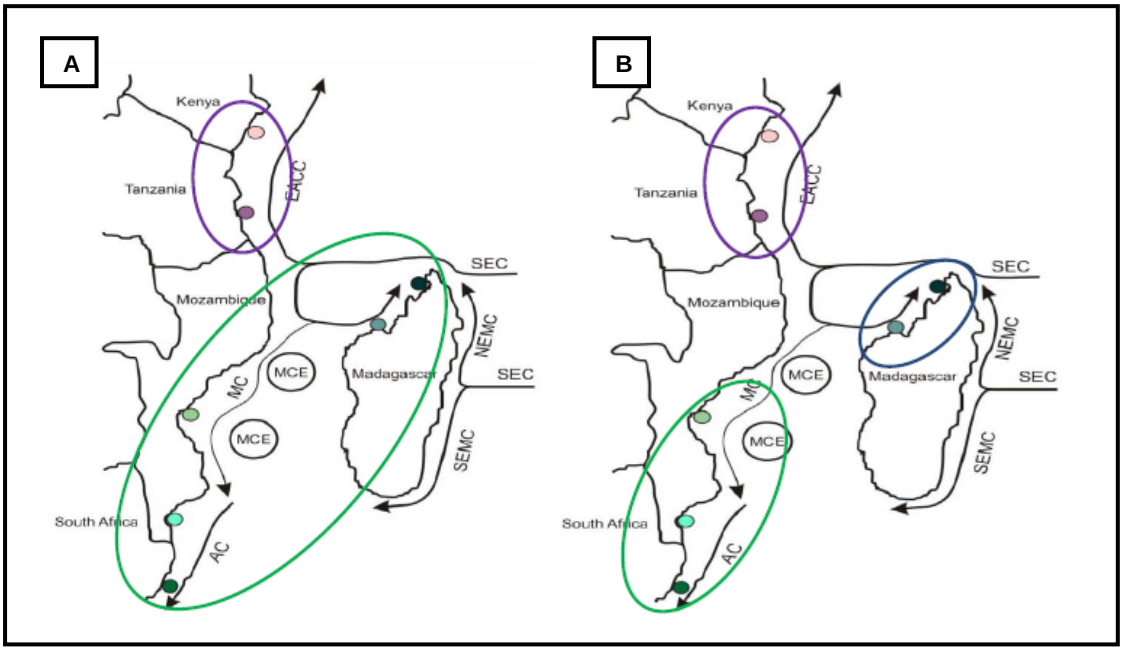


Fig. 3.4 Illustration of the geographic partitions that showed significant differentiation based on an Analysis of Molecular Variance (AMOVA) of cytochrome *b* data among seven localities of *Otolithes ruber* from the South West Indian Ocean. **A)** north vs south and **B)** three groups (north, Madagascar and south).

Pairwise F_{ST} values (Table 3.3) confirm the AMOVA results and show that Kenya and Tanzania samples are highly differentiated from the rest of the localities. In most of the pairwise comparisons, the two Madagascar localities are also significantly differentiated from the localities from southern Africa. There is a lack of significant differentiation between neighbouring locations. The test for isolation by distance did not show a statistically significant correlation between genetic and geographic distance (Fig. 3.5; $p = 0.094$), the null hypothesis of no relationship between genetic distance and physical distance of individuals was not rejected.

Table 3.3 Population differentiation among seven South Western Indian Ocean locations of *Otolithes ruber* is illustrated by pairwise F_{ST} values for mtDNA above the diagonal and for microsatellites below the diagonal. Significant values (after Bonferroni correction) are in bold; *: p < 0.05, **: p < 0.01, ***: p < 0.001. The locality designations follow Fig. 3.1.

	K	T	MNW	MW	MC	MS	SA
K	-	0.036	**0.684*** **	**0.718*** **	**0.443*** **	**0.689*** **	**0.518*** **
T	0.011	-	**0.774*** **	**0.809*** **	**0.531*** **	**0.783*** **	**0.617*** **
MNW	**0.035*** **	**0.034*** **	-	-0.020	**0.169*** **	**0.047***	0.126
MW	**0.059*** **	**0.062*** **	0.006	-	**0.213*** **	**0.069***	**0.172*** **
MC	**0.032*** **	**0.033*** **	**0.016***	0.018	-	0.107	-0.019
MS	**0.044*** **	**0.056*** **	0.010	**0.014***	0.003	-	0.072
SA	**0.048*** **	**0.057*** **	0.004	**0.016*** **	0.009	0.007	-

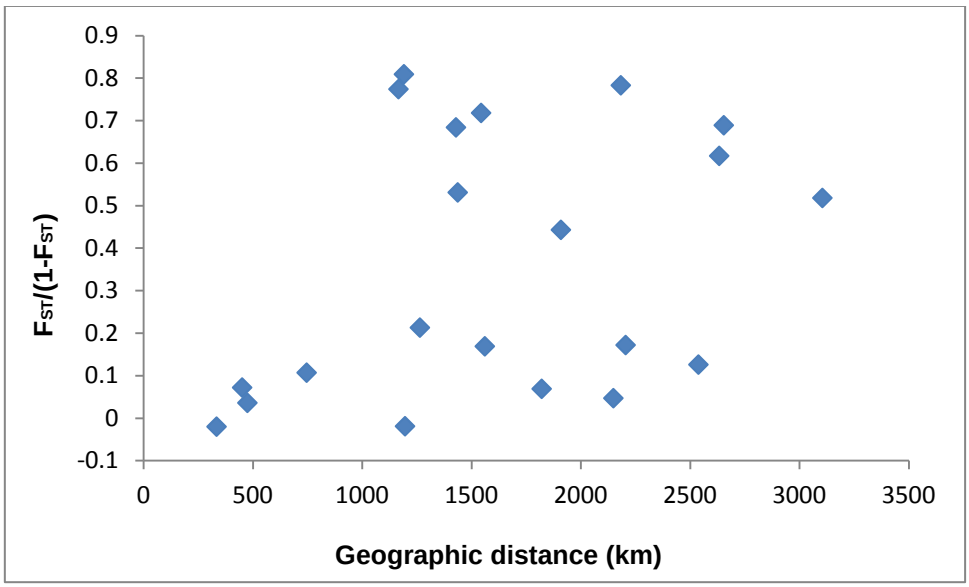
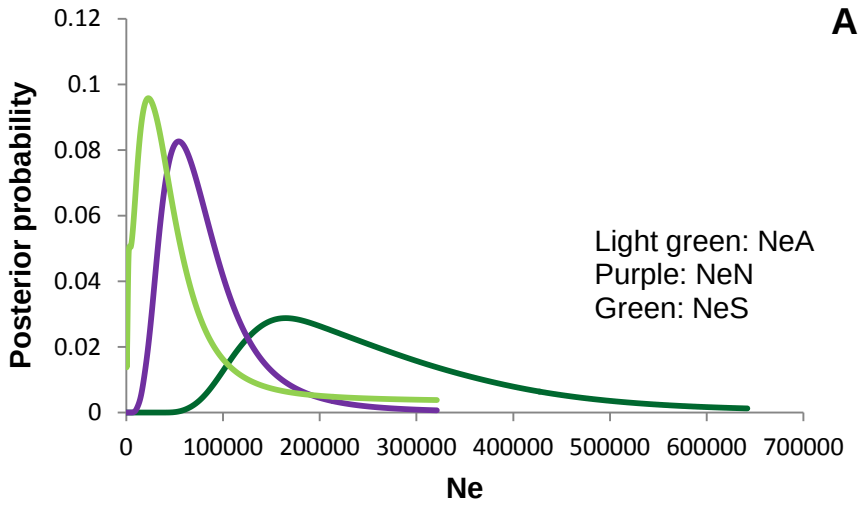
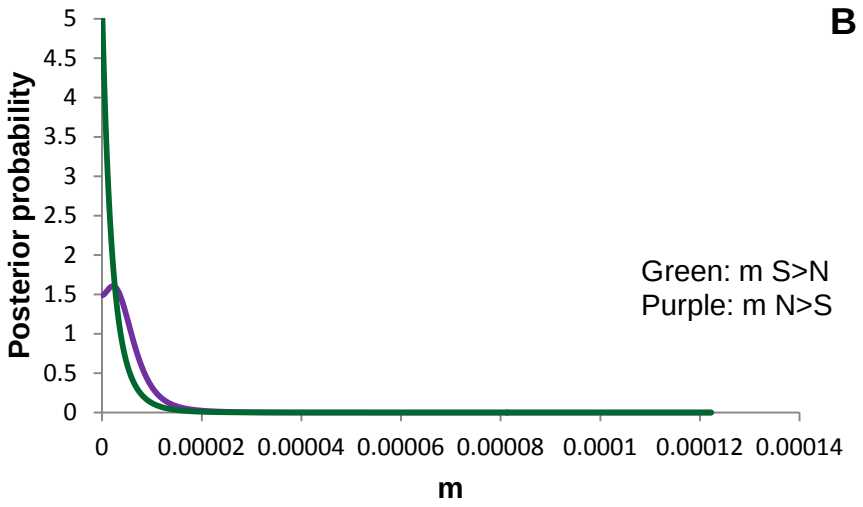


Fig. 3.5 Test for isolation by distance among seven South West Indian Ocean localities based on cytochrome *b* sequences of *Otolithes ruber*. The graph illustrates the relationship between the genetic distance as $F_{ST}/(1- F_{ST})$ and the geographic distance in km.

3.3.3 Historical demography and female gene flow

The coalescent-based method implemented in IMa was used to infer the pattern of historical female gene flow, effective female population sizes and the population divergence time; the estimated posterior distributions of parameters showed sharp peaks, returning to zero values (Fig. 3.6). Effective sample sizes in the analyses were good (ESS>50) with no significant autocorrelation values. Marginal peak locations were good enough for accurate estimates as indicated by the corresponding values not varying for the first half or the second half of the data set and for the full data set. The HiPt (the bin with the highest value in the histogram) values of each parameter were used for the calculations. The southern population (NeS = 164 779 individuals) was estimated to be larger than the northern population (NeN = 54 444 individuals) (Fig. 3.6A). Ne for the ancestral population was the lowest (NeA = 22 965 individuals). The descendant populations underwent expansions following the population split which occurred 48 731 years ago (Fig. 3.6C). Migration rate distributions (Fig. 3.6B) represented unidirectional gene flow from the northern population to the southern population. Calculation of the population migration rate ($2Nm = \Theta \times m/2$), which is the actual number of migrants entering each population (the effective rate at which genes come into a population), showed that approximately 2 to 3 migrants per generation are contributing to female gene flow from north to south, whereas it was negligible in the opposite direction. The trend of historical female population sizes and the existence of unidirectional gene flow from north to south were also confirmed with MIGRATE-n analyses (Appendix 3A).





