Population genetics of the cichlid, *Cynotilapia afra* (Günther 1894), in its native and introduced ranges in Lake Malawi

PhD Thesis
for the degree of
DOCTOR OF PHILOSOPHY
from University of Hull

By

Hastings K. Zidana

BSc. (University of Malawi, Malawi)
MSc. (National Taiwan Ocean University, Taiwan)

January, 2010

Abstract

Species introductions can provide unplanned and occasionally replicated experiments that can be studied to understand fundamental ecological and evolutionary processes associated with range expansions in the natural world. The cichlid species flock of Lake Malawi consists of an estimated 451-800 species and is a textbook example of explosive speciation that has been studied as a model system of evolution in the past three decades. In addition, fish are of major socio-economic importance to Malawian people, and they form an important source (circa 70%) of animal protein in their diet. Furthermore, fisheries activity employs 3% of the country's population and contributes to 4% of the country's Gross Domestic Product (GDP).

This thesis studies a well documented, human mediated introduction of a rocky-shore, plankton-feeding cichlid fish *Cynotilapia afra* into Lake Malawi National Park. This introduction has important evolutionary and ecological consequences on the native populations of *Pseudotropheus zebra*, and here I investigate its population genetic impacts using contemporary molecular genetic tools and analyses. Three hypotheses were tested:

1. Introduction events are usually associated with a small founder population size, and the resulting genetic bottleneck is expected to reduce genetic variation of *C. afra* in the introduced range.
2. The invasive scenario during the introduction followed a stepping stone pattern, or alternatively, it occurred as several independent introductions of *C. afra* in Lake Malawi National Park.
3. Introgressive hybridisation between *C. afra* and *P. zebra* may have facilitated the introduction of the invading *C. afra* population and restored its depleted genetic variation associated with the founder event. Furthermore, the gene pool of *C. afra* has more non-native genetic material as compared to *P. zebra*.

Samples were collected from six native and four introduced populations of *C. afra*, as well as three native populations of *P. zebra*. The latter species is from a different genus, although laboratory experiments indicate that both species hybridise in laboratory conditions. Sequence variation in the mitochondrial DNA (mtDNA) control region was analysed using 15 individuals per sample population and 60 individuals per sample population were

genotyped at six microsatellite loci. These data were analysed to test the three hypotheses and identify potential source populations, infer introduction patterns (stepping stone or independent), and deduct whether introgressive hybridisation may have facilitated the founder event and subsequent establishment of *C. afra* in the invasive range in Lake Malawi Natural Park. The three data chapters in this thesis discuss the findings of the mtDNA sequence study (Chapter 2) and the microsatellite study (Chapter 3). In Chapter 4, I analyse the microsatellite data in further detail and consider the role of introgression by using Bayesian analysis tools.

The mtDNA study presented in Chapter 2 reveals that *C. afra* and *P. zebra* mtDNA sequences show high levels of lineage sorting (i.e. the DNA sequences of both species are remarkably distinct). This finding is in sharp contrast to previous studies on Lake Malawi rock-dwelling cichlids which have shown that cichlid species share the same or very similar mtDNA haplotypes. Furthermore, the introduced populations showed a higher sequence and haplotype diversity than their native counterparts. This analysis suggests that elevated gene diversity was largely due to *C. afra* populations being founded by individuals from several genetically distinct and geographically separate populations. In Chapter 4, I discuss the role of introgressive hybridisation with native *P. zebra*, and its impact on mtDNA variation in the introduced *C. afra* gene pool.

In Chapter 3, I show that in contrast to the signal obtained from the mtDNA, the genetic variation at the microsatellite loci exhibited a significant reduction in the introduced range. Introduced *C. afra* populations have a lower mean effective number of alleles (n_e) than *C. afra* populations in their native range. I use an approximate Bayesian analysis and show compelling evidence that at least two independent introductions have contributed to the introduced *C. afra* gene pool, a conclusion that is supported by high probability values. This conclusion differs from that of previous studies which suggested a stepping stone introduction pattern around Thumbi West Island. Surprisingly, a population of *C. afra* at Domwe Island was founded by a source population from Thumbi West Island, and this stepping stone introduction pattern is supported with a high probability (95%).

Microsatellite analysis furthermore suggests that the founder event of *C. afra* in Lake Malawi National Park was associated with strong genetic drift associated with a genetic bottleneck. I was not able to detect this signal from the mtDNA genetic marker alone, which showed an increase in genetic variation at the mtDNA due to different source populations

contributing to the founder event (see Chapter 2). These combined studies reported in Chapter 2 and 3 thus demonstrate that microsatellites may be well-suited to investigate questions related to conservation issues such as bottlenecks associated with founder events, while mtDNA is more suited to reveal the evolutionary processes and establish different source populations that have contributed to the introduction.

In Chapter 4, I analyse the level of genetic differentiation at microsatellite loci, and show that the introduced *C. afra* and native *P. zebra* populations at Thumbi West Island are genetically more similar (G'_{ST}=0.36±0.05) than the species-pair at Otter point (G'_{ST}=0.94±0.18) and Domwe Island (G'_{ST}=0.55±0.09). In addition, *C. afra* and *P. zebra* at Thumbi West Island showed a lower genetic distance than allopatric *C. afra* or *P. zebra* populations from Otter point and Domwe Island. Further analysis using a Bayesian assignment approach supports previous findings and demonstrates the likelihood of introgressive hybridisation between an introduced *C. afra* and a native *P. zebra* population at Thumbi West Island. No evidence of introgression is found at Otter point and Domwe Island, where the *C. afra* and *P. zebra* populations show distinct genetic structure. The occurrence of introgressive hybridisation at Thumbi West between species from distinct genera shows that translocations can have a dramatic impact even on the gene pools of heterospecific recipient populations. The results from this work have crucial implications in evolution of cichlid fishes and in invasion biology when predicting the evolution of invasiveness.

In summary, the thesis shows that hybridisation, as well as the introduction of multiple genetically differentiated source populations has increased the genetic diversity of introduced *C. afra* populations, and this may have facilitated their establishment in Lake Malawi National Park. Translocation of cichlid species in Lake Malawi can have a dramatic impact even on heterospecific gene pools.

Acknowledgements

I wish to express my sincere gratitude to my supervisors: Dr Cock van Oosterhout, Dr Bernd Hänfling and Prof George Turner, who have contributed tirelessly towards the design and implementation of the work of this thesis. They provided comments on each and every chapter of this thesis time and again. Their patience and trust I will always remember and make use of in years to come.

I am thankful to financial support from: Commonwealth Scholarship Fund, Department of Fisheries Malawi, Department of Biological Sciences in the University of Hull, Biology Department at Chancellor College in Malawi, and Department of Parks and Wildlife in Malawi. Without their support this work could have not been possible.

I would also like to thank Dr Martin Genner and Dr Domino Joyce for their support with sample collection and comments on the draft manuscripts of the main chapters of this thesis. Ad Konings for fruitful discussions on introduced species and permission to use any of the pictures in his authoritative book on Lake Malawi rock dwelling cichlids. The comments and personal communications that we exchanged helped to improve the work of this thesis.The fisherman in Malawi who helped during sample collection; we braved the high waves from the south-easterly wind (mwera) along Ntekete and Makanjira collecting *Cynotilapia afra* samples after a sleepless night at Chiofu Bay camp, I am thankful for that.

All members of the Molecular Ecology group at the University of Hull, who provided support and comments during lab talk presentations on the main chapters and practical help during lab work. Their efforts made a difference towards reorganization and technical work of this thesis.

Last but not least, I would like to thank my family both in the United Kingdom (UK) and Malawi for their continuing support and encouragements. Their effort and financial assistance during sampling in Malawi and studying in the UK is greatly appreciated.

Contents

CHAPTER 1

1. GENERAL INTRODUCTION

1.1 *Cichlid fish as model system of rapid speciation in evolutionary studies and geological history of Lake Malawi*

Lake Malawi is situated in the southern most part of the African rift valley system. It is the ninth largest lake in the world, third largest in Africa with a surface area of 30,800 km^2 and a maximum depth of over 700m (Delvaux 1995). The Lake Malawi rift basin started to develop about 8.6 million years ago (MYA), but deepwater conditions were acquired only 4.5 MYA (Sturmbauer *et al.* 2001). Geological inferences suggests that Lake Malawi was severely affected by a change to a drier climate in the late Pliocene/early Pleistocene, resulting in a drop of the lake level by approximately 250-500m compared to its present level from 0.57 to 1.6 MYA (Delvaux 1995; Cohen *et al.* 2007) and even as recent as 25,000 to 135,000 years ago (Scholz & Rosendahl 1988). Most of the southernside of Lake Malawi was completely dry. The lake is fresh water with pH ranges of 7.8 to 8.5 and surface water temperature varies between 23-28°C (Konings 2008). Population genetic studies have shown that dynamics of diversification events in African cichlid fishes are likely to be connected to these lake level fluctuations (Sturmbauer & Meyer 1992; Johnson *et al.* 1996; Sturmbauer *et al.* 2001).

The cichlid fishes are a text book example of explosive radiation. In the past three decades they have been used as a model system of rapid speciation in evolution studies (see reviews by Kornfield & Smith 2000; Kocher 2004; Seehausen 2006). The African Great Lakes Victoria, Tanganyika and Malawi contain the richest lacustrine fish faunas in the world. In terms of cichlids alone there have been estimates of 451-800 species in Lake Malawi, 447-535 in Lake Victoria (and many more in neighbouring lakes) and 162-184 in Lake Tanganyika (Genner *et al*. 2004; Konings 2008).

The lakes vary in size, shape, age and water parameters and the high diversity of cichlid fishes inhabiting them have been derived from different riverine ancestors (Salzburger *et al.* 2005). The cichlid fishes of Lake Tanganyika show a wide range of reproductive

strategies not found in the younger and less genetically diverse cichlid faunas of the other two lakes. However, all three lakes are dominated by cichlid fish, which exhibit striking parallels in morphology (Young *et al.* 2009), behaviour (Blais *et al.* 2009), life histories (Duponchelle *et al.* 2008) and even opsin amino acid sequences (Sugawara *et al.* 2005).

Several hypotheses have been suggested to explain the mechanisms behind explosive cichlid speciation: morphology, particularly of pharyngeal jaw apparatus (Fryer 1959; Liem 1980), sexually-selected colour variation (Kosswig 1947; Dominey 1984; van Oppen *et al.* 1998), habitat specialisation and fragmentation (Fryer 1959; Ribbink *et al.* 1983), hybridisation (Salzburger *et al.* 2002, Seehausen 2004) and visually-based mate choice in clear water habitats (Seehausen *et al.* 1997; Terai *et al.* 2006). Roles for both allopatric (Fryer 1959; Ribbink *et al.* 1983; van Oppen *et al.* 1997) and sympatric (Seehausen & van Alphen 1997; Shaw *et al*. 2000) speciation modes have been proposed.

There is a high level of geographic colour variation among the rock-dwelling cichlids (Fryer 1959; Ribbink *et al.* 1983) which is associated with strong genetic structuring (van Oppen *et al.* 1997; Rico *et al.* 2003) and partial reproductive isolation (Knight & Turner 2004). This suggests a major role for geographic isolation in diversification and evolution (Ribbink *et al.* 1983; Reinthal 1993; Genner *et al.* 2004a & b). Differences in parasite fauna are also known among species and habitats (Blais *et al.* 2007; Amin *et al*. 2008; Bray *et al.* 2006).

1.2 *General biology of rock-dwelling cichlids "mbuna" group in Lake Malawi*

The name "*mbuna*" comes from a "*Chitumbuka*" language spoken in northern Malawi, and the term has no formal taxonomic validity (Genner & Turner 2005). The group exhibit high levels of endemism to individual islands and stretches of rocky habitat along the lake shoreline. Sandy shorelines and deep water act as boundaries to dispersal (Ribbink *et al.* 1983; Reinthal 1993; Genner *et al.* 2004a & b).

Females are mouth brooders: they keep their offspring in their mouths for an estimated period of 20-30 days before they are released as free-swimming juveniles (Kellog *et al*. 1995; Genner & Turner 2005). Mostly, males are territorial and females are suggested to breed throughout the year, often brooding young in a secluded cave under the rocks, which acts as a refuge from predators (Hert 1992; Kellog *et al*. 1995; Robinson & Ribbink 1998; Konnings 2008). Molecular genetic data has revealed existence of multiple paternity in

several species (Kellogg *et al*. 1995; Blais *et al.* 2009). Mortality is high amongst the free-swimming juvenile stage, estimated to be more than 95% within the first 40 days after being released from females mouth (Trendall 1988).

Most mbuna are sexually dimorphic with males attaining larger sizes, longer pelvic, dorsal and anal fins, and exhibiting brighter colours (Seehausen & Schluter 2004; Genner & Turner 2005). Several factors have been suggested to be involved in mate recognition and female choice including colour, territory quality, male size, pattern symmetry, courtship vigour, length and condition of fins, courtship sounds and pheromones (Genner & Turner 2005; Genner *et al.* 2008; Young *et al.* 2009c).

Diet is suggested to be broad for most species. They feed on primarily algae, plankton, benthic invertebrates or fish fry (Trewavas 1935; Fryer 1959; Konnings 2008). The anatomical features associated with feeding habits are pharyngeal jaws for prey processing, and teeth and outer jaw structures that are used for prey handling. Some species appear primarily adapted for combing algae to collect loose strands and periphyton from rock substrates, while others appear better adapted for collecting zooplankton from open waters (Reinthal 1990). Morphological trophic structures are used in taxonomy studies as a main characteristic to distinguish among several of the mbuna genera, for example *Cynotilapia* has unicuspid teeth in outer rows, while *Pseudotropheus* typically has bicuspid teeth (Trewavas 1935; Fryer 1959; Konnings 2008) (Fig. 1.1).

Many mbuna have small fragmented geographic ranges and narrow habitat preferences and all lack a pelagic dispersal phase (Genner & Turner 2005). Despite that, mbuna are widely distributed across Lake Malawi's rocky shores. For example, *Cynotilapia afra* and *Pseudotropheus zebra* are distributed from Ikombe in the northernmost part of the lake, to Mbenji Island and Makanjira point in the south (Kassam *et al.* 2005; Konings 2008). (Appendix IV.1).

1.3 *Study species:* Cynotilapia afra *and* Pseudotropheus zebra *in Lake Malawi*

The genus *Cynotilapia* is among one of the eleven described mbuna cichlid genera (*Cyathochromis*; *Genyochromis*; *Gephyrochromis*; *Iodotropheus*; *Labeotropheus*; *Labidochromis*; *Maylandia*; *Melanochromis*; *Petrotilapia* and *Pseudotropheus*) in Lake Malawi. *Cynotilapia* means dog-tilapia because of its possession of unicuspid canine teeth on the outer lower jaw (Konings 2008) (Fig.1.1).

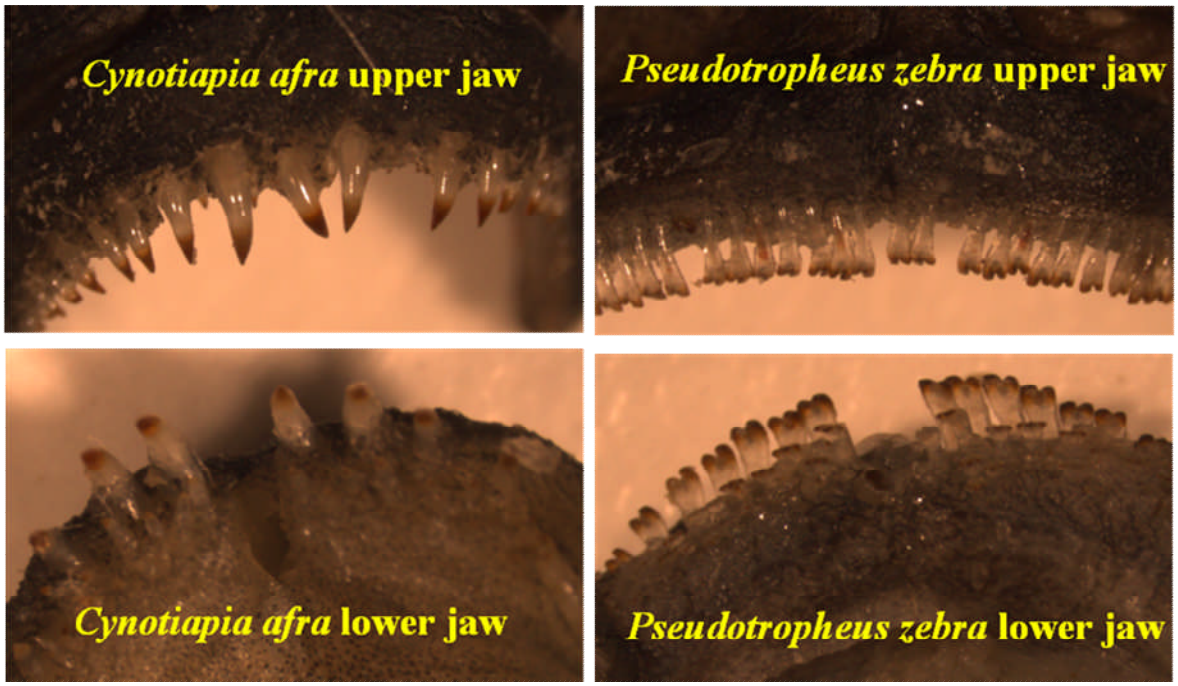

Fig.1.1. Showing differences in the tooth morphology of the outer upper and lower jaws between *C. afra* (uncispid) and *P. zebra* (bicuspid). This is the main physical difference which differentiates the two species.

Cynotilapia afra is endemic to Lake Malawi and indigenous to northern and central parts (Fig. 1.2 & Appendix Fig. IV.1). It prefers deeper regions (>50m) with sediment free rocky habitats and feeds on zooplankton in the water column (Genner *et al.* 1999; Konings 2008). Most adult "mbuna" cichlids are restricted to rocky habitats (Ribbink *et al.* 1983; Hert 1992), but *C. afra* is found in open waters (Konings 2008). It has a wide but fragmented distribution range: on the eastern side of the lake it is found from Ikombe to Lumbaulo, Cobwe to Mbweca, Tumbi Point to Makanjira and on the Islands of Likoma and Chizumulu (Konings 2008). On the western shore line it is found on the rocky shores of the Island of Mbenji and Jalo reef and from Kande Island to Ngara (Konings 2008) (Fig. 1.2 & Appendix Fig. IV.1).

The breeding males show a great variability in colouration exhibiting blue barred, yellow, orange, white, or black dorsal fin and forehead (Konings 2008, H. Zidana personal observation at Nkhata Bay) (Fig. 1.2). Females have a light blue-brown colour and do not have a vertical blue black barring (Konings 2008).

There are two species which have been described under the genus *Cynotilapia*: *C. afra* (Günther 1894) and *C. axelrodi* (Burgess 1976), and phylogenetic studies suggest that

the genus *Cynotilapia* is closer to *Pseudotropheus* among the rock dwelling cichlid group (Kassam *et al.* 2005).

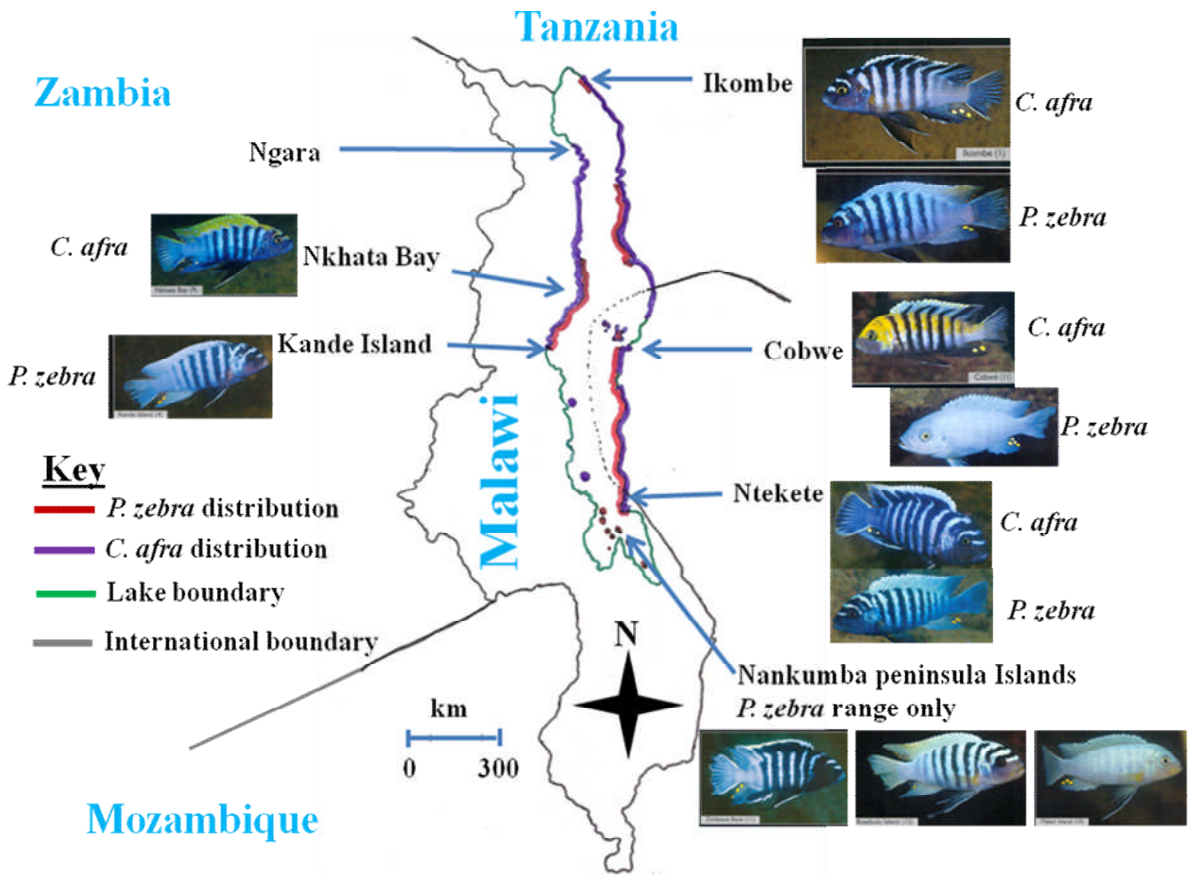

Fig. 1.2. Map of Malawi with neighbouring country boundaries, showing the natural distribution ranges of *C. afra* (purple line) and *P. zebra* (red line) in Lake Malawi. All pictures are of males in their breeding colours (pictures taken from Ad Konings 2008, with permission).

The type species *Psuedotropheus williamsi* was the first to be described in the genus *Psuedotropheus* (Regan 1922). The genus contains twenty five described species and some are suggested to be dissimilar to the type species *P. williamsi* (Trewavas 1984). Two other subgenera *Maylandia* (Meyer & Foerster 1984) and *Metriaclima* (Stauffer *et al.* 1997) have been suggested but they are still contested (Konings & Stauffer 2006). All three genera are used in the literature, but in this thesis I have followed the genus name of the type species, *Pseudotropheus*.

The species studied in this thesis is *P. zebra* (Konings 2008). *P. zebra* habitat preference is clear, sediment free rocky coast and the species does not have a continuous distribution (Ribbing *et al.* 1983; Konings 2008). On the eastern shoreline natural distribution is from Ikombe to Lumbila, Lundu to Mbamba Bay, the Islands of Likoma and Chizumulu, Mara Point to Makanjira, the Nankumba peninsula Islands and further south at Boadzulu Island (Ribbink *et al.* 1983; Konings 2008). On the eastern shore line the distribution ranges from Cape Manulo to Kande Island and Namalenje Island (Ribbink *et al.* 1983; Konings 2008) (Fig. 1.2 & Appendix Fig. IV.1). In the natural distribution range *P. zebra* occurs sympatrically with *C. afra* (Ribbink *et al.* 1983; Kassamu *et al.* 2005). The breeding males show a great variability in colouration, with a vertical blue barred body, and a white and black banded dorsal fin (Fig. 1.2). Similar colourations are also exhibited in *C. afra* populations, for example *C. afra* at Ntekete and *P. zebra* at Zimbabwe Rock (Nankumba Peninsula area) all show a black banded dorsal fin (Fig.1.2). Nevertheless, all *P. zebra* populations that were used in this thesis possessed a white dorsal fin and all but Nkhata Bay *C. afra* populations had a black banded dorsal fin.

1.4 *Fish introductions in African Lakes and "mbuna" translocations in Lake Malawi*

There are three conservation issues concerning the management of fish resources in African lakes: introduction of fish species, impoundment and heavy exploitation (Pitcher & Hart 1995). All three factors have contributed to large sustained changes in species composition of African lakes, but in this thesis my interest is the impact of introduced fish species.

The introduction of fish from their endemic localities to new environments has been an ongoing activity in the world (Welcome 1988). A synopsis of worldwide fish introductions documents a total of 1354 introductions of 237 species into 140 countries dating from middle of the 19th century (Welcome 1988; Pitcher & Hart 1995). African lakes have experienced 147 (11%) of the total world introductions (Pitcher & Hart 1995), but have received a high profile due the well documented predation characteristics of the introduced Nile perch (*Lates niloticus*) in Lake Victoria (Witte & Goudswaard 1985; Ohwayo 1990). There are several reasons why fish are introduced to new environements and introduction for aquaculture purposes ranks highly in African Lakes (Table 1.1).

Table 1.1 Reasons given for fish introductions to African lakes.

Reason for introduction	Number	Percentage (%)
Aquaculture	72	49
Sport fishing	26	18
Create fishery	23	16
Control other fish	9	6
Control weed	6	4
Control mosquito	6	4
Control snail	2	1
Ornamental	1	1
Accidental	1	1
No reason given	1	1

The percentage is out of 147 introductions (Table taken from Pitcher & Hart 1995)

Fish introductions to African Lakes can be beneficial, for example, with reports of increased fish food production due to introduction of *Oreochromis niloticus* and *Limnothrissa miodon* (Pitcher & Hart 1995). But there are also detrimental impacts, for example, predation on native haplochromines and tilapia fish species in Lake Victoria by *Lates niloticus* (Witte & Goudswaard 1985; Ohwayo 1990). The cichlids of Lake Victoria have not co-evolved with *L. niloticus* and this explains their vunerability to predation (Pitcher & Hart 1995). Another conservation concern of species introductions is hybridisation with natives, which may compromise the native gene pool and increase competition for available resources (Ohwayo 1990; Canonico 2005).

Lake Malawi has not experienced introduction of fish species from outside the Lake. Here the main concern is intralacustrine introductions, mainly of the mbuna cichlids which are targeted by aquarium traders (Pitcher & Hart 1995; Turner 1995; Turner 1996; Genner e*t al*. 2006). During the past three decades, over 20 species of mbuna have been introduced to Thumbi West Island which is located in Lake Malawi National Park, while 12 species have been introduced to Nkhata Bay in northern part of the Lake (Ribbink *et al.* 1983; Munthali 1998; Genner *et al.* 2006). The introduced mbuna cichlid species compositions were well documented in a survey of Lake Malawi rock dwelling cichlids (Ribbink *et al.* 1983), but not much work has been done to assess the genetic diversity (Streelman *et al.* 2004; Zidana *et al.*

2009) and ecological interactions with natives (Munthali 1998; Stauffer *et al.* 1996;Young *et al.* 2009). The introduced populations have shown increased condition factor and fecundity compared to their native range (Munthali 1998) and their genetic variation is not depauperated (Streelman *et al.* 2004; Zidana *et al.* 2009).

1.5 *Establishment of Lake Malawi National Park*

The Lake Malawi National Park was established in 1980 as the World's first National Park targeted principally to conserve the freshwater fish, which are mainly of the "mbuna" group (Hough 1989). The park was designated a UNESCO World Heritage site in 1984 and it covers 9,400ha mostly in the southern part of the Lake Malawi (Hough 1989). Lake Malawi National Park is situated around Nankumba peninsula and other disjunct mainland areas (Fig.1.2). Habitat types vary from rocky shorelines to sandy beaches and from wooded hillsides to swamps and lagoons.

The management plan details 4 conservation zones within the park: special zone, wilderness zone, natural zone and general zone (Glenfell 1993). Islands and lacustrine areas are designated wilderness zones where no fishing is permitted, while terrestrial areas are designated natural zones. There is a management fishing zone of 2 km off the mainland shoreline where fishing is prohibited. There are two conservation concerns: the possible impact of introduced fish species on natives and illegal fishing by local people around the park area (Croft 1981). The waters of Lake Malawi National Park are among the most impacted with fish introductions (Genner *et al.* 2006), suggesting the need for biological studies to monitor the situation. The results from this thesis will provide crucial information that may be utilized during policy design and conservation management of fish stocks within the park.

1.6 *Evolutionary and ecological genetics of introduced populations*

The rate of human-mediated species introductions has been increasing worldwide (Mack *et al.* 2000; Allendorf & Luikart 2007). Identifying the source of introduced populations is an important prerequisite for development of effective prevention and management strategies (Kolar & Lodge 2002; Rollins *et al.* 2006). Knowledge of the origins of introduced species can also provide insight into their biological history and genetic

evolution (Collins *et al.* 2002; Downie 2002; Kolbe *et al*. 2004; Goolsby *et al.* 2006). Comparions can be made between introduced species and natives on phenotypic traits, allele frequencies and genetic diversity among introduced and native populations (Kolbe *et al.* 2004; Wares *et al.* 2005). Ecological studies designed to explore determinants of invasion success often require information on the source of introduced individuals to facilitate the selection of appropriate contrast groups (Colautti *et al.* 2004; Hierro *et al.* 2005).

The impact of introduced species on native species, communities and ecosystems has been widely recognized for decades (Elton 1958, Vitousek & Walker 1989; Rhymer & Simberloff 1996; Sakai *et al.* 2001). To date, the literature on introduced species has overwhelmingly emphasized their negative effects on biodiversity (Levin *et al.* 1996; Rhymer & Simberloff 1996; Frankham & Ralls 1998; Huxel 1999; reviewed in Vellend *et al.* 2007) and the economy (Pimentel 2000).

An introduction event may be followed by a rapid evolutionary change within a founder population, resulting in a population that is genetically diverse over space and time (Reznick *et al.* 2000; Mooney & Cleland 2001; Lee 2002; Cowx 2004; Lambrinos 2004, Strauss *et al.* 2006). Genetic characteristics of founder populations can be an important determinant of establishment success and range expansion after an introduction event (Tsutsui *et al.* 2000; Ellstrand & Schierenbeck 2000; Filchak *et al.* 2000; Carroll *et al.* 2001).

Evolutionary genetics studies play an important role in understanding the genetic changes and characteristics that determine establishment success and range expansion of introduced species. The classic symposium on “Genetics of colonizing species” (Baker & Stebbins 1965) has inspired many studies to investigate genetic mechanisms responsible for evolution of introduced populations. Studies have shown that sufficient additive genetic variance is needed for evolutionary adaptation in response to new environments (Pappert *et al.* 2000; Carroll *et al.* 2001). The underlying mechanism is that dominance and epistatic variance within small populations could be converted into additive genetic variance through genetic drift, and this may provide novel genetic substrate for selection to work on (Bryant & Meffert 1996; Cheverud *et al.* 1999; Soltis & Soltis 2000). This implies that a genetic bottleneck in a founder population might expose nonadditive genetic variance to selection and contribute to rapid rates of evolution observed during range expansion in introduced populations (Rieseberg *et al.* 1999; Soltis & Soltis 2000).

The differences in environment between native and introduced ranges can lead to diversifying selection through interaction of genotypic and environmental tradeoffs which

may sometimes lead to divergence of phenotypes (Filchak *et al.* 2000; Weinig 2000). Thus, small numbers of genes could have crucial impact on establishment and range expansion of introduced populations (Paterson *et al.* 1995; Krieger & Ross 2002). Another factor suggested to provide important genetic substrate for selection to work on in introduced populations is genomic rearrangements, such as chromosome inversions through transposable elements (Prevosti *et al.* 1988; Biemont *et al.* 1999). In an extreme case genetic drift alone has promoted success of an introduced population of Argentine ant *Linepithema humile* (Tsutsui *et al.* 2000). A loss of genetic diversity in *L. humile* at microsatellite loci was associated with reduced intraspecific aggression among spatially separate nests, which led to the formation of interspecifically dominant supercolonies (Tsutsui *et al.* 2000).

Hybridisation, either interspecific or among well-differentiated populations of the same species and through multiple introductions may be an important stimulus for the evolution of introduced species (Soltis & Soltis 2000; Ellstrand & Schierenbeck 2000; Kolbe 2004). Hybridisation can lead to adaptive evolution in a number of ways such as: evolutionary novelty whereby hybridisation between parental species and their first or second generation offspring may result in gene admixture, a phenomenon called introgression (Rieseberg & Wendel 1993; Arnold 1997). Through introgression extreme phenotypes from hybrid populations may be produced, a mechanism called transgressive segregation (Abbott 1992; Rieseberg 1999). Transgressive segregation gives rise to extreme phenotypes that may be better adapted to a new environment than any of the parental species (Abbott 1992; Arnold 1997; Neuffer *et al*. 1999; Ellstrand & Schierenbeck 2000). Previous studies have suggested several genetic factors that could be responsible for creation of transgressive phenotypes but the following stands out: expression of rare recessive alleles (Rick & Smith 1953) and complementary gene action (Vega & Frey 1980; de Vicente &Tanksley 1993). The mechanism is that recombination through hybridisation may generate increased genetic variation in hybrid lineages and this may be responsible for the evolutionary success of introduced populations (Neuffer *et al*. 1999; Kolbe *et al.* 2004). Recombination enables selection to purge some of the mutations that were fixed in the parental species, which reduces the frequency of detrimental alleles (Lande 1995; Richards 2000; Charlesworth & Willis 2009). Other mechanisms which have been suggested to stimulate evolutionary potential through hybridisation are fixed heterosis (hybrid vigour) which is usually expressed among polyploids and may provide a fitness boost (Thompson 1991; Abbott 1992; Ellstrand & Schierenbeck 2000). The polyploids may have characteristics that enable them to occupy a

niche that was not available to the parents (Mooney & Cleland 2001; Ellstrand & Schierenbeck 2000). Hybridisation that has resulted in creation of polyploid crops with different phenotypes to either of the parents is common among agricultural crops, for example: wheat, after millennia of hybridisation and modification by humans, has strains that are tetraploid (four sets of chromosomes) with the common name of durum or macaroni wheat, and hexaploid (six sets of chromosomes) with the common name of bread wheat (Lagudah *et al.* 1991; reviewed by Jules 2009).

Founder populations usually have small effective population sizes so are susceptible to inbreeding, because they accumulate detrimental mutations through a phenomenon called genetic load (Ellstrand & Schierenbeck 2000; Lopez *et al.* 2009). Hybridisation can afford them an opportunity to reduce this mutational load because different mutations are fixed in the two hybridizing species. Hybridisation among differentiated populations of the same species may act to promote evolution of introduced species in the same way as hybridisation among distantly related species (Ellstrand & Schierenbeck 2000). The potential for hybridisation between two distantly related taxa may last for many millions of years after they have diverged suggesting that post-zygotic isolation evolves slowly. A previous study on birds has suggested that bird species can produce fertile hybrids if they diverged up to 7-17 MYA, with complete hybrid sterility setting in if the species diverged more than around 11-55 MYA (Price & Bouvier 2002). This is because differences that evolve in allopatric populations are reinforced by divergent selection arising from ecological pressures and not from reproductive (genetic) incompatibility (Noor 1995; Kelly & Noor 1996). Ecological studies have suggested factors like broad environmental tolerance, plasticity and a lack of competition also contribute to the establishment success of introduced populations (Strayer 1999; Lee & Bell 1999; Weinig 2000; Hendry 2000; Colautti *et al.* 2004).

In invasion biology and evolution, a genetic paradox exists: how do introduced populations, whose genetic variation is expected to be impacted by population bottlenecks, persist and adapt in a new environments (Sakai *et al.* 2001; Allendorf & Lundquist 2003; Roman & Darling 2007). The classical studies of Wright (1931) and Mayr (1954) suggest that when a population goes through a small bottleneck, genetic variability of the population is expected to decline rapidly. Thereafter, when population sizes increase, variability starts to increase due to new mutations. Rapid evolution associated with an introduction event provides the necessary substrate for natural selection (Ellstrand & Schierenbeck 2000; Mooney & Cleland 2001; Lee 2002).

1.7 *Establishment success and spread of introduced populations*

The colonisation of new habitats requires that the first arrivals initiate new populations. Hermaphroditic species and selfing species such as many plants are good colonizers because they have no problem finding a mate to reproduce with (Whittier & Limpus 1996; Simberloff *et al.* 2000; Sakai *et al.* 2001). For example, most strains of marine algae *Caulerpa taxifolia* are not invasive, but one introduced population of this species from an aquarium in 1984 established itself in the wild and spread to more than 6,000 hectares, outcompeting native species and reducing biological diversity in the northwestern Mediterranean (Jousson *et al.* 2004). It was found that this introduced population was able to reproduce asexually, whereas the natives reproduce sexually (Jousson *et al.* 2004).

One common feature of introduced populations is a lag time between arrival and initial spread or the onset of rapid population growth (Sakai *et al.* 2001). The lag time is expected if evolutionary change such as adaptation to new environment, evolution of invasive life history characteristics and purging of genetic load responsible for inbreeding depression is an important part of establishment process (Mack *et al.* 2000; Parker *et al.* 2003; Allendorf & Luikart 2007) (Fig. 1.3).

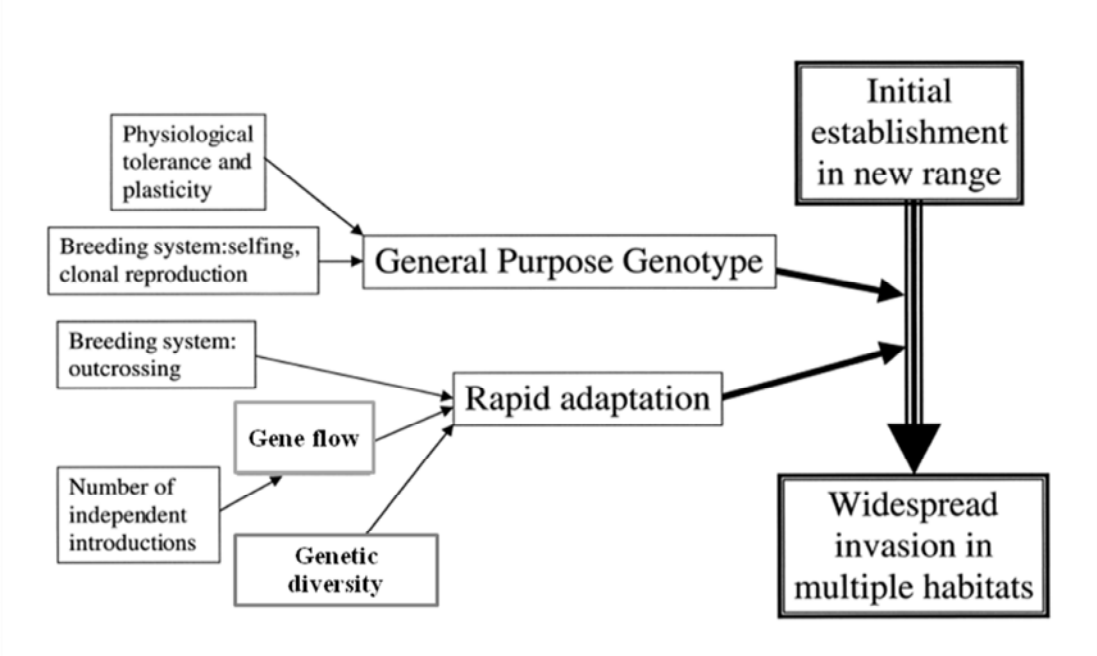

Fig. 1.3. The factors which influence the process by which an introduced species moves from initial establishment in a new range to wide spread invasion of multiple habitats. Two altenative, but not mutually exclusive, mechanisms are presented: rapid adaptation and general purpose genotype. Characteristics that influence the invading species are breeding system and number of introductions (from Parker *et al.* 2003).

Propagule pressure can be defined as the quality, quantity and frequency of invading organisms (Richardson *et al.* 2000; Kolar & Lodge 2001). It is suggested to be the most important factor predicting whether the introduced populations will become established in a new environment or will go extinct (Kolar & Lodge 2001; Parker *et al.* 2003). Species introduced in large quantities and consistent quantitites prove more likely to survive, whereas species introduced in small numbers with only a few release events are more likely to go extinct (Leung *et al.* 2004; Lockwood 2005). The genetics of an introduced population will be affected by propagule pressure, in that a great number of founders would be expected to reduce the effect of a population genetic bottleneck and different releases may have different source populations which may hybridize (Wright 1931; Mayr 1954; Allendorf & Luikart 2007). Hybridisation between individuals from genetically divergent source populations will result in increased genetic variation (Ellstrand & Schierenbeck 2000; Mooney & Cleland 2001; Lee 2002) (Fig. 1.4).

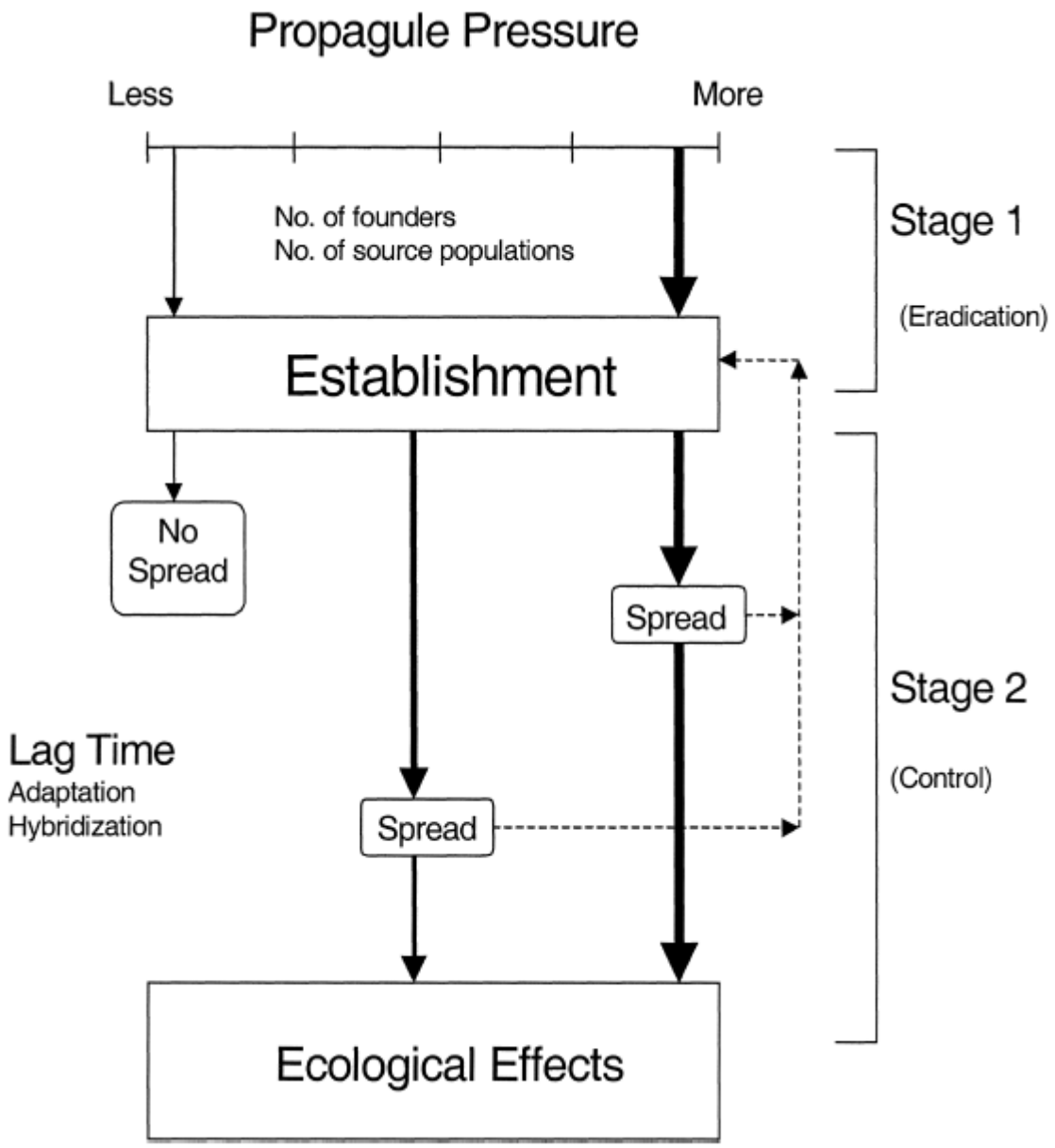

Fig. 1.4. The two stages of invasion that generally coincide with different management responses (see next section in this chapter). Propagule pressure is a continuum with greater pressure leading to an increased chance of establishment and spread with shorter lag times. If spread involves small groups of dispersing individuals, each group must be able to establish in a different area. Establishment or subsequent spread may be inhibited where groups reach the limits of particular environment conditions (from Allendorf and Lundiquist 2003).

After an introduced population has managed to establish in a new range, the next stage is spread (Fig1.4), and the time taken at this stage may vary, it may depend on genetic variation in the founder population or environmental factors in invasive range (Kolar &

Lodge 2001; Sakai *et al.* 2001). During the spread stage, if an introduced species becomes abundant resulting in negative effects upon natives it is then referred to as invasive (Richardson *et al.* 2000).

1.8 *Use of population genetics in management and control of introduced populations*

Currently, most introductions of non-native populations are human-mediated. The underlying social and economic factors are often as critical as biological and genetic factors (Sakai *et al.* 2001). Plants and animals have been introduced as forage, fibre, medicine, for ornamental purposes, game hunting, etc (Simberloff 1996; Pimentel *et al.* 2000). The best way to reduce the probability that an introduced population becomes established and spreads is to eradicate it (stage 1, Fig. 1.4) before it has sufficient time to evolve any adaptations that may allow successful competition with natives (Simberloff 1996; Myers *et al.* 2000; Allendorf and Lundiquist 2003; Allendorf & Luikart 2007). The potential ecological and social implications of an eradication project make them controversial and making a decision to attempt eradication is not simple so alternative approaches should be considered in some situations (Myers *et al.* 2000). Population genetics can be used to identify isolated reproductive units that are appropriate groups for eradication on the basis of patterns of genetic divergence (Calmet *et al.* 2001; Robertson & Gemmell 2004). Little genetic differentiation between spatially isolated populations may indicate significant gene flow, while significant differentiation between adjacent populations indicates limited dispersal (Lowe *et al.* 2004; Allendorf & Luikart 2007).

Population genetics studies of the introduced species could help our understanding of the evolutionary potential which may exist within a popultion and predictions can be made on the establishment success and range expansion of the introduced population in new environment (Paterson *et al.* 1995; Parker *et al.* 2003; Rollins *et al.* 2006; Rollins *et al.* 2009). The reduced genetic variation due to the genetic bottleneck experienced by a founder population will be a good indicator of a small effective population size. This is expected to reduce the amount of variation at adaptive loci, implying increased lag time before the expected population boom (Allendorf & Luikart 2007; Mack *et al.* 2000). An elevated genetic variation in the introduced population as compared to natives is an indicator of multiple sources of introductions (Kolbe *et al.* 2004; Zidana *et al.* 2009). Increased genetic variation would indicate substantial amounts of adaptive genetic variation so that the

populations would easily escape the effects of control agents (Allendorft & Lundquist 2003; Müller-Schärer *et al.* 2004; Messing *et al*. 2006). The evolution of resistance to insecticides and herbicides has increased rapidly in many species and reducing this may require an understanding of the origin, selection and spread of resistant genes (Denholm *et al.* 2002; Müller-Schärer *et al.* 2004).

Another option when introduced populations become established is to control them (Stage 2, Fig 1.4), (Allendorf and Lundiquist 2003). Genetics should play a central role as an explicit analysis of the genetic structure of introduced populations may allow development of more effective management policies and control of introduced populations (Sakai *et al.* 2001; Müller-Schärer *et al.* 2004; Allendorf & Luikart 2007). The conservation managers may recommend increasing fishing effort towards introduced populations to reduce their census population sizes.

Attempts have been made to control introduced populations using genotypes that could subsequently render the pest innocuous. Another approach has been to release genotypes with chromosomal aberrations whose subsequent segregation would result in reduced fertility (Foster *et al.* 1972). Other efforts include using transgenic sterile insects which are homozygous for repressible female specific lethal effects (Thomas *et al.* 2000).

1.9 *Microsatellite and mitochondrial DNA (mtDNA) genetic markers*

Genetic markers approache began to be used in fisheries in the 1950s. Initial studies were based on blood group variants, primarily in tunas, salmonids and cod that successfully demonstrated the existence of genetic variation which could be used in analyses of population structure (Ward & Grewe 1994). Further advances in protein technology in 1960s made it possible to study variation in amino acid sequences of homologous proteins and in 1980s discovery of *in vitro* DNA fragment amplification using the polymerase chain reaction (PCR) procedure, made it possible to acquire enough amounts of DNA to study variation in homologous DNA sequences (Mullis & Faloona 1987; Saiki *et al.* 1988; Sambrook *et al.* 1989; Ferguson 1994). Over the years, several molecular tools and methods have been designed and studies of population genetics have increasingly depended on these advances to investigate protein and DNA variation within and among populations (Sambrook *et al.* 1989; Lowe *et al.* 2004). Among the most widely used tools are molecular genetic markers isolated from microsatellite DNA and mitochondrial DNA (mtDNA) regions.

Microsatellites, or simple sequence repeats, are short stretches of repeat regions of tens to hundreds of base pairs of DNA composed of mononucleotide, dinucleotide, trinucleotide, or tetranucleotide repeats arranged in tandem (Tautz 1989; Park & Moran 1994; Jarne & Lagoda 1996). They are randomly distributed along the chromosome, but are underrepresented in centromeric and telomeric regions (Wong *et al.* 1990; Wintero *et al.* 1992). Each microsatellite locus is an array of tandem repeats that is flanked by a unique sequence (Britten & Khone 1968; Estoup 1993). A primer can be designed to complement the flanking region and hence microsatellite loci can be amplified by use of PCR (Jarne & Lagoda 1996; Glenn & Schable 2005).The designed microsatellite primers can also sometimes amplify across a number of closely or distantly related species (van Oppen *et al*. 1997b; Kellogg *et al*. 1995; Primmer *et al.* 1996; Taylor *et al*. 2002).

Microsatellites are highly susceptible to length mutation due to slippage during DNA replication, a characteristic which enables loci to be highly polymorphic and possess high heterozygosity (Rassmann *et al.* 1991; Wright & Bentzen 1994; reviewed in Schlötterer 2000). A single microsatellite locus can have more than 25 alleles and heterozygosity exceeding 90% (Wright & Bentzen 1994). Several mutation models have been suggested for microsatellite loci: infinite allele model (IAM, Kimura & Crow 1964), stepwise mutation model (SMM, Ohta & Kimura 1973) and two phase mutation model (TPM, Di Rienzo *et al.* 1994). In all these models, the mutation rate is several orders of magnitude higher than insertion or deletion (indel) mutations, which makes microsatellites ideal markers for many population genetic analyses.

Another advantage of microsatellites in population genetics is that it is a co-dominant marker and is (typically) inherited in a Mendelian fashion (Tautz 1989; Jarne & Lagoda 1996). Therefore both parental genotypes are passed over to the next generation, unlike mtDNA which is haploid and inherited from the maternal line (Birky 1995).

The characteristics responsible for biology and evolution of microsatellite loci have made it possible for this marker to be used successfully in populations genetics of different taxa to address questions involving: parentage analysis (Parker & Kornfield 1994; Teixeira & Bernasconi 2007; Blais *et al.* 2009; Yin *et al.* 2009); population structure (Bowcock *et al.* 1994; Ferguson 1994; van Oppen *et al.* 1997; Arnegard *et al.* 1999; Danley, 2000); migration or dispersal (Markert *et al*. 1999; Kim & Sappington 2006; Rollins *et al.* 2009); speciation process (Rico *et al*. 2003; Smith *et al*. 2003); genetic mapping (Dietrich *et al.* 1994; Dib *et al.* 1996); to identify sources of introduced populations (Estoup *et al.* 2001; Estoup & Clegg

2003); genetic bottleneck detection (Garza & Williamson 2001; Waldick *et al.* 2002; Schultz *et al.* 2009) and to detect hybridisation (Estoup *et al.* 1999; Valbuena-Carabaña *et al.* 2005; Streelman *et al*. 2004).

One potential drawback of using microsatellite markers is the presence of "null" alleles that fail to amplify during PCR (van Oppen *et al*. 1997; Markert *et al*. 1999). Sequencing studies have indicated that substitution and indel mutations in microsatellite flanking regions sequences occur at non negligible rates (Angers & Bernatchez 1997; Grimaldi & Crouau-Roy 1997). Such variation in the microsatellite sequences of flanking regions may prevent the primer annealing to template DNA during amplification of the microsatellite locus by PCR, resulting in a null allele (Callen *et al.* 1993; Glenn & Schable 2005). Other possible causes of microsatellite null alleles includes: allelic drop outs due to stochastic sampling errors from low template DNA concentrations and quality (Miller & Waits 2003), or resulting from preferential amplification of small alleles (Wattier *et al*. 1998) and slippage during PCR amplification (Gagneux *et al*. 1997; Shinde *et al.* 2003). The presence of null alleles in population genetic studies can cause heterozygote deficiency potentially biasing analyses (Shaw *et al.* 1999). The presence of null alleles causes deviations from Hardy-Weinberg proportions similar to those caused by effects of inbreeding and assortative mating (van Oosterhout *et al.* 2004), implying that data should be interpreted carefully. It is recommended to use, population genetics software which can identify the presence of null alleles and short allele dominance (large allele dropout) in data sets (van Oosterhout *et al.* 2004).

Measures of genetic variation in population genetics studies may also be affected by size homoplasy. This is due to the nature of mutations in the microsatellite repeat region, in which loss or gain of a variable number of repeat units generates alleles identical in state but not in descent (Viard *et al.* 1998; van Oppen *et al.* 2000; reviewed in Estoup *et al.* 2002). Size homoplasy reduces observed number of alleles and levels of heterozygosity within populations (Estoup *et al.* 1995; Jarne & Lagoda 1996; Angers & Bernatchez 1997; Viard *et al.* 1998; Taylor *et al.* 1999; Angers *et al.* 2000). Size homoplasy may also affect the statistical power of data generated from microsatellite markers, since a higher level of polymorphism results in a higher power of exact tests on genotypic linkage disequilibrium and Hardy-Weinberg equilibrium (Raymond & Rousset 1995; Hubisz *et al.* 2009).

Ascertainment bias is the phenomenon that microsatellite loci are generally more polymorphic in the species for which they were isolated, and that they can be comparatively

invariant in distantly related species (Ellegren *et al.* 1995). This is another setback when using microsatellite markers, as populations may be incorrectly found to exhibit low genetic variation (Ellegren *et al.* 1995; reviewed in Garner *et al.* 2005). However, success of cross-species microsatellite amplification is generally related to the evolutionary distance from the focal species. For species diverging 10-20 million years ago, approximately 25% of primer sets will amplify polymorphic loci (Gemmell *et al.* 1997; Primmer *et al.* 1996).

Some of the most studied DNA sequences within population and evolutionary genetics are those from mitochondrial DNA (Avise 1994; Park & Moran 1994). Since mtDNA is located in the cytoplasmic organelles of eukaryotic cells it is physically separated from nuclear DNA (Avise 1986; Birky 1995). The physical separation from billions of other nucleotides in the nuclear genome and relatively small size in animals (15 to 20 kilo base pairs) makes it easy for mtDNA molecules to be isolated for population genetics studies (Avise 1986; Park & Moran 1994). The mtDNA codes for genes that are vital for the respiration functions of the cell. It is maternally inherited through ovum cytoplasm in most organisms, which is a rare exception among the genetic markers used in evolution and population genetics studies (Brown *et al*. 1979; Birky 1995; Ballard & Whitlock 2004). The mtDNA is haploid in state with no recombination between mtDNA genomes such that each mitochondrion contains one type of DNA with a few notable exceptions (Avise 1986; Bentzen *et al.* 1988; Park & Moran 1994). These characteristics combine to reduce the mtDNA effective population size to one quarter of the nuclear DNA of the same organism (Nei & Tajima 1981). A smaller effective population size means that genetic drift can cause frequency differences between isolated populations more readily in mtDNA than in nuclear DNA (Avise 1994).

In many organisms mtDNA evolves rapidly, in part due to lack of sequence repair mechanism for mutations that arise during replication (Wilson *et al.* 1985; Ballard *et al.* 2004). This characteristic provides mtDNA sequences with high variability compared to most from nuclear DNA (e.g. single nucleotide polymorphisms, SNPs), and is utilized to investigate questions on population genetic structure (Moran *et al.* 1994; Jamandre *et al.* 2009). The mitochondrial genome is divided into different regions (Fig. 1.5), which evolve at different rates. They have been used to address different questions at various taxonomic levels: Cytochrome *b* gene and Nicotinamide Adenine Dehydrogenous (NADH) genes sequences evolve relatively fast and they have been used to investigate genetic differences at population level (Park *et al.* 1993; Park & Moran 1994), whereas ribosomal genes evolve

more slowly and are particularly useful to investigate genetic differentiation at family level (Milinkovitch *et al.* 1993).

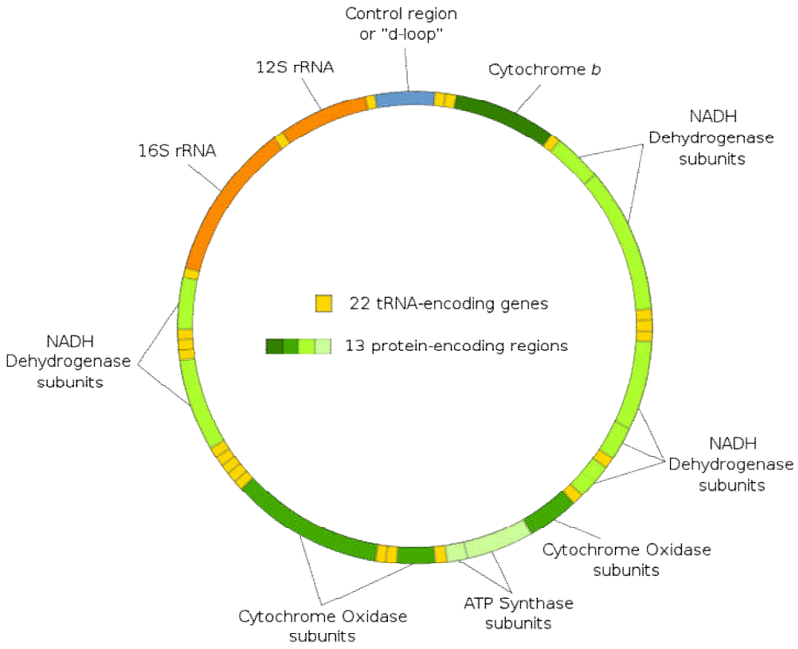

Fig. 1.5. Organisation of the animal mitochondrial DNA (mtDNA) genome, showing different regions usually used as molecular genetic markers in population genetics studies (from public domain: http://commons.wikimedia.org/wiki/File:Mitochondrial_DNA_it.svg).

In evolution and conservation genetic studies the mtDNA maker has been used to investigate questions related to: introgressive hybridisation (Ferris *et al.* 1983; Tegelstrom 1987; Wilson & Bernatchez 1998; Hird & Sullivan 2009), phylogenetic relationships (Moran *et al.* 1994; Moran & Kornfield 1993; Phillipsen & Metcalf 2009) and identifying the source of introduced populations (Downie 2002; Zardus & Hadfield 2005; de Carvalho *et al.* 2009; reviewed in Muirhead *et al.* 2008).

Animal mtDNA has proved to be an ideal marker for phylogenetic analysis, and approximately 70% of all studies with phylogeographic component use information from this molecule (Lowe *et al.* 2004). One of the largest problems with mtDNA markers in phylogenetic studies, however, is the retention of ancestral polymorphisms in closely related species due to incomplete lineage sorting (Moran & Kornfield 1993; Parker & Kornfield 1997). The use of a more distant relative as an outgroup increases the ability to detect such ancestral polymorphisms (Charlesworth *et al.* 2005). Migratory patterns and population structure have been investigated using mtDNA markers (Hänfling *et al.* 2002; Zardus &

Hadfield 2005; Muirhead *et al.* 2008). However, mtDNA differentiation may sometimes be produced by lack of female dispersal or male-biased dispersal and selection in other taxa. This should be considered when interpreting results. Sex-biased dispersal has effects on the dynamics and genetics of the species (see review Prugnolle & de Meeus 2002). Under sex-bised dispersal pattern, one sex is philopatric (i.e. individuals of this sex stay or return to their natal site or group to breed) while the other is more prone to disperse. In cichlids the males are territorial, guarding their breeding ground all year round (Hert 1992; Kellog *et al*. 1995; Robinson & Ribbink 1998; Konnings 2008), suggesting that the system may be following a female-biased dispersal pattern. Since allele frequencies are equally randomised between males and females in a population it is possible to detect contrasted pattern of differentiation depending on the genetic marker used (microsatellite or mtDNA) and sex sampled. Therefore considerable care should be taken when interpreting genetic differentiation estimates in populations that are affected by sex-biased dispersal.

The mtDNA haplotype diversity is usually positively correlated both with number of samples collected within a habitat, and with the number of locations surveyed (Muirhead *et al.* 2008). Exceptions to this generality include constrained genetic or species diversity in challenging environments and genetic diversity of some invading populations which may be represented by a single haplotype (Cristescu *et al.* 2001; Vasquez *et al.* 2005). This emphasizes the importance of the sampling strategy for data interpretation. Although the exact strategy depends on the nature of the investigation, effective strategies may include sampling within and among populations, and sampling within and among genomes (Lowe *et al.* 2004).

There are exceptions to strict maternal inheritance of mtDNA: there is evidence of some incorporation of male mitochondria “paternal leakage” in mice (Gyllensten *et al.* 1991) and humans (Awadalla *et al.* 1999). Some species show double uniparental inheritance of mtDNA (Sutherland *et al.* 1998). This may lead to heteroplasmy, which is the presence of more than one mitochondrial genotype within an individual. Effective neutrality is one of the major assumptions underlying most analyses of mtDNA population data, especially when mtDNA diversity patterns are interpreted in terms of demography, migrations and founding events (reviewed by Galtier *et al.* 2009). The assumption is that, the mitochondrial genome essentially undergoes neutral or deleterious mutations and adaptive mutations that spread through positive selection and are thought to be very rare. The deleterious changes are rapidly removed by purifying selection without affecting much of the adptive diversity at linked sites,

therefore the observable variation would only reflect neutral processes and this is in agreement with the neutralist theory of molecular evolution (Kimura 1983). However, a study on mtDNA analysis on synonymous/nonsynonymous ratio appeared to be consistent with the prevalence of purifying selection (Weinreich & Rand 2000). Episodes of positive selection in mitochondrial protein coding gene regions have been detected in several groups of animals including primates (Grossman *et al.* 2004), agamid lizards and snakes (Castoe *et al.* 2009) and in humans (Balloux *et al.* 2009).

1.2 Objectives and outline

The Lake Malawi fish are of major socio-economic importance to Malawian people: as a source of 70% of animal protein in the diet, it contributes 4% of the Gross Domestic Product (GDP) and gives employment to 3% of the country's population. The biodiversity of Lake Malawi is extraordinary - it is suggested to have the greatest wealth of endemic fish species of all the lakes in the world. The vast majority of this ichthyofauna belongs to the cichlidae family, which is currently a textbook example of explosive speciation and is at the forefront of evolutionary research. Many species of cichlids, especially mbuna, are popular in the specialized aquarium trade. Given the profound importance of cichlids to people of Malawi and scientific community as a whole, introduction of cichlid species within Lake Malawi and its consequences is of great conservation concern.

The overall objective of this thesis is to investigate the population genetics of *C. afra*in native and introduced ranges in Lake Malawi. *C. afra* was introduced into the south-western part of the lake in 1960s by an aquarium trader who established this business on the shore line. This thesis used this well documented introduction event to investigate two important questions in conservation and evolution studies: What are the factors that facilitate establishment success of introduced species? How do small founder populations whose genetic variation is expected to be affected by genetic bottleneck and drift, still evolve and successfully establish in new environments?

The results from this thesis will have crucial implications on the conservation and evolutionary genetics of Lake Malawi cichlids. This thesis will be of interest to reasearch workers in biology, zoology, ecology and evolution of fish. The aims of this thesis will be achieved in the following 5 chapters: Chapter 1 has described general issues which have been studied in cichlid evolution, general biology of rock-dwelling cichlids, fish introductions,

implications of introduced species on evolution and conservation and genetic markers which have been used in invasion biology studies. Chapter 2 investigates the population genetics of *C. afra* in introduced range and native ranges using the mtDNA control region. Chapter 3 compares the suitability of two genetic markers, microsatellite and mtDNA when investigating founder event in populations with complex demographic history like multiple founder sources. Chapter 4 investigates the impact of introduced *C. afra* on native *Psuedotropheus zebra* population, and assesses evidence of introgressive hybridisation. Chapter 5 summarises and highlights the results in a general discussion.

1.2.1 *mtDNA analysis in introduced and native ranges-investigation in Chapter 2*

The hypothesis that an introduction event is usually associated with a small founder population size which is expected to be affected by reduced genetic variation due to effects of a genetic bottleneck and drift in the introduced range was investigated. Many studies have shown that introduced populations persist and adapt in new environments, suggesting that introduction events may not usually be associated with reduced genetic variation. One hundred and fifty individuals of *C. afra* from four introduced populations and six native populations were sampled. The mtDNA control region was fully sequenced to investigate the genetic variation in *C. afra* populations in introduced and native ranges. Using descriptive population genetic measures calculations on nucleotide and haplotype diversity as employed in DNAsp software were investigated. A parsimony network analysis was used to investigate haplotype sharing and infer sources of the introduced population.

1.2.2 *Suitability of microsatellite and mtDNA genetic markers when investigating founder events with complex demographic history-investigation in Chapter 3*

Microsatellite and mtDNA genetic markers exhibit different biology and evolutionary history. The two markers have been widely used in population genetics studies investigating founder events and colonisation history in different organisms. Six hundred *C. afra* samples were genotyped at six microsatellite loci. The collection localities are the same as those used in mtDNA analysis in chapter 2. The mtDNA control regions results were reanalyzed to compare results with microsatellite markers. Descriptive genetic variation measures were calculated. The effective number of alleles and halpotypes, genetic differentiation among *C.*

afra populations (G'_{ST}) and analysis of molecular variance (AMOVA) as employed in ARLEQUIN software were investigated. Previous studies have suggested a stepping stone colonisation pattern of *C. afra* in Lake Malawi National Park. In this study, an approximate Bayesian computation (ABC) procedure was used to investigate two competing hypotheses in colonisation history of *C. afra* in Lake Malawi National Park: a stepping stone introduction pattern and an independent introduction pattern.

1.2.3 *Introgressive hybridisation between* Cynotilapia afra *and* Pseudotropheus zebra-*investigation in Chapter 4*

The hypothesis that introgressive hybridisation is more likely to occur in the direction from the more common native species into the gene pool of the rare, introduced species was tested. *P. zebra* and *C. afra* naturally occur sympatrically within Lake Malawi. The two species have common habitat preferences for sediment free rocky waters, and all possess the blue body colour and vertical black banding also called (BB) morphology. The introduction of *C. afra* to an area with abundant *P. zebra* individuals like Lake Malawi National Park, suggests possible hybridisation candidates. Two hundred and sixteen samples of introduced *C. afra* and native *P. zebra* from the introduced range were analysed using six microsatellite loci. Descriptive population genetics measures including mean number of alleles (*A*), and observed and expected heterozygosity (H_O and H_E) were investigated. Among population genetic differentiation measure (G'_{ST}) from sympatric populations was also calculated. A Bayesian clustering method was used to investigate admixture of genetic material between *C. afra* and *P. zebra*, and admixture proportions were calculated. A factor component analysis was used to explore the contribution of genetic material from sympatric *C. afra* and *P. zebra* populations in a two dimensional scale.

Chapter 2

2. Analysis of michondrial DNA (mtDNA) in introduced and native populations of *Cynotilapia afra* in Lake Malawi

Analysis of mtDNA in introduced *Cynotilapia afra* published in: Zidana H, Turner GF, van Oosterhout C, Hänfling B (2009) Elevated mtDNA diversity in introduced populations of *Cynotilapia afra* (Günther 1894) in Lake Malawi National Park is evidence for multiple source populations and hybridisation. *Molecular Ecology,* **18,** 4380-4389.

2.1 Introduction

Biological invasions provide the opportunity to observe evolution in action as such events may result in significant demographic changes. Furthermore, interactions between the invasive and native species in the invaded community may have both ecological and genetic repercussions, and rapid evolutionary changes will occur before a new equilibrium is reached. Recent molecular evidence suggests that many invasive species do not show the reduced genetic variability expected of populations with relatively few founders. Genetic diversity may be retained or even enhanced by multiple introductions from genetically divergent source populations (Kolbe *et al*. 2004; Hänfling 2007; Roman & Darling 2007) or by introgressive hybridisation with native species (Ellstrand & Schierenback 2000).

Two important evolutionary consequences of multiple introductions and introgression have been suggested. Firstly, these processes may dilute the founder effects and reduce the risk of inbreeding depression (Novak & Mack 1993; Kolbe *et al*. 2004; reviewed in Dlugosch & Parker 2008). Secondly, the admixture of gene pools may lead to novel gene combinations with advantageous effects in the new environment (Hänfling 2007). Likewise, the production of novel gene combinations during hybridisation has been proposed as a major factor contributing to natural adaptive radiation, particularly by African cichlid fishes (Seehausen 2004).

The African great Lakes Victoria, Tanganyika and Malawi have the richest lacustrine faunas in the world and their cichlid fishes are a text book example of explosive radiation, with a recent estimate of 451-600 species in Lake Malawi, 447-535 in Lake Victoria (and many more in neighbouring lakes) and 162-184 in Lake Tanganyika (Genner *et al*. 2004a). The lakes vary in size, shape, age and water parameters and the high diversity of cichlid fishes inhabiting them have been derived from different riverine ancestors (Salzburger *et al.* 2005). The cichlid fishes of Lake Tanganyika show a wide range of reproductive strategies not found in the younger and less genetically diverse cichlid faunas of the other two lakes. However, all three lakes are dominated by cichlid fish, which exhibit striking parallels in morphology (Young *et al.* 2009a), behaviour (Blais *et al.* 2009), life histories (Duponchelle *et al.* 2008) and even in opsin amino acid sequences (Sugawara *et al.* 2005).

These cichlid fishes have served as a model system in studies of evolution (see reviews by Kornfield & Smith 2000; Kocher 2004; Seehausen 2006). Several hypotheses have been suggested to explain the mechanisms behind explosive cichlid speciation: morphology, particularly of the pharyngeal jaw apparatus (Fryer 1959; Liem 1980), sexually-selected colour variation (Kosswig 1947; Dominey 1984; van Oppen *et al.* 1998), habitat specialisation and fragmentation (Fryer 1959; Ribbink *et al.* 1983), hybridisation (Salzburger *et al.* 2002, Seehausen 2004) and visually-based mate choice in clear water habitats (Seehausen *et al.* 1997; Terai *et al.* 2006). Equally, roles for both allopatric (Fryer 1959; Ribbink *et al.* 1983; van Oppen *et al.* 1997) and sympatric (Seehausen & van Alphen 1997; Shaw *et al*. 2000) speciation modes have been proposed to play a major role. Notably, rocky shore cichlid fishes show high levels of geographic colour variation (Fryer 1959; Ribbink *et al.* 1983) associated with strong genetic structuring (van Oppen *et al.* 1997; Rico *et al.* 2003) and partial reproductive isolation (Knight & Turner 2004) suggesting a major role for geographic isolation in diversification and also a significant degree of local endemism and evolutionarily significant beta-diversity (Ribbink *et al.* 1983; Reinthal 1993; Genner *et al.* 2004a & b). Differences in parasite fauna are also known among species and habitats (Blais *et al.* 2007; Amin *et al*. 2008; Bray *et al.* 2006). Given the high level of endemism anthropogenic translocations of fishes within Lake Malawi are a cause for concern (Ribbink *et al*. 1983; Trendall 1988; Streelman *et al*. 2004; Genner *et al*. 2006; Young *et al*. 2009b).