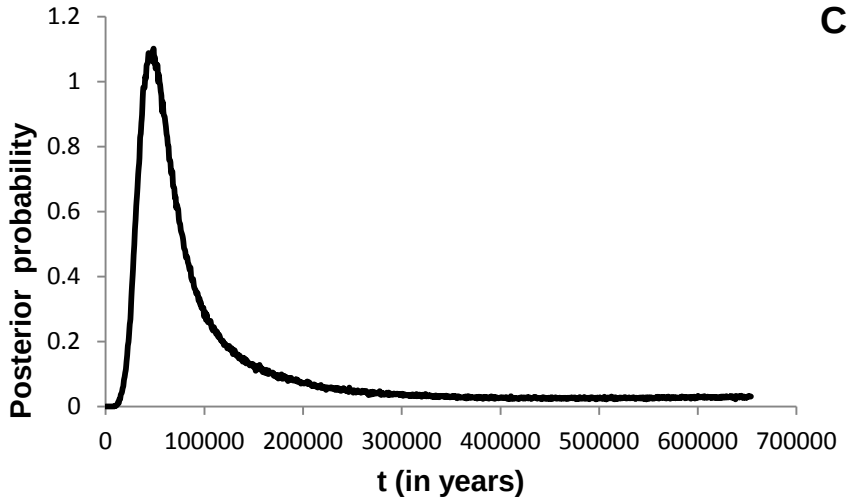


Fig. 3.6 IMa marginal posterior probability distributions for the population parameters based on cytochrome *b* sequences of *Otolithes ruber* of the northern (Kenya and Tanzania) and southern (South Africa, Mozambique and Madagascar) haplogroups in the South West Indian Ocean. **A)** Effective population sizes in individuals/generation (Ne=Θ/4μ): south population (NeS), north population (NeN) and ancestral population (NeA). **B)** Migration rates (m= *m*xμ): m S>N is the migration rate from south to north, m N>S is the migration rate in the opposite direction. **C)** Divergence time (t=*t*/μ) in years.

3.3.4 Microsatellites data analysis and summary statistics

Seven polymorphic microsatellite loci were genotyped in 221 *O. ruber* individuals, with a total of 84 observed alleles. Testing all marker pairs for linkage disequilibrium showed a significant value $p < 0.05$ for 16 cases, all involving different pairs of loci with no specific pattern that would be of concern. No evidence of null alleles or large allele dropouts was detected. Only one of the loci, Soc243, showed null alleles (Oosterhout value: 0.157) due to heterozygote deficiencies. However, it did not affect subsequent analyses since the frequency is less than 0.2 (Dakin & Avise 2004). Summary statistics for each marker and each location were produced and are represented in Appendix 3B. All loci were used for further analyses. Summary statistics over all loci are indicated in Table 3.4. Allelic richness had the same magnitude among locations. Average expected heterozygosity (H_E) ranged from 0.604 to 0.670. The F_{IS} values ranged from -0.029 to 0.106, with two values representing deviations from HWE, the F_{IS} values did not show any pattern that would be of concern for the individual loci (see Appendix 3B). All seven localities showed fairly high levels of genetic variation (based on allelic richness and heterozygosity).

Table 3.4 Summary statistics of genetic variation in seven localities of *Otolithes ruber* averaged over seven microsatellite loci. N: Number of samples, AR: Allelic richness, H_E: Expected heterozygosity, H_O: Observed heterozygosity, F_{IS}: Fixation index, with significance of *: $p < 0.05$, **: $p < 0.01$.

Locations	N	AR	H_E	H_O	F_{IS}
Kenya	30	7.25	0.623	0.610	0.021
Tanzania	30	7.45	0.604	0.581	0.038
Madagascar northwest	29	8.11	0.670	0.643	0.040
Madagascar west	30	7.99	0.657	0.628	0.045
Mozambique central	29	8.65	0.635	0.569	**0.106****
Mozambique south	33	8.60	0.643	0.593	**0.079***
South Africa	40	7.86	0.627	0.645	-0.029

3.3.5 Genetic differentiation based on microsatellite data

Genetic differentiation between localities based on F_{ST} values is represented in Table 3.3. Pairwise F_{ST} values were mostly significant, and ranged from 0.007 to 0.057. Localities K and T were significantly differentiated from the rest of the locations, represented by high F_{ST} values ranging from 0.032 to 0.059 and from 0.034 to 0.062, respectively. The lowest values included comparison among the locations in the southern part of the region, with only three of the values related to Madagascar remaining significant after Bonferroni correction. The F_{ST} values based on the microsatellite loci are much lower than those estimated based on the mtDNA sequences.

Factorial Correspondence Analysis showed spatial differentiation with a level of mixing when samples were designated as north and south (Fig. 3.7A; the groupings as defined for the AMOVA were used). In this analysis, axis 1 fully explained the genetic variation. Mixing and clustering of genotypes was clearer when the samples were designated as three groups (Fig. 3.7B). In the latter analysis, both axes 1 and 2 contributed to the genetic variation in the data set.

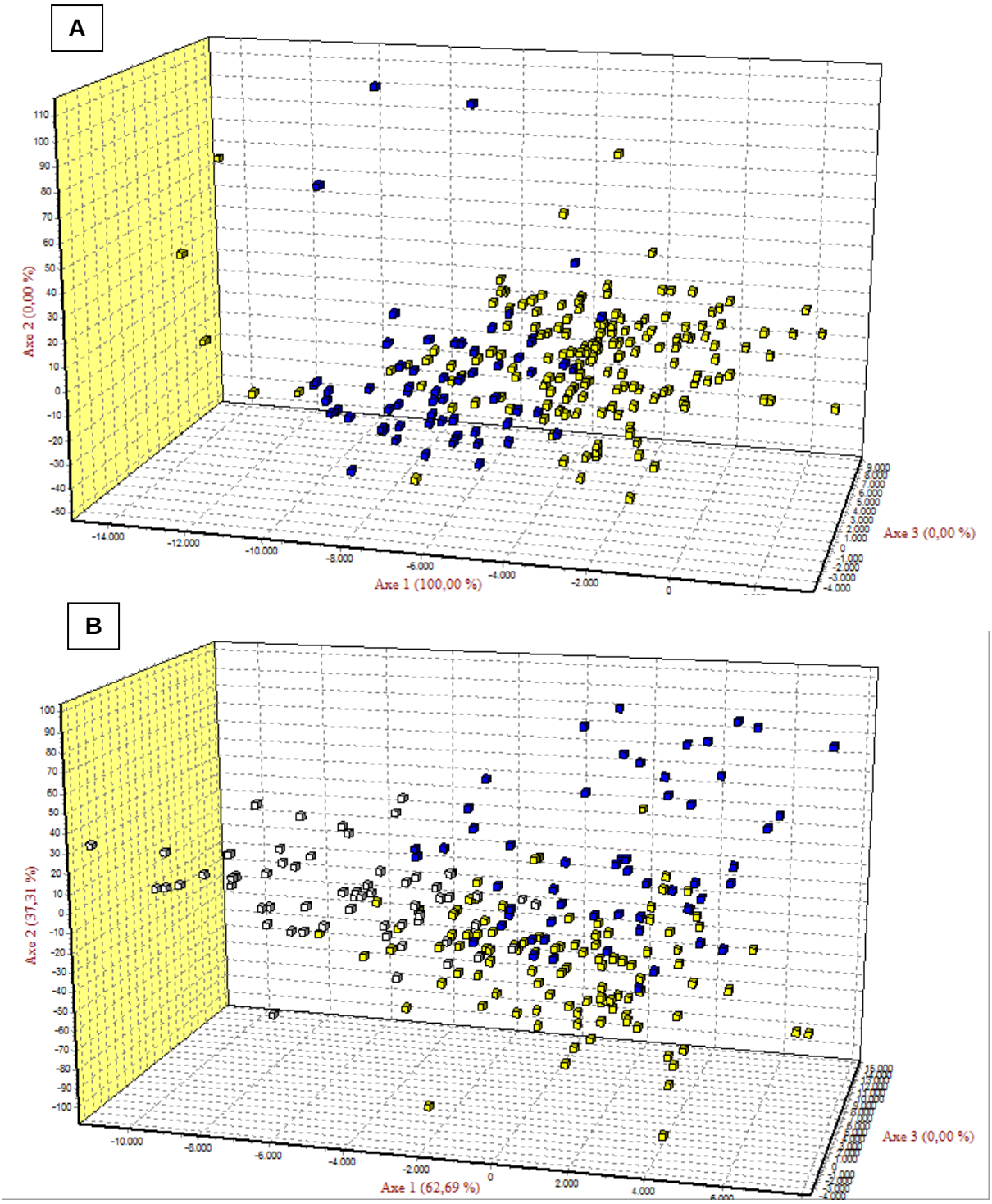


Fig. 3.7 Factorial Correspondence Analysis of variation among *Otolithes ruber* microsatellite genotypes from the South West Indian Ocean. A) North vs South: Blue (North: K, T), yellow (South: SA, MS, MC, MNW, MW). B) Three groups: White (North: K, T), yellow (South: SA, MS, MC), blue (Madagascar: MNW, MW).

The STRUCTURE analysis indicated that the most likely number of populations is two, as the largest rate of change (ΔK) was at K = 2 (this value of K also showed the highest likelihood and smallest standard deviation of the likelihood) (Fig. 3.8). Figure 3.9 illustrates the clustering of all the individuals at K=2; the major separation is between the northern locations (Kenya and Tanzania) versus the remainder of the locations. South Africa, Mozambique and Madagascar samples clustered with high probability (q ≥ 80%) in one cluster; some Kenya and Tanzania samples were not only assigned to a single population (q < 80%) but to both populations.