The recent and rapid evolution of cichlid fishes in the younger lakes, such as Malawi and Victoria renders the exact reconstruction of species trees from molecular data difficult: many rock-dwelling cichlids retain shared ancestral mtDNA polymorphisms, even where

morphological differentiation is clear and reproductive isolation demonstrated or suspected (Moran & Kornfield 1993; Parker & Kornfield 1997; Sturmbauer *et al.* 2001). However, mitochondrial DNA sequence data can be highly informative in a population genetic framework. This approach has long been employed in studies of humans and other organisms (Bandelt *et al.* 1999; Pakendorf & Stoneking 2005; Rowold *et al.* 2007; Atkinson *et al.* 2007; Krystufek *et al.* 2007) and more recently has proved useful with cichlid fishes from Central America (Barluenga *et al.* 2007) and East Africa (Verheyen *et al.* 2003; Genner *et al.* 2007 Koblmüller *et al.* 2007b). Furthermore, mitochondrial DNA has proven to be a particularly powerful genetic marker to trace source populations of recent invasions in lizard (Kolbe *et al.* 2004), crab (Hänfling *et al.* 2002; Roman 2006), fish (Lindholm *et al.* 2005; Azzurro *et al.* 2006) and snails (Chuong *et al.* 2008). The aim of this study was to employ population genetic analyses of mitochrondrial DNA to investigate the population genetic consequences of human-mediated translocation of a rock-dwelling cichlid fish, *C. afra*into the southern-west arm of Lake Malawi. Such data will not only provide a population genetic frame work to study the evolutionary consequences of introductions, admixture and hybridisation in cichlids but also have conservation implications since inferences about spread and colonisation pathways of introduced cichlids in Lake Malawi National Park will be made.

2.2 Materials and methods

2.2.1 *Study species and area*

*C. afra*is a small plankton-feeding fish from the 'mbuna' group confined to clear-water rocky habitats (Fryer 1959; Ribbink *et al.* 1983; Konings 2008). It is naturally widely distributed within Lake Malawi (Ribbink *et al.* 1983; Konings 2008). It is believed that the *C. afra* species complex is indigenous to the northernmost 80% of the lake, from the far north at Ikombe to Mbenji Island and Ntekete in the south (Konings 2008). At several localities within this range, populations of *C. afra* occur sympatrically with populations of its putative sister species the *P. zebra* complex (Kassam *et al.* 2005). In the late 1970s/early 1980s, *C. afra* has been reported to have spread around the Thumbi West Island following a documented release of fish destined for the aquarium fish export trade in the 1960s (Ribbink *et al.* 1983) and has subsequently been found on other nearby rocky shores (Fig. 1). Since its introduction, *C. afra* is believed to have undergone rapid microevolutionary change as well as

introgressive hybridisation with *P. zebra* (Streelman *et al.* 2004). Thumbi West Island now forms part of the Lake Malawi National Park (Trendall 1988; Streelman *et al*. 2004). The Lake Malawi National Park, established in 1980 as the World's first national park targeted principally on freshwater fish, and designated a UNESCO World Heritage Site in 1984, covers 9,400ha in the southern part of the lake (Hough 1989). Its waters are among the most heavily impacted by translocations, with at least 13 taxa of cichlid fishes being introduced to Thumbi West Island alone (Genner *et al.* 2006).

2.2.2 *Sampling*

Samples of *C. afra* were collected from six localities within the native range of the species: Mara Rocks; Nkhata Bay; Likoma Island; Mbenji Island; Chiofu Bay; and Ntekete Rocks and from all four known sites inhabited by introduced populations: Domwe Island; Otter Point and two locations from Thumbi West Island (Table 1 & Fig. 1). Specimens of *P. zebra* were collected from the four sites known to be colonised by the introduced populations (Fig. 2.1). Fifteen individuals per site were sampled from each population.

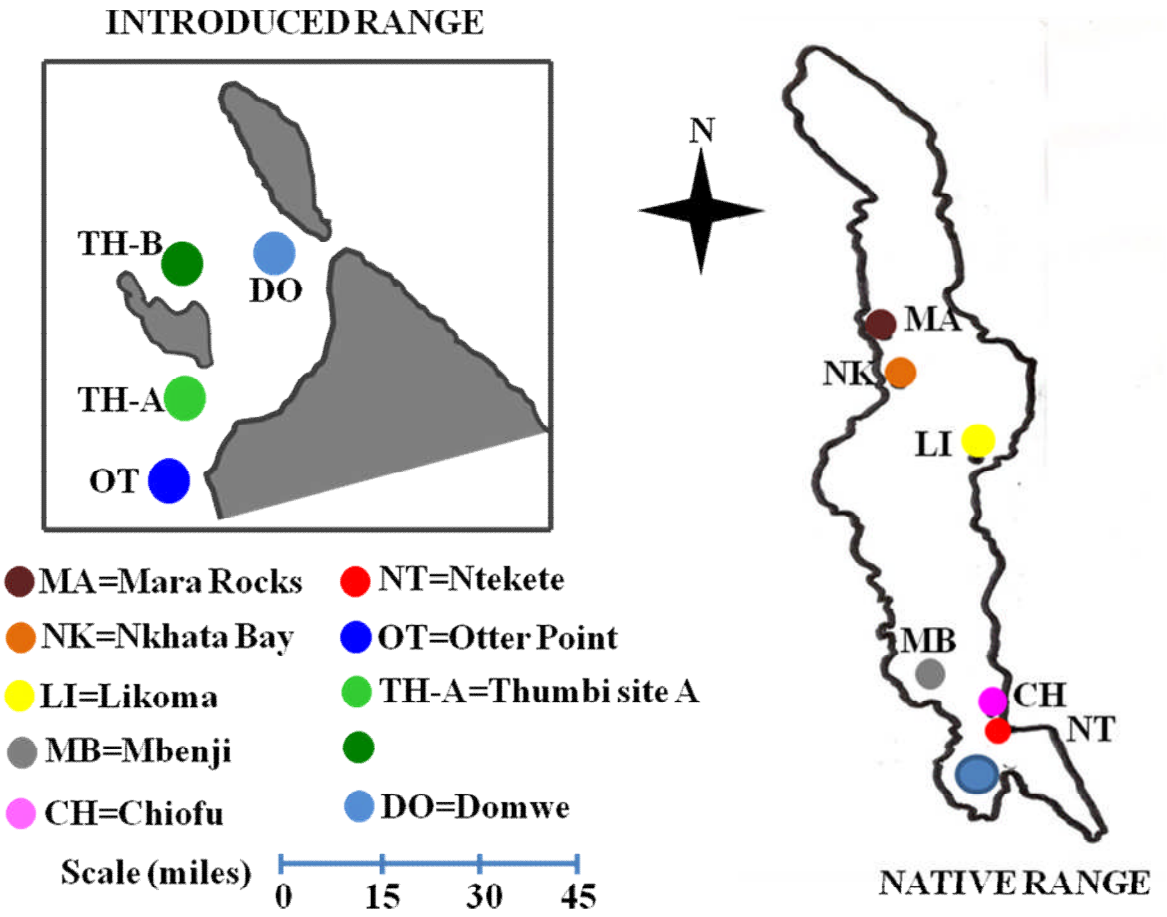

Fig. 2.1. Lake Malawi map showing 10 geographic locations sampled in this study, from native (MA=Mara, NK=Nkhata Bay, LI=Likoma, MB=Mbenji, CH=Chiofu, NT=Ntekete) and introduced (DO=Domwe, TH-A=Thumbi site A, TH-B=Thumbi site B and OT=Otter Point) ranges of *C. afra* populations. The colour coded circles corresponds to mtDNA haplotype identification (see Fig. 2.3).

A non-invasive sampling method was employed, collecting live fish samples using monofilament nets while SCUBA diving. A small tissue sample was taken by fin clipping the posterior (soft-rayed) dorsal fin, which regenerates within a few days. To ensure reliable species identification, all samples were collected from territorial males in full breeding colour and jaw teeth of all specimens were examined using a hand lens, as the possession of unicuspid teeth is diagnostic for the genus *Cynotilapia* (Trewavas 1935, Fryer 1959). All live specimens were released back into the water. Fin clips were preserved in 100% ethanol before molecular analysis in the laboratory.

2.2.3 Molecular analysis

Total genomic DNA was extracted using Promega WIZARD genomic DNA purification kit. The mitochondrial DNA (mtDNA) of the entire control region was amplified using forward primer HapThr-2+4 (5'–CCTACTCCCAAAGCTAGGATC-3') and reverse primer Fish12s (5'-TGCGGAGACTTGCATGTGTAAG-3'), (Joyce *et al*. 2005). Polymerase chain reaction (PCR) was performed in a 25 µL reaction consisting of 2 µL DNA template, 1X PCR buffer, dNTPs (0.1mM), primer (0.4mM), $MgCl_2$ (1.2mM stock), PCR water and 0.25 Units *Taq*. The PCR running conditions were as follows, 3 min at 94°C, 45 sec at 94°C and 50 sec at 56°C for 34 times, then 1 min 30 sec at 72°C, and followed by 10 min of 72°C. PCR products were purified and sequenced in both directions using the professional sequencing service of MACROGEN Centre (Pathfinder in Genomics Research, 908 World Meridian Centre, Seoul, 153-023 Korea).

2.2.4 *Data analysis*

The software CodonCode Aligner (Ewing *et al*. 1998) was used to edit individual sequences and to align forward and reverse sequences of each individual. The resulting consensus sequences of all individuals were aligned using ClustalW (Thompson *et al.* 1997) in combination with MEGA version 4.1 (Tamura *et al*. 2007) and submitted to GenBank (accessions GQ380500-GQ380547). The alignment was checked by eye and was trimmed to a uniform length of 910bp including gaps, which were treated as point mutations during analysis. DnaSP version 4.5 (Rozas *et al*. 2003) was used to make inferences of parameters involving within-population genetic diversity mtDNA haplotype diversity (*H*), nucleotide diversity (π) and between-population measures (F_{ST} and D_A (Nei 1987)). The mtDNA number of haplotype and standard errors (SE) were estimated by a jackknife procedure (Nei & Jin 1989; Lynch & Crease 1990).

To estimate the contribution of genetic introgression between *C. afra* and *P. zebra* within the introduced range, I analysed mtDNA sequences of *P. zebra*, identifying haplotypes shared with *C. afra* and calculating their contribution to nucleotide (π) and haplotype (*H*) diversity using DnaSP version 4.5 (Rozas *et al*. 2003).

To explore the relationships between the introduced population of *C. afra* and possible source populations, I used two approaches. Genetic distances among populations

(D_A) were calculated using DnaSp version 4.5 (Rozas *et al*. 2003), which were then used in XLSTAT version 2008 (Addinsoft 2008) to configure coordinates for a multidimensional scaling (MDS) approach. Secondly, a network of haplotypes was constructed using a median-joining algorithm in the Network software version 4.5.10 (Bandelt *et al*. 1999).

2.3 **Results**

2.3.1 *Within-population genetic diversity*

A total of forty-eight haplotypes were found across the whole data set of 195 sequences, 150 from *C. afra* and 45 from *P.zebra* (Table 2.1).

Table 2.1: Sampled populations, sample sizes and summary statistics. All sequence statistics are based on a full mtDNA control region sequence (910bp) alignment from 195 sequences.

Species	Locality	Co-ordinates	N sequenced	Total haplotypes	Haplotypes unique to population	Haplotype diversity (*H*)	Nucleotide diversity (π)
Introduced							
C. afra	Domwe Island	13º 59′ 85.5″S, 34º 50′ 32.6″E	15	3	0	0.65	0.0019
	Thumbi West Island	14º 01′ 50.5″S, 34º 49′ 44.4″E	15	6	3	0.76	0.0007
	Thumbi West Island (B)	14º 01′ 08.0″S, 34º 49′ 19.7″E	15	6	3	0.82	0.0030
	Otter Point	14º 02′ 31.6″S, 34º 49′ 37.6″E	15	8	4	0.90	0.0033
Native							
C. afra	Mara	11º 14′ 44.8″S, 34º 16′ 44.5″E	15	3	2	0.23	0.0003
	Nkhata Bay	11º 36′ 47.1″S, 34º 28′ 23.1″E	15	4	3	0.67	0.0007
	Likoma	12º 01′ 48.9″S, 34º 43′ 46.5″E	15	5	3	0.63	0.0012
	Mbenji	13º 26′ 31.8″S, 34º 29′ 29.1″E	15	4	2	0.47	0.0010
	Chiofu	13º 32′ 01.8″S, 34º 51′ 86.6″E	15	5	5	0.71	0.0011
	Ntekete	13º 37′ 72.4″S, 34º 51′ 25.7″E	15	4	4	0.37	0.0019
P. zebra	Domwe Island	13º 59′ 85.5″S, 34º 50′ 32.6″E	15	8	6	0.91	0.0036
	Thumbi West Island	14º 01′ 50.5″S, 34º 49′ 44.4″E	15	5	1	0.63	0.0016
	Otter Point	14º 02′ 31.6″S, 34º 49′ 37.6″E	15	4	2	0.37	0.0006
Total			195	48	-	-	-

The introduced populations of *C. afra* at Thumbi West Island site B and Otter Point had higher mtDNA haplotype and nucleotide diversity compared to the native populations (Table 2.1 & Fig. 2.2). The introduced population of *C. afra* at Otter Point exhibited the highest mtDNA haplotype (H=0.90) and nucleotide (π=0.0030) diversity, while the native population at Mara Rocks had the lowest mtDNA haplotype (H=0.23) and nucleotide (π=0.0003) diversity (Table 2.1 & Fig. 2.2). In the native *P. zebra* group, Domwe Island population had a higher haplotype (H=0.91) and nucleotide (π=0.0040) diversity, while Otter Point population had the lowest (H=0.37) and (π=0.0006) diversity (Table 2.1).

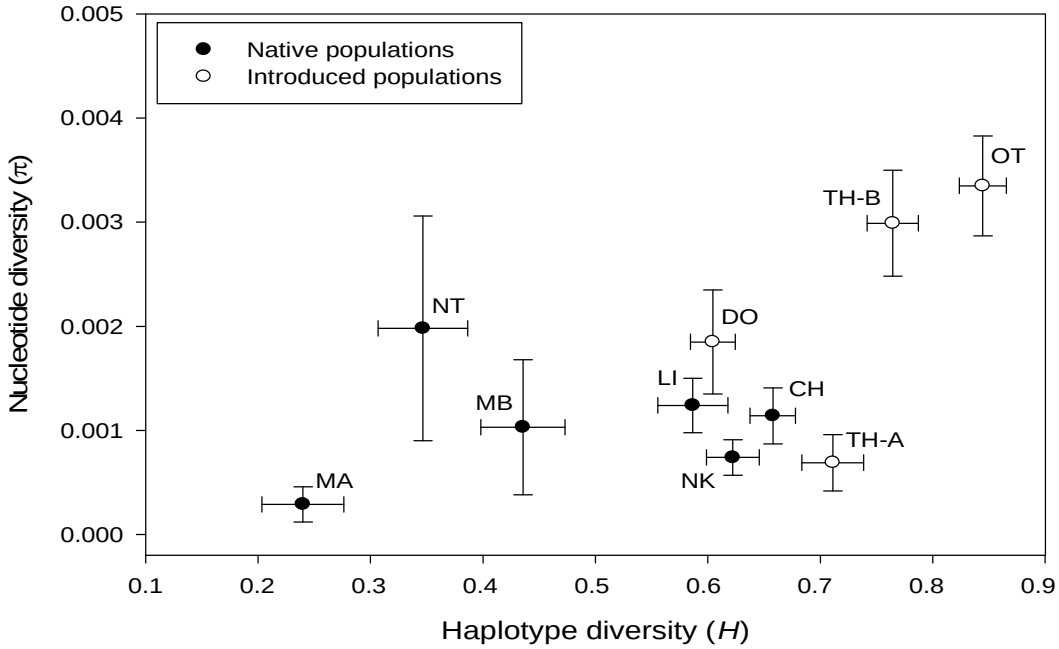

Fig. 2.2. Haplotype versus nucleotide diversity (mean and SE) for the native (solid circles) and introduced (open circles) *C. afra* populations. Standard errors of *H* were calculated using a jackknife procedure. (The abbreviations for populations are given on Fig. 1).

2.3.2 *Among-population genetic diversity*

The network analysis revealed substantial lineage sorting and no shared haplotypes among most native populations. Only the two samples from the north-eastern part of the lake (Mara and Nkhata Bay) appear to be closely related and even share their most common haplotype (Fig. 2.3). The native populations Ntekete and Chiofu appear to be particularly distinct and share haplotypes neither with introduced populations nor with other native populations. However, two highly divergent haplotypes were found in Ntekete in low frequencies indicating immigration from a genetically divergent conspecific population or mitochondrial introgression (Fig. 2.3).

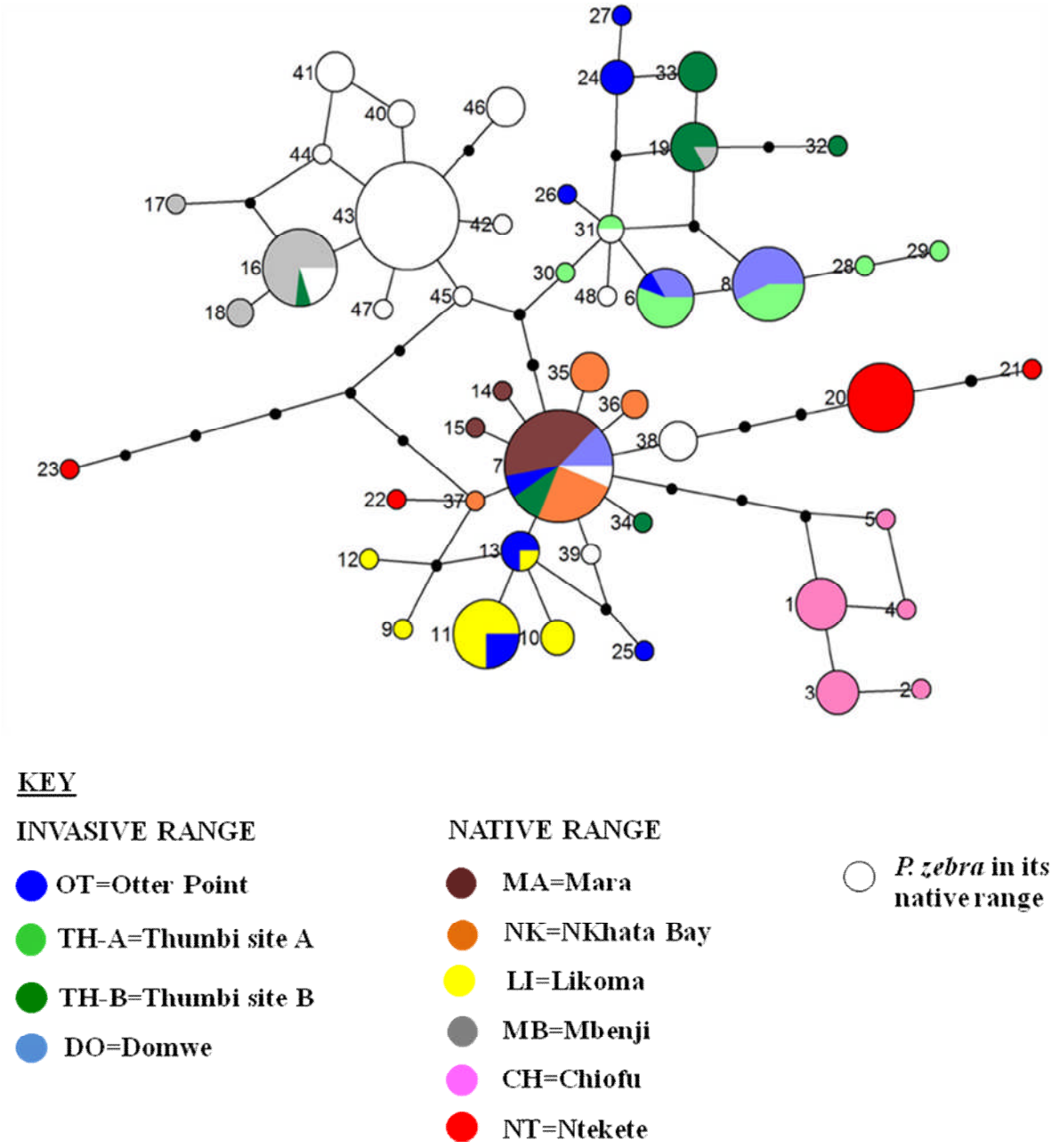

Fig. 2.3. A parsimony network of *C. afra* and *P. zebra* haplotypes (*H*) based on mtDNA sequences and each of the 48 haplotypes is numbered. Each circle represents a single haplotype and its diameter is proportional to the number of individuals with that haplotype. The colour codes represent the source of haplotype basing on geographic location within Lake Malawi a black dot (•) represents unsampled haplotypes and open circle (○) = *P. zebra* haplotypes sampled from the range of introduced *C. afra*.

In contrast, haplotypes from most introduced populations are widely scattered across the network and a number of haplotypes are shared with other introduced and native populations. Haplotype H7 was common at both Nkhata Bay and Mara Rocks and was shared

with three of the introduced samples. Haplotypes H11 and H13 otherwise unique to Likoma Island were shared with the Otter Point population. The Mbenji population shares two haplotypes with introduced populations, but surprisingly, one of these haplotypes is also shared with *P. zebra* (H16). A substantial cluster of haplotypes (top right in Fig. 2.3) was shared among introduced populations but only rarely found in the native *C. afra* populations. This is evidence of one or more unsampled source populations.

Only three haplotypes (H7, H16 and H31) were shared between *C. afra* and *P. zebra* in the range of the introduced species (Fig. 2.3), so that proportionally the genetic diversity contributed through hybridisation was: H=0.52% and π =1.9%. In other words, we can estimate that >98% of the genetic diversity in the introduced *C. afra* populations is of conspecific origin, and <2% from hybridisation with *P. zebra*.

Pairwise genetic differentiation (both D_A and F_{ST}) was lower for comparisons among introduced populations than native populations (Fig. 2.4).

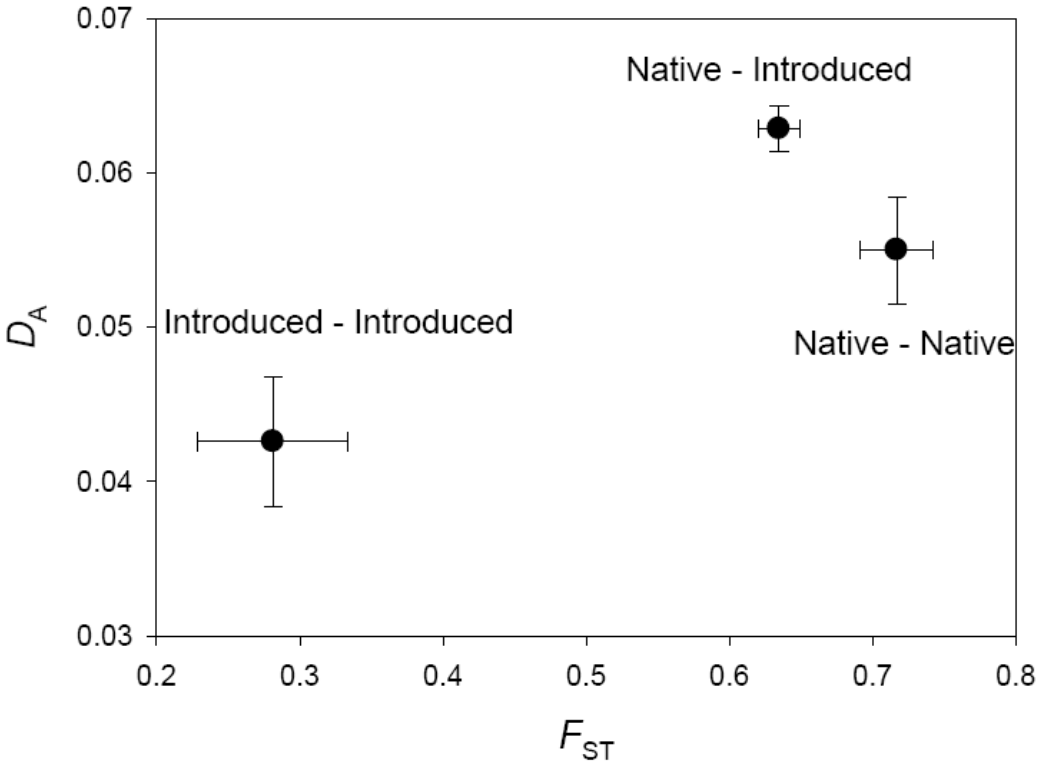

Fig. 2. 4. Haplotype versus nucleotide differentiation (mean and SE) for comparisons of categories of *C. afra* populations (Introduced - Introduced, Native - Introduced and Native - Native). Standard errors were calculated by jackknifing the pairwise population comparisons.

This is consistent with the MDS plot which shows less scatter for the introduced populations than for the native populations (Fig. 2.5). Furthermore, the MDS plot shows that the Ntekete population, although geographically closest to the introduced populations, is genetically distinct from the introduced populations and therefore unlikely to have contributed to the introduction. Relationships among the native populations were not predictable from geographic distances among them. For example, the neighbouring populations at Ntekete and Chiofu were genetically different, while those of the more distant populations at Nkhata Bay and Mara Rocks were genetically very similar (Fig. 2.1 & 2.5).

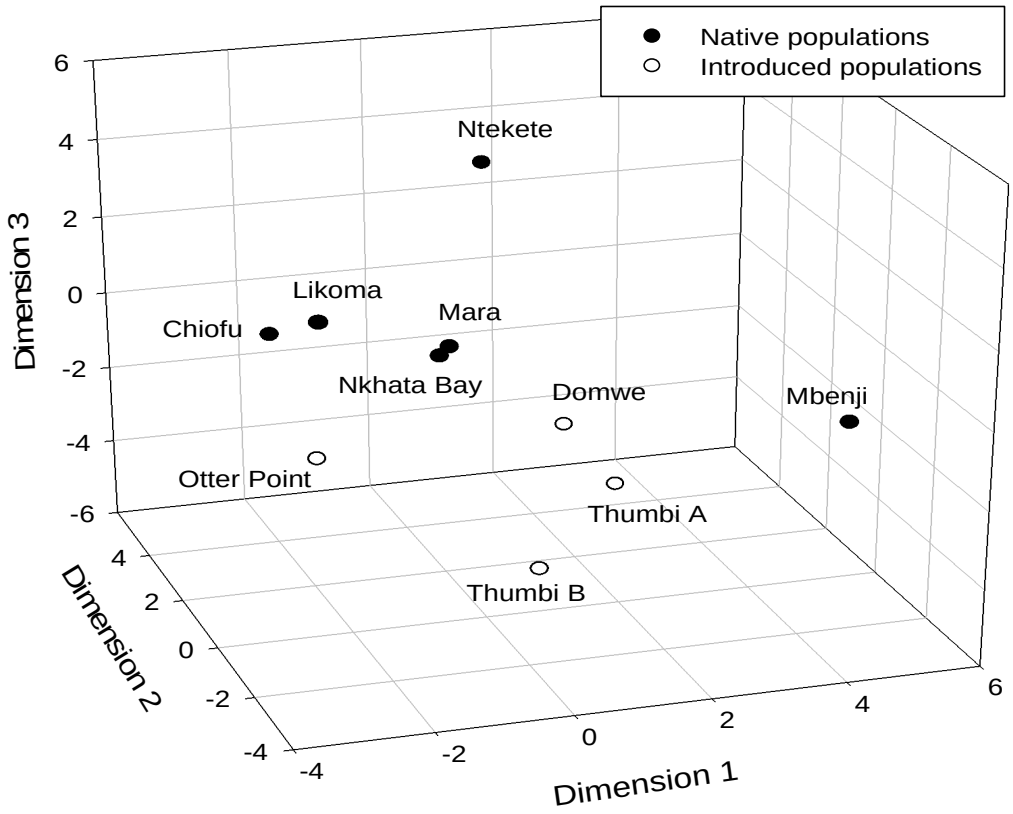

Fig. 2.5. Multidimensional scaling (MDS) plot of genetic distances (D_A) for the native (solid circles) and introduced (open circles) *C. afra* populations.

2.4 **Discussion**

Here I analysed the full control region of mitochondrial DNA of several native and introduced populations of the cichlid fish *C. afra* that was introduced in Lake Malawi National Park in the 1960s. The study shows that introduced *C. afra* populations have high mtDNA variation and also demonstrate evidence of introgressive hybridisation with native *P. zebra*.

2.4.1 *Population structure of native* C. afra

This study has demonstrated a high level of mtDNA sequence variation among the native populations of *C. afra*. This contrasts sharply with previous studies of Lake Malawi cichlids, where levels of lineage sorting are very low not only among intra-specific populations, but even among species from different genera (Moran & Kornfield 1993; Parker & Kornfield 1997; Sturmbauer *et al.* 2001). The presence of incomplete lineage sorting among cichlids taxa, has been attributed to recent contact among populations during low lake level stands (Owen *et al.* 1990; Fryer 1997), recent and rapid speciation (Moran & Kornfield 1993), persistently large population sizes and introgressive hybridisation on secondary contact (Streelman *et al.* 2004).

These results suggest a different recent history for *C. afra* populations, perhaps with persistent geographical isolation and recent genetic bottlenecks. In contrast to many other mbuna species, *C. afra* appears to be subject to local genetic drift and shows high genetic differentiation across populations. This hypothesis may be investigated further using microsatellite markers in chapter 3 of this thesis. The results are unexpected, given that *C. afra* often exists in large population densities, forms large feeding schools in the water column and penetrates to deeper water than many other mbuna (at least 50m; Konings 2008). These traits would suggest that *C. afra* is a good disperser across stretches of deep water. Nevertheless, the high levels of genetic differentiation found in the present study suggest that such potentially favourable conditions for dispersal do not result in frequent gene flow. This conclusion is consistent with its failure to naturally colonise the southern arms of the lake following the last lake level rise. At the time of the first thorough survey of the southern part of the lake in the late 1970s, no *C. afra* was recorded from any other site in the south-eastern or south-western arms of the lake, apart from sites with likely anthropogenic introduction

(Ribbink *et al.* 1983). The southern arms of the lake are relatively shallow and are believed to have been dry during low water stands of the lake around 25,000-70,000 years ago (Scholz & Rosendahl 1988; Cohen *et al.* 2007) or perhaps even as recently as 200 years ago (Owen *et al.* 1990).

2.4.2 *Colonisation history of introduced populations*

All four invasive populations were characterised by elevated levels of nucleotide and/or haplotype diversity compared to the native populations, which is consistent with a scenario of admixture among genetically divergent source populations (Kolbe *et al.* 2004, Roman & Darling 2007). The results furthermore show that genetic differentiation among introduced populations is lower than among native populations, which is also expected under such a scenario if the same source populations have contributed to the gene pool of different introduced populations. The contrasting pattern of genetic differentiation (D_A and F_{ST}) may be due to how these measures are derived, the former being an estimate of net nucleotide differentiation measure. The admixture events between different source populations could have happened either in the holding facilities of the aquarium trader who received stock from a number of geographical localities and/or within the lake when the fish from different localities were released on the shoreline.

The results provide some indication which native populations could have contributed to the gene pool of the introduced populations. Populations from the northern part of Lake Malawi (Mara, Likoma and Nkhata Bay) and to a smaller extent from the central west coast (Mbenji) share haplotypes with at least one introduced population and are therefore potential source populations. By contrast, haplotypes from two nearby south-eastern populations (Chiofu and Ntekete) were not found in the introduced populations. These localities are close to the introduction site but were not regularly visited by aquarium exporters until recently.

It seems likely that at least one additional genetically differentiated native population not sampled in this study must have contributed to the gene pool of the introduced populations. Comparison of this data with 5125 sequences all from African cichlids accessed in the GenBank did not shed more light on the un-sampled population. I suggest that future studies should also investigate the possible contribution from other populations of *Cynotilapia* in the south eastern arm part of the lake, such as *Cynotilapia sp.* "Chinyankhwazi" and *Cynotilapia sp*. "Chinyamwezi". Overall this analysis provides clear

evidence that all introduced *C. afra* populations were founded by an admixture from genetically differentiated source populations.

I furthermore suggest that two independent introductions could explain the genetic differentiation between populations on the northern (Thumbi B) and southern shores of Thumbi West Island (Thumbi A). Streelman *et al.* (2004) had previously found a similar pattern of divergence in both male colour phenotype and microsatellite allele frequencies, but suggested that this was due to rapid evolution following a single colonisation event, perhaps associated with hybridisation with *P. zebra*. This study supports the hypothesis that *C. afra* hybridised with *P. zebra*, but indicates that such introgression is unlikely to be a major cause of differentiation among populations from northern and southern shores of Thumbi West. Rather, it seems plausible that *C. afra* was introduced into two sites on Thumbi West Island, but like the Otter Point populations, the northern Thumbi West population was not recorded by Ribbink *et al.* perhaps because it was small and of restricted distribution at the time of their study. This hypothesis may be investigated further, genetic markers like microsatellites using Approximate Bayesian Computations (ABC) have proved to be useful in answering questions of competing introduction scenario and have described colonisation history of toad *Bufo marinus* (Estoup *et al*. 2001, 2004), bird *Zosterops lateralis* (Estoup & Clegg 2003), corn rootworm *Diabrotica virgifera virgifera* (Miller *et al.* 2005), fruit flies *Drosophila melanogaster* (Thornton & Andolfatto 2006) and *D. subobscura* (Pascaul *et al.* 2008).

The results suggest that *C. afra* may not in fact be spreading out from its original introduction site, as initially feared by conservation biologists. This is consistent with the poor dispersal capacity of *C. afra* as suggested by the marked genetic structure among native populations. It appears that the populations at Otter Point and Thumbi West (A and B sites) probably represent at least three separate introductions. Furthermore, the Domwe Island population does not seem to have been founded by the population from the neighbouring north shore of Thumbi W, but more likely from the population from the southern shore. This suggests anthropogenic transportation, perhaps by some of the many local fishing crews that illegally operate their nets in the area, *C. afra* being a mid-water feeder is particularly prone to being caught in the “Chirimila” nets that target the zooplankton-feeding *Copadichromis* and *Mchenga* locally known as “utaka” and the *Engraulicypris sardella* “Usipa” schools.

2.4.3 *Genetic variation and establishment success*

The analysis suggests that the establishment success of *C. afra* in the southern part of the lake was associated with, and possibly facilitated by, the relatively high genetic diversity of the founder populations, probably in part due to the multiple origins of the founders. Other studies in plants *Phalaris arundinacea* (Lavergne & Molofsky 2007 and review Ellstrand & Schierenbeck 2000), crab *Carcinus maenas* (Roman 2006) and lizard *Anolis sagrei* (Kolbe *et al*. 2004) have suggested that increased genetic diversity of introduced species may be associated with invasion success. Multiple introductions or introduction from different origins is likely to offset the effects of random genetic drift and loss of genetic variation commonly associated with founder events (Kolbe *et al*. 2004; Hänfling 2007; Roman & Darling 2007).