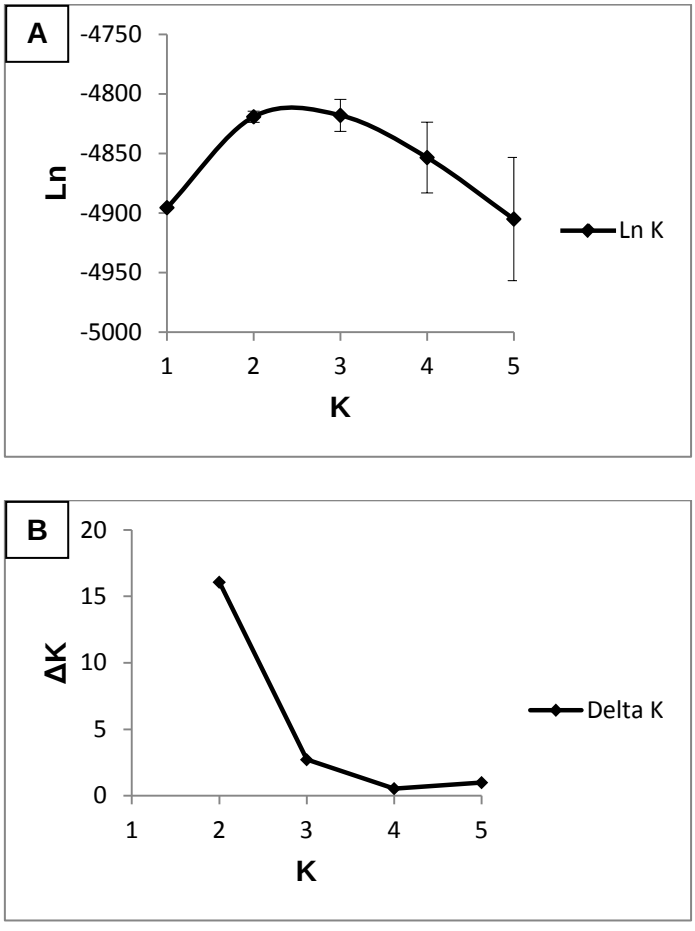


Fig. 3.8 Inference of the mostly likely number of populations (K) based on STRUCTURE analyses of *Otolithes ruber* microsatellite genotypes. The X-axes represent the number of populations. Graph A represents the log probability of the data (LnK) at each value of K and the standard deviation of each estimate based on 10 iterations. Graph B represents the relative rate of change in the log probability of the data (ΔK).

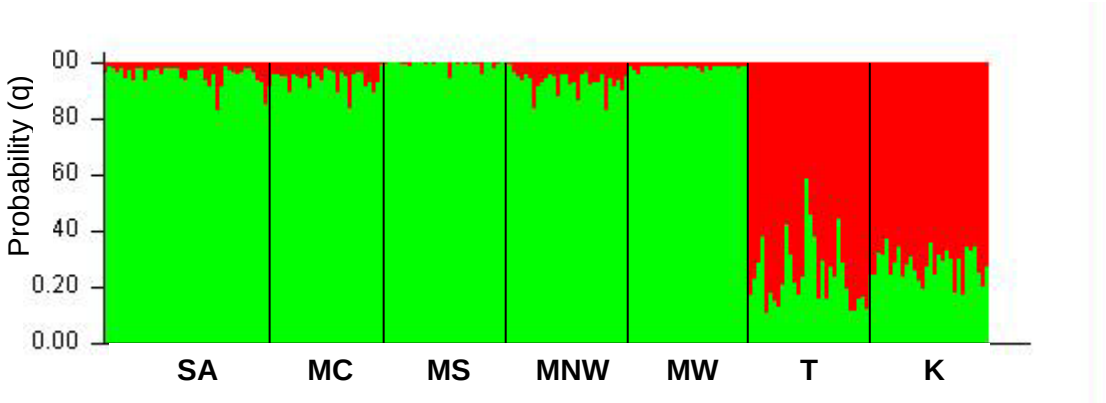


Fig. 3.9 STRUCTURE clustering analysis of 221 *Otolithes ruber* microsatellite genotypes from the South West Indian Ocean. Each vertical bar represents the posterior probability (q) of an individual being assigned to one of two (K=2) clusters.

The 10 STRUCTURE iterations at K = 2 yielded consistent assignment of individuals to the two clusters. Less consistency was evident for the K = 3 runs. Appendix 3C illustrates that the South Africa and Mozambique samples largely assign to a single cluster. Madagascar represents a second cluster, but only the MW individuals were consistently assigned with high probability to the green cluster (q≥ 80%). Kenya individuals appear genetically distinct while some Tanzanian genotypes appear to be admixed of the southern African population and the northern population (Appendix 3C). Running STRUCTURE excluding the Kenyan and Tanzanian samples showed the separation of Madagascar samples from the rest. However, it was not consistent for all the 10 iterations.

An AMOVA based on the microsatellite data and the groups as reflected in the mtDNA analyses (Table 3.2, Fig. 3.4) showed that most of the overall variation is due to within-population variation (Table 3.5). The between-group variance components contributed little to the overall variance, but were significant for the scenario of three groups.

Table 3.5 An Analysis of Molecular Variance based on seven microsatellite loci according to *Otolithes ruber* groups clustering in the north and the south of the South West Indian Ocean region (see Fig. 3.4). GV%: Percentage contribution to the overall genetic variance, df: Degrees of freedom. Significant values are shown in bold; *: p<0.05, **: p<0.01, ***: p<0.001.

Source of variation	df	GV%	Fixation index	p-value
(K, T**) (**MNW, MW, MC, MS, SA**)**	1	3.79	F_{CT}: 0.0379	0.0518
Among populations within groups	5	0.97	F_{SC}: 0.0101	**0.0000*** **
Within populations	435	95.24	F_{ST}: 0.0476	**0.0000*** **
(K, T**) (**MNW, MW**) (**MC, MS, SA**)**	2	2.75	F_{CT}: 0.0275	**0.0088** **
Among populations within groups	4	0.69	F_{SC}: 0.0071	**0.0108***
Within populations	435	96.56	F_{ST}: 0.0344	**0.0000*** **

As for the mtDNA, IBD based on microsatellites was not significant (Fig. 3.10, p = 0.052). The observed differentiation in *O. ruber* based on both marker types was not due to geographic distance between the sampling locations.

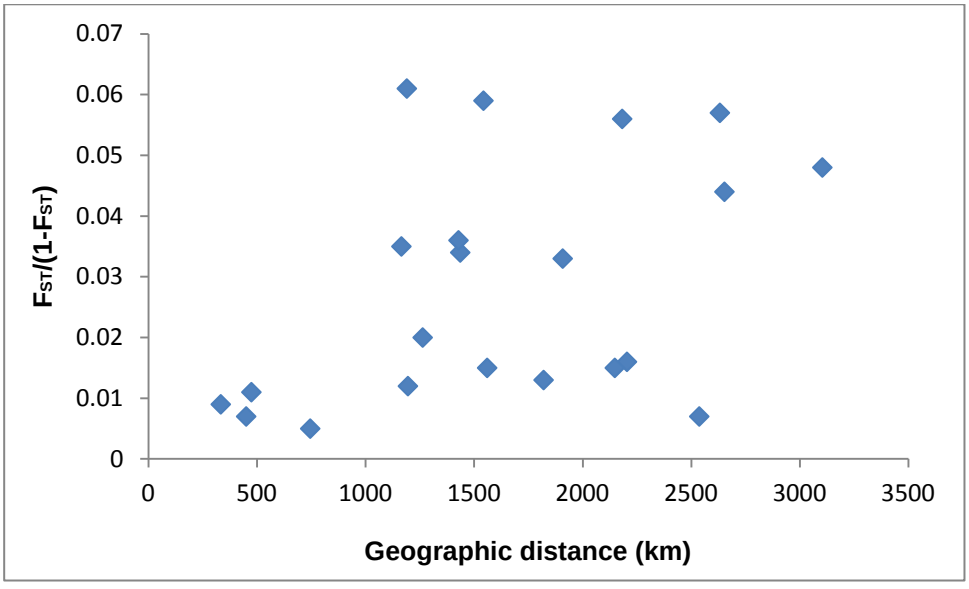
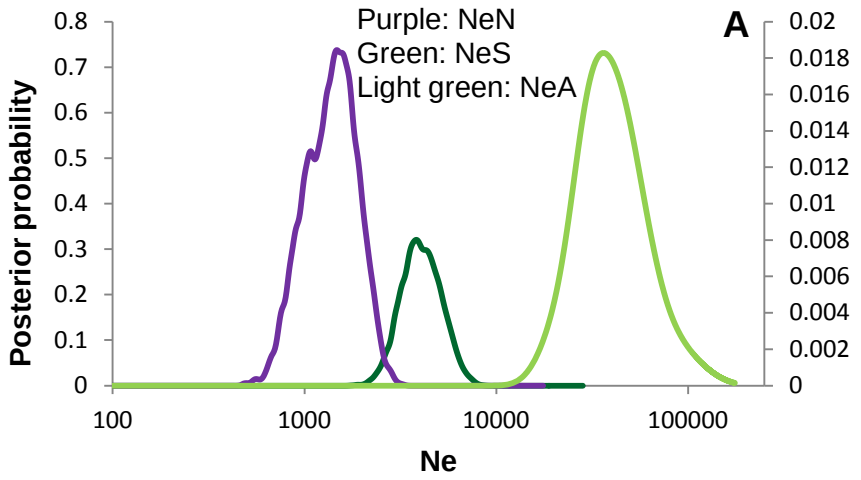
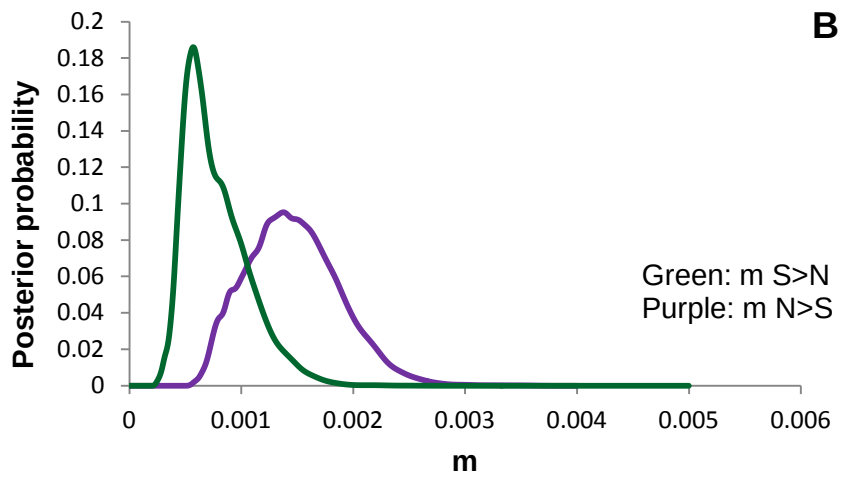


Fig. 3.10 Test for isolation by distance based on microsatellite genotypes of *Otolithes ruber* from seven South West Indian Ocean localities.

3.3.6 Historical demography and nuclear gene flow

Compared to mtDNA, IMa results based on microsatellites were varied for each estimated population parameter (Fig. 3.11). Optimization was challenging (computation times of more than two months) to obtain a fixed set of parameters for the actual run. Optimization was dependent on the number of individuals and the numbers of loci used in the study; more data require more time to test the possible scenarios. The model inferred a large ancestral population with an effective size of NeA = 36 190 individuals, divided into two populations: A southern population (NeS = 3 828 individuals) and northern population (NeN = 1 481 individuals) (Fig. 3.11A). A very recent population split was estimated at 7940 years before present (Fig. 3.11C). Bidirectional, asymmetrical gene flow was observed between the southern and northern populations, with 2.5 times greater gene flow from north to south than in the opposite direction (Fig. 3.11B). Population migration rate calculations suggested a contribution of 32 migrants per generation into the southern population from the north and 5 migrants per generation in the opposite direction.





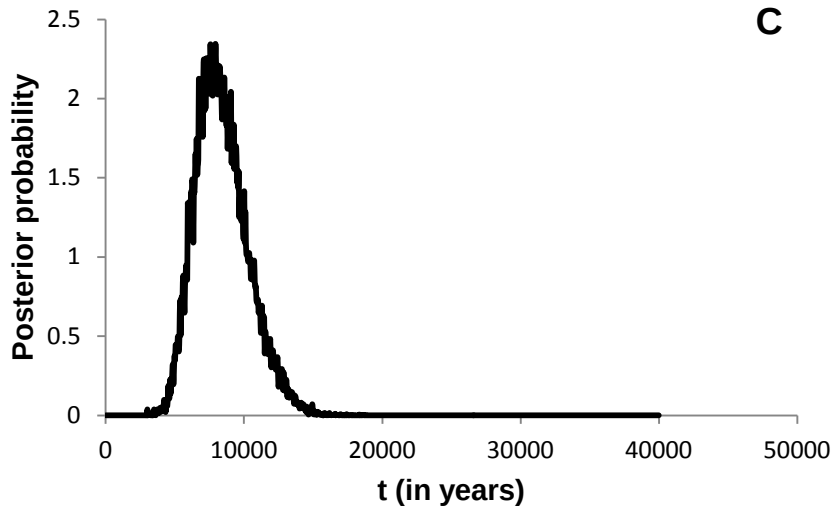


Fig. 3.11 IMa marginal posterior probability distributions for the population parameters based on microsatellite genotypes of *Otolithes ruber* of the northern population (Kenya and Tanzania) and the southern population (South Africa, Mozambique and Madagascar) in the South West Indian Ocean. **A)** Effective population sizes in individuals/generation (Ne=Θ/4μ): south population (NeS), north population (NeN) and ancestral population (NeA). **B)** Migration rates (m= *m*xμ): m S>N is the migration rate from south to north, m N>S is the migration rate in the opposite direction. **C)** Divergence time (t=*t*/μ) in years.

3.4 Discussion

Regional connectivity among *Otolithes ruber* from seven localities in the South West Indian Ocean, from Kenya to South Africa, was investigated through the analysis of mitochondrial DNA cytochrome *b* sequence variation and variation at seven microsatellite loci. Both marker types showed that gene flow in *O. ruber* in the SWIO is restricted and, thus, that the species is not represented by a single population in the region. Both marker types provided strong evidence for at least two subpopulations, one in the north (Kenya and Tanzania) and the other in the south (Madagascar, Mozambique and South Africa) of the region. This reduced connectivity between the north and south of the SWIO likely reflects a signature of historical divergence with ongoing gene flow.

3.4.1 The pattern of genetic differentiation

Population differentiation was examined using multiple analytical approaches. Kenya and Tanzania consistently differentiated from the remainder of the localities as represented by the significant pairwise F_{ST} values for both marker types. The mtDNA haplotype network, AMOVA analyses and clustering based on microsatellite genotypes all confirmed the observed pattern of distinction between the north and the south populations. In some of the analyses, Madagascar could also be differentiated as a third subpopulation. It was shown that population structuring was not a result of physical geographic distances between sampling locations, as isolation by distance failed to explain the genetic differentiation in *O. ruber*.

The partitioning of a tropical population, Kenya and Tanzania, from the subtropical southern population(s) could be explained by the oceanography of the SWIO region, features of the biology and demography of the species. The South Equatorial Current (SEC) acts as a likely mechanism of general larval transportation from an as yet unsampled source population in the Western Indian Ocean (WIO) to the East African mainland, whereas the East African Coastal Current (EACC) is thought to transports the larvae along the coast in a northerly direction (Visram *et al.* 2010). These currents are expected to separate *O. ruber* populations off Tanzania and Kenya from the populations off Mozambique and South Africa (Fratini *et al.* 2010). The action of EACC is strong during the Southeast Monsoon while reduced during the Northeast Monsoon (Ragionieri *et al.* 2010). The Northeast Monsoon takes place from December to April (Benny 2002), which coincides with the reported spawning season of *O. ruber* in spring and summer, September to February for the southern African population (Fennessy 2000). When the EACC current action is reduced, larvae may be transported

southwards via the other branch of the SEC, which flows into the Mozambique Channel. The northern part of the channel is dominated by the large anticyclonic cell of the Comoros Basin (Kai & Marsac 2010). The action of this cell and the extent of the continental shelf off Madagascar may act to retain larvae off northwestern Madagascar. Although there is mean southward flow of the Mozambique Current along the continental slope (see Fig. 3.1), the flow of the current is strongly influenced by a complex and variable system of cyclonic and anticyclonic eddies (Matano *et al*. 2008; Kai & Marsac 2010). The eddy system in the channel transports larvae southwards, but eddies can also leads to local larval retention within populations as suggested by previous studies (Hauser & Carvalho 2008 and references therein).

The Mozambique Channel's circulation mechanisms are also likely involved in reduced connectivity in the SWIO. Such study which investigated mtDNA variation in the green turtle (*Chelonia mydas*) found two populations in the North Mozambique Channel and South Mozambique Channel, respectively (Bourjea *et al.* 2007). Restricted gene flow was also represented between two separate genetic stocks from the islands of Europa and Juan de Nova. Unusual sea surface temperatures at the south end of the channel, the presence of eddy and combined current effects could be involved in the observed differentiation in this species (Bourjea *et al.* 2007). A similar conclusion was also presented by Muths *et al.* (2011) for the squirrelfish (*Myripristis berndti*) from *Cyt b* and microsatellite data. They illustrated genetic differentiation of Europa Island from the northern localities, while the island of Juan de Nova showed intermediate representation, either from the northern or southern populations based on the marker type (Muths *et al.* 2011). Significant differentiation among the northern populations was also indicated suggesting that complete homogeneity is not the case in the Mozambique Channel (Muths *et al.* 2011). Ridgway *et al.* (2001) demonstrated the absence of genetic differentiation for the coral species *Pocillopora verrucosa* from the WIO based on allozyme data, and a follow-up study using microsatellites suggested a contradictory result, with gene flow between South Africa and southern Mozambique being limited (Ridgway *et al.* 2008).

Furthermore, analyses of cytochrome oxidase subunit *1* data showed two evolutionary distinct groups in the mangrove crab (*Neosarmatium meinerti*) (Ragionieri *et al.* 2010). Differentiation between the East African mainland metapopulation and the Western Indian Ocean island group was suggested to have occurred during the early Pleistocene due to a fragmentation event (Ragionieri *et al.* 2010). Their analyses indicated restricted gene flow between the African mainland and Seychelles due to the absence of stepping stone populations between the sites. However, genetic homogeneity was found from northern

Kenya to South Africa, a region in which mangroves are variably distributed, representing stepping stone populations (Ragionieri *et al.* 2010). A similar study on the parrotfish (*Scarus ghobban*) showed significant genetic structuring between Kenya and Seychelles vs. Mauritius and Tanzania based on the mitochondrial control region, again emphasizing the role of the SEC and the EACC, respectively (Visram *et al.* 2010). In addition to these studies, our findings also show that the shallow water demersal species *O. ruber* occurs as two distinct populations in the north and south of the region, linked to oceanographic features. Moreover, the northern Mozambique region may be less suitable habitat for the species because of extensive coral reefs and island systems with clear water. The species is less frequently recorded from the Cabo Delgado province (northern Mozambique) compared to the Sofala Bank (central Mozambique) (pers. comm. Dr Atanasio Brito, Deputy Director, Institudo de Investigacao Pesquera, Maputo).

3.4.2 Historical demography and gene flow

The historical demography of *O. ruber* and the star-shaped mtDNA haplotype network suggest a recent population expansion. Historical population sizes based on mtDNA show a small ancestral population size compared with the two daughter populations, reflecting population size increase after the split 48 731 years ago. Effective population size appears to be considerably larger in the south of the region; this may relate to the greater extent of the continental shelf off western Madagascar, and the wide Sofala Bank and Delagoa Bight off Mozambique (Kai & Marsac 2010). The estimated time of split corresponds with a period that had an impact on marine species distributions, shaping of phylogeographic structure within species and leaving genetic signatures of fluctuations in population size (Lessios 2008; Hoareau *et al.* 2012). Following an extended glacial period (190-130 ka), during which sea levels dropped to 120 meters below present levels, temperatures and the extent of suitable habitat increased, leading to the expansion of populations (see Visram *et al*. 2010 and references therein). Obura (2012) suggested that the steep slope of the East African and Madagascan coasts would have been minimally impacted by changes in sea level and proposed that the northern Mozambique Channel (and perhaps the WIO as a whole) would have served as a refugial region for shallow water tropical species. The lack of suitable habitat and possible confinement to the North Mozambique Channel refugia are the features that led to the smaller population size of the northern population.

Further analyses are required at a finer scale of resolution to tease apart evolutionary history (including colonization patterns) and ongoing ecological processes. Polymorphism levels

observed by both marker types were high in all localities, but some differences between locations need to be explored further. Tanzanian samples showed the lowest haplotype diversity based on mtDNA, which may reflect a bottleneck event prior to the population expansion, and the recent introduction of new nuclear DNA alleles through gene flow. The history of Tanzanian *O. ruber* may, thus, be different compared to the remainder of the region. Overall, there was a strong signal of an expansion, yet South Africa and Mozambique may have experienced a more stable long-term history.

The difference in effective population size of the species in the north and south of the SWIO may impact on the estimation of gene flow between these regions. Nuclear data showed a different pattern of population sizes to that based on mtDNA, with current effective population sizes being much smaller than the ancestral population. However, in agreement with the mtDNA inference, the microsatellite data also indicate that the southern population is larger than the northern population. These small population sizes indicate that the demographic history of *O. ruber* is marked by recent reductions in Ne. This might suggest that genetic drift shaped the pattern observed from microsatellite data in addition to periods of restriction in gene flow. The time of the population split based on microsatellites was about 8000 years ago, which fits with the early Holocene period when sea levels were rising after a drastic sea level drop during the Last Glacial Maximum (LGM) (Waelbroeck *et al.* 2002; Hoareau *et al.* 2012).