Of course, it is possible that other factors may facilitate an invasion, such as the availability of a suitable ecological niche. Most Malawian rock-dwelling cichlids feed at least partly on benthic prey, such as loose or attached algae and associated fauna. By contrast, *C. afra* seems to feed exclusively on plankton in the water column (Konings 2008). Unlike other mbuna that are generally solitary, territorial and tightly bound to the substrate, *C. afra* typically forms large feeding schools that rise high up into the water column. It is often extremely abundant in suitable habitats (Ribbink *et al.* 1983), as indeed it is now on Thumbi West Island. A recent study of species distributions and abundances on Thumbi West Island attempted to estimate levels of competition among rocky shore cichlid fish species, finding generally higher levels of competitive exclusion between species that had not naturally co-evolved (Young *et al.* 2009b). A striking exception was *C. afra* which, despite its abundance, appeared to have no impact on any of the indigenous species, suggesting that its niche was largely vacant prior to its introduction.

2.5 **Summary**

Genetic variation in many invasive species shows little or no signs of a founder event, suggesting that high genetic diversity may facilitate establishment success. The rockyshore, plankton-feeding cichlid fish *C. afra*is endemic to Lake Malawi, but
naturally absent from many suitable sites. In the 1960s, this species was introduced to the southern areas of the lake, presumably as a result of the aquarium fish trade. It has now

become established on a number of rocky areas within the Lake Malawi National Park.

Here, I have analysed DNA sequence variation in the mitochondrial control region of six native and four introduced populations of *C. afra*, and three populations of a phylogenetically related and hybridizing *P. zebra*. In contrast to previous studies of Lake Malawi rock dwelling cichlids, network analyses suggested that native populations of *C. afra* showed high levels of lineage sorting in mtDNA.

Introduced populations showed higher sequence and haplotype diversity than their native counterparts. These analyses suggested that the elevated gene diversity was largely attributed to the fact that the introduced *C. afra* populations were derived from several genetically distinct and geographically separate populations, and to a lesser extent because of introgressive hybridisation with native *P. zebra*.

The establishment and spread of *C. afra* in Lake Malawi National Park may be partly because of its ability to occupy a vacant ecological niche, but it may also have been facilitated by its enhanced genetic diversity.

Chapter 3

3 Comparison of microsatellite and mtDNA genetic diversity in introduced and native populations of *Cynotilapia afra*

Comparison of microsatellite and mtDNA to investigate founder event (manuscript in preparation): Zidana *et al.* (2009) Differences in suitability of microsatellite and mtDNA to investigate impact of founder event on genetic diversity.

3.1 **Introduction**

Population genetics, in the broad sense, is the study of allele and gene frequency distributions in and among subdivided populations (Lowe *et al.* 2004; Dyck *et al.* 2005). It encompasses estimations of variation in terms of allelic and genotypic frequencies. With appropriate data, the application of population genetic theory allows estimation of gene flow within and among populations, estimation of effective population sizes, testing for occurrence of genetic bottlenecks, testing rapid expansions of population size and pairwise genetic distances (Lowe *et al.* 2004; Dyck *et al.* 2005; Allendorf &Luikart 2007).

Genetic diversity is a key factor enabling adaptation, and therefore survival, of natural populations in changing environments (Frankham 1998; Allendorf &Luikart 2007). Conversely, limited genetic diversity may reduce the potential for populations to adapt in the long term but loss of genetic diversity can also more immediately lead to decreased fitness within populations, due to inbreeding depression (Lande 1995; Frankham 2005; Lavergne & Molofsky 2007). Therefore, evaluations of genetic diversity are common in population genetics and are particularly important in conservation genetic studies working on introduced populations (Frankham 2005). In principle, genetic diversity at loci with functional importance, such as protein-coding genes, RNA-coding or regulatory sequences, is what affects a population's ability to respond to selection (Nei 1987; Avise 1994). Ultimately, researchers would like to measure genetic variability across all or at least a significant

fraction of such functionally important loci, which is an unrealistic scenario today but it may be potentially possible within a few years. While awaiting technology for large-scale genomic analysis of multiple individuals within a population geneticists have primarily used neutral genetic markers to infer the levels and patterns of genetic diversity (Sambrook *et al.* 1989; Lowe *et al.* 2004). Among the most widely used are molecular genetic markers isolated from microsatellite DNA and mitochondrial DNA (mtDNA) regions (Avise 1994; Jarne & Lagoda 1996). These markers have been used extensively in studying evolutionary processes associated with biological introductions to investigate issues associated with founder events such as the amount of genetic variation within introduced populations, compared to their native counterparts (Angers & Bernatchez 1997; Lindholm *et al.* 2005; Roman 2006; Brown & Stepien 2009). The biological and evolutionary processes that affect the microsatellite and mtDNA sequences are different (Wilson *et al.* 1985; Tautz 1989; Avise 1994; Wright & Bentzen 1994; Ballard *et al.* 2004). A microsatellite locus is usually diploid and may be inherited from both parents, while mtDNA is haploid and inherited from the female parent (Avise 1986; Park & Moran 1994; Birky 1995). This suggests some differences may exist in the way these two regions of DNA respond to demographic changes and the stochastic events that are usually associated with founder populations. Therefore the use of multiple markers has been recommended in conservation and evolutionary genetics studies (Takezak & Nei 1996; Nauta & Weissing 1996; Colson & Goldstein 1999; Ting *et al.* 2000; Selkoe & Toonen 2006; Zachos *et al.* 2009), although many studies still make conclusions based on inferences from one molecular marker (Roman & Palumbi 2003).

Introduced species are well placed for studying rapid evolutionary changes in population and conservation genetics (Hänfling 2007; Roman & Darling 2007). However efficient statistical inferences are often difficult to achieve when dealing with populations with complex demographic history, for example, multiple founders from genetically differentiated source populations (Beaumont *et al.* 2002; Allendorf &Luikart 2007). The approximate Bayesian computation (ABC) inference on the basis of summary statistics is well suited to testing complex scenarios that arise in population genetics (Beaumont *et al.* 2002; Pascual *et al.* 2008). Properties of the posterior distribution of a parameter, such as its mean or density curve, are approximated by fitting a local linear regression of simulated parameter values on simulated summary statistics, and then substituting the observed summary statistics into the regression equation (Beaumont *et al.* 2002).

When enough background information is available on historical and demographic statistics from an introduced population the ABC approach can be used successfully to investigate colonisation history (Estoup *et al.* 2001; Pascual *et al.* 2008; Cornuet *et al.* 2008). Reconstruction of colonisation histories may provide the opportunity to investigate the genetic consequences of founder events as well as the evolutionary trajectories responsible for established of introduced populations in an environment (Mayr 1954). The establishment success of introduced species may depend on their ability to evolve in response to the selection pressures in their new environment (Lee 2002). Genetic drift may also play an important role during establishment and even determine colonisation or invasion success (Tsutsui *et al.* 2000). Consequently, when a species independently invades several areas, the outcome of the colonisation and the convergence of results in the different areas may depend upon the specific genetic pool of the founding population, as well as the population dynamics and selective processes associated with colonisation. To understand the evolutionary genetics of colonisation one must identify the most likely source of colonizers, the levels of genetic diversity of both introduced and native populations, the geographical pathways of spread, and the ability of the populations to evolve in novel environments (Frankham 1998; Allendorf &Luikart 2007).

The introduction event of *Cynotilapia afra* into Lake Malawi National Park is well documented (Ribbink *et al.* 1983; Konings 2008). Previous studies have investigated the source of founding populations (Zidana *et al.* 2009) and genetic diversity of introduced *C. afra* populations at microsatellite loci (Streelman *et al.* 2004) and mtDNA control region (Zidana *et al.* 2009).

This study analyses the microsatellite genetic diversity of introduced and native *C. afra* populations using population genetic analyses and a newly developed ABC approach to investigate the competing hypotheses of introduction history patterns (independent or stepping stone) in Lake Malawi National Park. I have analysed data on the mtDNA control region (described in chapter 2), and compared the results on effective number of haplotypes with alleles, and also compared standardised genetic differentiation measures between the two markers. Multiple introduction events associated with genetic drift and a bottleneck may have generated a contrasting pattern of genetic variation at microsatellite and mtDNA loci in the introduced range. Finally these results are used to interpret the suitability of microsatellite and mtDNA markers for investigating the impact of founder events on the genetic diversity of introduced populations.

3.2 **Materials and methods**

3.2.1 *Sampling*

Samples of *C. afra* were collected from six localities within native range: Mara Rocks; Nkhata Bay; Likoma Island, Ntekete, Chiofu and Mbenji Island and from all four known sites inhabited by introduced populations: Domwe Island; Otter Point and two locations from Thumbi West Island (site A and B). Sixty individuals were collected from each locality (Fig. 2.1).

A non-invasive sampling method was used by collecting live fish samples using monofilament nets while SCUBA diving. A small tissue sample was taken by fin clipping the posterior (soft-rayed) dorsal fin, which regenerates within a few days. To ensure reliable species identification, all samples were collected from territorial males in full breeding colour and jaw teeth of all specimens were examined using a hand lens, as possession of unicuspid teeth is diagnostic for the genus *Cynotilapia* (Trewavas 1935, Fryer 1959). All live specimens were released back into the lake. Fin clips were preserved in 99% ethanol before molecular analysis in the laboratory.

3.2.2 *DNA extraction and molecular data analysis*

Total genomic DNA was extracted from fin clips using Promega WIZARD genomic DNA purification kit. Samples were screened at 6 microsatellite loci using published primers: Ppun32, Ppun21, Ppun7 and Ppun5 (Taylor *et al*. 2002), Pzeb3 (van Oppen *et al*. 1997) and UHN154 (Kellogg *et al*. 1995). The 10 µL polymerase chain reaction (PCR) consisted of 1µL DNA template, primer (1mM), dNTPs (0.1mM), 0.5 Units of *Taq* polymerase, 1X PCR buffer and $MgCl_2$ (1.25mM). The PCR amplification was carried out on a Peltier Thermal Cycler 200 using the following parameters: 3 min 94°C 1 cycle, followed by 94°C for 30 s, annealing temperature ranged from 54 to 62 °C for 30 s and 72 °C for 30 s for 27 cycles, followed by 1 cycle of 10 min at 72 °C. Two loci Ppun32 and Ppun5 were amplified using two stage protocols, thus 4 cycles at an initial high annealing temperature, followed by 29 and 26 cycles respectively at lower annealing temperature. PCR products were genotyped on

a Beckman Coulter CEQ^{TM} 8000 machine. The genotyped microsatellite alleles were tested for data consistency and presence of null alleles using MICRO-CHECKER version 2.2.3 (van Oosterhout *et al*. 2004).

To determine the suitability of microsatellite markers to investigate the impact of founder event on genetic diversity of introduced populations, I have compared the results of this study with mitochondrial DNA (mtDNA) genetic variation from the full sequences of the control region (Zidana *et al.* 2009).

3.2.3 *Genetic variation within and among* C. afra *populations*

The genetic diversity within populations of *C. afra* was calculated as mean number of effective alleles (n_e; Kimura & Crow 1964) using allele frequencies and expected heterozgosity (H_E) values generated from ARLEQUIN version 3.1 (Excoffier *et al.* 2005). The effective number of alleles (n_e), is considered as a better estimate other than mean number of alleles within a population since the latter depends on data sample size (Tajima *et al.* 1994). General Linear Model (GLM) was used to test the significance of results as employed in MINITAB version 12.1 (MINITAB Inc.).

Pairwise genetic differentiation was quantified as G_{ST} (Nei 1978) using FSTAT version 1.2 (Goudet 1995). Theoretically, mtDNA effective population size is one-fourth that of microsatellite loci and this may increase estimates of genetic differentiation in mtDNA as compared to microsatellite loci (Excoffier *et al.* 1992; Jarne & Lagoda 1996; Rubinsztein *et al.* 1995; Vigouroux *et al.* 2003). Therefore, to compare the genetic differentiation measures between the two markers the genetic distance values were further standardised (G'_{ST}) according to Hedrick (2005). The significance of genetic differentiation was tested using Kruskal-Wallis test as employed in MINITAB version 12.1 (MINTAB Inc.). The relationship between mtDNA and microsatellite differentiation was further investigated using analysis of molecular variance (AMOVA) framework as implemented in ARLEQUIN 3.1 (Excoffier *et al.* 2005). In AMOVA framework analysis, the hierarchical level of genetic structures was investigated among *C. afra* populations from three levels of genetic structure which consisted of: native-native (N-N), native-introduced (N-I) and introduced-introduced (I-I) population genetic structures. The significance of variation was tested using 10000 permutations.

3.2.4 *Identification of source of introduced* C. afra *populations*

The sources of introduced populations were determined by comparing the genetic structure of native against introduced *C. afra* populations using STRUCTURE v2.3 (Pritchard *et al*. 2000; Hubisz *et al.* 2009). The STRUCTURE software uses Bayesian cluster algorithms to assign individuals into predetermined number of K genetic clusters. The most probable number of genetic clusters in the data set is determined by the least negative value of the posterior probability (Pritchard *et al*. 2000). In this study the most likely number of genetic clusters was determined using a statistic delta K which is based on the rate of change in the log probability of sampled data between successive K values (Evanno *et al.* 2005). In this model STRUCTURE accurately detects the uppermost hierarchical level of genetic structure for the tested scenario (Evanno *et al.* 2005). The appropriate run length and model settings were chosen by running several simulations at each K with different lengths and settings, the admixture model with location prior (Hubisz *et al.* 2009) using (500,000 burn-in and 1,000,000 Markoc Chain Monte Carlo (MCMC)) was the best fit and is the one I used for this study. The location prior allows STRUCTURE to detect genetic structuring with a few microsatellite alleles, at lower levels of divergence and less data samples without being bias towards detecting false structure (Hubisz *et al.* 2009). In this model a statistic "*r*" with a range of 0-1 is used to determine if the sample location represents informative genetic structures, 0=informative and 1=not informative (Hubisz *et al.* 2009).

The data was further analysed using a relatively new method in population and conservation genetics studies, an approximate Bayesian computation (ABC) inference (Beaumont *et al.* 2002). The competing hypotheses in introduction pattern (independent and stepping stone) of *C. afra* into Lake Malawi National Park was investigated using DIY ABC software v 7.1 (Cornuet *et al*. 2008).

Previous studies have suggested Thumbi west Island site A as an initial point of *C. afra* introduction into Lake Malawi National Park (Ribbink *et al.* 1983; Streelman *et al.* 2004). I have tested three introduction patterns: two stepping stone colonisation patterns (Scenarios 1: native to introduced population 1 and to introduced population 2; Scenario 2: native to introduced population 2 and to introduced population1) and one independent introduction pattern (Scenario 3: native to introduced population 1 and native to introduced

population 2). The populations (1 to 3) in this analysis are also representing locations of this introduction event (Fig. 3.1).

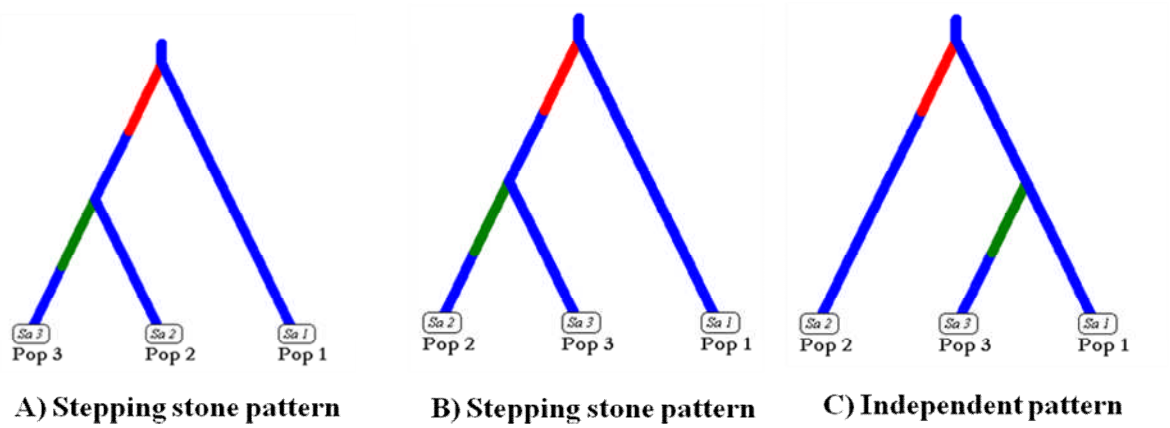

Fig. 3.1. ABC output for introduction scenarios considered for the inference of the introduction pattern of *C. afra* in Lake Malawi National Park. Red and green parts indicate the various bottlenecks occurring in the first few generations following introductions. Pop 1=native populations, Pop 2 and 3 represents introduced populations.

The demographic and historical information of this introduction event was specified by six variable parameters: stable effective population size (*Ns*), effective number of founders (N_f1 and N_f2) during the first and second introduction event respectively, duration of bottleneck (*DB*) and generation time (*t*).

The dispersal pattern was based on Thumbi West site A as an initial point of introduction and then dispersal to other nearby localities (Thumbi west Island site B, Otter point and Domwe) within the park. The native populations were pooled as one sample basing on the results from mtDNA which showed that this introduction event was associated multiple sources (Zidana *et al.* 2009), therefore the four sample sets are: (1) a pool of native populations, Thumbi West Island site A and Thumbi site B, (2) a pool of native populations + Thumbi site A + Otter Point, (3) a pool of native populations + Thumbi site A + Domwe Island and (4) a pool of native populations + Thumbi site B + Domwe Island.

The posterior probabilty values are expressed as a percentage that represents the most probable scenario among the three competing ones analysed here were determined using approximate Bayesian computations (ABC) with prior information. The prior distribution information used in this analysis were based on the known dates of introduction, time scale since the year of introduction (*t*=1-40), based on an introduction event of 1960s (Ribbink *et*

al. 1983; Konings 2008; Young *et al.* 2009), founder size of introduced populations (*N*=10-2000) and stable effective population size of source population (*Ns*=100-100000) were based on estimates from other mbuna cichlids and studies which have investigated the distribution of introduced fish populations at Thumbi West Island (Ribbink *et al.* 1983; Trendall 1988; van Oppen *et al.* 1997; Won *et al.* 2005; Young *et al.* 2009). To my knowledge there is no other study to date that has tested or estimated duration of bottleneck in Lake Malawi cichlid populations, therefore with the help of DIY ABC framework I estimated a duration of 5 generations (*DB*=1-10 years) (Table 3.2).

DIY ABC software makes inferences about each of the parameters based on posterior distributions probabilities (Cornuet *et al*. 2008). The ABC provides accurate estimates of posterior probabilities with as few as five loci and 30 individuals in a data set (Guillemaud *et al.* 2009). In the estimation stage, I have used a local linear regression which is suggested to provide a better approximation of posterior distributions (Beaumont *et al.* 2002) and an estimate of the most supported introduction scenario among the three tested was selected using the highest posterior probability estimated percentage (> 95%) from the analysis (Beaumont *et al.* 2002; Cornuet *et al*. 2008).

3.3 **Results**

3.3.1 *Microsatellite diversity within populations*

The introduced *C. afra* populations showed on average a slightly lower number of effective alleles across loci (mean (±SE), n_e =16.50 ± 1.34)) than natives n_e =19.42 ± 0.79)), p=0.08 (Fig. 3.2). The introduced population at Thumbi West Island site A, exhibited greater genetic diversity than other populations in the introduced range, and this was comparable to populations from the native range (Fig. 3.2). In the native range, the Nkhata Bay population exhibited the highest effective number of alleles (n_e = 14.64 ± 1.91).

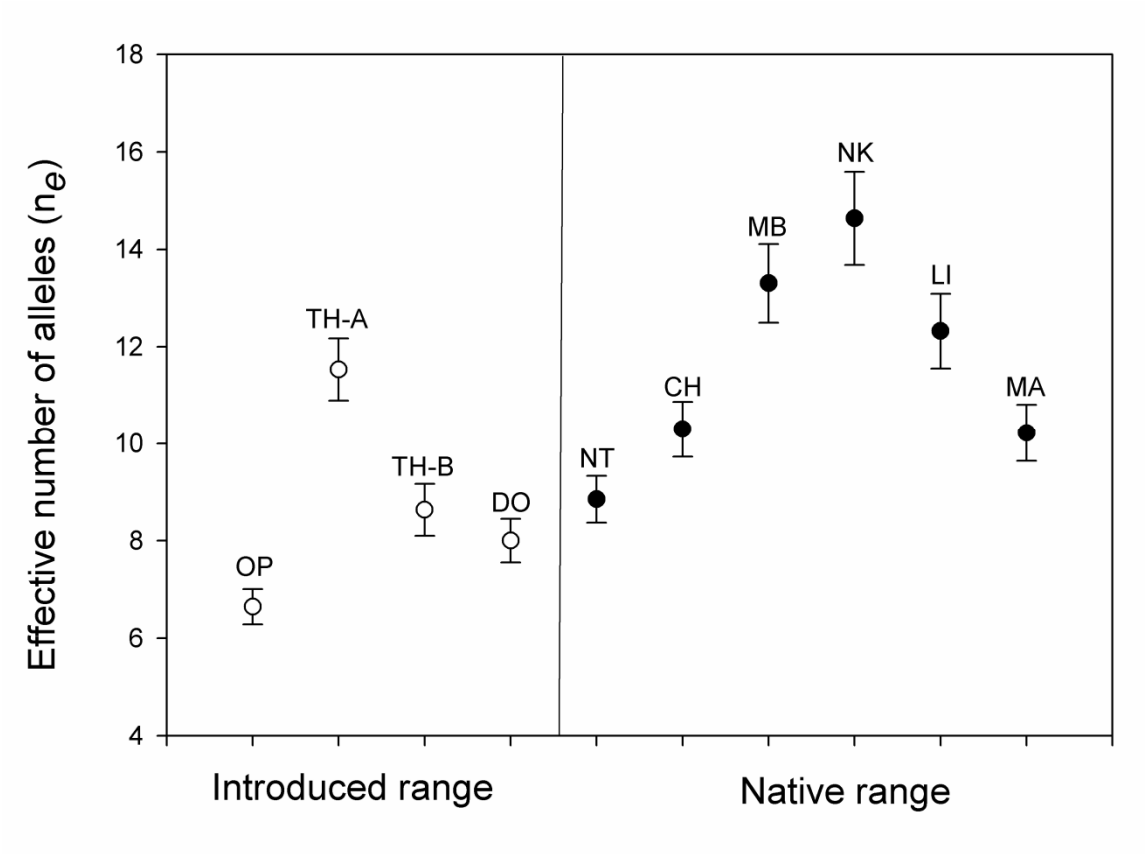

Fig. 3.2. Genetic diversity of *C. afra* populations expressed as (mean (SE)) effective number of alleles (n_e) for native (solid circles) and introduced (open circles) populations. Standard errors were calculated using a jackknife procedure. Populations identities, TH-A and B = Thumbi west Island site A and B, OP = Otter point, DO = Domwe Island, CH = Chiofu, NT = Ntekete, MB = Mbenji, NK = Khata bay, MA = Mara rocks, LI = Likoma.

Genetic diversity observed within *C. afra* populations was compared to results from a simulated founder event. The size (number of founders = N_f) of source populations was derived for populations with known allele or mtDNA haplotype frequency distributions (Chapter 2). The source populations were modelled to contribute founder individuals proportional to the observed mtDNA haplotype frequency distribution in the introduced *C. afra* population. The demographic parameters (N_f and the time of introduction, t) were based on analysis of microsatellite variation using DIY ABC software version 7.1 (Cornuet *et al.* 2008). The mtDNA data was reanalysed to investigate the effective number of haplotypes within *C. afra* populations in native and introduced ranges (Appendix I).

The results showed a contrasting pattern between microsatellite and mtDNA among population genetic diversity (Fig. 3.3). The introduced populations exhibited a low mean effective number of alleles at microsatellite loci and an increased number of haplotypes at mtDNA control region. The contrasting pattern of genetic variation within populations at microsatellite and mtDNA control region is exhibited in both observed and simulated data sets (Fig. 3.3).

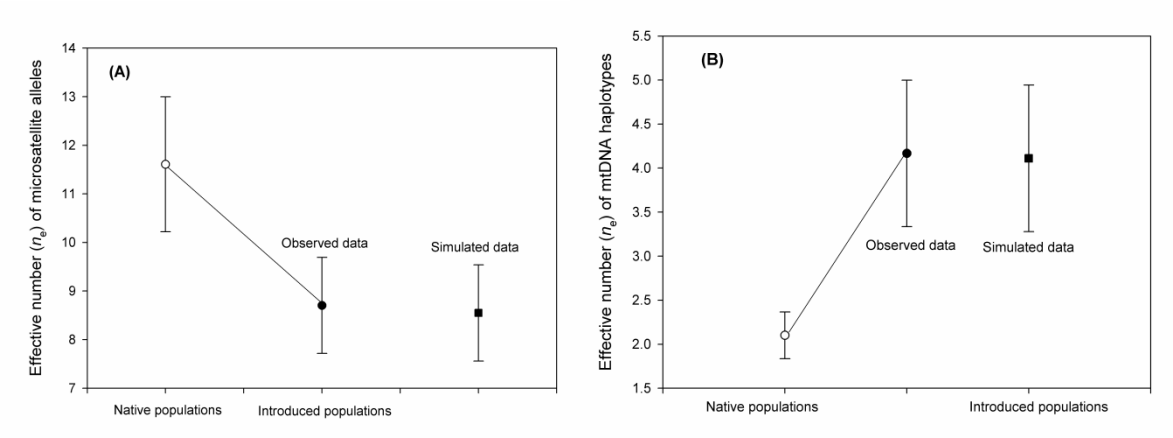

Fig. 3.3. The contrasting pattern of genetic variation from effective number of alleles (mean and SE) using data from (A) microsatellites and (B) the mtDNA control region, in groups of introduced and native *C. afra* populations. Two kinds of data were used in this analysis, the open and closed circle data point=observed data estimates, while closed box=simulated data estimates basing on statictics from microsatellite DIY ABC and mtDNA analysis.

3.3.2 *Microsatellite diversity among populations*

There were no significant differences in microsatellite differentiation (G'_{ST}) among groups of: native-native (N-N), native-introduced (N-I) and introduced-introduced (I-I) comparisons (Kruskal-Wallis test, H=1.18, d.f.=2, p=0.57), but mtDNA control region showed significant differences in genetic differentiation (G'_{ST}) among groups: N-N, N-I and I-I for the mtDNA control region (Kruskal-Wallis test, H=8.67, d.f.=2, p=0.01) (Fig. 3.4).

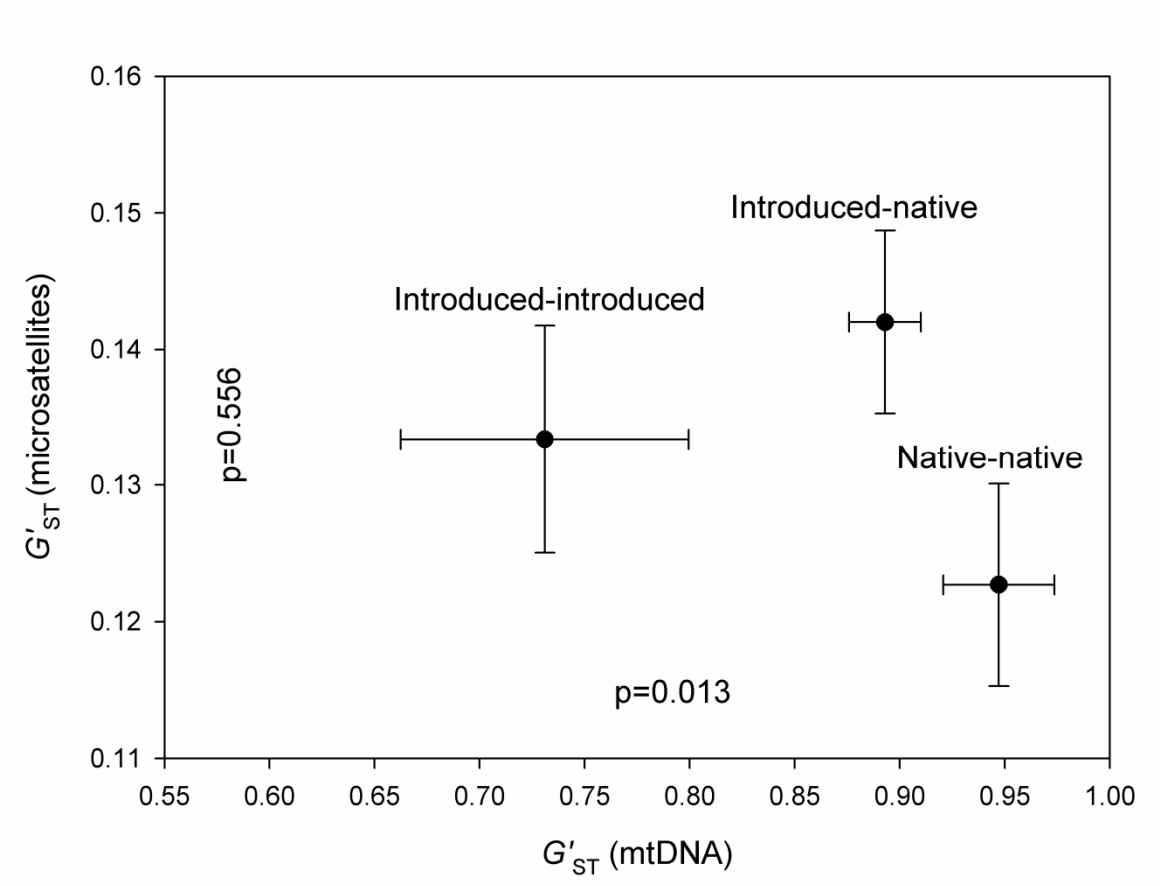

Fig. 3.4. Mean (SE) of the genetic differentiation among groups of native-native (N-N), native-introduced (N-I) and introduced-introduced (I-I) comparisons for mtDNA (X-axis) and microsatellites (Y-axis). Standard errors were calculated using a jackknife procedure.

AMOVA results showed that the genetic differentiation at microsatellite loci was mainly explained by the genetic structure level within populations (96.58%, $p<0.001$), while mtDNA was explained by the differences at within groups level (49.81 %, $p<0.001$) of differences (Fig. 3.5). This was consistent with the genetic differentiation G'_{ST} results in (Fig. 3.4) which showed no significant differences among *C. afra* groups (N-N), (N-I) and (I-I) at the microsatellite level.

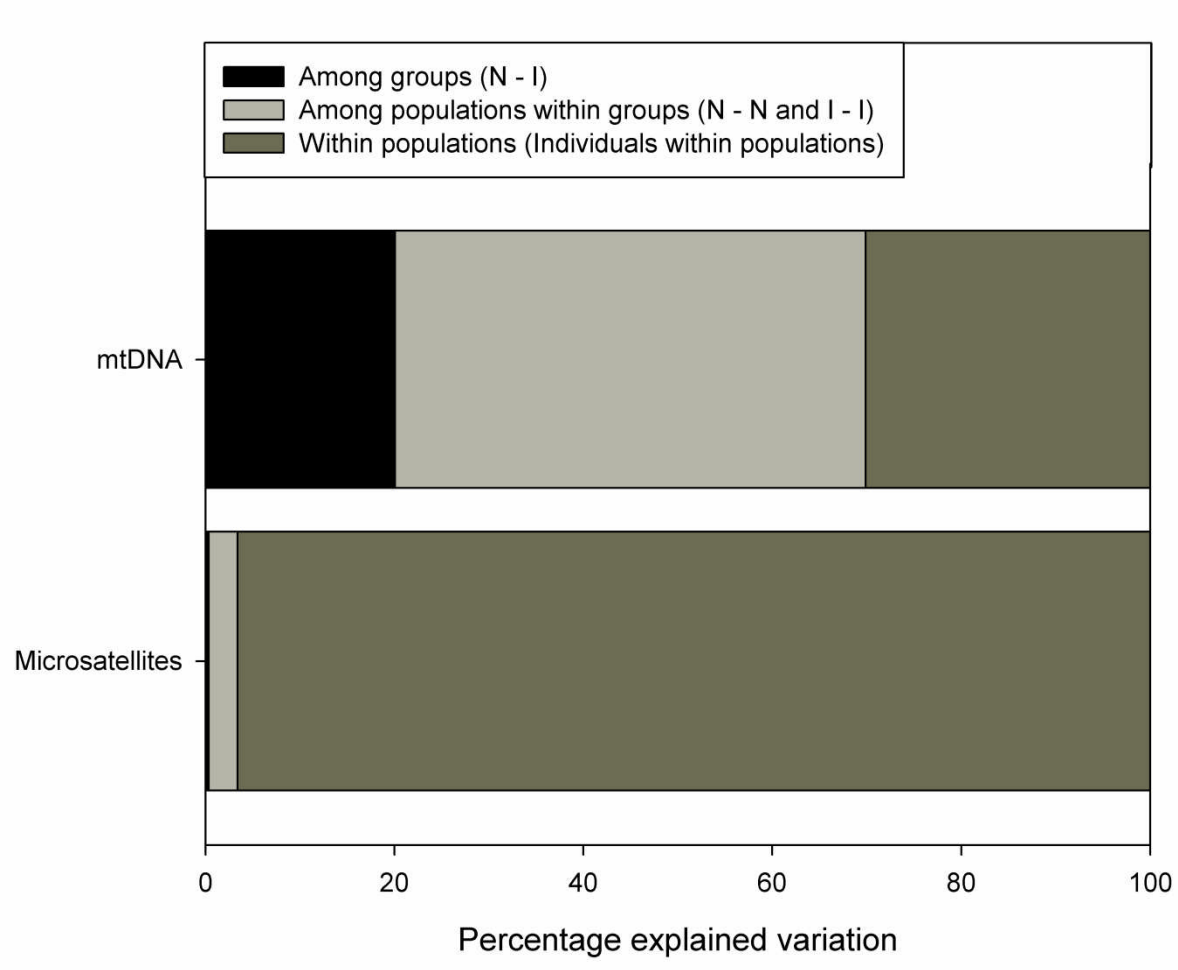

Fig. 3.5. The analysis of molecular variance (AMOVA) at microsatellite and mtDNA control region genetic structure levels analysed are native vs native (N-N), native vs introduced (N-I) and introduced vs introduced (I-I) among *C. afra* populations in introduced and native ranges of Lake Malawi

3.3.3 *Identification of source of introduced C. afra populations*

The STRUCTURE analysis revealed four genetic clusters (K=4), assigned into clusters of: (Otter point), (Thumbi West Island site B and Domwe), (Thumbi West Island site A, Ntekete, Likoma and Mara rocks) and (Chiofu, Mbenji and Nkhata Bay) from analysis of *C. afra* populations (Fig. 3.6). The results showed that sample locations were representative of genetic clusters (r=0.49) and that the use of admixture model with assistance of location prior improved the signal in genetic clustering of individuals (Fig. 3.6b). The four genetic clusters were tested for validation using an F_{ST}-based AMOVA and the results showed significant differences among the clusters $p<0.05$ (Fig. 3.6).

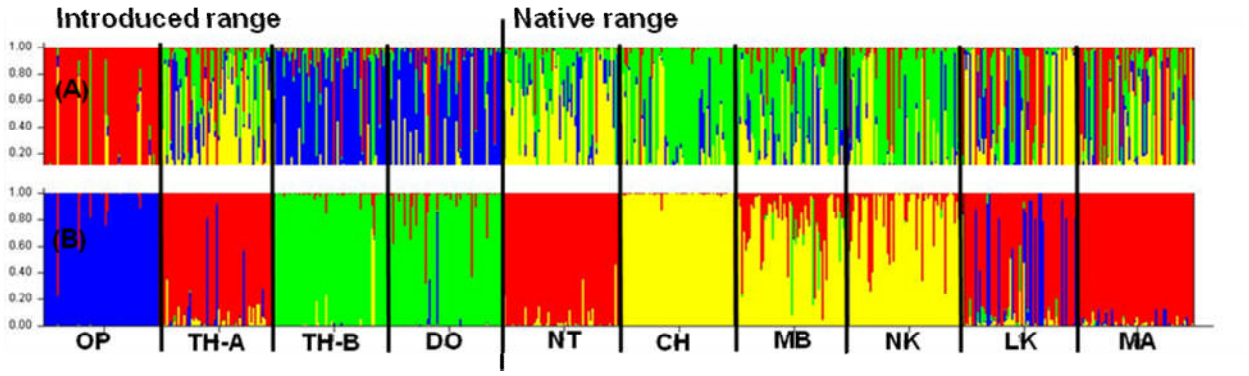

Fig. 3.6. Population structure of *C. afra* in introduced and native range in Lake Malawi, analysed by STRUCTURE software, K=4 (A) without location prior and (B) with location prior. Each individual sample is represented by a thin colour line and each colour represents a different genetic cluster. The thick black line separates the populations from different sites, population identities are listed below each segment (OP=Otter Point, TH-A=Thumbi West Island site A, TH-B=Thumbi West Island site B, DO=Domwe Island, NT=Ntekete, CH=Chiofu, MB=Mbenji, NK=Nkhata Bay, LK=Likoma and MA=Mara Rocks).