Taken together, the rapidly evolving microsatellites represent the recent population demography, while mtDNA tracks changes in Ne over a longer time frame, subsequent to the first expansion event. Uncertainty regarding microsatellite mutation rate and the best mutation model for these loci should be considered while interpreting the results. The mutation rate of microsatellite loci ranges between 10^{-2} and 10^{-6} (Selkoe & Toonen 2006); 10^{-4} is the averaged mutation rate commonly used (also in this study). The averaged mutation rate may lead to over- or underestimation of population parameters. Additionally, IMa applies only the SMM mutation model (Estoup *et al.* 2002) which might not be the appropriate model for all microsatellite loci.

The mtDNA indicated unidirectional historical female gene flow from north to south, while the microsatellites showed bidirectional nuclear gene flow between the two populations. The level of gene flow, however, was higher from north to south. Differences in the pattern of gene flow inferred from nuclear versus mitochondrial DNA most likely reflect sex-biased dispersal within species (Okumuş & Çiftci 2003). For *O. ruber,* male-biased migration might be a life history trait linked to avoidance of non-random mating, however, there is no

published evidence of this in the species or in other sciaenids. The pattern also suggests that males have more impact on connectivity in comparison to females, whereas sedentary females have implications for sustainable harvesting. Furthermore, the microsatellite results provide evidence of ongoing gene flow between the southern and northern population. No migratory behaviour has been recorded in *O. ruber*, and long distance movements are unlikely in this small-sized demersal species; the eggs and larvae are almost certainly pelagic, as seen in other sciaenids such as *A. japonicus* (Froese & Pauly 2013). Thus, connectivity in the case of *O. ruber* may largely be the result of passive larval dispersal by the prevailing ocean currents (refer to Fig. 1.5), since there is much more gene flow from north to south rather than in the opposite direction. If there is a spawning area off North West Madagascar, south to north gene flow can be explained by the reduced current action during the different monsoon periods. However, a tagging recapture study of the species would provide independent empirical data on the movement patterns of adults.

Sciaenids mainly reproduce in estuaries, bays or the open sea, where currents play a role in the dispersal of eggs and larvae (Santos *et al.* 2006). The Thukela Bank off South Africa and the Sofala Bank in central Mozambique are both shallow (< 50m depth) nursery areas for *O. ruber* (Fennessy 2000), but the location of spawning areas further north and in Madagascar is not known. The degree of larval dispersal and adult population dynamics can be influenced by coastline topography as it affects oceanographic features (Nicastro *et al.* 2008). For example, pelagic larvae of deep-water demersal tuskfish (*Brosme brosme*) are retained in gyres and eddies surrounding banks and islands, thereby limiting gene flow and leading to genetic differentiation (Knutsen *et al.* 2009). There is currently no information available to indicate whether a similar scenario may apply to eddies in the WIO or whether the eddies in the northern Mozambique Channel disperse larvae southwards. Based on flow patterns (e.g. de Ruiter *et al.* 2002), the latter possibility appears likely.

Several larger sciaenids such as red drum (*Sciaenops ocellatus*), silver kob (*Argryosomus inodorus*) and spotted seatrout (*Cynoscion nebulosus*) migrate considerable distances, as shown by recapture studies (Mercer 1984a, b; Griffiths 1997). King fish (*Macrodon ancylodon*) from Brazil, which is a sciaenid with limited migratory behaviour, has two distinct evolutionary lineages representing tropical and subtropical groups (Santos *et al.* 2006). The temperature variations and the physical separation of the South Equatorial Current into the North Brazil and Brazil Currents might act as barriers to gene flow between the tropical and subtropical clades of this species (Santos *et al.* 2006). Another sciaenid, the meagre (*A. regius*), also showed at least two genetically distinct units, Atlantic and Mediterranean, based on microsatellite data due to vicariance, species biology and glacial cycles (Haffray *et*

al. 2012). A recent study based on the mtDNA control region demonstrated two genetic units (Brazil and Argentina) in the southern king weakfish (*M. articauda*). The authors related the observed spatial genetic variation to historical colonization from south to north, followed by expansion (Rodrigues *et al.* 2013). Although not based on genetic studies, sub-population structuring was also observed in the sciaenid mulloway (*A. japonicus*) from Australia (Silberschneider & Gray 2008) and in the silver kob (*A. inodorus*) (Griffiths 1997) from South Africa based on biological parameters such as reproduction, growth and morphometric differences, respectively.

Information based on diet can also be important in understanding the distribution of exploited species (Phelan *et al.* 2008). Sciaenids frequently associate with penaeid prawns, to the extent that they are considered indicators of potential prawn fishing grounds (Pauly & Neale 1985). *Otolithes ruber* feeds selectively at various life stages (Pillai 1983). Young individuals prefer zooplankton and pelagic biota in surface waters, followed by a change to teleosts and, particularly, penaeid prawns as they mature (Bapat & Bal 1952; Nair 1980; Pillai 1983; Fennessy 2000). The most abundant penaeid prawn species in the region is *Penaeus indicus*, which dominates catches in Malindi-Ungwana Bay (Kenya), the Rufiji Delta (Tanzania), Maputo Bay and the Sofala Bank (Mozambique), the Thukela Bank in South Africa and most of the west coast of Madagascar (Fennessy *et al.* 1994, Teikwa & Mgaya, 2003; De Sousa *et al.* 2006; Caveriviere *et al.* 2008, Munga *et al.* 2013). *Otolithes ruber* forms a substantial part of the catch along with *P. indicus* in all these areas (Olbers & Fennessy, 2007; Fennessy *et al*. 2008), and this co-occurrence partly explains the observed distribution of *O. ruber.* A genetic study on *P. indicus* from these areas may prove helpful in elucidating drivers of gene flow in the SWIO, and biological studies of the species from Madagascar, Kenya and Tanzania would advance understanding of phylogeographic patterns.

Critical to our interpretation of the gene flow results will be future inclusion of additional localities, since unsampled (so-called “ghost”) populations may be responsible for observed gene flow between the northern and southern localities. Rather than direct movement from East Africa to southern Africa, there may be mixing of these populations at the gyres in the northern Mozambique Channel and significant mixing by eddies in the southwest of the channel. How the currently analysed *O. ruber* subpopulations relate to the populations from the broader Indo-West Pacific distribution of the species, and the southwest and east coasts of Madagascar would refine our understanding of connectivity in the region.

3.4.3 Conservation and management implications

The impact of fishing on marine ecosystems has become an increasing concern, either via a direct impact by removal of species or via an indirect impact such as habitat alteration, and changes in abundance of prey or predator species (Stobutzki *et al.* 2003). Unwanted and discarded species are the victims of bycatch and it was estimated that one-third of global fishery discards is due to prawn trawl fisheries (Stobutzki *et al.* 2003). *Otolithes ruber* is one example of a bycatch species. Large-sized individuals are used for sale, whereas smaller ones are discarded due to their non-commercial value; all are killed during the trawling process (Olbers & Fennessy 2007). This is a concern as the smaller individuals, which have not matured sexually, will then not have the opportunity to spawn and, thus, cannot make a contribution to the existing populations.

Fish stocks or populations are described as the basic biological units in fisheries management and assessments (Reiss *et al.* 2009). Sustainable harvesting of fish stocks and preventing depletion of these stocks is the prime goal of management in fisheries (Reiss *et al.* 2009). Local reduction of populations and reduced productivity are the result if spatial features of population structure and management units are disregarded (Worm *et al.* 2006). Importantly, the migration pattern between two putative populations is of substantial interest to managers; in the case of low migration they can be regarded as demographically independent, while high migration suggests a single demographic unit (Waples *et al.* 2008). Separate management plans should be considered for each stock/population to protect the genetic integrity represented by the distinct populations.

Our findings demonstrated the existence of two *O. ruber* populations in the SWIO region, hence, the northern and southern populations need to be managed separately to maintain the observed genetic variation. Even though the analyses show unidirectional historical gene flow and bidirectional ongoing gene flow between the populations, the magnitude of gene flow is small and genetic distinctiveness is retained. Both populations have unique historical variation and Madagascar might also present as a separate population (which requires further work using more markers). The critical point is that the northern population is only weakly replenished from the southern population. Hence, the northern population should have priority for management. There is an indication that the Madagascan population is somewhat distinct. Owing to the physical separation of this country from the mainland, management of *O. ruber* in this country should be distinct from that of the other SWIO countries.

A suggested conservation recommendation for managing and protecting fish species is the introduction of fishery closures (Silberschneider *et al.* 2009). Closure action might be considered as additional management aid to protect spawning, to improve reproductive capacity and to give individuals in separate populations a chance to attain maturity (Phelan *et al.* 2008). For *O. ruber*, an alternative might be the periodic closure of the prawn trawl fishery since, as mentioned earlier, *O. ruber* forms a large component of the bycatch in both Kenya and Tanzania. In this regard, both of these countries have closed their prawn trawl fisheries at times in recent years (albeit owing to clashes with artisanal fisheries; Fennessy & Isaksen 2007). In the case of Tanzania, the fishery is still closed. A further alternative is the introduction of trawl bycatch reduction devices (e.g. Broadhurst 2000) which reduce catches of unwanted fish bycatch.

3.4.4 Conclusion

This is the first genetic study conducted for *O. ruber*, revealing its demographic history, genetic differentiation and connectivity patterns among SWIO localities. *Otolithes ruber* comprises two structured populations in the north and the south of the region, with ongoing gene flow between them. Although the populations are distinguishable from each other, the microsatellite analyses suggest greater connectivity in the region. Even though there is bidirectional movement, it is not symmetrical; there appears to be more movement from north to south. Genetic connectivity is an important measure as it helps define management units and assists in protecting biodiversity by supporting management plans (Palumbi 2003). Based on the combined mtDNA and microsatellite data from this study, it is recommended that *O. ruber* should be managed as at least two independent units in the SWIO. Fisheries stock assessments and biological studies should be undertaken for these identified subpopulations.

3.5 References

Avise JC, Nelson WS, Sibley CG (1994) DNA sequence support for a close phylogenetic relationship between some storks and New World vultures. *Proceedings of the National Academy of Sciences of the USA*, **91**, 5173-5177.

Bapat SV, Bal DV (1952) The food of some young fishes from Bombay. *Proceedings of the Indian Academy of Sciences,* **35**, 78-92.