The approximate Bayesian computation (ABC) analysis results strongly supported an independent introduction event of *C. afra* at Thumbi west Island site A and B with an estimated probability value of 97% (sample set 1) (Fig. 3.7 & Table 3.1). Domwe Island populations was highly supported by a stepping stone introduction event with an estimated probability value of 95%, surprisingly the analysis showed that the source population was Thumbi west Island site A and not site B which is geographically closer (Fig. 3.7 & Table 3.1). The results were not conclusive on the introduction pattern and a source population for *C. afra* population at Otter Point.

Table 3.1: The approximate Bayesian computation (ABC) estimated posterior probability values for competing introduction scenarios (independent and stepping stone) pattern for *C. afra* introduction into Lake Malawi National Park

	Posterior probability		
Sample set (N + Intr1 + Intr2)	Scenario 1 N > Intr1> Intr2	Scenario 2 N > Intr2 >Intr1	Scenario 3 N > Intr1 and then N > Intr2
1) N + THA + THB	0.03	0.00	0.97*
2) N + THA + OP	0.78	0.00	0.22
3) N + THA + DO	0.95*	0.00	0.05

Note: Posterior probabilities were computed from the (n_δ) smallest Euclidian distances 50,000 using logistic regression for all scenarios. Intr1 and Intr2=Introduced population 1 and 2 respectively. Scenario 1and 2=stepping stone introduction pattern; Scenario 3=independent introduction pattern. The stars (*) are showing a significant support for the tested scenario. N=native; THA=Thumbi west site A; THB=Thumbi west site B; OP=Otter point; DO=Domwe Island.

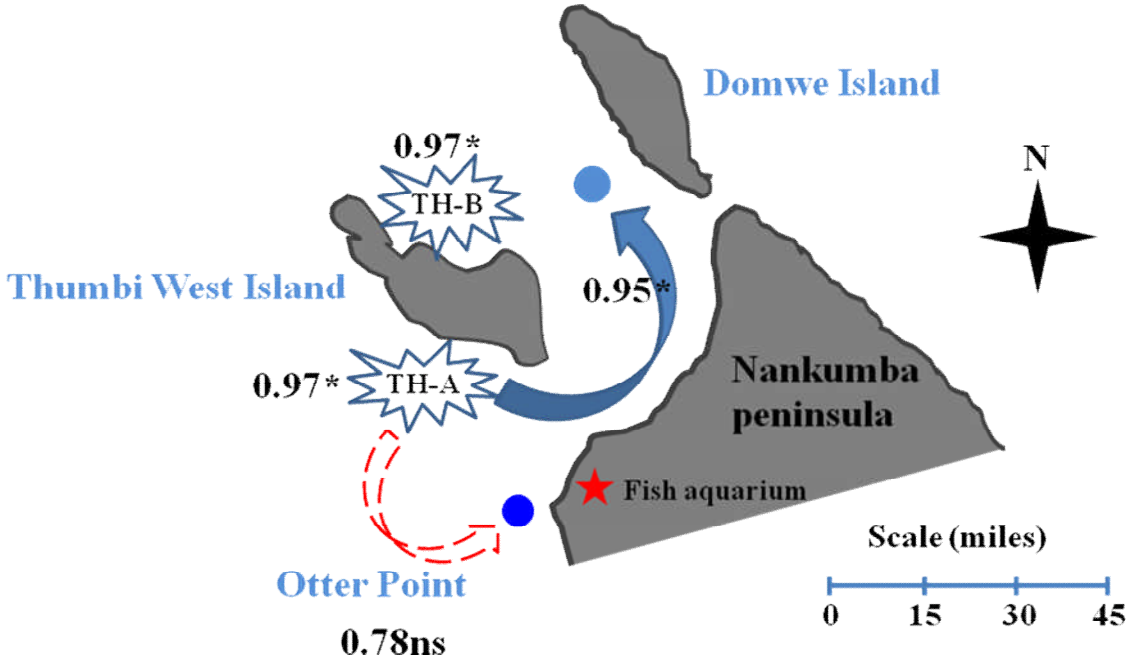

Fig. 3.7. Study area showing the introduction pattern of *C. afra* at four locations (Thumbi West Island site A and B (TH-A and TH-B), Otter Point and Domwe Island within Lake Malawi National Park with the support of posterior probabilities from DIY ABC analysis. The red star at Nankumba peninsula is a location of an aquarium for a fish trader that was used to hold the putative founder populations. The two competing introduction scenarios are represented by the big open blue stars (independent) and curved arrows (stepping stone) introduction pattern.

Using an independent introduction pattern (scenario 3) and sample set 1 (Natives + Thumbi A + Thumbi B) which exhibited highest (97%) probability support values among the three scenarios which were tested in this study (Table 3.1 & Fig. 3.7), I estimated supported values using point estimates of 2.5 and 97.5% quantiles, for the analysed demographic parameters: number of founders, duration of bottleneck and time since this introduction event occurred. The results showed that this introduction event was associated with a small duration of genetic bottleneck 2-10 years after introduction (Table 3.2).

Table 3.2: The prior distribution and supported values for demographical parameters describing the introduction pattern model of *C. afra* in Lake Malawi National Park

Parameter	Prior distributions	Range of supported values (2.5 and 97.5% quantiles)
Ns	(100, 100000)	19945 to 96667
N_f1	(10, 2000)	87 to 7247
N_f2	(10, 2000)	181 to 4124
DB	(1,10)	2 to 10
t1	(1,40)	8 to 49
t2	(1,40)	3 to 49

Note: *Ns* = stable effective population size; N_f1 and N_f2 = effective number of founders in introduced population 1 and 2 respectively; *DB* = duration of bottleneck and *t1* and *t2* are time 1 and 2 after introduction event in years. Supported values were generated from sample set 1 (Natives + Thumbi A + Thumbi B) and scenario 3 (independent introduction pattern) which exhibited highest (97%) probability support values.

The summary below shows the suitability of microsatellite and mtDNA marker suitability differences when investigating impact of founder events on genetic diversity. The summary shows that microsatellite markers are useful in detecting genetic bottleneck and drift in introduced range. The mtDNA marker is good in detecting genetic differentiation among populations in native range and founder event associated with multiple sources (Table 3.3).

Table 3.3: A summary to compare and contrast the colonisation history and genetic variation signal obtained from microsatellite and mtDNA genetic markers with their implication on conservation genetics signal

	Molecular genetic marker used	
Parameter investigated	Microsatellite loci	mtDNA (D-loop)
Variation within population		
Introduced range	Low	High
Native range	High	Low
Variation among populations (G_{ST})		
Introduced range	Moderate higher	Low
Native range	Moderate lower	High
Demographic inference (colonisation history)		
Introduced range	Good	Good
Native range	-	-
Bottleneck detection		
Introduced range	Yes	No
Native range	-	-
Multiple source detected		
Introduced range	Yes (STRUCTURE analysis)	Yes (Network analysis)
Native range	-	-
Conservation genetics implications		
Introduced range	Good (expose effects of drift)	Bad (Over estimate)
Native range	Good (high genetic variation)	Bad (under estimate)

3.4 **Discussion**

In this study I analysed microsatellite genetic variation within and among *C. afra* populations in its introduced and native ranges in Lake Malawi. The results were compared with reanalysed data of mtDNA control region, effective number of haplotypes and standardised measure of genetic differentiation.

This study shows that introduced *C. afra* populations have reduced genetic variation at microsatellite loci compared to native populations. This is in contrast with mtDNA results which showed elevated genetic variation in the introduced range compared to the native range. Comparison of observed against simulated data sets also showed a similar contrasting pattern of genetic variation between microsatellites and the mtDNA control region.

3.4.1 *Within population genetic diversity*

The relatively high genetic variation at the mtDNA control region in introduced populations compared with native populations may be explained by multiple genetically differentiated founder populations (Soltis & Soltis 2000; Ellstrand & Schierenbeck 2000; Kolbe 2004), while decreased diversity in native *C. afra* populations at microsatellite loci may be explained by a small founder population size which is associated with a genetic bottleneck and drift (Allendorf & Luikart 2007; Mack *et al.* 2000; Ellstrand & Schierenbeck 2000; Lopez *et al.* 2009). This is consistent with the expectation that founder populations usually have small effective population sizes that are susceptible to inbreeding and that they accumulate detrimental mutations in a phenomenon called genetic load (Tsutsui *et al.* 2000; Dlugosch & Parker 2008; Peacock *et al.* 2009; Valade *et al.* 2009). Founder populations with reduced genetic variation may have undergone a lag period between arrival, initial spread and the onset of rapid population growth (Sakai *et al.* 2001; Allendorf & Luikart 2007). This is expected if evolutionary change, such as adaptation to new habitat, evolution of invasive life history characteristics, and purging of genetic load responsible for inbreeding depression is an important part of the establishment process (Mack *et al.* 2000; Parker *et al.* 2003; Allendorf & Luikart 2007).

The contrasting pattern of genetic variation within *C. afra* populations in native and introduced ranges between microsatellite loci and the mtDNA control region may suggest that microsatellite markers are probably best suited to investigating levels of genetic variation

in founder events associated with complex population demographic history in conservation genetics studies.

The propagule pressure, which includes the number of individuals released during a founder event and also the number of release events (Richardson *et al.* 2000; Kolar & Lodge 2001), is suggested to be one of the most important factors predicting whether the introduced population will become established and spread to nearby localities (Kolar & Lodge 2001; Parker *et al.* 2003). This has been demonstrated in studies from many different taxa including: birds (Griffith *et al.* 1989), spiders (Schoener & Spiller 1995), mammals (Forsyth & Duncan 2001) and insects (Berggren 2001). The mechanism behind establishment success through increased propagule pressure is that increased founder population size may decrease the effects of demographic stochasticity, in turn decreasing the effects of loss of genetic diversity through drift, and decreasing inbreeding depression (Mooney & Cleland 2001; Allendorf & Luikart 2007). The increased genetic variation at Thumbi West Island may be an indication of a large founder size from multiple genetically differentiated source populations, which may be facilitating their establishment success. Another alternative explanation for establishment success of *C. afra* at Thumbi West Island may be the availability of suitable ecological niches that were not being utilised by native species. The rock-dwelling cichlid group in the African Great Lakes is well known for its feeding specialisations (Fryer 1959; Danley *et al*. 2000; Konings 2008). An ecological study on competitive exclusion has shown that introduced *C. afra* populations in Lake Malawi National Park have not had a significant impact on native species despite their abundance (Young *et al.* 2009).

3.4.2 *Among population genetic diversity*

This study has shown a contrasting pattern of genetic differentiation between microsatellite and mtDNA markers, with microsatellite markers showing no genetic differentiation among groups of native-native, native-introduced and introduced-introduced. The mtDNA marker is haploid in state with no recombination, unlike microsatellites that experience recombination (Avise 1986; Bentzen *et al.* 1988; Park & Moran 1994). This characteristic enables mtDNA to have a smaller effective population size, implying that genetic drift can cause frequency differences between isolated populations more readily in mtDNA than in nuclear DNA (Avise 1994). In many organisms mtDNA evolves more rapidly than nuclear DNA. This may be in part due to lack of sequence repair mechanisms in

mtDNA for mutations that arise during replication (Brown *et al.* 1979; Wilson *et al.* 1985; Ballard *et al.* 2004). The characteristics of mtDNA when compared to microsatellite loci and the results from this study suggest that mtDNA may be well suited to investigate genetic differentiation among populations.

The contrasting pattern of genetic variation as revealed by microsatellite and mtDNA markers in this study has also been confirmed by modelling the effects of founder events using similar demographic scenarios as that encountered in our observed data set. This may suggest that the signal obtained in the observed data set may not be due to stochastic events alone but may be a true reflection of evolutionary processes underlying the mtDNA and microsatellite sequences.

3.4.3 *Population genetic structure and introduction history of C. afra into Lake Malawi National Park*

Population genetic structure analysis showed Otter Point, Thumbi West Island site B and Domwe populations had distinct genetic structures compared to the introduced population at Thumbi West Island site A and the native populations. This may indicate an introduction event associated with a genetic bottleneck and drift at these localities. This is consistent with the low level of genetic variation exhibited by these populations. Microsatellite analysis confirms that most of the founder events may be associated with the effects of random genetic drift and loss of genetic variation (Hänfling 2007; Roman & Darling 2007).

The use of approximate Bayesian computation has made it possible to successfully reconstruct the introduction history of *C. afra* in Lake Malawi National Park. The results from this study support an independent introduction pattern of *C. afra* at the two sites of Thumbi West Island (A=southern side and B=northern side) and a stepping stone pattern for Domwe Island. These results are consistent with the mtDNA haplotype diversity, which showed differences between the populations of Thumbi West Island site A and B (Zidana *et al.* 2009). Previous authors have only speculated on the sources and initial point of this introduction event of *C. afra* populations in Lake Malawi National Park (Ribbink *et al.* 1983; Stauffer *et al.* 1996; Streelman *et al.* 2004; Konnings 2008). The approximate Bayesian computation approach has been used successfully in biological invasion studies to investigate sources and routes of colonisation in several taxa: the toad *Bufo marinus* (Estoup *et al*. 2001);

the bird *Zosterops lateralis* (Estoup & Clegg 2003), corn rootworm *Diabrotica virgifera virgifera* (Miller *et al.* 2005); and fruit flies *Drosophila melanogaster* (Thornton & Andolfatto 2006) and *D. subobscura* (Pascaul *et al.* 2008).

The introduction of the *C. afra* population at Domwe Island from Thumbi West Island site A was notable considering that geographically Domwe Island is much closer to Thumbi West Island site B. This may indicate human-mediated translocation from Thumbi West Island site A to Domwe Island. Unlike many rock-dwelling cichlids in Lake Malawi which are confined to rock substrate, *C. afra* usually forms school in the water column for plankton feeding (Fryer 1959; Konnings 2008). This may expose them to the nets of illegal fishing crews in the National Park that target zooplankton-feeding fish that form schools from the *Copadichromis* and *Mchenga* genera locally known as "utaka", and the *Engraulicypris sardella* species, known locally as "Usipa".

3.5 **Summary**

Despite being used extensively in studies of conservation and evolutionary genetics, the biological and evolutionary processes affecting microsatellite and mtDNA sequences have notable differences. These differences have profound implications on the use and interpretation of data generated from these two markers.

In this study, I have used a well documented introduction event of a zooplankton feeding cichlid, *C. afra*, in Lake Malawi National Park, to investigate the suitability of microsatellite and mtDNA markers in studies related to levels of genetic diversity after a founder event. I genotyped six microsatellite loci from 600 individuals in six native and four introduced *C. afra* populations within Lake Malawi. The comparison has been made with the mtDNA control region by reanalysing data generated from the same populations used in this study.

In contrast to mtDNA, microsatellites exhibited reduced genetic variation in the introduced range. There was no significant genetic differentiation among groups of native-native, native-introduced and introduced-introduced at microsatellite loci, while there is a significant difference at the mtDNA level. When combined, these results suggest a founder event associated with a genetic bottleneck and drift. It also indicated the suitability of microsatellite markers in conservation genetic studies to identify the loss of genetic diversity in populations with complex demographic history like multiple founders from genetically

distinct sources. At the same time this study confirms the suitability of an mtDNA marker when investigating questions related to differences at the population level such as phylogenetics. This study has also reconstructed the introduction history of *C. afra* into Lake Malawi National Park, by supporting with high probability an independent introduction event at Thumbi West Island and stepping stone introduction event at Domwe Island.

Chapter 4

4 Introgressive hybridisation between *Cynotilapia afra* and *Pseudotropheus zebra* in Lake Malawi National Park

Introgressive hybridisation between *C. afra* and *P. zebra* (manuscript in preparation): Zidana H *et al.* (2009) Introgressive hybridisation between introduced *Cynotilapia afra* and native *Pseudotropheus zebra* in Lake Malawi National Park.

4.1 Introduction

Introductions of new species provide unplanned and replicated experiments that can be used to better understand fundamental processes of ecology and evolution. The introduction of organisms by humans has increased the rate of species hybridisation in the world (Allendorf 2001; Sakai *et al.* 2001). Hybridisation as a process in evolution and ecology plays two significant roles: as a source of new variation in adaptive evolution (Abbott 1992; Mooney & Cleland 2001; Ellstrand & Schierenbeck 2000; Lee 2002), and as a threat to species integrity which may sometimes be associated with extinctions (Huxel 1999). Incoporation of non-native genetic material into the genome may increase the risk of extinction of the native population through outbreeding depression (Edmands 1999; Montalvo and Ellstrand 2001; Marr *et al.* 2002; Gilk *et al.* 2004).

In the plant Kingdom, hybridisation is suggested to have contributed to the origin of numerous species (Abbott 1992). The first or second generations of hybrids in many plants are fertile, allowing a backcross of gene flow with parents a phenomenon called "introgression" (Rieseberg & Wendel 1993; Arnold 1997). The offspring from introgressed parents posses a complex gene pool depending on the pathway of introgression (Abbott 1992;

Beaumont *et al.* 2001; Roques *et al.* 2001). Within such hybrid populations, phenotypes that are extreme or overstated when compared to parental lines may be produced, a mechanism called 'transgressive segregation' (Rieseberg 1999; Seehausen 2004). Previous studies have suggested several genetic factors that could be responsible for creation of transgressive phenotypes, but expression of rare recessive alleles (Rick & Smith 1953; Ellstrand & Schierenbeck 2000; Lee 2002) and complementary gene action (Vega & Frey 1980; de Vicente &Tanksley 1993; Ellstrand & Schierenbeck 2000; Lee 2002) stand out as the primary explanations. This could be a crucial mechanism in adaptive evolution and divergence of populations.

Hybridisation has been suggested as one of the factors to explain the genetic paradox in invasion biology, namely, how do newly founded populations overcome low genetic diversity and expected low evolutionary potential, typically associated with extinction risk and become invasive? (Ellstrand & Schierenbeck 2000; Frankham 2005 and reviewed in Roman & Darling 2007). The genetic admixture from genetically differentiated multiple sources is suggested to be among the possible solutions (Kolbe *et al*. 2004). Other explanations that hybridisation provides are heterosis resulting from fixed heterozygosity in hybrid genotypes (Thompson 1991), increased genetic variation providing more opportunities for natural selection to bring about adaptive evolutionary changes (Stebbins 1969; Neuffer *et al*. 1999; Ellstrand & Schierenbeck 2000; Mooney & Cleland 2001; Lee 2002; Kolbe *et al.* 2004) and dumping of genetic load through recombination which creates genotypes with reduced deleterious mutations (Ellstrand & Schierenbeck 2000; Mack *et al.* 2000; Parker *et al.* 2003; Allendorf & Luikart 2007).

Hybridisation has been reported within many animal lineages, usually between closely-related taxa (Arnold 1997), sometimes after human-mediated introductions (Goodman *et al.* 1999) and in hybrid zones (Barton & Hewitt 1989). In African cichlid fishes, hybrids have been observed in the wild (Stauffer *et al.* 1996; Seehausen *et al.* 1997; Koblmüller *et al.* 2007) and genetic evidence for continued interspecific gene flow has been documented (Samonte *et al.* 2007). The rapid speciation of African cichlid fishes from the Lakes of Malawi, Tanganyika and Victoria has served as a model system in studies of evolution (see reviews by Kornfield & Smith 2000; Kocher 2004; Seehausen 2006) and hybridisation has been suggested to have played a role in explosive speciation (Salzburger *et al.* 2002; Smith *et al.* 2003; Seehausen 2004; Schliewen & Klee 2004; Schelly *et al.* 2006).

Advances in molecular genetics techniques and tools have provided an opportunity to investigate evolutionary changes in newly founded populations and explore the possibility of genetic admixture between native and introduced populations. Microsatellite genetic markers have been fundamental in this regard and have been used successfully to investigate introgressive hybridisation events from several taxa including fish (Roques *et al.* 2001; Rüber *et al.* 2001; Pritchard *et al.* 2007), plants (Abbott 1992; Rieseberg & Wendel 1993) and animals (Goodman *et al.* 1999; Beaumont *et al.* 2001). Introgressive hybridisation studies are used to generate knowledge to understand different processes in several different disciplines of science. In medicine, knowledge has been gained on inheritance mechanisms of complex genetic diseases (Chakraborty & Weiss 1988; McKeigue 1997). In biological conservation genetics introgression studies have provided information which can be used to monitor the efficiency of a restocking program (Giuffra *et al.* 1996), and to evaluate the risk of genetic admixture between introduced and native populations or hatchery and wild populations (Cornuet *et al*. 1986; Wayne & Jenks 1991; Gotteli *et al.* 1994; Ellstrand *et al.* 1999). In evolutionary genetics, introgression studies have provided insights on genetic changes in founder populations (Ellstrand & Schierenbeck 2000; Lee 2002).

Introgression studies may be carried out with different aims for example evaluation of the number and duration of admixture events (Roberts 1955; Roberts & Hiorns 1962), to identify putative parental populations (Glass 1955; Haig *et al.* 1997; Nielsen *et al.* 1997; Reed *et al.* 1997) and to investigate different contributions of parental genomes to the admixed population (Roberts & Hiorns 1965; Szathmary & Reed 1978; Parra *et al.* 1998). The last point is most informative for characterization of the current state of a putative introgressive hybridisation event (Chakraborty 1986), and this is one of the aims of this study. Introgressive hybridisation has been mainly studied in closely related taxa: (Beaumont *et al.* 2001; Roques *et al.* 2001; Rüber *et al.* 2001; Pritchard *et al.* 2007), but it is less common to study the process among more distantly related taxa (Greenfield & Deckert 1973).

In this study I investigated evidence of a putative introgressive hybridisation event between two cichlid genera, *Cynotilapia* and *Pseudotropheus*, during a secondary contact event after a human-mediated introduction in the southern part of Lake Malawi. The populations of *Cynotipia afra* and *Pseudotropheus zebra* occur sympatrically in northern and central parts of Lake Malawi (Kassam *et al.* 2005) and no evidence of hybridisation has been recorded in their natural distribution range. In 1960s, *C. afra* was introduced into the southern

part of the lake probably by an aquarium trader, an area which includes Lake Malawi National Park (Ribbink *et al.* 1983; Konings 2008). Since the introduction, the first evidence of hybridisation between *C. afra* and *P. zebra* was reported in 1996 based on morphology (Stauffer *et al.* 1996), and later in 2004 based on morphology and analysis of four microsatellite loci (Streelman *et al.* 2004). Survey on rock dwelling cichlids in Lake Malawi suggested *C. afra* to be restricted to the island of introduction, Thumbi West Island (Ribbink *et al.* 1983; Stauffer *et al.* 1996; Streelman *et al.* 2004), but since then *C. afra* has also been reported on the nearby rocky habitats of Otter Point and Domwe Island (Zidana *et al.* 2009; Young *et al.* 2009).

The availability of *C. afra* in nearby rocky areas provided an opportunity of a replicated natural experiment whereby a rare event in animals, introgressive hybridisation between two taxa can be studied. The hypotheses, that occurrence of relatively fewer individuals of a given species in an area with large abundance of a second one should increase the probability of introgression towards the least abundant species was tested (Payne & Ni 1982; Ni & Sandeman 1984; Arnold *et al.* 1993). Following results from previous studies, I conducted a wide scale sampling in all sites known to be inhabited by introduced *C. afra* and native *P. zebra*, Thumbi West Island, Otter Point and Domwe Island. In this study microsatellite genetic markers were used to investigate the population genetics of the two taxa, and analysis was done using Bayesian analytical procedures and Factor Component analysis (FCA) to investigate non-native genetic material within the samples. The results are discussed with regards to evidence of introgressive hybridisation between taxa and selection forces underlying introgressive hybridisation. The results from this study have important implications for evolutionary processes involved during an introgressive hybridisation event, and for conservation and management of the Lake Malawi fish fauna.

4.2 **Materials and methods**

4.2.1 *Sample species and study area*

The Lake Malawi National Park, established in 1980 as the World's first national park targeted principally to conserve freshwater fish, and designated a UNESCO World Heritage site in 1984, covers 9,400ha in the southern part of the lake (Hough 1989). Its waters are

among the most heavily impacted by translocations, with at least 13 taxa of cichlid fishes being introduced to Thumbi West Island alone (Genner *et al.* 2006). *C. afra* was introduced in 1960s in the park by an aquarium trader who was closing his business along the shores of Lake Malawi. The populations of *C. afra* and *P. zebra* occur sympatrically in several localities within their natural distribution ranges: from northernmost part of the Lake, Ikombe to Mbenji Island and Makanjira point to Ntekete in the south (Kassam *et al.* 2005; Konnings 2008). The Nankumba peninsula divides the southern tip of Lake Malawi into parts of (south west arm=Cape Maclear) and (south east arm=Monkey bay) (Appendix IV. 1).

The similarities in morphology and habitat preferences between *C. afra* and *P. zebra* suggest them as possible candidates for hybridisation following a secondary contact event. Habitat preference for both species is clear water rocky habitats, with *C. afra* feeding almost exclusively in open waters on plankton, while *P. zebra* will additonally feed on epilithic algae (Fryer 1959; Ribbink *et al.* 1983; Konings 2008). Both species have vertical black bars against a light blue base colour. The only notable morphological difference between the species is tooth dentition; *P. zebra* has bicuspid teeth in the first tooth row while *C. afra* has unicuspid teeth in the first tooth row (Fryer 1959; Ribbink *et al.* 1983; Konings 2008) (Appendix V.1).

4.2.2 *Sampling*

Sixty individuals of both *C. afra* and *P. zebra* were collected from each of the four known sites inhabited by introduced populations of *C. afra* in Lake Malawi National Park: Domwe Island; Otter Point and Thumbi West Island (southern part = site A). Domwe Island and Otter Point are located 2.5 to 1.5 km away from Thumbi West Island respectively. *C. afra* was collected at ~10m depth from Domwe Island and Otter Point while *P. zebra* was collected at 3-5m depth. At Thumbi West Island both species were collected at 3-5m depth. During sampling period (2006-2008) I did not observe individuals with intermediate dorsal fin colour and tooth dentition morphology as reported in previous studies. Basing on dorsal fin colour and tooth dentition, 72% were intermediates from collected samples of 1996 survey (Stauffer *et al.* 1996), basing on same physical characteristics, 60% were were intermediates in samples of 2004 survey (Streelman *et al.* 2004). Since these previous estimates were based on morphological characteristics alone, they could not estimate the direction of this introgression event. A native *C. afra* population from Mara rocks, which is

hundreds of km away from Nankumba peninsula was also sampled as a positive control for this analysis (Appendix IV.1).

I used a non-invasive sampling method collecting live fish samples using monofilament nets while SCUBA diving. A small tissue sample was taken by fin clipping the posterior (soft-rayed) dorsal fin, which regenerates within a few days. To ensure reliable species identification, all samples were collected from territorial males in full breeding colours and jaw teeth of all specimens were examined using a hand lens, as possession of unicuspid teeth is diagnostic for the genus *Cynotilapia* and bicuspid teeth for genus *Pseudotropheus* (Trewavas 1935; Fryer 1959; Konings 2008). All live specimens were released back into the water and fin clips were preserved in 99% ethanol before molecular analysis in the laboratory.

4.2.3 *DNA extraction and molecular data analysis*

Total genomic DNA was extracted from fin clips using Promega WIZARD genomic DNA purification kit. Samples were genotyped at 6 microsatellite loci using published primers: Ppun32, Ppun21, Ppun7 and Ppun5 (Taylor *et al*. 2002), Pzeb3 (van Oppen *et al*. 1997) and UHN154 (Kellogg *et al*. 1995). The 10 µL polymerase chain reaction (PCR) consisted of 1µL DNA template, primer (1mM), dNTPs (0.1mM), 0.5 Units of *Taq* polymerase, 1X PCR buffer and $MgCl_2$ (1.25mM). The PCR amplification was carried out on a Peltier Thermal Cycler 200 using the following parameters: 3 min 94°C 1 cycle, followed by 94°C for 30 s, annealing temperature ranged from 54 to 62 °C for 30 s and 72 °C for 30 s for 27 cycles, followed by 1 cycle of 10 min at 72 °C. Two loci Ppun32 and Ppun5 were amplified using two stage protocols, thus 4 cycles at an initial high annealing temperature, followed by 29 and 26 cycles respectively at lower annealing temperature. PCR products were genotyped on a Beckman Coulter CEQ^{TM} 8000 machine and alleles were scored manually with the help of Beckman Coulter CEQ^{TM} software. Data was checked for consistency and the presence of null alleles using MICRO-CHECKER version 2.2.3 (van Oosterhout *et al*. 2004).

4.2.4 *Genetic variation within C. afra and P. zebra populations*

Genetic diversity within populations was expressed in mean number of alleles across loci (*A*), observed heterozygosity (H_O) and expected heterozygosity (H_E; Nei 1978), analysed using ARLEQUIN 3.1 (Excoffier *et al*. 2005). In all analyses, significance of variation across sites was tested using general linear model (GLM) procedure as employed in MINITAB version 12.1 (MINTAB Inc.). While, mean difference significance between allopatric populations was tested using 2 sample *t*-test calculated in MINITAB version 12.1 (MINITAB Inc.). The standard errors (SE) were calculated using a resampling jackknife procedure.

4.2.5 *Genetic variation among C. afra and P. zebra populations*

The pairwise genetic differentiation for sympatric populations was quantified as G_{ST} (Nei 1978) using FSTAT version 1.2 (Goudet 1995). To estimate the maximum possible limit of G_{ST} values among population groups (Thumbi West Island, Otter Point and Domwe Island), the genetic distance values were standardised (G'_{ST}) according to Hedrick (2005). The genetic differentiation (G'_{ST}) among allopatric populations of *C. afra* and native *P.zebra* in introduced range were calculated and compared with results from sympatric populations.

4.2.6 *Introgression analysis and admixture estimation between C. afra and P. zebra*

To investigate evidence of introgressive hybridisation between *C. afra* and *P. zebra* the microsatellite data were analysed in NewHybrids 1.1 (Anderson & Thompson 2002). NewHybrids assigns individual genotypes into four categories: pure parental genotypes, F1 generation, F2 generation and backcross. Only individuals assigned with 95% certainly to one of the four groups were used in the analysis (Anderson & Thompson 2002).

To further investigate any evidence of genetic admixture between sympatric populations of *C. afra* and *P. zebra*, I employed a Bayesian cluster algorithm using STRUCTURE v2.3 (Pritchard *et al*. 2000; Hubisz *et al.* 2009). In this analysis the sympatric populations represent pure samples that may contain a genetic admixture of the two parental taxa, while *C. afra* from Mara Rocks was used as a negative control containing a pure allopatric genetic material. STRUCTURE software assigns individuals into predetermined number of K genetic clusters in this case it was set at K=2. Prior to analysis test runs were

employed to estimate appropriate run lengths and model settings by running several simulations of K=2 with different lengths and model settings. The initial runs indicated low levels of population divergence therefore the option to use a model informed by location prior which is appropriate under such conditions (Hubisz *et al.* 2009) was used.

The STRUCTURE software was used further to estimate admixture proportions (Pritchard *et al.* 2000; Beaumont *et al.* 2001) from sympatric populations of *C. afra* and *P. zebra* individuals at each of the four sites. STRUCTURE estimates 95% probability intervals for individual admixture proportions (*q*) of each individual within a sample, where no pure native reference population is available (Beaumont *et al.* 2001; Hansen *et al.* 2001; Pritchard *et al.* 2007). In essence, the model considered *C. afra* and *P. zebra* as pure individuals which were then used to estimate admixture proportions of *P. zebra* genes in *C. afra* genomes. An individual was considered to contain non-native genetic material when its estimated probability was <0.99 following (Pritchard *et al.* 2007). Finally, I used permutation of 1000 iterations, resampling stats excel add-in version 3.2 to test if mean admixture proportions differed at the four locations.

The possibility of genome admixture was further investigated by analysing factors that explain the genotypes of introduced *C. afra* and native *P. zebra* populations in two dimensional space using Factorial Correspondence Analysis (FCA) procedure as employed in GENETIX version 4.5 (Belkhir *et al.* 2002). FCA is an exploratory technique which allows investigation of correspondence between rows (individuals) and columns (alleles) in a two-way table (Axis 1 and 2). The alleles which exhibit the strongest non-random association with groups of individuals contribute most to the axes (Roques *et al.* 2001). This technique is an appropriate method with which to visualize genetic relationships among populations with complex histories (Streelman *et al.* 2004).

4.3 Results

4.3.1 *Genetic variation within C. afra and P. zebra populations*

The populations at Thumbi West site exhibited significantly higher mean (± SE) number of alleles for *C. afra* = 19.83 (± 1.93) and *P. zebra* = 20.50 (± 2.51) than other sites tested, $P < 0.001$ (Fig. 4.1). There were no significant differences in allelic richness between

populations of *P. zebra* from Domwe Island and Otter Point (t = 2.40, P = 0.06). Similar results were observed in mean differences of expected heterozygosity (H_E) (Appendix II.1).

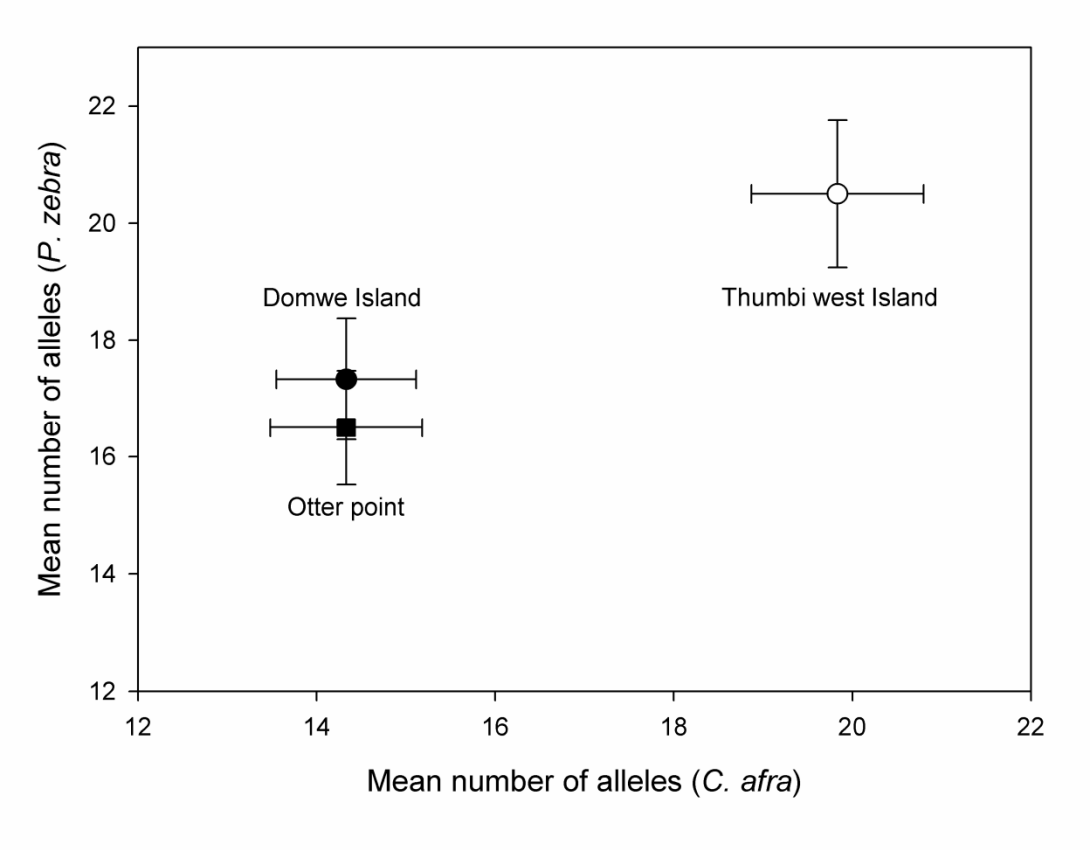

Fig. 4.1. Mean number of alleles (A) ± SE within *C. afra* and *P. zebra* populations, Otter Point = closed box, Thumbi west Island = open circle and Domwe Island = closed circle. Standard errors were calculated by Jackknifing procedure.