Bay LK, Choat JH, Van Herwerden L, Robertson DR (2004) High genetic diversities and complex genetic structure in an Indo-Pacific tropical reef fish (*Chlorurus sordidus*): Evidence of an unstable evolutionary past? *Marine Biology,* **144**, 757-767.

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics,* **22**, 341-345.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences USA,* **98**, 4563-4568.

Beerli P, Palczewski (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics,* **185**, 313-326.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05 logciel sous windows pour la genetique des populations, Laboratorie Genome, Populations, Interactions, CNRS UMR 5171, Universite de Montpellier II, Montpellier (France).

Benny PN (2002) Variability of western Indian Ocean currents. *West Indian Ocean Journal of Marine Science,* **1**, 81-90.

Bourjea J, Lapègue S, Gagnevin L, Broderick D, Mortimer JA, Ciccione S, Roos D, Taquet C, Grizel H (2007) Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. *Molecular Ecology,* **16**, 175-186.

Brash JM, Fennessy ST (2005) A preliminary investigation of age and growth of *Otolithes ruber* from KwaZulu-Natal, South Africa. *Western Indian Ocean Journal of Marine Science,* **4**, 21-28.

Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *Journal of Biogeography,* **39**, 12-30.

Broadhurst MK (2000) Modifications to reduce bycatch in prawn trawls: A review and framework for development. *Review of Fish Biology and Fisheries,* **10**, 27-60.

Caveriviere A, Chaboud C, Rafalimanana T (éds.) (2008). Les crevettes côtières de Madagascar. *Biologie, exploitation, gestion*. IRD Editions : p 350.

Clement M, Posada D, Crandall KA (2000) TCS: A computer program to estimate gene genealogies. *Molecular Ecology,* **9**, 1657-1659.

Craig MT, Eble JA, Bowen BW, Robertson DR (2007) High genetic connectivity across the Indian and Pacific Oceans in the reef fish *Myripristis berndti* (Holocentridae). *Marine Ecology Progress Series*, **334**, 245-254.

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

de Ruijter WPM, Ridderinkhof H, Lutjeharms JRE, Schouten MW, Veth C (2002) Observations of the flow in the Mozambique Channel. *Geophysical Research Letters*, **29**, 1502.

de Sousa LP, Brito A, Abdula S, Caputi N (2006) Research assessment for the management of the industrial shallow-water multi-species shrimp fishery in Sofala Bank in Mozambique. *Fisheries Research,* **77**, 207-219.

DiBattista JD, Berumen ML, Gaither MR, Rocha LA, Eble JA, Howard Choat J, Craig MT, Skillings DJ, Bowen BW (2013) After continents divide: Comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography*, **40**, 1170-1181.

Durand JD, Collet E, Chow S, Guinand B, Borsa P (2005) Nuclear and mitochondrial DNA markers indicate unidirectional gene flow of Indo-Pacific to Atlantic bigeye tuna (Thunnus obesus) populations, and their admixture off southern Africa. *Marine Biology*, **147**, 313-322.

Durgadoo JV, Ansorge IJ, Lutjeharms JRE (2010) Oceanographic observations of eddies impacting the Prince Edward Islands, South Africa. *Antarctic Science,* **22**, 211-219.

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

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology,* **14**, 2611-2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver 3.11: An integrated software package for population genetics data analysis. *Evolution Bioinformatics,* **1**, 47-50.

Faulks LK, Gilligan DM, Beheregaray LB (2010) Clarifying an ambiguous evolutionary history: Range-wide phylogeography of an Australian freshwater fish, the golden perch (*Macquaria ambigua*). *Journal of Biogeography,* **37**, 1329-1340.

Fennessy ST (2000) Aspects of the biology of four species of Sciaenidae from the east coast of South Africa. *Estuarine, Coastal and Shelf Science,* **50**, 259-269.

Fennessy ST, Villacastin C, Field JG (1994) Distribution and seasonality of ichthyofauna associated with commercial prawn trawl catches on the Tugela Bank of Natal, South Africa. *Fisheries Research,* **20**, 263-282.

Fennessy ST, Isaksen B (2007) Can bycatch reduction devices be implemented successfully on prawn trawlers in the Western Indian Ocean? *African Journal of Marine Science*, **29**, 453-463.

Fennessy ST, Vincent X, Budeba Y, Mueni EM, Gove DZ (2008) An update on initiatives to reduce prawn trawl bycatch in the Western Indian Ocean. *Western Indian Ocean Journal of Marine Science,* **7**, 217-222.

Fitzpatrick JM, Carlon DB, Lippe C, Robertson DR (2011) The West Pacific diversity hotspot as a source or sink for new species? Population genetic insights from the Indo-Pacific parrotfish *Scarus rubroviolaceus*. *Molecular Ecology,* **20**, 219-234.

Fontaine MC, Tolley KA, Michaux JR, Jr AB, Ferreira M, Jauniaux T, Llavona A, Ozturk B, Ozturk AA, Ridoux V, Rogan E, Sequeira M, Bouquegneau JM, Baird SJE (2010) Genetic and historic evidence for climate-driven population fragmentation in a top cetacean predator: The harbour porpoises in European water. *Proceedings of the Royal Society B: Biological Sciences,* **277**, 2829-2837.

Fratini S, Ragionieri L, Cannicci S (2010) Stock structure and demographic history of the Indo-West Pacific mud crab *Scylla serrata*. *Estuarine, Coastal and Shelf Science,* **86**, 51-61.

Froese R, Pauly D (2013) Fishbase. Editors. 2013. FishBase. World Wide Web electronic publication. www.fishbase.org, version (02/2013).

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics,* **147**, 915-925.

Gaither MR, Bowen BW, Bordenave TR, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT (2011) Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific barrier with contemporary overlap in the coral triangle. *BMC Evolutionary Biology,* **11**, 1-15.

Galarza JA, Carreras-Carbonell J, Macpherson E, Pascual M, Roques S, Turner GF, Rico C (2009) The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of the National Academy of Sciences of the United States of America,* **106**, 1473-1478.

Gonzalez EG, Zardoya R (2007) Relative role of life-history traits and historical factors in shaping genetic population structure of sardines (*Sardina pilchardus*). *BMC Evolutionary Biology,* **7**, 197.

Goudet J (2002) FSTAT, a program to estimate and test gene diversities and fixation indices (Version 2.9.3.2). Available from http://www.unil.ch/izea/softwares/fstat/html.

Griffiths MH (1997) The life history and stock separation of silver kob, *Argyrosomus inodorus*, in South African waters. *Fishery Bulletin,* **95**, 47-67.

Haffray P, Malha R, Sidi MOT, Prista N, Hassan M, Castelnaud G, Karahan-Nomm B, Gamsiz K, Sadek S, Bruant JS, Balma P, Bonhomme F (2012) Very high genetic fragmentation in a large marine fish, the meagre *Argyrosomus regius* (Sciaenidae, Perciformes): Impact of reproductive migration, oceanographic barriers and ecological factors. *Aquatic Living Resources,* **25**, 173–183.

Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acid Symposium Series,* **41**, 95-98.

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology,* **66**, 591-600.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J*ournal of Molecular Evolution,* **22**, 160-174.

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. *Fish and Fisheries,* **9**, 333-362.

Heemstra P, Heemstra E (2004) *Coastal fishes of Southern Africa* NISC, SAIAB.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences USA,* **104**, 2785-2790.

Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after Avise, 2000. *Molecular Phylogenetic and Evolution,* **54**, 291-301.

Hoareau TB, Boissin E, Berrebi P (2012) Evolutionary history of a widespread Indo-Pacific goby: The role of Pleistocene sea-level changes on demographic contraction/expansion dynamics. *Molecular Phylogenetics and Evolution,* **62**, 566-572.

Hoareau TB, Boissin E, Paulay G, Bruggemann JH (2013) The Southwestern Indian Ocean as a potential marine evolutionary hotspot: perspectives from comparative phylogeography of reef brittle-stars. *Journal of Biogeography,* **40**, 2167-2179.

Huijbers CM, Nagelkerken I, Debrot AO, Jongejans E (2013) Geographic coupling of juvenile and adult habitat shapes spatial population dynamics of a coral reef fish. *Ecology,* **94**, 1859–1870.

Kai ET, Marsac F (2010) Influence of mesoscale eddies on spatial structuring of top predators' communities in the Mozambique Channel. *Progress in Oceanography,* **86**, 214-223.

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE,* **5**, e8594.

Kimura M, Ohta T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the United States of America,* **75**, 2868-72.

Knutsen H, Jorde PE, Sannæs H, Hoelzel AR, Bergstad OA, Stefanni S, Johansen T, Stenseth NC (2009) Bathymetric barriers promoting genetic structure in the deepwater demersal fish tusk (*Brosme brosme*). *Molecular Ecology,* **18**, 3151-3162.

Lessios HA (2008) The Great American schism: Divergence of marine organisms after the rise of the central American Isthmus. *Annual Review of the Ecology, Evolution, and Systematics,* **39**, 63–91

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics,* **25**, 1451-1452.

Lutjeharms JRE (2006) The Coastal Oceans of South-Eastern Africa (15,W), pp. 783-834. In: AR Robinson, KH Brink (eds.) The Sea: The Global Coastal Ocean Interdisciplinary Regional Studies and Synthesis. Harvard University Press.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research,* **27**, 209-220.

Martínez P, González EG, Castilho R, Zardoya R (2006) Genetic diversity and historical demography of Atlantic bigeye tuna (*Thunnus obesus*). *Molecular Phylogenetics and Evolution,* **39**, 404-416.

Matano RP, Beier EJ, Strub PT (2008) The seasonal variability of the circulation in the South Indian Ocean: Model and observations. *Journal of Marine Systems,* **74**, 315-328.

McCusker MR, Bentzen P (2010) Historical influences dominate the population genetic structure of a sedentary marine fish, Atlantic wolfish (*Anarhichas lupus*), across the North Atlantic Ocean. *Molecular Ecology,* **19**, 4228–4241.

Mercer LP (1984a) A biological and fisheries profile of red drum *Sciaeniops ocellatus.* Science report, North Carolina department of Natural Resources Community Development, division of Marine Fisheries, No. 41.

Mercer LP (1984b) A biological and fisheries profile of spotted seatrout *Cynoscion nebulosus*. Science report, North Carolina department of Natural Resources Community Development, division of Marine Fisheries, No. 40.

Munga CN, Mwangi S, Ong'anda H, Ruwa R, Manyala J, Groeneveld JC, Kimani E, Vanreusel A (2013) Species composition, distribution patterns and population structure of penaeid shrimps in Malindi-Ungwana Bay, Kenya, based on experimental bottom trawl surveys. *Fisheries Research,* **147**, 93–102.

Muths D, Tessier E, Gouws G, Craig M, Mwale M, Mwaluma J, Mwandya A, Bourjea J (2011) Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. *Marine Ecology Progress Series,* **443**, 167-180.

Muths D, Gouws G, Mwale M, Tessier E, Bourjea J (2012) Genetic connectivity of the reef fish *Lutjanus kasmira* at the scale of the Western Indian Ocean. *Canadian Journal of Fisheries and Aquatic Sciences,* **69**, 842-853.

Nair SKV (1980) Food and feeding habits of *Otolithes ruber* (Schneider) at Calicut. *Indian Journal of Fisheries,* **26**, 133-139.

Navaluna NA (1982) Morphometrics, biology and population dynamics of the croaker fish, *Otolithes ruber,* pp.38-55. In: *Small-scale fisheries of San Miguel Bay, Philippines: Biology and stock assessment*, *ICLARM Technical Reports 7*, (eds Pauly D, Mines AN), pp. 124. Institute of Fisheries Development and Research, College of Fisheries, University of the Philippines in the Visayas, Quezon City, Philippines; International Center for Living Aquatic Resources Management, Manila, Philippines; and the United Nations University, Tokyo, Japan.

Nicastro KR, Zardi GI, McQuaid CD, Teske PR, Barker NP (2008) Coastal topography drives genetic structure in marine mussels. *Marine Ecology Progress Series,* **368**, 189-195.

Obura D (2012) Evolutionary mechanisms and diversity in a western Indian Ocean center of diversity. Proceedings of the 12th International Coral Reef Symposium, Cairns, Australia.

Okumuş I, Çiftci Y (2003) Fish population genetics and molecular markers: II- molecular markers and their applications in fisheries and aquaculture. *Turkish Journal of Fisheries and Aquatic Sciences,* **3**, 51-79.

Olbers JM, Fennessy ST (2007) A retrospective assessment of the stock status of *Otolithes ruber* (Pisces: Sciaenidae) as bycatch on prawn trawlers from KwaZulu-Natal, South Africa. *African Journal of Marine Science,* **29**, 247-252.

Palumbi SR, Martin A, Romano S, McMillan WO, Stice L, Grabowski G, (1991) The simple fools guide to PCR. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu, Hawaii.

Palumbi SR (1996) Nucleic acids II: The polymerase chain reaction. In: *Molecular Systematics* (eds Hillis DM, Moritz C, Mable BK), pp. 205–247. Sinauer & Associates Inc., Sunderland, Massachusetts.

Palumbi SR (2003) Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications,* **13**, 146-158.

Pauly D, Neal R (1985) Shrimp vs. fish in Southeast Asian fisheries: The biological, technological and social problems. In: *Recursos Pesqueros Potenciales de Mexico: La Pesca Acompanante del Camaron,* pp. 487-510*. Progr. Univ. de Alimentos, Inst. Cienc. del Mar y Limnol., Inst. Nal. de Pesca. UNAM, Mexico D.F.*

Phelan MJ, Gribble NA, Garrett RN (2008) Fishery biology and management of *Protonibea diacanthus* (Sciaenidae) aggregations in far Northern Cape York Peninsula waters. *Continental Shelf Research,* **28**, 2143-2151.

Pillai PKM (1983) On the biometry, food and feeding and spawning habits of *Otolithes ruber* (Schneider) from Porto Novo. *Indian Journal of Fisheries,* **30**, 69-73.

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

Quartly GD, Srokosz MA (2004) Eddies in the southern Mozambique Channel. *Deep-Sea Research Part II: Topical Studies in Oceanography,* **51**, 69-83.

Ragionieri L, Cannicci S, Schubart CD, Fratini S (2010) Gene flow and demographic history of the mangrove crab *Neosarmatium meinerti*: A case study from the western Indian Ocean. *Estuarine, Coastal and Shelf Science,* **86**, 179-188.

Reiss H, Hoarau G, Dickey-Collas M, Wolff WJ (2009) Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries,* **10**, 361-395.

Rice WR (1989) Analyzing tables of statistical test. *Evolution,* **43**, 223-225.

Ridgway T, Hoegh-Guldberg O, Ayre D (2001) Panmixia in *Pocillopora verrucosa* from South Africa. *Marine Biology,* **139**, 175-181.

Ridgway T, Sampayo EM (2005) Population genetic status of the Western Indian Ocean: What do we know? *Western Indian Ocean Journal of Marine Science,* **4**, 1-9.

Ridgway T, Riginos C, Davis J, Hoegh-Guldberg O (2008) Genetic connectivity patterns of *Pocillopora verrucosa* in southern African Marine Protected Areas. *Marine Ecology Progress Series,* **354**, 161-168.

Rodrigues R, Santos S, Haimovici M, Saint-Paul U, Sampaio I, Schneider H (2013) Mitochondrial DNA reveals population structuring in *Macrodon atricauda* (Perciformes: Sciaenidae): A study covering the whole geographic distribution of the species in the southwestern Atlantic. *Mitochondrial DNA*, Early Online: 1–7.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics,* **145**, 1219-1228.

Rousset F (2008) Genepop'007: A complete reimplementation of Genepop software for Windows and Linux. *Molecular Ecology Resources,* **8**, 103-106.

Santos S, Hrbek T, Farias IP, Schneider H, Sampaio I (2006) Population genetic structuring of the king weakfish, *Macrodon ancylodon* (Sciaenidae), in Atlantic coastal waters of South America: Deep genetic divergence without morphological change. *Molecular Ecology,* **15**, 4361-4373.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics,* **152**, 1079-1089.

Schouten MW, De Ruijter WPM, Van Leeuwen PJ, Ridderinkhof H (2003) Eddies and variability in the Mozambique Channel. *Deep-Sea Research Part II,* **50**, 1987-2003.

Schultz N (1992) Preliminary investigations on the population dynamics of *Otolithes ruber* (Sciaenidae) on Sofala Bank, Mozambique. In: *Population dynamics of exploited fishes and crustaceans in Mozambique: Contributions from a course on the "Use of Computers in Fisheries Research",* (ed Pauly D), pp 41-49. Instituto de Investigaque Pesqueira, Maputo, Mozambique.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellites. *Ecology Letters,* **9**, 615-629.

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

Silberschneider V, Gray CA (2008) Synopsis of biological, fisheries and aquaculture-related information on mulloway *Argyrosomus japonicus* (Pisces: Sciaenidae), with particular reference to Australia. *Journal of Applied Ichthyology,* **24**, 7-17.

Silberschneider V, Gray CA, Stewart J (2009) Age, growth, maturity and the overfishing of the iconic sciaenid, *Argyrosomus japonicus*, in south-eastern, Australia. *Fisheries Research,* **95**, 220-229.

Stobutzki I, Jones P, Miller M (2003) A comparison of fish bycatch communities between areas open and closed to prawn trawling in an Australian tropical fishery. *ICES Journal of Marine Science,* **60**, 951-966.

Tajima F (1989) Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics,* **123**, 585-595.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony methods. *Molecular Biology and Evolution* (submitted).

Teikwa ED, Mgaya YD (2003) Abundance and reproductive biology of the penaeid prawns of Bagamoyo coastal waters, Tanzania. *Western Indian Ocean Journal of Marine Science*, **2**, 117-126.

Theisen TC, Bowen BW, Lanier W, Baldwin JD (2008) High connectivity on a global scale in the pelagic wahoo, *Acanthocybium solandri* (tuna family Scombridae). *Molecular Ecology*, **17**, 4233-4247.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research,* **22**, 4673-4680.

van der Elst R (1981) *A guide to the common sea fishes of southern Africa*, 1st edn. C. Struik Publishers, Cape Town.

van der Elst R, Everett B, Jiddawi N, Mwatha G, Afonso PS, Boulle D (2005) Fish, fishers and fisheries of the Western Indian Ocean: Their diversity and status. A preliminary assessment. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* **363**, 263-284.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes,* **4**, 535–538.

Viñas J, Gordoa A, Fernández-Cebrián R, Pla C, Vahdet Ü, Araguas RM (2010) Facts and uncertainties about the genetic population structure of Atlantic bluefin tuna (*Thunnus thynnus*) in the Mediterranean. Implications for fishery management. *Reviews in Fish Biology and Fisheries*, 1-15.

Visram S, Yang MC, Pillay RM, Said S, Henriksson O, Grahn M, Chen CA (2010) Genetic connectivity and historical demography of the blue barred parrotfish (*Scarus ghobban*) in the western Indian Ocean. *Marine Biology,* **157**, 1475-1487.

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

Waelbroeck C, Labeyrie L, Michel E, Duplessy JC, McManus JF, Lambeck K, Balbon E, Labracherie M (2002) Sea-level and deep water temperature changes derived from benthic foraminifera isotopic records. *Quaternary Science Reviews,* **21**, 295-305.

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity,* **89**, 438-450.

Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: How can we do it better? *Fish and Fisheries,* **9**, 423-449.

Was A, Gosling E, Hoarau G (2010) Microsatellite analysis of plaice (*Pleuronectes platessa L*.) in the NE Atlantic: Weak genetic structuring in a milieu of high gene flow. *Marine Biology,* **157**, 447-462.

White C, Selkoe KA, Watson J, Siegel DA, Zacheri DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences,* **277**, 1685-1694.

White TA, Stefanni S, Stamford J, Hoelzel AR (2009) Unexpected panmixia in a long-lived, deep-sea fish with well-defined spawning habitat and relatively low fecundity. *Molecular Ecology,* **18**, 2563-2573.

Williams ST, Jara J, Gomez E, Knowlton N (2002) The marine Indo-West Pacific break: Contrasting the resolving power of mitochondrial and nuclear genes. *Integrative and Comparative Biology,* **42**, 941-952.

Worm B, Barbier EB, Beaumont N, Duffy JE, Folke C, Halpern BS, Jackson JBC, Lotze HK, Micheli F, Palumbi SR, Sala E, Selkoe KA, Stachowicz JJ, Watson R (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science,* **314**, 787-790.

Yue GH, David L, Orban L (2007) Mutation rate and pattern of microsatellites in common carp *(Cyprinus carpio L.*). *Genetica,* **129**, 329–331.

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PD (2009) DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources,* **9**, 237-242.