4.3.2 *Genetic differentiation among C. afra and P. zebra populations*

The sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island showed significantly lower genetic differentiation across all the sites (G'_{ST} = 0.36 ± 0.05); Otter Point, (G'_{ST} = 0.94 ± 0.18) and Domwe Island (G'_{ST} = 0.55 ± 0.09), P = 0.02 (Fig. 4.2). There were no significant differences in genetic differentiation among sympatric populations at Thumbi West Island with allopatric populations *C. afra* (G'_{ST} = 0.61 ± 0.07; t = 1.13, P = 0.31) and *P. zebra* (G'_{ST} = 0.45 ± 0.07; t = -0.51, P = 0.63) (Fig. 4.2).

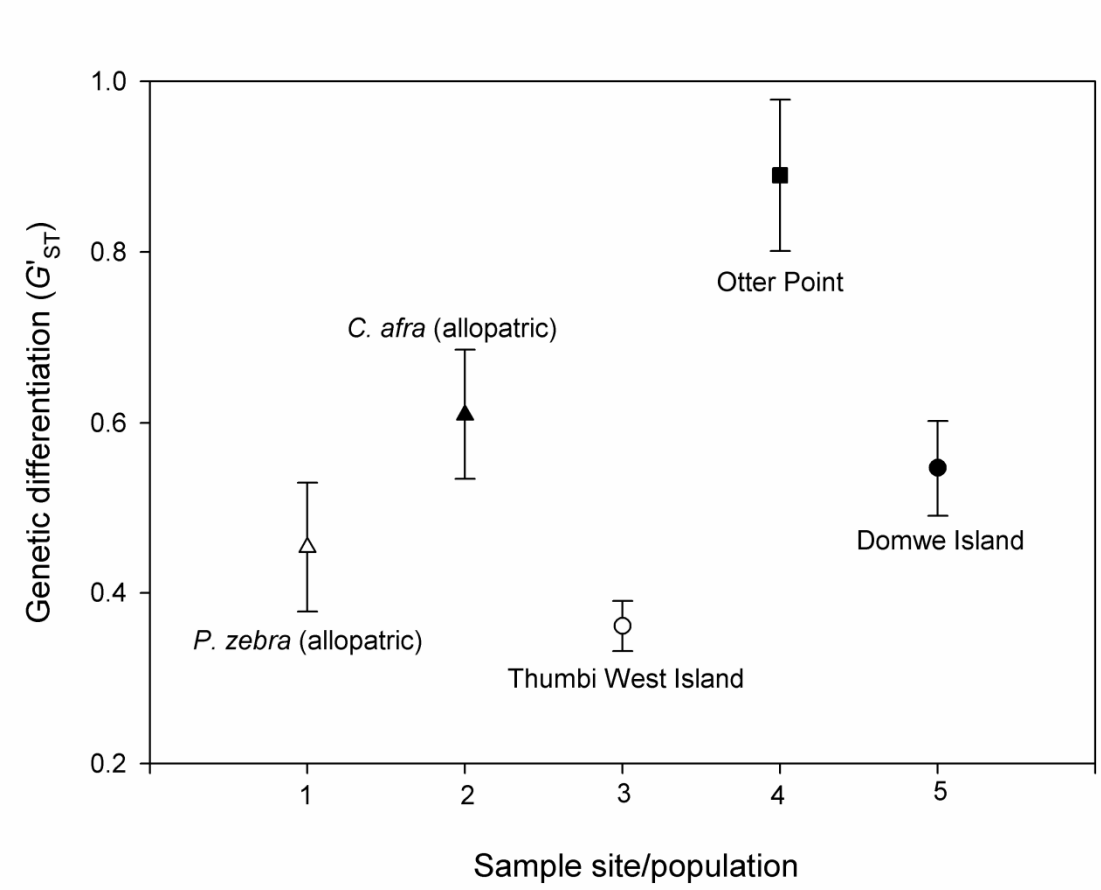

Fig. 4. 2. The microsatellite genetic differentiation standardised (G'_{ST}) ± SE among sympatric *C. afra* and *P. zebra* populations in introduced ranges: Thumbi West Island (open circle); Otter Point (closed box) and Domwe Island (closed circle) in Lake Malawi National Park. Comparison is also made among all allopatric native *P. zebra* (open triangle) and introduced *C. afra* (closed triangle) populations from Thumbi West, Otter Point and Domwe Island. The standard errors were calculated by Jackknifing procedure.

4.3.3 *Introgression analysis and admixture proportions between C. afra and P. zebra*

The analysis of sympatric populations of *C. afra* and *P.zebra* into different hybrid categories revealed all individuals at Thumbi West Island were not of pure breed, they were assigned with < 95% posterior probability. While, Otter Point showed pure *C. afra*=24% and *P. zebra*=43% with > 95% probability. Domwe Island were assigned with *C. afra* = 5% and 28% = *P. zebra* with > 95% (Fig 4.3). In all populations, the hybrid categories F1, F2 and backcross were assigned with < 95% probability (Fig. 4.3).

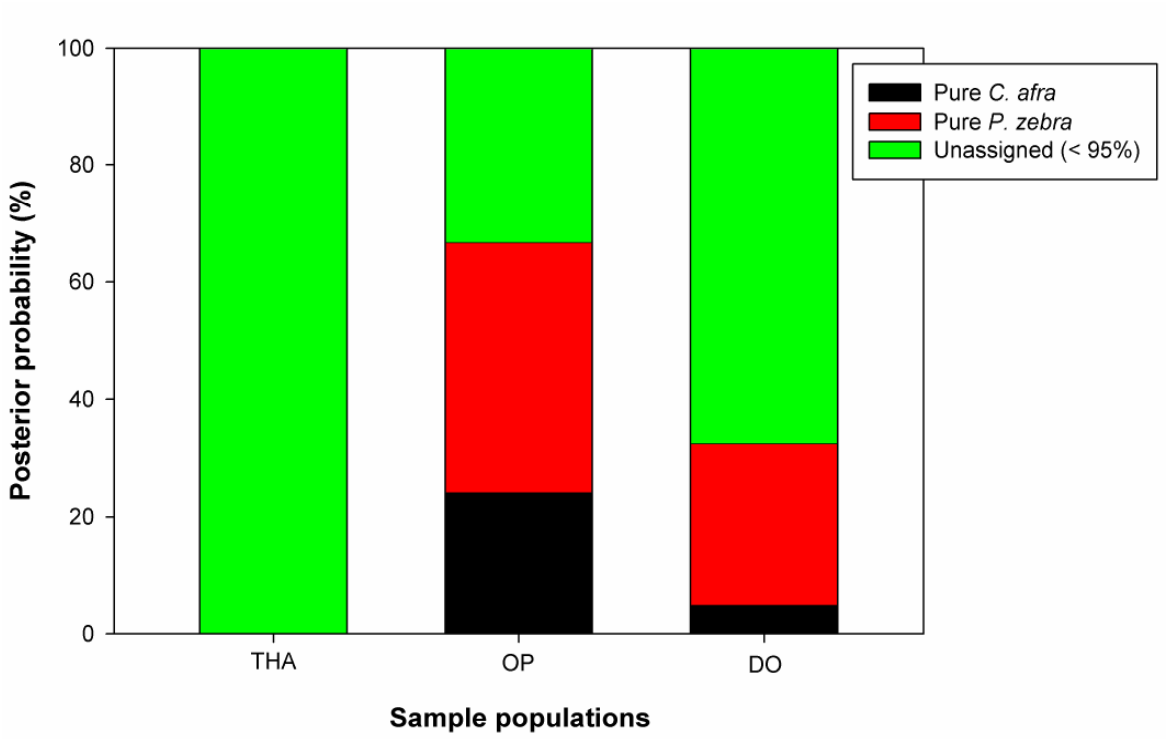

Fig. 4.3. The NewHybrid analysis showing partitions of individuals with (> 95%, assigned) and (< 95%, unassigned) posterior probabilities from the sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island (THA), Otter Point (OP) and Domwe Island (DO).

The comparsion of all models as employed in STRUCTURE showed admixture mode of ancestry with correlated allele frequency was the best fit for microsatellite data set in this study (Fig. 4.4B) and was used for subsequent analysis (Fig. 4.5). Model choice was based on assumption admixture happened when the *C. afra* population was still small. The results showed that over time, the fraction of *P. zebra* genes were homogenised across *C. afra* individuals (Fig. 4.4A).

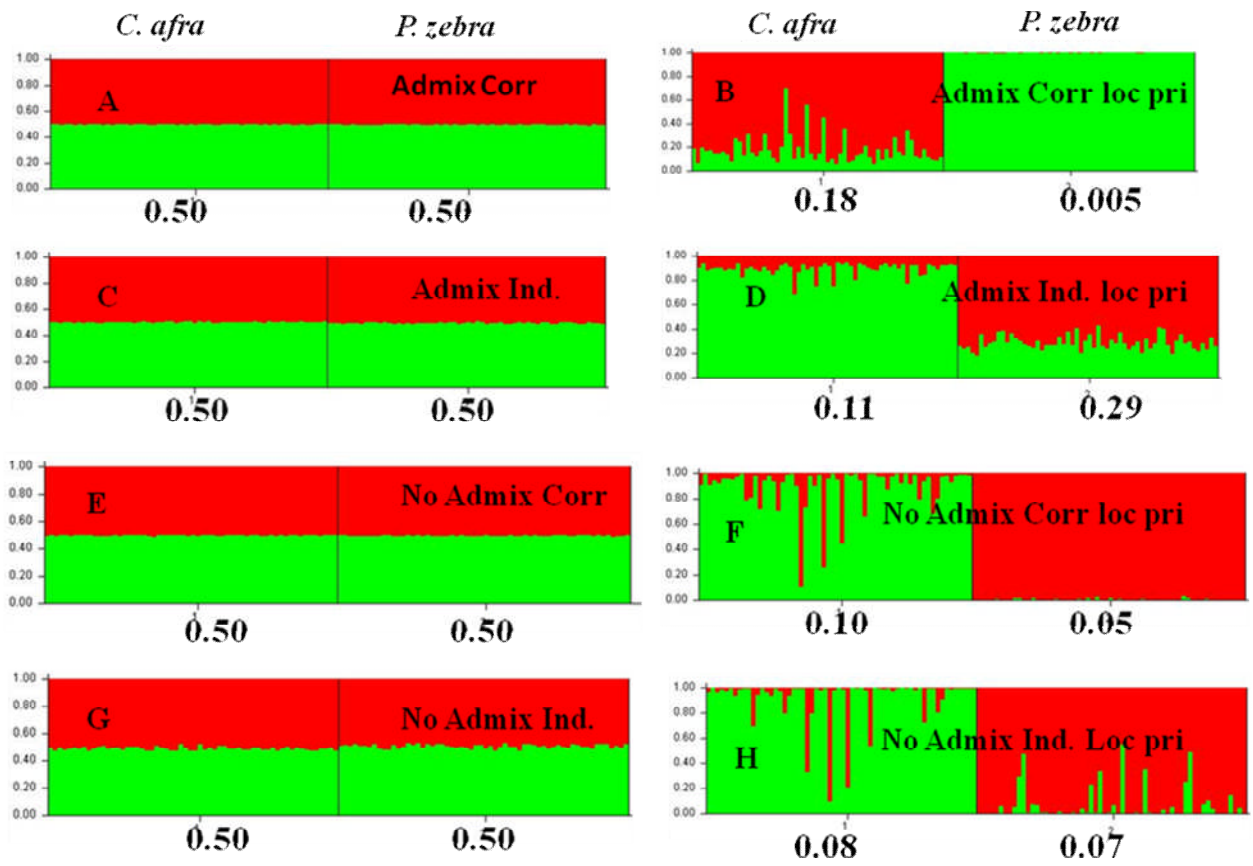

Fig. 4.4 Showing admixture proportions from sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island. Comparioson of all models as employed in STRUCTURE software: Admixture mode of ancestry and correlated allele frequency=Admx Corr; Admixture mode of ancestry and correlated allele frequency with location prior=Amdx corr Loc pri; Admixture mode of ancestry and independent allele frequency=Admx Ind.; Admixture mode of ancestry and independent allele frequency with location prior=Amdx Ind. Loc pri and similar codes for No admixture models. The admixture proportions are indicated at the bottom of each model.

The STRUCTURE analysis showed evidence of introgressive hybridisation between sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island (Fig. 4.5a), but distinct genetic structure at Otter Point and Domwe Island (Fig. 4.5 b & c). As expected the negative control showed no evidence of introgressive hybridisation (Fig. 4.5c).

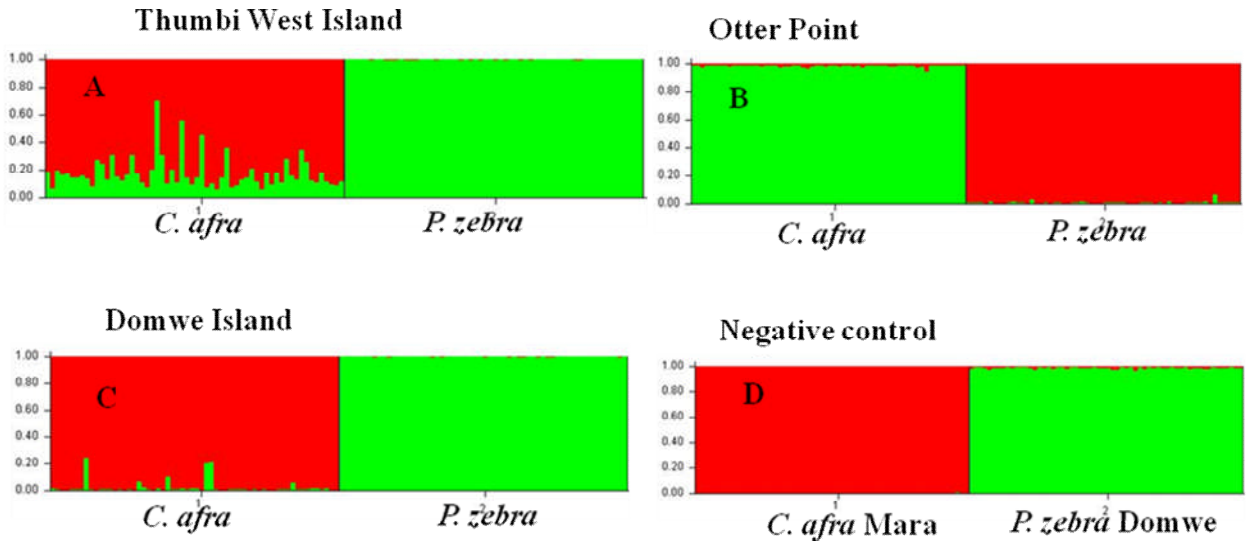

Fig. 4. 5. Population structure of sympatric populations of *C. afra* and *P. zebra* in Lake Malawi National Park, analysed by STRUCTURE software (K=2), using admixture mode of ancenstry and correlated allele frequency with location prior from four localities (a=Thumbi West Island; b=Otter point; c=Domwe Island and d=Negative control). Each individual sample is represented by a thin colour line and each colour represents a different genetic cluster. A thick black line separates *C. afra* and *P. zebra* populations and names of population are listed below each segment.

Analysis of admixture proportions showed significant differences of mean (SE) "*q*" values across all sites (P < 0.001), with highest values from Thumbi West Island (*C. afra* = 0.18 ± 0.02, *P. zera* <0.01 ± < 0.01); then Domwe Island (*C. afra* = 0.05 ± < 0.01, *P. zebra* = 0.01 ± < 0.01) and Otter Point (*C. afra* = 0.02 ± 0.01, *P. zebra* = 0.01 ± < 0.01) (Fig. 4.6). This was compatible with the reduced genetic differentiation of sympatric populations at Thumbi West (Fig. 4.2 & 3). There were no significant differences in admixture proportions between allopatric populations of *C. afra* at Domwe Island and Otter Point (N = 60, t = -1.48, P = 0.14) (Fig. 4.6). The overall mean (SE) admixture levels per site showed Domwe Island = 0.002 ± 0.001 and Otter Point = 0.002 ± 0.001, were significantly lower than Thumbi West island = 0.053 ± 0.012, (P < 0.001).

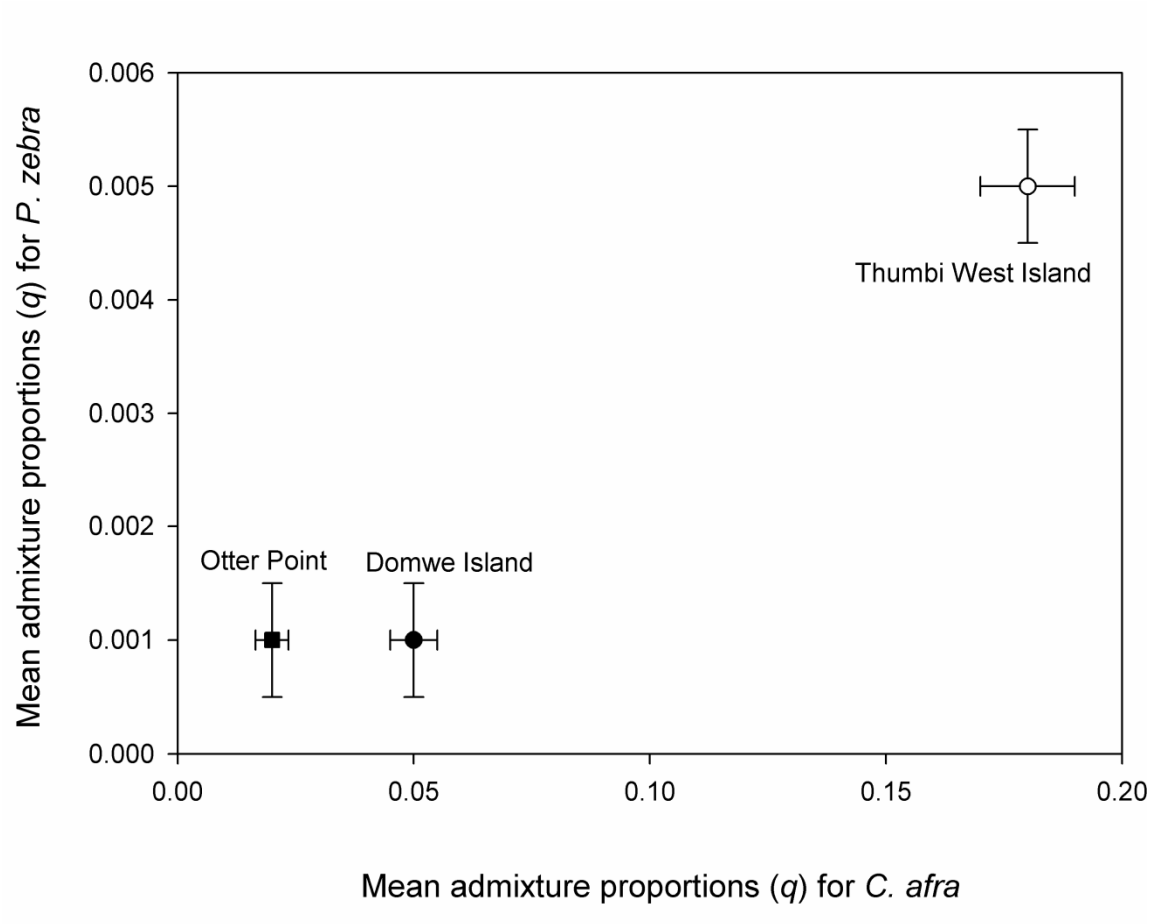

Fig. 4.6. Mean admixture proportions (*q*) in sympatric populations of *C. afra* and *P. zebra* from, Otter Point=closed box, Thumbi West Island=open circle and Domwe Island=closed circle. The standard errors were calculated by Jackknifing procedure.

The factor component analysis (FCA) clustered the admixed genomes of sympatric *C. afra* and *P. zebra* at Thumbi West Island between allopatric populations of (*C. afra*=left side and *P. zebra*=right side) from Otter Point and Domwe Island (Fig. 4.7). The FCA analysis correlates with the results of introgressive hybridisation at Thumbi West Island (Fig. 4.5), and consistence with results of reduced genetic differentiation among sympatric populations at Thumbi West Island as compared to Otter Point and Domwe Island (Fig. 4.2 & 3).

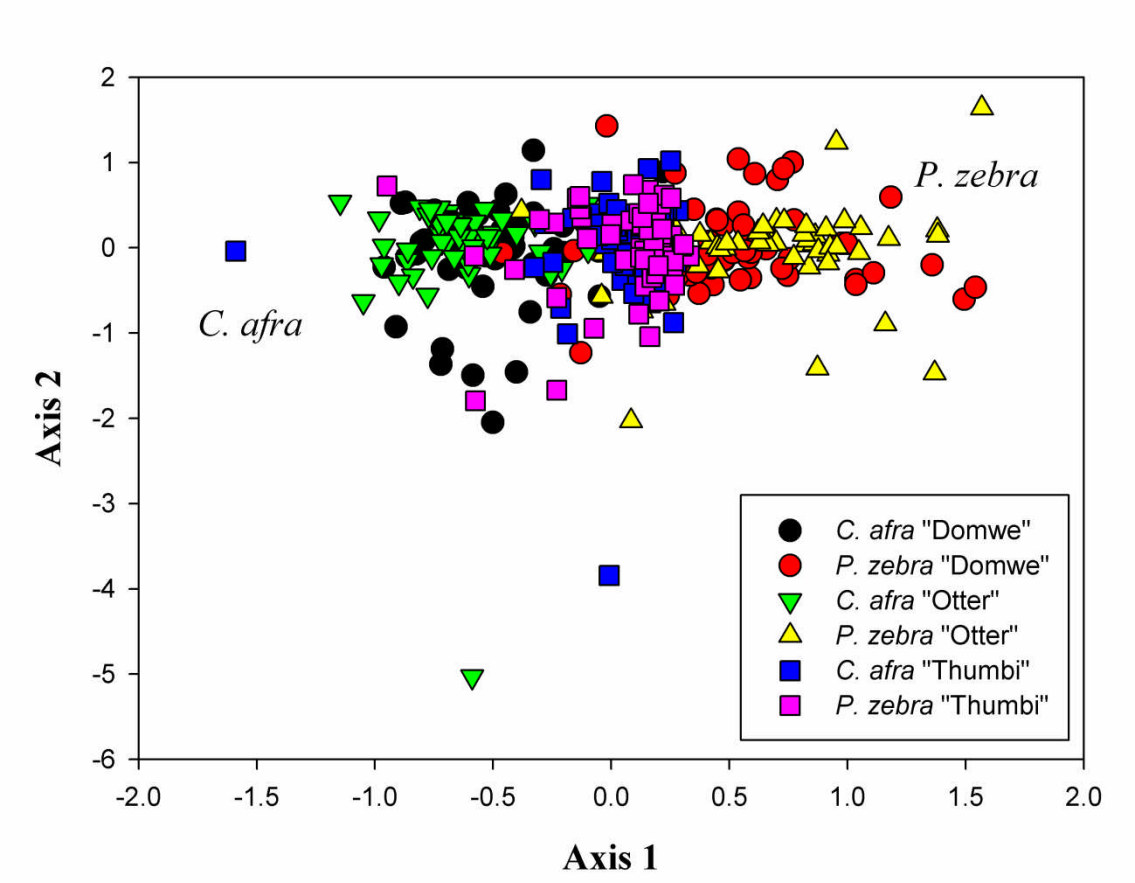

Fig. 4.7. Factorial correspondence analysis (FCA) analysis plot clustering sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island as intermediates of allopatric populations of each taxa from Otter point and Domwe Island.

4.4 **Discussion**

This study presents evidence for a case of introgressive hybridisation between two different genera *C. afra* and *P. zebra* during a secondary contact event, through human induced introduction. This microsatellite analysis from three different localities known to have presence introduced *C. afra*, in combination with analytical approaches allowed a dissection of the dynamics of introgressive hybridisation between these taxa. This study showed *C. afra* have significantly more non-native genetic material as compared to *P. zebra* at all localities and Thumbi West Island was significantly more affected as compared to the other sites tested. In most instances hybridisation will not affect taxon that has coexisted naturally for thousands of generations (Allendorf & Luikart 2007). However, when premating barriers are weak hybridisation is likely to occur (Allendorf & Luikart 2007).

4.4.1 *Genetic variation within populations*

Thumbi West Island populations showed the highest genetic variation across all sampled populations. This may be partly explained by high presence of non-native genetic material at Thumbi West Island, which has increased the genetic variation within *C. afra* and *P. zebra* populations as compared to Otter Point and Domwe Island. The higher genetic variation in sympatry populations as compared to allopatry may be an indication of non-native genetic material because hybridisation increases genetic variability within populations (Roques *et al.* 2001).

4.4.2 *Genetic differentiation among populations*

Sympatric populations of *C. afra* and *P. zebra* showed a significantly lower genetic differentiation at Thumbi West Island than at other locations. Usually rock dwelling cichlids exhibit significant genetic isolation even among conspecific taxa at a fine geographical scale (van Oppen *et al.* 1997 & 1998; Rico & Turner 2002). The significantly low genetic differentiation at Thumbi West Island as compared to Otter Point and Domwe Island, suggests that there is high gene flow between *C. afra* and *P. zebra* at Thumbi West Island. Studies of hybrid zones have shown a general pattern that exchange of genes from one species to another is expected to decrease divergence between them (Roques *et al.* 2001). Therefore, the reduced genetic differentiation between sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island as compared to other locations may be explained by a hybridisation event after a secondary contact.

4.4.3 *Introgression and admixture proportions*

Comparatively, there were significantly high introgression levels at Thumbi West Island as compared to Domwe Island and Otter Point. At all sites, introduced *C. afra* populations exhibited increased levels of non-native genetic material as compared to native *P. zebra*. Therefore, I accept the hypothesis that the occurrence of a few individuals of a given species in an area with large abundance of a second one should increase the probability

of introgression towards the least abundant species which was tested in this study (Payne & Ni 1982; Ni & Sandeman 1984; Arnold *et al.* 1993).

The introgression level of 17% in *C. afra* at Thumbi West Island, is much lower as compared to previous studies working on introduction event of Thumbi West Island populations, in a survey of 1996 they reported 72% intermediates in their sample basing on dorsal fin colour morphology and teeth dentition (Stauffer *et al.* 1996), and basing on the same traits in 2001 survey 60 % were reported as intermediates in *C. afra* alone (Streeman *et al.* 2004). This suggests selection against hybrids may be involved in maintaining the observed genetic pattern of introgression in Lake Malawi National Park. The contribution of colour morphology on reproductive isolation has been suggested in Lake Malawi cichlids (van Oppen *et al.* 1997; Rico *et al.* 2003; Knight &Turner 2004).

In nature, several biotic and abiotic factors have been suggested to facilitate natural hybridisation including: unbalanced population density for mates (Rieseberg *et al*. 1999; Stauffer *et al.* 1996), turbid water (Seehausen *et al*. 1997; Streelman *et al.* 2004), lake level fluctuations leading to secondary contact (Rüber *et al.* 2001), the faunal translocations to new environments (Ellstrand & Schierenbeck 2000; Stauffer *et al.* 1996; Streelman *et al*. 2004) and lack of competitors and predators (Mooney & Cleland 2001; Lee 2002). Two ecological factors have been suggested to facilitate establishment success of introduced *C. afra* at Thumbi West Island: high nutrient availability in introduced range as compared to their native ranges (Munthali & Ribbink 1998; Young *et al.* 2009) and availability of unutilized niches, for example occupying deep water level habitat. Hybridisation with natives has also been suggested to be among the contributing factors for establishement success (Stauffer *et al.* 1996; Streelman *et al.* 2004).

Another factor may be lack of feeding niche differentiation, as is the case between *C. afra* and *P. zebra* at Thumbi West Island as compared to Otter Point and Domwe Island. Both species at Thumbi West Island were collected at 3-5m, while at Otter point and Domwe Island *C. afra* was collected at much deeper water level ~10m during sampling period of this study. The differences may be due to feeding habit preferences *C. afra* feeds on plankton in open waters while *P. zebra* feeds on loose algae sifting small insects and plants on rock substrate (Fryer 1959; Ribbink *et al*. 1983; Konings 2008). The low water level has previously been suggested to facilitate natural hybridisation in aquatic organisms (Greenfield & Deckert 1973; Rüber *et al.* 2001). Depth distributions have been suggested as an important factor in maintaining species diversity among species of the rock-dwelling cichlid community

(Albertson 2008; Genner et al. 2004; Ribbink *et al.* 1983; Sturmbauer *et al.* 2001). Since, the rise and fall of water level in African Lakes is very common (Sturmbauer *et al.* 2001), natural hybridisation among cichlids may be a common feature in areas with low water level, such that it may be among the common evolutionary forces generating cichlid biodiversity in African Lakes.

4.4.4 *Conservation management and ecological implications*

The identification and management of introgressed (admixed) populations within taxa of conservation concern presents a number of challenges. Individuals containing non-native genetic material may be morphologically indistinguishable from those containing an intact native genome (Daniels *et al.* 1998; Weigel *et al.* 2002; Chan *et al.* 2006).

Decisions regarding treatment of introgressed individuals or populations are best made on a case-by-case basis (Allendorf *et al.* 2001, 2004). In North America, in response to problem of introgressive hybridisation of cutthroat trout, populations are divided into management categories according to proportion of non-native genetic material (Pritchard *et al.* 2007). All populations which exhibit less than 1% introgression are in 'Core Conservation' category, while most agencies in other countries include less than 10% in this category (Pritchard *et al.* 2007). If the level is more than 10% they fish are termed "Sportfish" and managed as non-natives. In this study, all native *P. zebra* populations have showed less than 1% non-native genetic material, while *C. afra* from Otter and Domwe showed 2% and Thumbi West was 17%. However, since admixture proportions in Lake Malawi National Park have shown a decreasing pattern over the years, monitoring should be continued before any conclusions are made towards conservation of these populations.

In ecology, theoretical and empirical studies identify niche separation between hybrids and parental genotypes as a single most important factor favouring hybrid establishment, without this hybrids may face competition with the parents (Riesberg 1999). These predictions are supported by reports that most stabilized introgressed or hybrid species are ecologically divergent with respect to their parental species (Abbott 1992; Arnold 1997; Rieseberg 1997). This supports the observation of decreasing pattern of introgression at Thumbi West Island. Alternatively, it may suggests that due to feeding niche differentiation at Domwe and Otter Point, if there is increased propagule pressure it may consequently lead to a well established introgression which may persists for a foreseeable future. This may

create a cichlid conservation management challenge in this park. Therefore, any human-mediated species translocation within Lake Malawi National Park and other areas with similar ecological conditions should be avoided.

4.5 **Conclusion**

Thousands of cichlid species have evolved within the great lakes of Tanganyika, Malawi and Victoria in the last 2 million years (Sturmbauer & Meyer 1993; Turner *et al.* 2001; Salzburger *et al.* 2005; Genner *et al.* 2007). This study has added to the accumulating evidence of introgressive hybridisation in different genera of cichlid fishes: *Sebastes* (Roques *et al.* 2001); *Eretmodini* (Rüber *et al.* 2001); *Neolamprologus* (Salzburger *et al.* 2002); *Lepidiolamprologus* (Schelly *et al.* 2006) and *Lamprologini* (Koblmüller *et al.* 2007).

4.6 **Summary**

Globally, incidences of hybridisation are increasing due to anthropogenic activities such as habitat modification, and the intentional or accidental introduction of new species. Introduction of new species provide unplanned and replicated experiments that can be used to better understand fundamental ecological and evolutionary process in natural world.

Here I have used a well documented human introduction of a zooplanktivorous cichlid *C. afra*into Lake Malawi National Park to investigate if there is any evidence of introgressive hybridisation with a native *P. zebra* population.

I sampled sympatric populations of introduced *C. afra* and native *P. zebra* from three localities within the park (Thumbi West Island, Otter point and Domwe Island), as well as a native population of *C. afra* from Mara Rocks as a control. Six microsatellite markers from 60 individuals from each sampled population were genotyped. Measures of genetic differentiation and a Bayesian cluster analysis were employed to investigate genetic relationships among sympatric and allopatric populations of *C. afra* and *P. zebra* from the three locations.

The results showed that standardised genetic differentiation expressed as G'_{ST} between sympatric populations of *C. afra* and *P. zebra* populations differed significantly (P = 0.02) across locations (Thumbi West Island: G'_{ST}=0.36±0.05, Otter point: G'_{ST}=0.94±0.18

and Domwe Island: G'_{ST}=0.55±0.09). Further analysis showed evidence of introgressive hybridisation between an introduced *C. afra* and native *P. zebra* populations, with significantly high levels being observed at Thumbi West Island as compared to Otter Point and Domwe Island.

This introgressive hybridisation between *C. afra* and *P. zebra* may have been facilitated partly by low water level at Thumbi West Island, that do not allow the two species to feed separately as is the case with Otter Point and Domwe Island locations.

Thousands of cichlid species have evolved within the great lakes of Tanganyika, Malawi and Victoria in the last 2 million years, this study has added to the evidence that natural hybridisation even among related cichlids may have played a role in this explosive speciation event.

The results from this work have profound implications on evolution of cichlid fishes and conservation biology when strategising on management of introgressed populations.

Chapter 5

5.1 **Summary of findings**

5.1.1 *Mitochondrial DNA analysis*

The introduced *C. afra* populations showed higher nucleotide and haplotype diversity than their native counterparts.

The elevated genetic diversity appears to be largely attributed to the fact that introduced *C. afra* populations were derived from several genetically distinct and geographically separate populations, and to a lesser extent, because of introgressive hybridisation with native *P. zebra*.

5.1.2 *Microsatellite analysis in comparison to mtDNA genetic variation*

Microsatellites showed reduced genetic variation in the introduced range, which suggests a founder event associated with a genetic bottleneck and a random loss of microsatellite alleles.

Comparing the demographic signal between microsatellite and mtDNA markers revealed a contrasting pattern of genetic differentiation among groups of native-native (N-N), native-introduced (N-I) and introduced-introduced (I-I) *C. afra* populations: there was no significant genetic differentiation observed for the microsatellite loci, while the mtDNA control region showed high levels of genetic divergence.

The contrasting pattern of results of genetic differentiation among groups genetic suggest that inferences on population or phylogenetic relationships, founder events and genetic bottlenecks would benefit from using multiple genetic markers as these genetic loci uncover distinct demographic and phylogenetic signals.

Reconstruction of the introduction history of *C. afra* into Lake Malawi National Park was supported by an independent introduction event on southern and northern side of Thumbi West Island and a stepping stone introduction event at Domwe Island, with Thumbi West Island southern side *C. afra* population as its source.

5.1.3 *Introgressive hybridisation and admixture proportions analysis*

The sympatric populations of *C. afra* and *P. zebra* at Thumbi West Island showed significantly higher levels of genetic variation than Otter Point and Domwe Island *C. afra* and *P. zebra* populations.

The sympatric populations at Thumbi West Island showed significantly reduced genetic differentiation across sites tested, Thumbi West Island: G'_{ST}=0.36±0.05; Otter point: G'_{ST}=0.94±0.18 and Domwe Island: G'_{ST}=0.55±0.09.

Further analysis showed evidence of introgressive hybridisation between the two distantly related species; *C. afra* and *P. zebra*. Thumbi West Island showed significantly higher levels of admixture proportions than Otter Point and Domwe Island.

Admixture proportion analysis showed significantly higher levels of non-native genetic material in *C. afra* than *P. zebra* in all the sites.

Introgressive hybridisation between *C. afra* and *P. zebra* may have been facilitated partly by lack of conspecifics during the initial period of this founder event. In addition, the shallow water at Thumbi West Island means that both species have to stay in the same area and feed in the same habitat, unlike in their native areas or at the Otter Point and Domwe Island locations, which have similar habitat to the native areas.

5.2 General Discussion

5.2.1 *Implications on evolutionary and conservation genetics*

Cichlid fishes found in the lakes of Africa have served as model systems for the study of evolution, due to thousands of cichlid species that have evolved within these Great Lakes in the last 2 million years (Sturmbauer & Meyer 1993; Turner *et al.* 2001; Salzburger *et al.* 2005; Genner *et al.* 2007). The cichlid fishes are confined to discrete lacustrine environments and their origination is bounded by geological features, these groups provide models with which to study evolution (Kornfield & Smith 2000). There is general consensus that almost all of the endemic cichlid species and genera of Lakes Malawi and Victoria originated within the lakes proper (Moran & Kornfield 1993; Parker & Kornfield 1997; Sturmbauer *et al.* 2001).

The recent and rapid speciation of cichlid fishes in the younger lakes, such as Malawi and Victoria renders the exact reconstruction of species trees and phylogenetic history from molecular data difficult: many rock-dwelling cichlids retain shared ancestral polymorphisms, even where morphological differentiation is clear and reproductive isolation is demonstrated or suspected (Moran & Kornfield 1993; Parker & Kornfield 1997; Sturmbauer *et al.* 2001). Several previous studies have recommended the use of multiple genetic markers when inferring evolutionary processes and population history (Takezak & Nei 1996; Nauta & Weissing 1996; Colson & Goldstein 1999; Ting *et al.* 2000; Selkoe & Toonen 2006; Zachos *et al.* 2009). The use of multiple markers, for example in this thesis microsatellites and the mtDNA control region, has provided an opportunity to investigate evolutionary processes and reconstruct some of the historical events of cichlid evolution.

In invasion biology a long standing issue is: how do introduced populations, whose genetic variation is likely to have been depleted by population bottlenecks, persist and adapt to new conditions (Sakai *et al.* 2001; Frankham & Ballou 2002; Allendorf & Lundquist 2003)? What are the evolutionary and population genetic processes that facilitate establishment of the founder populations, render non-native species invasive, and result in range expansion? Studies in conservation genetics have demonstrated that reduced genetic variation due to genetic bottlenecks and drift limits the ability of a founder population to adapt in a new environment and that the small effective population size increases risk of extinction (Frankham & Ralls 1998; Frankham & Ballou 2002; Allendorf & Lundquist 2003).

Nevertheless, there is accumulating evidence, including the results from this study, which suggests that many introduced species experiencing these similar environmental conditions during initial introductions still persist, expand their ranges, evolve rapidly and even become invasive (Kolbe *et al.* 2004; Lavergne & Molofsky 2007; Dlugosch & Parker 2008; Zidana *et al*. 2009).

Several factors have been suggested to explain mechanisms underlying establishment success of introduced populations: admixture from multiple genetically differentiated sources and hybridisation with natives may elevate genetic variation in founder populations (Ellstrand & Schierenbeck 2000; Lee 2002; Kolbe *et al*. 2004). This mechanism has been demonstrated in this thesis, whereby a founder population composed of individuals from multiple genetically differentiated native *C. afra* populations has resulted in increased genetic diversity levels in introduced range, as demonstrated by mtDNA control region. These results may also help us to understand how genetic variation increases in invasive populations.

Hybridisation of introduced populations with natives may also facilitate establishment success (Thompson 1991; Ellstrand & Schierenbeck 2000; Lee 2002). Hybridisation initially results in hybrid vigour or heterosis due to the fact that recessive deleterious mutations that are fixed in the parental species are in a heterozygous state (Thompson 1991). In subsequent generations, the genotypes recombine and selection can remove part of this mutational load (Ellstrand & Schierenbeck 2000; Mack *et al.* 2000; Parker *et al.* 2003; Allendorf & Luikart 2007). Increased genetic variation through hybridisation may also provide an opportunity for adaptive evolutionary changes as the genetic variation in the hybrids is increased compared to that in the native gene pools (Stebbins 1969; Neuffer *et al*. 1999; Ellstrand & Schierenbeck 2000; Mooney & Cleland 2001; Lee 2002; Kolbe *et al.* 2004).

The investigation of the mtDNA control region in this thesis has showed evidence of hybridisation between introduced *C. afra* populations and native *P. zebra* in introduced range of *C. afra*. In this study, for the first time the mtDNA control region has shown that native *C. afra* populations are genetically differentiated. Admixture of the *C. afra* source populations has increased the mtDNA genetic variation in the introduced range.

The microsatellite loci showed a contrasting pattern of genetic variation when compared to the mtDNA control region. There was no significant genetic differentiation among groups of native vs native; native vs introduced and introduced vs introduced at microsatellite loci as compared to mtDNA which showed significant differentiation. This is consistent with previous studies that have suggested the use of multiple genetic markers when

inferring evolutionary processes and population history (Takezak & Nei 1996; Nauta & Weissing 1996; Colson & Goldstein 1999; Ting *et al.* 2000; Selkoe & Toonen 2006; Zachos *et al.* 2009).

Estimates of population parameters such as structure, number of migrants and effective population size are highly dependent on the mutational model assumed for the molecular markers of choice (van Oppen *et al.* 2000). The inference from a microsatellite marker may be affected by presence of null alleles (van Oppen *et al*. 1997; Markert *et al*. 1999) and size homoplasy (Viard *et al.* 1998; Shaw *et al.* 1999; van Oppen *et al.* 2000; reviewed in Estoup *et al.* 2002). Microsatellite alleles of the same size (electromorphs) can arise from mutational events outside of the repeat or by interrupting perfect repeat producing alleles that are not identical by descent (van Oppen *et al.* 2000; reviewed in Estoup *et al.* 2002). Size homoplasy is expected to increase with time of divergence among populations and taxa (Estoup *et al.* 1995). Most of Lake Malawi cichlids cichlid fishes are endemic and estimated to have originated within the last 700,000 to 2 million years ago, but a study on sequence analysis of alleles from a complex microsatellite locus (Pzeb4) among 11 closely related species, from this lake showed extensive size homoplasy (van Oppen *et al.* 2000). The effect of size homoplasy on the resolution of population structure has been demonstrated in three invertebrate species, the honey bee *Apis mellifera*, the bumble bee *Bombus terrestris*, and the freshwater snail *Bulinus truncates* (Viard *et al.*1998). In Viard *et al.* study size homoplasy was detected mostly among, but not within populations. The low genetic variation at microsatellite loci found in this thesis is consistent with random genetic drift associated with a founder event. Possibly, only a small number of founders was responsible for the introduction event resulting in a loss of microsatellite alleles (Tsutsui *et al.* 2000; Peacock *et al.* 2009; Valade *et al.* 2009). This appears plausible considering that *C. afra* was introduced into Lake Malawi National Park through an aquarium trader (Ribbink *et al.* 1983; Konings 2008), which suggests that only a relatively small number (N≤1000) of fish were released at that point in time.

Conservation genetic studies usually aim to test for low genetic variation in populations to identify species with the potential to become extinct (Allendorf 2001). In this thesis microsatellite markers are found to be to be an important tool in conservation genetic studies of populations with a complex introduction history, for example with multiple genetically differentiated founders. However, the detection of genetic differentiation both in the native and introduced range in the mtDNA control region suggests that mtDNA is a

preferred marker in phylogenetic studies. Lowe *et al.* (2004) reported that 70% of all studies with a phylogeographic component analysed mtDNA loci. The reason for this preference is understandable, given that in this study, with a modest sample size of N=15 individuals per population, mtDNA sequences analysis uncovered a more detailed demographic scenario than the analysis on six microsatellite loci and much larger sample size per population (N=60).

The advances in population genetic data analysis methods, for example the recent development of Approximate Bayesian Computations (ABC) software for microsatellite data analysis are particularly powerful tools for reconstructing the colonisation history of taxa (Estoup *et al*. 2001; Beaumont *et al.* 2002). ABC employs a Bayesian framework for example incorporating prior information to output an "approximate" posterior probability distribution of the tested scenario or model (Beaumont *et al.* 2002). A full Bayesian approach conducts Markov Chain Monte Carlo (MCMC) simulations to obtain exact posterior probability of the raw sample (e.g. allele number and frequency distribution) using thousands of simulated data sets (e.g. genealogies) for population model under consideration (admixture or non admixture ancenstry) (Pritchard *et al.* 2000; Anderson & Thompson 2002). The Bayesian approach to population genetics allows for computations using complex models that could not be achieved using other statistical approaches (Beaumont & Rannala 2004). Bayesian methods have been used successfully to reconstruct colonisation history in populations with complex demographic histories (Estoup *et al*. 2001; Estoup & Clegg 2003; Miller *et al.* 2005; Pascaul *et al.* 2008). An introduced *C. afra* population in Lake Malawi National Park, has multiple founders from genetically differentiatiated populations (Zidana *et al.* 2009). Previous studies have suggested the southern side of Thumbi West Island as an initial point of introduction of *C. afra* into Lake Malawi National Park (Ribbink *et al.* 1983; Stauffer *et al.* 1996), which then dispersed to nearby rocky areas (Streelman *et al.* 2004). Using ABC approach in this thesis, I have successfully reconstructed the competing introduction history (independent or stepping stone) pattern of *C. afra* in Lake Malawi National Park. In contrast to previous suggestions, this study supported two independent introduction scenarios at Thumbi West Island and a stepping stone pattern at Domwe Island, with a population from the southern side of Thumbi West Island as a source population. The results of this study support the importance of microsatellite loci as a useful tool in invasion biology studies and conservation genetics, in particular for the detection of genetic bottlenecks and reconstructing colonisation history at the same time.

Intentional or accidental introduction of organisms and habitat modifications by humans has increased the rate of species hybridisation in the world (Allendorf 2001; Sakai *et al.* 2001). Previous studies hypothesised that occurrence of relatively fewer individuals of a given species in an area with a more abundant (native) species should increase the probability of introgression into the genepool of the rarer species (Payne & Ni 1982; Ni & Sandeman 1984; Arnold *et al.* 1993). However, species that have coexisted naturally for thousands of generations will not hybridize due to reinforced reproductive isolation (Allendorf & Luikart 2007). However, cichlids are suggested to be of recent origin in the African Great Lakes - they have evolved within the last 2 million years (Sturmbauer & Meyer 1993; Turner *et al.* 2001; Salzburger *et al.* 2005; Genner *et al.* 2007), suggesting that there is high potential for hybridisation. The human mediated introduction of *C. afra* into Lake Malawi National Park poses a conservation concern of hybridisation with closely related taxa.

Sympatric populations of *C. afra* and native *P. zebra* at all localities known to have introduced populations of *C. afra* in Lake Malawi National Park; Thumbi West Island, Otter Point and Domwe Island have been compared at six microsatellite loci. This study has found evidence of introgressive hybridisation between both taxa. The populations of *C. afra* showed higher levels of admixture proportions at all the sites than *P. zebra*, accepting the hypothesis that should secondary contact occur, non availability of conspecifics may facilitate hybridisation (Payne & Ni 1982; Ni & Sandeman 1984; Stauffer *et al.* 1983; Arnold *et al.* 1993).

Introgressive hybridisation has been mainly studied in closely related taxa (Beaumont *et al.* 2001; Roques *et al.* 2001; Rüber *et al.* 2001; Pritchard *et al.* 2007), whereas only few studies analysed distantly-related species and species from distinct genera (Greenfield & Deckert 1973). This study demonstrates an event of introgression between two species in the genera *Cynotilapia* and *Pseudotropheus* in Lake Malawi National Park. Nevertheless, this study supports the evidence of introgressive hybridisation in other genera of cichlid fishes: *Sebastes* (Roques *et al.* 2001); *Eretmodini* (Rüber *et al.* 2001); *Neolamprologus* (Salzburger *et al.* 2002); *Lepidiolamprologus* (Schelly *et al.* 2006) and *Lamprologini* (Koblmüller *et al.* 2007). This is inconsistent with strong outbreeding depression, a phenomenon often observed when crossing diverged species, or at least it indicates that despite a loss in fitness associated with hybridisation, introgression does occur in nature.

5.2.2 Implications on conservation management

In 1980, part of Lake Malawi was designated a National Park, due to the importance of the freshwater fish that are found there. The park is under constant threat of fish introductions, and the availability of aquarium trade holding facilities along the shores of the park increases the probabilty of fish escapes. Currently there are at least 13 taxa of cichlid fishes which have been introduced into Lake Malawi National Park (Genner *et al.* 2006).

The results of this study suggest that *C. afra* may not in fact be spreading out from its original introduction site, as initially feared by conservation biologists working with Lake Malawi National Park and Fisheries Department in Malawi. Nevertheless, it is recommended that monitoring of introduced populations in Lake Malawi National Park should be continued, as evolutionary processes are still unfolding. Most of the aquarium fish awaiting transportation to various parts within and outside Malawi have holding facilities, for example, plastic tanks within the lake. Fish escaping such holding facilities is a potential liability, resulting in translocation and introgressive hybridisation, with deleterious conservation genetic implications. Indeed, hybridisation is listed as one of the main reasons for species extinction by the International Union for Conservation of Nature (IUCN) (Frankham *et al.* 2002). Therefore it is recommended that all aquarium fish businesses along the shores of Lake Malawi should construct sufficient fish holding facilities outside of the lake waters, on the shores.

Genes, species and ecosystem are three primary levels of biodiversity recognized by IUCN (Allendorf & Luikart 2007). A fourth level not mentioned by IUCN is the population (Luck *et al.* 2003). For successful biodiversity conservation all four levels must be considered because they are all interrelated and complimentary (Luck *et al.* 2003). Populations may be designated as conservation units (CUs) depending on the associated conservation goals. An evolutionary approach to conservation typically involves maintaining the genetic diversity of a species, or the potential for future genetic divergence or speciation (Allendorf & Luikart 2007). Goals have been formalized using concepts such as: minimum viable populations (MVP) and evolutionarily significant units (ESUs) (Luck *et al.* 2003). MVP size is the minimum size of population that will be viable in the long term, with a probability of extinction of 1% in 1000 years (Shaffer 1981; Nunney & Campbell 1993). The effective population size may be estimated using population genetics analysis, but environmental stochasticity and catastrophes should also be considered. ESUs are populations with

independent evolutionary dynamics (Moritz 1994; Cody *et al.* 1996; Crandall 2000; Luck *et al.* 2003). A classic Mendelian population is a reproductive community of sexually reproducing and cross-fertilizing individuals which share a common gene pool (Dobzhansky 1950). The use of neutral microsatellite loci will generally identify larger ESUs than would be identified using loci under strong, geographically varying selection (Luck *et al.* 2003). The success of Lake Malawi National Park as a conservation tool of freshwater fish depends on well defined conservation management strategies and policies. Under the current Lake Malawi National Park Act, the management plan details 4 conservation zones within the park: special zone, wilderness zone, natural zone and general zone (Glenfell 1993). Any area of the lake that falls within a margin of 2km from the shoreline is a "wilderness" no fishing zone, however, conservation strategies may be more efficient if they are well defined according to biodiversity levels explained above. To achieve this, conservation genetic studies must be undertaken to characterize fish populations in the National Park, and their ecology must be known. Lastly, I would like to say that future conservation genetics research should focus on producing an improved conservation and management strategy for Lake Malawi National Park.

In summary, this thesis was motivated by an urge to facilitate the formulation of conservation and management strategies of freshwater cichlids which are represented by more than 30 species in Lake Malawi National Park. In the context of conservation genetics and use of molecular techniques to investigate parameters that are of much interest to conservation and fisheries resource managers.The results of this thesis has contributed to our understanding of the factors rensponsible for establishment success and spread of introduced populations. The cichlid fish are used as a model organism in studies of evolutionary biology research and this thesis has contributed to this important field, mainly on evolutionaly processes involved during a secondary contact event.

5.2.3 Future research work

Since, *C. afra* and *P. zebra* occur sympatrically in most of their natural distribution range, an interesting avenue for future investigation is to identify any evidence of natural hybridisation in localities with similar shallow water as that found in Thumbi Island and contrast results from with localities with deep waters like Nkhata Bay. A particularly promising site is Chisumulu Island in the northern part of the Lake Malawi. While in this

thesis, I have investigated the possibility of hybridisation between introduced *C. afra* and native *P. zebra*, future work should investigate other potential candidates of hybridisation like another introduced population of *Pseudotropheus callainos.* These investigations will deepen the understanding on the factors and mechanisms responsible for establishment and spread of introduced *C. afra* in Lake Malawi National Park. Furthermore, the high levels of introgression at Thumbi West Island, as compared to other areas such as Otter Point and Domwe Island, warrants further studies to investigate the evolutionary and population genetic processes during and after an introgression event.

Mate choice experiments in the laboratory may be investigated to infer behavioural processes involved during a secondary contact and deduce reinforcement mechanisms and any possibility of trangressive phenotypic traits. Reinforcement is an increase in premating reproductive isolation between taxa resulting from selection against hybrids. The cichlids easily interbreed in the laboratory and this makes them possible candidates for reinforcement studies. Comparison between pure and hybrid fish fitness traits like: body size, fecundity, age at maturity, survival rate and growth rate can be made to determine any selection forces against hybrids.

While, studies on competitive exclusion has been done there is still a lot we can learn from introduced populations in Lake Malawi National Park. Future work should investigate any evidence of evolution of increased competitive ability (EICA) in introduced populations of *C. afra* in Lake Malawi National Park. Studied on diseases and parasites affecting the natives as compared to introduced populations can be carried out to investigate EICA and inferences can also be made using molecular genetic markers like major histocompatibility complex (MHC). The major histocompatibility complex (MHC) is a large genomic region or gene family found in most vertebrates. It is the most gene-dense region of the mammalian genome and plays an important role in the immune system and autoimmunity. The high levels of MHC diversity are an indication of high immune diversity in the population, suggesting that a population might have been previously exposed to a high number of parasites or diseases. If levels of parasites infecting *C. afra* and *P. zebra* in native and introduced ranges were investigated, then their impact on establishment success of introduced *C. afra* in Lake Malawi National Park can be assessed.

References

Abbott RJ (1992) Plant invasion, interspecific hybridisation, and the evolution of new plant taxa. *Trends in Ecology and Evolution,* **7,** 401-405.

Addinsoft (2008) *XLSTAT 2008: Data analysis and statistics with Microsoft Excel*. Paris, France.

Albertson RC, Kocher TD (2005) Genetic architecture sets limits on transgressive segregation in hybrid cichlid fishes. *Evolution,* **59,** 686-690.

Albertson RC (2008) Morphological Divergence Predicts Habitat Partitioning in a Lake Malawi Cichlid Species Complex. *Copeia,* **3,** 690-699.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution,* **16,** 613-622.

Allendorft FW, Lundquist LL (2003) Introduction: population biology, evolution and control of invasive species. *Conservation Biology,* **17,** 24-30.

Allendorf FW, Luikart G (2007) *Conservation and the genetics of populations*. Blackwell Publishing, Oxford, UK.

Amin OM, van Oosterhout C, Blais J *et al.* (2008). On the ecology and host relationships of *Acanthogyrus (Acanthosentis) tilapiae* (Acanthocephala: Quadrigyridae) from Cichlids in Lake Malawi. *Comparative Parasitoogy,* **75,** 278-282.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics,* **160,** 1217-1229.

Angers B, Bernatchez L (1997) Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution,* **14,** 230-238.

Angers B, Estoup A, Jarne P (2000) Microsatellites size homoplasy, SSCP, and population structure: a case study in the freshwater snail *Bulinus truncatus*. *Molecular Biology and Evolution,* **17,** 1926-1932.

Arnegard ME, Markert JA, Danley PD *et al.* (1999) Population structure and colour variation of the cichlid fish *Labeotropheus fuelleborni* Ahl along a recently formed archipelago of rocky habitat patches in southern Lake Malawi. *Proceedings of the Royal Society of London Series B:,* **266,** 119-130.

Arnold ML (1997) *Natural Hybridisation and Evolution*. Oxford University Press, New York.

Atkinson QD, Gray RD, Drummond AJ (2007) mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Molecular Biology and Evolution,* **25,** 468-474.

Avise JC (1986) Mitochondrial DNA and the evolutionary genetics of higher animals and discussion. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences,* **312,** 325-342.

Avise JC (1994) *Molecular Markers, Natural History and Evolution.* Chapman and Hall, New York.

Awadalla P, Eyrewalker A, Smith JM (1999) Linkage disequilibrium and recombination in hominid mitochondrial DNA. *Science,* **286,** 2524-2525.

Azzurro E, Golani D, Bucciarelli G, Bernardi G (2006) Genetics of the early stages of invasion of the Lessepsian rabbitfish *Siganus luridus*. *Journal of Experimental Marine Biology and Ecology,* **333,** 190-201.

Baker HG, Stebbins GL (1965) *The Genetics of Colonizing Species*, Academic Press.

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

Balloux F, Handley LJ, Jombart T *et al.* (2009) Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation. *Proceedings of the Royal Society B,* **276,** 3447-3455.

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution,* **16**, 37-48.

Barluenga M, Stölting KN, Salzburger W *et al.* (2007) Sympatric speciation in Nicaraguan Crater Lake cichlid fish. *Nature,* **439,** 719-723.

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. *Nature,* **341,** 497-503.

Beaumont M, Barratt EM, Gottelli D *et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology,* **10,** 319-336.

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics,* **162,** 2025-2035.

Beaumont MA, Rannala B (2004) The Bayesian revolution in genetics. *Nature Reviews Genetics,* **5,** 251-261.

Belkhir K, Borsa P, Chikhi L *et al.* (1996) *GENETIX, v. 4.0.* Logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Bensch S, Irwin D, Irwin JH *et al.* (2006) Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Molecular Ecology,* **15,** 161-171.

Bentzen P, Leggett WC, Brown GC (1988) Length and restriction site heteroplasmy in the mitochondrial DNA of American shad (*Alosa sapidissima*). *Genetics,* **118,** 509-18.

Besnard G, Henry P, Wille L *et al.* (2007) On the origin of the invasive olives (*Olea europaea* L., Oleaceae). *Heredity,* **99,** 608-619.

Biemont C, Vieira C, Borie N, Lepetit D (1999) Transposable elements and genome evolution: the case of *Drosophila simulans*. *Genetica,* **107,** 113-120.

Birky CW (1995) Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences USA,* **92,** 11331-11338.

Blais J, Rico C, van Oosterhout C *et al.* (2007) MHC adaptive divergence between closely related and sympatric African cichlids. *PLoS ONE,* **2,** e734.

Blais J, Plenderleith M, Rico C *et al.* (2009). Assortative mating among Lake Malawi cichlid fish populations is not simply predictable from male nuptial colour. *BMC Evolutionary Biology,* **9,** e53.

Bowcock AM, Ruiz-Lineares A, Tonfohrde J *et al.* (1994) High resolution of human evolutionary trees with polymorphic microsatellites. *Nature,* **368,** 455-457.

Bray RA, van Oosterhout C, Blais J, Cable J (2006). *Astiotrema turneri* n. sp (Digenea: Plagiorchiidae) from cichlid fishes (Cichlidae: Perciformes) of Lake Malawi, south-eastern Africa . *Zootaxa,* **19,** 43-58.

Bray RA, van Oosterhout C, Blais J, Cable J (2006). *Astiotrema turneri* n. sp (Digenea: Plagiorchiidae) from cichlid fishes (Cichlidae: Perciformes) of Lake Malawi, south-eastern Africa. *Zootaxa,* **19,** 43-58.

Britten RJ, Khone D (1968) Repeated sequences in DNA. *Science,* **161,** 529-540.

Brown WM, George M, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences USA,* **76,** 1967-1971.

Brown JE, Stepien CA (2009) Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. *Molecular Ecology,* **18,** 64-79.

Bryant EH, Meffert LM (1996) Nonadditive genetic structuring of morphometric variation in relation to a population bottleneck. *Heredity*, **77,** 168-176.

Burgess WE (1976) Studies on the family Cichlidae: 7. *Cynotilapia axelrodi*, a new species of mbuna from Lake Malawi (Pisces: Cichlidae). *Tropical Fish Hobbyist,* **25,** 37-44.

Callen DF, Thompson AD, ShenY *et al.* (1993) Incidence and origin of null alleles in the $(AC)_n$ microsatellite markers. *American Journal of Human Genetics,* **52,** 922-927.

Calmet C, Pascal M, Samadi S (2001) Is it worth eradicating the invasive pest *Rattus norvegicus* from Molène archipelago? Genetic structure as a decision-making tool. *Biodiversity and Conservation,* **10,** 911-928.

Canonico GC, Arthington A, Mccrary JK, Thieme M (2005) The effects of introduced tilapias on native biodiversity. *Aquatic Conservation: Marine and Freshwater Ecosystems* **15**, 463-483.

Carroll SP, Dingle H, Famula TR, Fox CW (2001) Genetic architecture of adaptive differentiation in evolving host races of the soapberry bug, *Jadera haematoloma*. *Genetica*, **112,** 257-272.

Castoe TA, de Koning AP, Kim HM *et al.* (2009) From the Cover: evidence for an ancient adaptive episode of convergent molecular evolution. *Proceedings of the National Academy of Sciences of the USA,* **106,** 8986-8991.

Chakraborty R (1986) Gene admixture in human populations: models and predictions. *Yearbook of Physical Anthropology,* **29,** 1-43.

Chakraborty R, Weiss KM (1988) Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proceedings of National Academy of Sciences of the USA,* **85,** 9119-9123.

Chan C, Ballantyne KN, Aikman H *et al.* (2006) Genetic analysis of interspecific hybridisation in the world's only Forbes' parakeet (*Cyanoramphus forbesi*) natural population. *Conservation Genetics,* **7,** 493-506.

Charlesworth B, Bartolome C, Noël V (2005) The detection of shared and ancestral polymorphisms. *Genetical Research,* **86,** 149-157.

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. *Nature Reviews Genetics,* **10,** 783-796.

Cheverud JM, Routman EJ (1996) Epistasis as a source of increased additive genetic variance at population bottlenecks. *Evolution,* **50,** 1042-1051.

Chuong T, Kenneth HA, Cowie RH (2008) Lack of mitochondrial DNA diversity in invasive apple snails (Ampullariidae) in Hawaii. *Malacologia,* **50,** 351-357.

Cody ML, Overton JMC (1996) Short-term evolution of reduced dispersal in island plant populations. *Journal of Ecology*, **84,** 53-61.

Cohen AS, Stone JR, Beuning KRM *et al.* (2007) Ecological consequences of early Late Pleistocene megadroughts in tropical Africa. *Proceedings of the National Academy of Sciences USA,* **104,** 16416-16421.

Colautti RI, Ricciardi A, Grigorovich IA, MacIsaac HJ (2004) Is invasion success explained by the enemy release hypothesis? *Ecology Letters*, **7,** 721-733.

Collins TM, Trexler JC, Nico LG, Rawlings TA (2002) Genetic diversity in a morphologically conservative invasive taxon: multiple introductions of swamp eels to the southeastern United States. *Conservation Biology*, **16,** 1024-1035.

Colson I, Goldstein DB (1999) Evidence for Complex Mutations at Microsatellite Loci in Drosophila. *Genetics,* **152,** 617-627.

Cornuet J-M, Daoudi A, Chevalet C (1986) Genetic pollution and number of mating in a black honey bee (*Apis mellifera mellifera*) population. *Theoretical and Applied Genetics,* **73,** 223-227.

Cornuet JM, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics,* **24,** 2713-2719.

Cox GW (2004) *Alien Species and Evolution*, Island Press, USA.

Crandall K, Bininda-Emonds ORP, Macec GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecolology and Evolution,* **15,** 290-295.

Cristescu MEA, Hebert PDN, Witt JDS *et al.* (2001) An invasion history for *Cercopagis pengoi* based on mitochondrial gene sequences. *Limnology and Oceanography,* **46,** 224-229.

Croft TA (1981) Lake Malawi: A Case Study in Conservation Planning. *Parks,* **6,** 7-11.

Daniels MJ, Balharry D, Hirst D *et al.* (1998) Morphological and pelage characteristics of wildliving cats in Scotland: implications for defining the 'wildcat.' *Journal of Zoology London,* 244, 231-247.

Danley PD, Markert AJ, Arnegard ME, Kocher TD (2000) Divergence with gene flow in the rock-dwelling cichlids of Lake Malawi. *Evolution,* 54, 1725-1737.

de Carvalho DC, de Oliveira DAA, dos Santos JE (2009) Genetic characterization of native and introduced populations of the neotropical cichlid genus *Cichla* in Brazil. *Genetics and Molecular Biology,* **32,** 601-607.

Delvaux D (1995) Age of Lake Malawi (Nyasa) and water level fluctuations. *Royal Museuam for Central Africa-Tervuren, Beligium,* **1995-1996,** 99-108.

Denholm I, Devine GJ, Williamson MS (2002) Insecticide resistance on the move. *Science,* **297,** 2222 -2223.

De Vicente MC, Tanksley SD (1993) QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics,* **134,** 585-596.

Dib C, Faure S, Fizames C *et al.* (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature,* **380,** 152-154.

Dietrich WF, Miller JC, Steen RG *et al.* (1994) A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Genetics,* **7,** 220-245.

Di Rienzo A, Peterson AC, Garza JC *et al.* (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences USA,* **91,** 3166–3170.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology,* **17,** 431-449.

Dominey WJ (1984) Effects of sexual selection and life history on speciation: Species flocks in African cichlids and Hawaiian Drosophila. In: *Evolution of fish species flocks* (eds. Echelle AA & Kornfield I), pp. 231-249. Orono Press, Maine, USA.

Dobzhansky T (1950) Mendelian populations and their evolution. *The American Naturalist,* **84,** 401-418.

Downie DA (2002) Locating the sources of an invasive pest, grape phylloxera, using a mitochondrial DNA gene genealogy. *Molecular Ecology,* **11,** 2013-2026.

Duponchelle F, Paradis E, Ribbink AJ, Turner GF (2008) Parallel life history evolution in mouthbrooding cichlids from the African Great Lake. *Proceedings of the National Academy of Sciences USA,* **105,** 15475-15480.

Durka W, Bossdorf O, Prati D, Auge H (2005) Molecular evidence for multiple introductions of garlic mustard (*Alliaria petiolata*, Brassicaceae) to North America. *Molecular Ecology,* **14,** 1697-1706.

Dyck VA, Hendrichs J, Robinson AS (2005) *Sterile insect technique: principles and practice in area-wide integrated pest management*. Springer, Netherlands.

Eales J, Thorpe RS, Malhotra A (2008) Weak founder effect signal in a recent introduction of Caribbean Anolis. *Molecular Ecology,* **17,** 1416-1426.

Edmands S (1999) Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution,* **53,** 1757-1768.

Ellegren H, Primmer CR, Sheldon BC (1995) Microsatellite 'evolution': directionality or bias? *Genetics*, **11,** 360-362.

Ellstrand NC, Prentice HC, Hancock JF (1999) Gene flow and introgression from domesticated plants into their wild relatives. *Annual Review of Ecology and Systematics,* **30,** 539-563.

Ellstrand NC, Schierenbeck KA (2000) Hybridisation as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences USA,* **97,** 7043-7050.

Elston RC (1971) The estimation of admixture in racial hybrids. *Annals of Human Genetics,* **35,** 9-17.

Elton CS (1958) *The ecology of invasion by animals and plants*. Chapman & Hall, London.

Estoup A, Presa P, Krieg F *et al.* (1993) (CT) and (GT) microsatellites: a new class of genetic markers for *Salmo trutta* L. (brown trout). *Heredity,* **71,** 488-496.

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

Estoup A, Angers B (1998) Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In: *Advances in Molecular Ecology* (ed. Carvalho G), pp. 55-86. NATO press, Amsterdam.

Estoup A, Cornuet JM, Rousset F, Guyomard R (1999) Juxtaposed microsatellite systems as diagnostic markers for admixture: theoretical aspects. *Molecular Biology and Evolution,* **16,** 898-908.

Estoup A, Wilson IJ, Sullivan C *et al.* (2001) Inferring population history from microsatellite and enzyme data in serially introduced Cane Toads, *Bufo marinus*. *Genetics*, **159,** 1671-1687.

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.

Estoup A, Clegg SM (2003) Bayesian inferences on the recent island colonisation history by the bird *Zosterops lateralis lateralis*. *Molecular Ecology,* **12,** 657- 674.

Estoup A, Beaumont M, Sennedot F *et al*. (2004) Genetic analysis of complex demographic scenarios: spatially expanding populations of the cane toad, *Bufo marinus*. *Evolution,* **58,** 2021- 2036.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology,* **14**, 2611-20.

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Research,* **8,** 175–185.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics,* **131,** 479-491.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version. 3.1): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatic Online,* **1**, 47-50.

Ferguson A (1994) Molecular genetics in fisheries: current and future perspective. *Reviews in Fish Biology and Fisheries,* **4,** 379-383.

Ferris SD, Sage RD, Huang CM *et al*. (1983) Flow of mitochondrial DNA across a species boundary. *Proceedings of the National Academy of Sciences, USA,* **80,** 2290-2294.

Filchak KE, Roethele JB, Feder JL (2000) Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature,* **407,** 739-742.

Fonseca DM, LaPointe DA, Fleischer RC (2000) Bottlenecks and multiple introductions: population genetics of the vector of avian malaria in Hawaii. *Molecular Ecology,* **9,** 1803-1814.

Forsyth DM, Duncan RP (2001) Propagule size and the relative success of exotic ungulate and bird introductions to New Zealand. *The American Naturalists,* **157,** 583-595.

Foster GG, Whitten MJ, Prout T, Gill R (1972) Chromosome rearrangements for the control of insect pests. *Science,* **176,** 875-880.

Frankham R, Ralls K (1998) Conservation biology: inbreeding leads to extinction. *Nature,* **392,** 441-442.

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge, United Kingdom.

Frankham R (2005) Invasion biology: resolving the genetic paradox in invasive species. *Heredity,* **94,** 385-385.

Fryer G (1959) The trophic interrelations and ecology of some littoral communities of Lake Nyasa with especial reference to the fishes, and a discussion of the evolution of a group of rock-frequenting Cichlidae. *Proceedings of the Zoological Society of London,* **132,** 153-281.

Fryer G (1997) Biological implications of a suggested Late Pleistocene desiccation of Lake Victoria. *Hydrobiologia,* **354,** 177–182.

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

Galtier N, Nabholz B, Glemin S, Hurst GDD (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology,* **18,** 4541-450.

Garner A, Rachlow JL, Hicks JF (2005) Patterns of genetic diversity and its loss in mammalian populations. *Conservation Biology,* **4,** 1215-1221.

Garza JC, Williamson EG (2001) Detection in reduction of population size using data from microsatellite loci. *Molecular Ecology,* **10,** 305-318.

Gemmell NJ, Allen PJ, Goodman SJ, Reed JZ (1997) Interspecific microsatellite markers for the study of pinniped populations. *Molecular Ecology,* **6,** 661-666.

Genner MJ, Turner GF, Barker S, Hawkins SJ (1999) Niche segregation among Lake Malawi cichlid fishes? Evidence from stable isotope signatures. *Ecology Letters,* **2,**185-190.

Genner MJ, Seehausen O, Cleary DFR *et al.* (2004a) How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? *J. Biogeography,* **31,** 93-102.

Genner MJ, Taylor MI, Cleary DFR *et al.* (2004b) Beta diversity of rock-restricted cichlid fishes from Lake Malawi: importance of environmental and spatial factors. *Ecogeography,* **27,** 601-610.

Genner, M.J. and Turner, G.F. (2005) The mbuna cichlids of Lake Malawi: a model for rapid speciation and adaptive radiation. *Fish and Fisheries* **6**, 1-34.

Genner MJ, Botha A, Turner GF (2006) Translocations of rocky habitat cichlid fishes to Nkhata Bay, Lake Malawi. *Journal of Fish Biology,* **69,** 622-628.