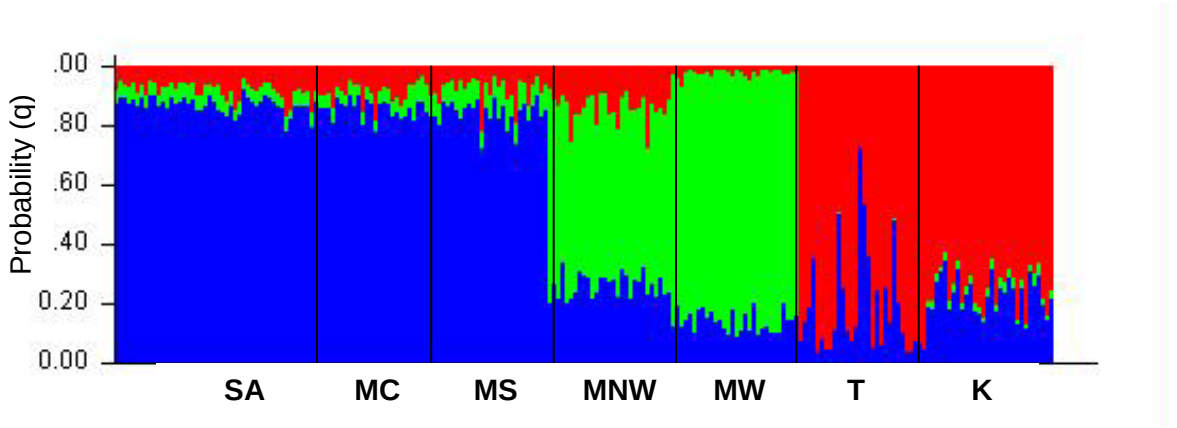
3.6 Appendices

Appendix 3A MIGRATE estimates for effective population size and gene flow between defined groups of *O. ruber*: 1) south and north, 2) south, Madagascar and north. S: South population, N: North population, MAD: Madagascar population. The Theta parameter is given by Θ = Neμ; M is the migration parameter (product of immigration rate and mutation rate).

Parameter	Estimated value	2.5%-97.5% confidence interval
Θ_S	0.00090	0 - 0.00260
Θ_N	0.00023	0 - 0.00200
$M_{S->N}$	0.3	0 - 242.0
$M_{N->S}$	133.7	0 - 436.7
Θ_S	0.00057	0 – 0.00220
Θ_{MAD}	0.00057	0.00227
Θ_N	0.00003	0.00200
$M_{MAD->S}$	0.3	414.0
$M_{N->S}$	107.7	456.0
$M_{S->MAD}$	0.3	378.7
$M_{N->MAD}$	0.3	265.3
$M_{S->N}$	0.3	243.3
$M_{MAD->N}$	0.3	258.7

Appendix 3B Microsatellite diversity indices for *Otolithes ruber* from seven locations in the South West Indian Ocean. N: Number of samples, H_E: Expected heterozygosity, H_O: Observed heterozygosity, F_{IS}: Fixation index, N_A: Number of alleles recovered in *O. ruber* from each locality.

	Soc508	**Soc826**	**Soc243**	**Soc232**	**UBA851**	**Soc140**	**JB09**
K (N)	30	30	30	30	30	30	30
H_e	0.8983	0.7439	0.4961	0.1261	0.4800	0.8672	0.6733
H_o	0.9000	0.6667	0.3000	0.1333	0.6000	0.9667	0.7000
N_A	13	12	3	3	2	13	8
F_{IS}	0.015	0.121	0.410	-0.040	-0.234	-0.098	-0.023
T (N)	30	30	30	30	30	30	30
H_e	0.8978	0.8011	0.2778	0.3139	0.3650	0.8578	0.6411
H_o	0.9333	0.6667	0.2000	0.3667	0.4000	0.9000	0.6000
N_A	11	12	2	4	3	15	9
F_{IS}	-0.023	0.184	0.296	-0.152	-0.079	-0.032	0.081
MC (N)	25	24	28	28	27	27	26
H_e	0.8944	0.9045	0.2283	0.1027	0.5425	0.8909	0.7944
H_o	0.7600	0.8333	0.2143	0.1071	0.4444	0.7778	0.8462
N_A	14	14	5	3	3	14	9
F_{IS}	0.170	0.100	0.080	-0.025	0.199	0.146	-0.046
MS (N)	33	33	33	33	33	33	33
H_e	0.8838	0.8691	0.2929	0.1428	0.5285	0.9105	0.8058
H_o	0.9091	0.7879	0.1212	0.1515	0.4545	0.8788	0.8485
N_A	16	13	5	3	3	17	9
F_{IS}	-0.013	0.109	0.596	-0.046	0.155	0.050	-0.038
SA (N)	39	39	40	40	40	40	39
H_e	0.8941	0.8373	0.1191	0.2659	0.5584	0.9069	0.7498
H_o	0.8974	0.8462	0.0750	0.3000	0.7250	0.9500	0.7179
N_A	16	12	3	4	3	17	7
F_{IS}	0.009	0.002	0.381	-0.116	-0.287	-0.035	0.055
MW (N)	29	29	30	30	30	30	30
H_e	0.8995	0.8424	0.0950	0.4428	0.5422	0.8711	0.8278
H_o	0.9310	0.8966	0.1000	0.4667	0.3667	0.8000	0.8333
N_A	16	11	2	3	4	13	9
F_{IS}	-0.017	-0.047	-0.036	-0.037	0.339	0.098	0.010
MNW (N)	27	25	28	29	28	25	26
H_e	0.9005	0.8152	0.2226	0.4370	0.5134	0.8832	0.8262
H_o	0.9259	0.8400	0.1071	0.5172	0.4286	0.8000	0.8846
N_A	14	10	3	3	3	15	10
F_{IS}	-0.009	-0.010	0.532	-0.167	0.183	0.114	-0.051



Appendix 3C STRUCTURE clustering analysis of 221 *Otolithes ruber* microsatellite genotypes from the South West Indian Ocean. Each vertical bar represents the posterior probability (q) of an individual assigning to one of three (K=3) clusters.

Chapter 4

Dissertation Conclusion

Knowledge of the connectivity among populations of a species is crucial for sustainable utilization and management. Connectivity among populations likely plays an important role in adaptation and long-term persistence. Genetic methods are applicable to fisheries management in order to assess the level of subpopulation connectivity and divergence - an important requirement if separate management is indicated (Reiss *et al.* 2009). Genetic connectivity is mostly explained by early life stage dispersal and retention in the light of biological and physical processes; however, there are also species demonstrating connectivity that relies on adult life stages (behavioural ecology of adults) (Frisk *et al.* 2013).

Improved molecular approaches and the availability of several molecular markers have simplified studies of marine species' distribution of populations/lineages and the genetic connectivity among populations. Before large-scale data generation is undertaken, critical selection of molecular markers is a necessity. As represented by the technical part of this investigation, a small pilot study is very helpful to save time and money (sequencing/genotyping all the loci in all the individual samples, doing analysis and removing those that are not informative, or that represent locus-related problems). Instead of screening all the available markers in all the specimens in this study, I was able to identify the most informative molecular tools for the study using a representative number of samples of the SWIO shallow water demersal sciaenid, *Otolithes ruber*. I also showed that a locus' polymorphism might differ from that expected, since some of the microsatellites in this study showed higher polymorphism in *O. ruber* compared to the source species. In cases where few molecular tools are available, testing of markers with no or low levels of polymorphism in their source species could be undertaken.

Environmental degradation and reductions in natural resources and biodiversity are reported in SWIO countries, including Kenya, Mozambique, Somalia, South Africa, Tanzania, Comores, Madagascar, Mauritius, Reunion, and the Seychelles (Berg *et al.* 2002). *Otolithes ruber* is one of the species exposed to overfishing in the region. The aim of this study was to investigate genetic variation in *O. ruber* in the SWIO and to assess the geographic structure of this variation using the mitochondrial *Cyt b* gene and seven polymorphic microsatellites. Data obtained from both markers demonstrated that populations of *O. ruber* are genetically diverse in the SWIO and not represented by a single genetic unit. There does not appear to

be a significant relationship between genetic and geographic distance. There is a strong signal for the presence of at least two populations with a low level of mixing between them (north = Kenya and Tanzania; south = Mozambique, South Africa and Madagascar). Although Madagascan samples were not differentiated from the rest to the same extent as Kenyan and Tanzanian samples, management may consider the existence of three subpopulations of *O. ruber* in the SWIO region. The lineages are distinguishable from each other and the microsatellite analyses suggest greater connectivity in the region compared to mitochondrial data. Although there is bi-directional movement, there appears to be more movement from north to south. Historical factors, possible biological differences between sexes and the oceanographic features in the region play a role in determining the observed genetic structuring in the species.

The primary aim of conservation and fisheries management is to ensure sustainable use of marine resources (Reiss *et al.*, 2009). Under such management, i.e. where a complex population is managed as a single population, decline in local populations, reduced productivity (Worm *et al.*, 2006) and extinction of subpopulations can consequence (Reiss *et al.*, 2009). Additionally, loss of genetic diversity is an impact of overexploitation in such population and, hence, the ability of the species to adapt and to evolve to new environmental circumstances is affected (Reiss *et al.*, 2009).

Given the concern regarding overexploitation of *O. ruber* and the presented results, the species needs separate management for the observed populations, especially for the northern and southern stocks. The northern population needs priority to be protected in terms of size, demography and the observed genetic diversity, since it does not receive many migrants from the south. The transboundary nature of both populations necessitates regional cooperation to ensure sustainable management of this living marine resource. More fine scale analysis should be conducted to characterize gene flow connections in the region. In particular, more widespread sampling off Madagascar could inform whether the species is represented by two or three stocks in the region. Stock assessments should be undertaken for independent populations and any differences in general biology explored. Connectivity within the WIO should be assessed on a finer scale. Broader connectivity within the Indo-West Pacific should also be investigated given the previously reported genetic distinction (Zemlak *et al.* 2009) between South African and Australian *O. ruber* populations.

4.1 References

Berg H, Francis J, Souter P (2002) Support to marine research for sustainable management of marine and coastal resources in the Western Indian Ocean. *Ambio,* **31**, 597-601.

Frisk MG, Jordaan A, Miller TJ (2013) Moving beyond the current paradigm in marine population connectivity: are adults the missing link? *Fish and Fisheries* doi: 10.1111/faf.12014.

Reiss H, Hoarau G, Dickey-Collas M, Wolff WJ (2009) Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries,* **10**, 361-395.

Worm B, Barbier EB, Beaumont N, Duffy JE, Folke C, Halpern BS, Jackson JBC, Lotze HK, Micheli F, Palumbi SR, Sala E, Selkoe KA, Stachowicz JJ, Watson R (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science,* **314**, 787-790.

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PD (2009) DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources,* **9**, 237-242.