Genner MJ, Seehausen O, Lunt DH *et al.* (2007) Age of Cichlids: New Dates for Ancient Lake Fish Radiations. *Molecular Biology and Evolution,* **24,** 1269-1282.

Genner MJ, Young KA, Haesler M, Joyce DA (2008) Indirect mate choice, direct mate choice, and species recognition in a bower-building cichlid fish lek. *Journal of Evolutionary Biology,* **21,** 1387-1396.

Gilk SE, Wang IA, Hoover CL, Smoker WW, Taylor SG, Gray AK, Gharrett AJ (2004) Outbreeding depression in hybrids between spatially separated pink salmon *Oncorhynchus gorbuscha,* populations: marine survival, homing ability, and variability in family size. *Environmental Biology of Fishes,* **69,** 287-297.

Giuffra E, Guyomard R, Forneris G (1996) Phylogenetic relationships and introgression pattern between incipient parapatric species of Italian brown trout (*Salmo trutta* L. complex). *Molecular Ecology,* **5,** 207-220.

Glass B (1955) On the unlikelihood of significant admixture of genes from the North American Indians in the present composition of the Negroes of the United States. *American Journal of Human Genetics,* **7,** 368-385.

Glenfell S (1993) *Lake Malawi National Park Management and Development Plan.* Department of National Parks and Wildlife, Lilongwe, Malawi.

Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. *Methods in Enzymology,* **295,** 202-222.

Goodman SJ, Barton NH, Swanson G *et al.* (1999) Introgression trough rare hybridisation: a genetic study of a hybrid zone between Red and Sika Deer (Genus *Cervus*) in Argyll, Scotland. *Genetics,* **152,** 355-371.

Goolsby JA, De Barro PJ, Makinson JR *et al.* (2006) Matching the origin of an invasive weed for selection of a herbivore haplotype for a biological control programme. *Molecular Ecology,* **15,** 287-297.

Gotteli D, Sillero-Zubiri C, Applebaum GD *et al.* (1994) Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. *Molecular Ecology,* **3,** 301-312.

Goudet J (1995) FSTAT version 1.2: a computer program to calculate Fstatistics. *Heredity*. **86,** 485-486.

Grant PR, Grant BR (1996) Speciation and hybridisation in island birds. *Philosophical Transactions of the Royal Society B:*, **351,** 765-772.

Griffith B, Scott JM, Carpenter JW, Reed C (1989) Translocation as a species conservation tool: status and strategy. *Science,* **245,** 477-480.

Grimaldi MC, Crouau-Roy B (1997) Microsatellite allelic homoplasy due to variable flanking sequences. *Journal of Molecular Evolution,* **44,** 336-340.

Grossman LI, Wildman DE, Schmidt TR, Goodman M (2004) Accelerated evolution of the electron transport chain in anthropoid primates. *Trends in Genetics,* **20,** 578-585.

Guillemaud T, Beaumont MA, Ciosi M *et al.* (2009) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity,* doi:10.1038/hdy.2009.92.

Günther A (1894) Second report on the reptiles, batrachians, and fishes transmitted by Mr. H. H. Johnston, C. B., from British Central Africa. *Proceedings of the General Meetings for Scientific Business of the Zoological Society of London,* **4,** 616-628.

Gyllensten U, Wharton D, Josefsson A, Wilson AC (1991) Paternal inheritance of mitochondrial DNA in mice. *Nature,* **352,** 255-257.

Haig SM, Gratto-Trevor CL, Mullins TD, Colwell MA (1997) Population identification of western hemisphere shorebirds throughout the annual cycle. *Molecular Ecology,* **6,** 413-428.

Hänfling B, Carvalho GR, Brandl R (2002) mtDNA sequences and possible invasion pathways of the Chinese mitten crab. *Marine Ecology Progress Series,* **238,** 307-310.

Hänfling B (2007) Understanding the establishment success of non-indigenous fishes: lessons from population genetics. *Journal of Fish Biology,* **71,** 115-135.

Hanis CL, Chakraborty R, Ferrell RE, Schull WJ (1986) Individual admixture estimates: disease associations and risk of diabetes and gallbladder disease among Mexican Americans of Starr County, Texas. *American Journal of Physical Anthropology,* **70,** 433-441.

Hansson B, Bensch S, Hasselquist D *et al.* (2000) Increase of genetic variation over time in a recently founded population of great reed warblers (*Acrocephalus arundinaceus*) revealed by microsatellites and DNA fingerprinting. *Molecular Ecology,* **9,** 1529-1538.

Harper GL, Maclean N, Goulson D (2008) Molecular evidence for a recent founder event in the UK populations of the Adonis blue butterfly (*Polyommatus bellargus*). *Journal of Insect Conservation,* **12,** 147-153.

Hedrick P W (2005) A standardized genetic differentiation measure. *Evolution,* **59,** 1633-1638.

Hendry AP, Wenburg JK, Bentzen P *et al.* (2000) Rapid Evolution of Reproductive Isolation in the Wild: Evidence from Introduced Salmon. *Science,* **290,** 516-518.

Hedrick P W (2005) A standardized genetic differentiation measure. *Evolution,* **59,** 1633-1638.

Hert E (1992) Homing and home-site fidelity in rock dwelling cichlids (Pisces: Teleostei) of Lake Malawi, Africa. *Environmental Biology and Fisheries* **33**, 229-237.

Hierro JL, Maron JL, Callaway RM (2005) A biogeographical approach to plant invasions: the importance of studying exotics in their introduced and native range. *Journal of Animal Ecology,* **93,** 5-15.

Hird S, Sullivan J (2009) Assessment of gene flow across a hybrid zone in red-tailed chipmunks (*Tamias ruficaudus*). *Molecular Ecology,* **18,** 3097-3109.

Hough J (1989) *Malawi's National Parks and Game Reserves*. Wildlife Society of Malawi, Blantyre, Malawi.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources,* **9,** 1322-1332.

Huxel GR (1999) Rapid displacement of native species by invasive species: effects of hybridisation. *Biological Conservation,* **89,** 143-152.

Jamandre BW, Durand JD, Tzeng WN (2009) Phylogeography of the flathead mullet *Mugil cephalus* in the north-west Pacific as inferred from the mtDNA control region. *Journal of Fish Biology,* **75,** 393-407.

Jarne P, Lagoda JL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution,* **11,** 424-429.

Johnson TC, Scholz CA, Talbot MR *et al.* (1996) Late Pleistocene desiccation of Lake Victoria and rapid evolution of cichlid fishes. *Science,* **273,** 1091-1093.

Joyce D, Lunt DH, Bills R *et al.* (2005) An extant fish radiation emerged in an extinct Pleistocene Lake. *Nature,* **435,** 90-94.

Jousson O, Pawlowski J, Zaninetti L *et al.* (2000) Invasive alga reaches California. Nature, **408,** 157-158.

Jules J (2009) *Plant Breeding Reviews. 31*. John Wiley & Sons, Inc. USA.

Kassam DD, Seki S, Rusuwa B, Ambali AJD, Yamaoka K (2005) Genetic diversity within the genus *Cynotilapia* and its phylogenetic position among Lake Malawi's mbuna cichlids. *African Journal of Biotechnology,* **4,** 1195-1202.

Kellogg, K.A., Markert, J.A., Stauffer, J.R. and Kocher, T.D. (1995) Microsatellite variation demonstrates multiple paternity in lekking cichlid fishes from Lake Malawi, Africa. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **260**, 79-84.

Kelly JK, Noor MAF (1996) Speciation by Reinforcement: A Model Derived from Studies of Drosophila. *Genetics*, **143,** 1485-1497.

Kim KS, Sappington TW (2006) Molecular genetic variation of boll weevil populations in North America estimated with microsatellites: Implications for patterns of dispersal. *Genetica,* **127,** 143-161.

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

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

Knight ME, Turner GF (2004) Laboratory mating trials indicate incipient speciation by sexual selection among populations of the cichlid fish *Pseudotropheus zebra* from Lake Malawi. *Proceedings of the Royal Society of London B:,* **271,** 675-680.

Koblmüller S, Duftner N, Sefc KM et al. (2007) Reticulate phylogeny of gastropod shell breeding cichlids from Lake Tanganyika the result of repeated introgressive hybridisation. *BMC Evolutionary Biology*, **7,** 7.

Koblmüller S, Sefc KM, Duftner N *et al.* (2007b) Genetic population structure as indirect measure of dispersal ability in a Lake Tanganyika cichlid. *Genetica,* **130,** 121-131.

Kocher TD (2004) Adaptive evolution and explosive speciation: The cichlid fish model. *Nature Reviews Genetics*, **5,** 288-298.

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology and Evolution,* 16, 199-204.

Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science,* **298,** 1233-1236.

Kolbe JJ, Glor RE, Schettino RE *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431,** 177-181.

Kolbe JJ, Glor RE, Schettino LR et al. (2007) Multiple sources, admixture, and genetic variation in introduced Anolis lizard populations. *Conservation Biology,* **21,** 1612-1625.

Konings A, Stauffer JR (2006) Revised diagnosis of *Metriaclima* with description of a new species (Teleostei: Cichlidae) from Lake Malawi National Park, Africa. *Ichthyological Explorations of Freshwaters,* **17,** 233-246.

Konings A (2008) *Malawi Cichlids in their Natural Habitat, 4th Edition*. Cichlid Press, El Paso, Texas.

Kornfield I, Smith P F (2000) African cichlid fishes: Model Systems for Evolutionary Biology. *Annual Reviews Ecology, Evolution and Systematics,* **31,** 163–196.

Kosswig C (1947) Selective Mating as a Factor for Speciation in Cichlid Fish of East African Lakes. *Nature,* **159,** 604-605.

Krieger MJB, Ross KG (2002) Identification of a major gene regulating complex social behavior. *Science,* **295,** 328-332.

Krystufek B, Buzan EV, Hutchinson WF, Hänfling B (2007) Phylogeography of the rare Balkan endemic Martino's vole, *Dinaromys bogdanovi*, reveals strong differentiation within the western Balkan Peninsula. *Molecular Ecology* **16**, 1221-1232.

Lagudah ES, Appels A, BrownnAHD, McNeil D (1991) The molecular-genetic analysis of *Triticum tauschii*, the D-genome donor to hexaploid wheat. Genome, 34, 375-386.

Lande R (1995) Mutation and Conservation. *Conservation Biology,* **9,** 782-791.

Lambrinos JG (2004) How interactions between ecology and evolution influence contemporary invasion dynamics. *Ecology,* **85,** 2061-2070.

Lavergne S, Molofsky J (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences USA,* **104,** 3883-3888.

Lee CE, Bell MA (1999) Causes and consequences of recent freshwater invasions by saltwater animals. *Trends in Ecology and Evolution,* **14,** 284-288.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution,* **17,** 386-391.

Leung B, Drake J, Lodge DM (2004) Predicting Invasions: propagule pressure and the gravity of allele effects. *Ecology,* 85**,** 1651-660.

Levin DA, Francisco-Ortega J, Jansen RK (1996) Hybridisation and the extinction of rate plant species. *Conservation Biology,* **10,** 10-16.

Liem KF (1980) Adaptive signicance of intra-and-interspecic differences in the feeding repertoire of cichlid fishes. *American Zoologist,* **20,** 295-314.

Lindholm AK, Breden F; Alexander HJ *et al.* (2005) Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulate* in Australia. *Molecular Ecology,* **14,** 3671-3682.

Lockwood JL, Cassey P, Blackburn T (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution,* **20,** 223-228.

Lopez S, Rousset F, Shaw FH *et al.* (2009) Joint Effects of Inbreeding and Local Adaptation on the Evolution of Genetic Load after Fragmentation. Conservation Biology, doi 10.1111/j.1523-1739.2009.01326.x.

Lowe A, Harris S, Ashton P (2004) *Ecological genetics: design, analysis, and application*. Blackwell Publishing, Oxford, UK.

Luck GW, Daily GC, Ehrlich PR (2003) Population diversity and ecosystem services. *Trends in Ecology and Evolution,* **18,** 331-336.

Lynch M, Crease TJ (1990) The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution,* **7,** 377-394.

Mack RN, Simberloff D, Lonsdale WM *et al.* (2000) Biotic invasions: causes, epidemiology, global consequences and control. *Ecological Applications*, **10,** 689-710.

Markert JA, Arnegard ME, Danley PD, Kocher TD (1999) Biogeography and population genetics of the Lake Malawi cichlid *Melanochromis auratus*: habitat transience, philopatry and speciation. *Molecular Ecology*, **8**, 1013-1026.

Marr AB, Keller LF, Arcese P (2002) Heterosis and outbreeding depression in descendants of natural immigrants to an inbred population of song sparrows (*Melospiza melodia*). *Evolution,* **56,** 131-142.

Mayr E (1954) Change of genetic environment and evolution. In: *Evolution as a Process*. (ed. Huxley J), pp157-180. Allen & Unwin, London.

McElroy DM, Kornfield I (1993) Novel jaw morphology in hybrids between *Pseudotropheus zebra* and *Labeotropheus fullerborni* (Teleostei: Cichlidae) from Lake Malawi, Africa. *Copeia,* 1993, 933-945.

McKeigue PM (1997) Mapping genes underlying ethnic differences in disease risk by linkage disequilibrium in recently admixed populations. *American Journal of Human Genetics,* **60,** 188-196.

McLean CJ, Workman PL (1973) Genetic studies on hybrid populations. II. Estimation of distribution of ancestry. *Annals of Human Genetics,* **36,** 459-465.

Messing RH, Wright MG (2006) Biological control of invasive species: solution or pollution? *Frontiers in Ecology and the Environment,* 4, 132-140.

Meyer MK, Foerster W (1984) Un nouveau *Pseudotropheus* du lac Malawi avec des remarques sur le complexe *Pseudotropheus-Melanochromis* (Pisces, Perciformes, Cichlidae). *Revue Française d'Aquariologie et de Herpetologie*, **10,** 107-112.

Milinkovitch MC, Orti G, Meyer A (1993) Revised phylogeny of whales suggested by mitochondrial ribosomal DNA sequences. *Nature*, **361,** 346-8.

Miller N, Estoup A, Toepfer S *et al.* (2005) Multiple transatlantic introductions of the western corn rootworm. *Science*, **310,** 992.

Montalvo AM, Ellstrand NC (2001) Nonlocal transplantation and outbreeding depression in the subshrub *Lotus scoparius* (Fabaceae). *American Journal of Botany*, **88,** 258-269.

Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. *Proceedings of the National Academy of Sciences USA,* **98,** 5446-5451.

Moran P, Kornfield I (1993) Retention of an ancestral polymorphism in the mbuna species • flock (Pisces: Cichlidae) of Lake Malawi. *Molecular Biology and Evolution,* **10,** 1015-1029.

Moran P, Kornfield I, Reinthal PN (1994) Molecular systematic and radiation of the haplochromine cichlids (Teleostei, Perciformes) of Lake Malawi. *Copeia,* **2**, 274-288.

Moritz C (1994) Defining 'evolutionarily significant units' for conservation. *Trends in Ecology and Evolution,* **9,** 373-375.

Muirhead JR, Gray DK, Kelly DW *et al.* (2008) Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology,* **17,** 1020-1035.

Müller-Schärer H, Schaffner U, Steinger T (2004) Evolution in invasive plants: implications for biological control. *Trends in Ecology and Evolution,* **19,** 417-422.

Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enymology,* **155,** 335-350.

Munthali SM, Ribbink AJ (1998) Condition and fecundity of translocated rock-dwelling cichlid fish in Lake Malawi. *Journal of the Zoological Society of London* **244**, 347-355.

Myers JHD, Simberloff AM, Kuris AM, Carey JR (2000) Eradication revisited: dealing with exotic species. *Trends in Ecology and Evolution,* **15,** 316-320.

Nauta MJ, Weissing FJ (1996) Constraints on Allele Size at Microsatellite Loci: Implications for Genetic Differentiation. *Genetics,* **143,** 1021-1032.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics,* **89,** 583-590.

Nei M, Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics,* **105,** 207-17.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia Univeristy Press, New York.

Nei M, Jin L (1989) Variances of the average numbers of nucleotide substitutions within and between populations. *Molecular Biology and Evolution,* **6,** 290-300.

Nielsen EE, Hansen MM, Loeschcke V (1997) Analysis of microsatellite DNA from old scale samples of Atlantic salmon Salmo solar: a comparison of genetic composition over 60 years. *Molecular Ecology,* **6,** 487-492.

Noor MA (1995) Reinforcement in speciation. *Trends in Ecology and Evolution,* **10,** 492.

Novak SJ, Mack RN (1993) Genetic variation in *Bromus tectorum* (Poaceae): comparison between native and introduced populations. *Heredity*, **71,** 167-176.

Neuffer B, Auge H, Mesch H (1999) Spread of violets in polluted pine forests: morphological and molecular evidence for the ecological importance of interspecific hybridisation. *Molecular Ecology,* **8,** 365-377.

Nunney L, Campbell KA (1993) Assessing minimum viable population size: demography meets population genetics. *Trends in Ecology and Evolution,* **8,** 234-239.

Ogutu-Ohwayo R (1990) The decline of the native fishes of lakes Victoria and Kyoga (East Africa) and the impact of introduced species, especially the Nile perch, *Lates niloticus*, and the Nile tilapia, *Oreochromis niloticus*. *Environmental Biology of Fishes* **27**, 81-96.

Ohta T, Kimura M (1973) The model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genetics research,* **22,** 201-204.

Owen RB, Crossley R, Johnson TC *et al.* (1990) Major low levels of Lake Malawi and their implications for speciation rates in cichlid fishes. *Proceedings of the Royal Society of London B,* **240,** 519-553.

Paetkau D, Waits LP, Clarkson Pl *et al*. (1997) An Empirical Evaluation of Genetic Distance statistics Using Microsatellite Data from Bear (Ursidae) Populations. *Genetics*, **147,** 1943-1957.

Pakendorf B, Stoneking M (2005) Mitochondrial DNA and human evolution. *Annual Review of Genomics and Human Genetics,* **6,** 165–83.

Pappert RA, Hamrick JL, Donovan LA (2000) Genetic variation in *Pueraria lobata* (Fabaceae), an introduced, clonal, invasive plant of the southeastern United States. *American Journal of Botany,* **87,** 1240-1245.

Park LK, Brainard MA, Dightman DA (1993) Low levels of intraspecific variation in the mitochondrial DNA of chum salmon (*Oncoryhnchus keta*). *Molecular Marine Biology and Biotechnology*, **2,** 362-70.

Park LK, Moran P (1994) Developments in molecular genetic techniques in Fisheries. *Reviews in Fish Biology and Fisheries,* **4,** 272-299.

Parker A, Kornfield I (1994) Polygynandry in *Pseudotropheus zebra,* a cichlid fish from Lake Malawi. *Environmental Biology of fishes,* **47**, 345-352.

Parker A, Kornfield I (1997) Evolution of the mitochondrial DNA control region in the mbuna (Cichlidae) species flock of Lake Malawi, East Africa. *Molecular Evolution,* **45**, 70-83.

Parker IM, Joseph Rodriguez J, Loik ME (2003) An evolutionary approach to understanding the biology of invasions: local adaptation and general-purpose genotypes in the weed *Verbascum thapsus*. *Conservation Biology,* **17,** 59-72.

Parra EJ, Marcini A, Akey J et al. (1998) Estimating African American admixture proportions by use of population-specific alleles. *American Journal of Human Genetics,* **63,** 1839-1851.

Pascual M, Chapuis MP, Mestres MF *et al.* (2008), Introduction history of *Drosophila subobscura* in the New World: a microsatellite-based survey using ABC methods. *Molecular Ecology,* **16,** 3069-3083.

Paterson AH, Schertz KF, Lin YR *et al.* (1995) The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. *Proceedings of the National Academy of Sciences USA***, 92,** 6127-6131.

Peacock MM, Beard KH, O'Neill EM *et al.* (2009) Strong founder effects and low genetic diversity in introduced populations of Coqui frogs. *Molecular Ecology,* **18,** 3603-3615.

Phillipsen IC, Metcalf AE (2009) Phylogeography of a stream-dwelling frog (*Pseudacris cadaverina*) in southern California. *Molecular Phylogenetics and Evolution,* **53,** 152-170.

Pimentel, D. Lach L, Zuniga R, Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. *Bioscience,* **50,** 53-65.

Pitcher TJ, Hart PJB (1995) Species changes and fisheries in African lakes: outline of the issues In: *The impact of species changes in African lakes* (eds. Pitcher TJ & Hart PJB), pp.1-16. Chapman and Hall, London.

Prevosti A, Ribo G, Serra L *et al.* (1988) Colonisation of America by *Drosophila subobscura*: experiment in natural populations that supports the adaptive role of chromosomal-inversion polymorphism. *Proceedings of the National Academy of Sciences USA*, **85,** 5597-5600.

Price TD, Bouvier MM (2002) The evolution of F1 postzygotic incompatibilities in birds. *Evolution,* **56,** 2083-2089.

Primmer CR, Moller AP, Ellegren H (1996) A wide-range survey of cross species microsatellite amplification in birds. *Molecular Ecology*, **5,** 365-378.

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

Pritchard VL, Jones K, Cowley DE (2007) Estimation of introgression in cutthroat trout populations using microsatellites. *Conservation Genetics,* **8,** 1311-1329.

Prugnolle F, de Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity,* **88,** 161-165.

Rassmann K, Schlötterer C, Tautz D (1991) Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, **12,** 113-118.

Raymond M, Rousset F (1995) Population genetics software for exact test and ecumenicism. *Journal of Heredity,* **86,** 248-249.

Reed TE (1969) Caucasian genes in American Negroes. *Science,* **165,** 762-768.

Reed JZ, Tollit DJ, Thompson PM, Amos W (1997) Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. *Molecular Ecology,* **6,** 225-234.

Reinthal, P.N. (1990b) The feeding habits of a group of herbivorous rock-dwelling cichlid fishes (Cichlidae: Perciformes) from Lake Malawi, Africa. *Environmental Biology and Fisheries* **27**, 215-233.

Reinthal P (1993) Evaluating biodiversity and conserving Lake Malawi cichlid fish fauna. *Conservation Biology,* **7,** 712-718.

Reznick, D.N. and Ghalambor, C.K. (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica,* **112,** 183-198.

Rhymer JM, Simberloff D (1996) Extinction by hybridisation and introgression. *Annual Review of Ecology, Evolution and Systematics,* **27,** 83-109.

Ribbink AJ, Marsh BA, Marsh AC *et al.* (1983) A preliminary survey of the cichlid fishes of rocky habitants in Lake Malawi. *South African Journal of Zoology*, **18**, 155-309.

Richards CM (2000) Inbreeding depression and genetic rescue in a plant metapopulation. *American Naturalists,* **155,** 383-394.

Richardson DM, Pysek P, Rejmanek M *et al.* (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions,* **6,** 93-107.

Rick CM, Smith PG (1953) Novel variation in tomato species hybrids. *American Naturalist,* **87,** 359-375.

Rico C, Bouteillon P, van Oppen MJH *et al.* (2003) No evidence for parallel sympatric speciation in cichlid species of the genus *Pseudotropheus* from north western Lake Malawi. *Journal of Evolutionary Biology,* **16,** 37-46.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison RG), pp. 70-109. Oxford University Press, New York.

Rieseberg LH, Van Fossen C, Desrochers AM (1995) Hybrid Speciation accompanied by genomic reorganization in wild sunflowers. *Nature,* **375,** 131-316.

Rieseberg LH, Archer MA, Wayne RK (1999) Transgressive segregation, adaptation and speciation. *Heredity*, **83**, 363-372.

Roberts DF (1955) The dynamics of racial admixture in American Negro: some anthropological considerations. *American Journal of Human Genetics,* **7,** 361-367.

Roberts DF, Hiorns RW (1962) The dynamics of racial admixture. American Journal of *Human Genetics,* **14,** 261-277.

Roberts DF, Hiorns RW (1965) Methods of analysis of the genetic composition of a hybrid population. *Human Biology,* **37,** 38-43.

Robertson BC, Gemmell N (2004) Defining eradication units to control invasive pests. *Journal of Applied Ecology,* **41,** 1042 -1048.

Robinson RL, Ribbink AJ (1998) Abundance and seasonality of temporary cichlid breeding territories on a rocky shore in Lake Malawi. *Ichthyological Exploration of Freshwaters,* **8,** 299-312.

Rollins LA, Woolnough AP, Sherwin WB (2006) Population genetic tools for pest management: a review. *Wildlife Research,* **33,** 251-261.

Rollins LA, Woolnough A, Wilton AN *et al.* (2009) Invasive species can't cover their tracks: using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. *Molecular Ecology*, **18,** 1560 -1573.

Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. *Proceedings of the Royal Society B-Biological Sciences,* **273**, 2453-2459.

Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution,* **22,** 9

Roques S, Sévigny JM, Bernatchez L (2001) Evidence for broadscale introgressive hybridisation between two redfish (genus *Sebastes*) in the North-west Atlantic: a rare marine example. *Molecular Ecology*, **10,** 149-165.

Rowold DJ, Luis JL, Terreros MC, Herrera RJ (2007) Mitochondrial DNA gene flow indicates preferred usage of the Levant Corridor over the Horn of Africa passageway. *Journal of Human Genetics*, **52**, 436–447.

Rozas J, Barrio JCS, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19,** 2496-2497.

Rüber L, Meyer A, Sturmbauer C, Verheyen E (2001) Population structure in two sympatric species of the Lake Tanganyika cichlid tribe Eretmodini: evidence for introgression. *Molecular Ecology*, **10**, 1207-1225.

Rubinsztein DC, Amos W, Leggo W *et al.* (1995) Microsatellite evolution-evidence for directionality and variation in rate between species. *Nature Genetics*, **10,** 337-343.

Saiki RK, Gelfand DH, Stofel S *et al.* (1988) Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239,** 487-490.

Sakai AK, Allendorf FW, Holt JS et al. (2001) The population biology of invasive species. *Annual Review of Ecology, Evolution, and Systematics,* **32,** 305-32.

Salzburger W, Baric S, Sturmbauer C (2002) Speciation via introgressive hybridisation in East African cichlids. *Molecular Ecology,* **11,** 619-125.

Salzburger W, Mack T, Verheyen E, Meyer A (2005a) Out of Tanganyika: genesis, explosive speciation, key-innovations, and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology,* **5,** 1-17.

Salzburger W, Mack T, Verheyen E, Meyer A (2005b) Genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evolutionary Biology,* **5,** 17.

Sambrook J, Fritch EF, Maniatis T (1989) *Molecular cloning.* Cold Spring Habour Laboratory Press, New York.

Samonte IE, Satta Y, Sato A et *al.* (2007) Gene flow between species of Lake Victoria haplochromine fishes. *Molecular Biology and Evolution,* 24**,** 2069-2080.

Schelly R, Salzburger W, Koblmüller S *et al.* (2006) Phylogenetic relationships of the lamprologine cichlid genus *Lepidiolamprologus* (Teleostei: Perciformes) based on mitochondrial and nuclear sequences, suggesting introgressive hybridisation. *Molecular Phylogenet Evolution,* **38,** 426-438.

Schliewen UK, Klee B (2004) Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Frontiers of Zoology,* **1,** 5

Schoener TW, Spiller DA (1987) High population persistence in a system with high turnover. *Nature,* **330,** 474-477.

Scholz CA, Rosendahl B (1988) Low Lake stands in Lakes Malawi and Tanganyika, delineated with multifold seismic data. *Science,* **240**, 1645-1648.

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

Schultz JK, Baker JD, Toonen RJ *et al.* (2009) Extremely Low Genetic Diversity in the Endangered Hawaiian Monk Seal (*Monachus schauinslandi*). *Journal of Heredity,* **100,** 25-33.

Seehausen O, van Alphen JJM (1997) Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecology Letters,* **2,** 262-271.

Seehausen O, van Alphen JJM, Witte F (1997) Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science,* **277**, 1808-1811.

Seehausen O (2004) Hybridisation and adaptive radiation. *Trends in Ecology and Evolution,* **19,** 198-207.

Seehausen, O and Schluter, D. (2004) Male-Male competition and nuptial colour displacement as a diversifying force in Lake Victoria cichlids fish. *Proceedings of the Royal Society of London Series B:,* **271**, 1345-1353.

Seehausen O (2004) Hybridisation and adaptive radiation. *Trends in Ecology and Evolution,* **19,** 198-207.

Seehausen O (2006) African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B-Biological Sciences,* **273**, 1987-1998.

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

Shaffer ML (1981) Minimum population sizes for species conservation. *Bioscience,* **31,** 131-134.

Shafland PL, Lewis WM (1984) Terminology associated with introduced organisms. *Fisheries,* **9,** 17-18.

Shaw PW, Pierce GJ, Boyle PR (1999) Subtle population structuring within a highly vagile marine invertebrate, the veined squid *Loligo forbesi*, demonstrated with microsatellite DNA markers. *Molecular Ecology,* **8,** 407-417.

Shaw PW, Turner GF, Idid MR *et al.* (2000) Genetic population structure indicates sympatric speciation of Lake Malawi pelagic cichlids. *Proceedings of the Royal Society of London* B, **267,** 2273-2280.

Shaw KL (2002) Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: What mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proceedings of the National Academy of Sciences USA,* **99,** 16122-16127.

Shinde D, Lai YL, Sun FZ, Arnheim N (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasilikelihood analysis: (CA/GT)(n) and (A/T)(n) microsatellites. *Nucleic Acids Research*, **31,** 974-980.

Simberloff D (1996) Impacts of Introduced Species in the United States. *Conesquences,* Volume 2, Number 2, 1-13.

Smith PF, Konings A, Kornfield I (2003) Hybrid origin of a cichlid population in Lake Malawi: Implications for genetic variation and species diversity. *Molecular Ecology*, **12**, 2497-2504.

Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences USA,* **97,** 7051-7057.

Stauffer JR, Bowers NJ, Kocher TD, McKaye KR (1996) Evidence of hybridisation between *Cynotilapia afra* and *Pseudotropheus zebra* (Telestoei: Cichlidae) following an intralacrustine translocation in Lake Malawi. *Copeia,* **1,** 203-208.

Stauffer JR, Bowers NJ, Kellogg KA, McKaye KR (1997) A revision of the blue-black *Pseudotropheus zebra* (Teleostei: Cichlidae) complex from Lake Malawi, Africa, with

a description of a new genus and ten new species. *Proceedings of the Academy of Natural Sciences of Philadelphia,* **148,** 189-230.

Stebbins, G. L. 1969. The significance of hybridisation for plant taxonomy and evolution. *Taxon,* **18,** 26-35.

Strauss SY, Lau JA, Carroll SP (2006) Evolutionary responses of natives to introduced species: what do introductions tell us about natural communities? *Ecology Letters,* **9,** 357-374.

Streelman JT, Gmyrek SL, Kidd MR *et al.* (2004) Hybridisation and contemporary evolution in an introduced cichlid fish from Lake Malawi National Park. *Molecular Ecology*, **13**, 2471-2479.

Strayer D (1999) Invasion of fresh waters by saltwater animals. *Trends in Ecology and Evolution,* **14,** 448-449.

Sturmbauer C, Meyer A (1992) Genetic divergence, speciation and morphological stasis in a lineage of African cichlid fishes. *Nature,* **358,** 578-581.

Sturmbauer C, Meyer A (1993) Mitochondrial phylogeny of the endemic mouthbrooding lineages of cichlid fishes from Lake Tanganyika in eastern Africa. *Molecular Biology and Evolution,* **10,** 751-768.

Sturmbauer C, Baric S, Salzburger W *et al.* (2001) Lake level fluctuations synchronize genetic divergences of cichlid fishes in African lakes. *Molecular Biology and Evolution,* **18,** 144-154.

Sugawara T, Terai Y, Imai H *et al.* (2005) Parallelism of amino acid changes at the RH1 locus affecting spectral sensitivity among deep-water cichlids from Lakes Tanganyika and Malawi. *Proceedings of the National Academy of Sciences USA*, **102,** 5448-5453.

Sutherland B, Stewart D, Kenchington ER, Zouros E (1998) The fate of paternal mitochondrial DNA in developing female mussels, *Mytilus edulus*: implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA. *Genetics,* **148,** 341-347.

Szathmary EJE, Reed TE (1978) Calculation of the maximum amount of gene admixture in a hybrid population. *American Journal of Physical Anthropology,* **48,** 29-34.

Tajima F, Tokunaga T, Miyashita NT (1994) Statistical method for estimating the effective number of alleles, expected heterozygosity and genetic distance in self-incompatibility locus. *Japanese Journal of Genetics,* **69,** 287-295.

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics,* **144,** 389-399.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution,* **24,** 1596–1599.

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

Taylor MI, Meardon F, Turner GF *et al.* (2002) Characterization of tetranucleotide microsatellite loci in a Lake Victorian, haplochromine cichlid fish: a *Pundamilia pundamilia* × *Pundamilia nyererei* hybrid. *Molecular Ecology Notes,* **2**, 443-445.

Tautz D (1989) Hypervariability of simple sequence as a general source for polymorphic DNA markers. *Nucleic Acids Research,* **16,** 6463-6471.

Tegelstrom H (1987) Transfer of mitochondrial DNA from the northern red-backed vole (*Clethrionomys rutilus*) to the bank vole (*Clethrionomys glareolus*). *Journal of Molecular Evolution,* **24,** 218-227.

Templeton AR (1980) The theory of speciation via the founder principle. *Genetics,* **94,** 1011-1038.

Terai Y, Seehausen O, Sasaki T *et al.* (2006) Divergent selection on opsins drives incipient speciation in Lake Victoria Cichlids. *PLoS Biology*, **4,** e433.

Teixeira S, Bernasconi G (2007) High prevalence of multiple paternity within fruits in natural populations of *Silene latifolia*, as revealed by microsatellite DNA analysis. *Molecular Ecology*, **16,** 4370 -4379.

Thomson JD (1991) The biology of an invasive plant: what makes *Spartina* so successful? *Bioscience,* **41,** 393-401.

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

Thornton K, Andolfatto P (2006) Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. *Genetics,* **172,** 1607-1619.

Ting CT, Tsaur SC, Wu CI (2000) The phylogeny of closely related species as revealed by the genealogy of a speciation gene, Odysseus. *Proceedings of the National Academy of Sciences USA*, **97,** 313-5316.

Trendall J (1988) The distribution and dispersal of introduced fish at Thumbi West Island in Lake Malawi, Africa. *Journal of Fish Biology*, **33,** 357-369.

Trewavas E (1935) A synopsis of the cichlid fishes of Lake Nyasa. *Annals and Magazine of Natural History*, **10,** 65.

Trewavas E (1984) Nouvel examen des genres et sous-genres du complexe *Pseudotropheus-Melanochromis* du lac Malawi (Pisces, Perciformes, Cichlidae). *Revue Française d'Aquariologie et de Herpetologie*, **10,** 97-106.

Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences USA*, **97,** 5948-5953.

Turner GF (1995) Management, conservation and species changes of exploited fish stocks in Lake Malawi. In: *The impact of species changes in African lakes* (eds. Pitcher TJ & Hart PJB), pp. 365-395. Chapman and Hall, London.

Turner, G.F. (1996) *Offshore Cichlids of Lake Malawi*, Cichlid Press. Ahornweg, Germany.

Valade R, Kenis M, Hernandez-Lopez A (2009) Mitochondrial and microsatellite DNA markers reveal a Balkan origin for the highly invasive horse-chestnut leaf miner *Cameraria ohridella* (Lepidoptera, Gracillariidae). *Molecular Ecology*, **18,** 3458-3470.

Valbuena-Carabaña M, González-Martínez SC, Sork VL et al. (2005) Gene flow and hybridisation in a mixed oak forest (*Quercus pyrenaica* Willd. and *Quercus petraea* (Matts.) Liebl.) in central Spain. *Heredity*, **95,** 457-465.

Valdes MA, Slatkin M, Freimert NB (1993) Allele frequencies at microsatellite loci: The stepwise mutation model revisited. *Genetics*, **133,** 737-749.

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.

van Oppen MJH, Turner GF, Rico C *et al.* (1997) Unusually fine-scale genetic structuring found in rapidly speciating Malawi cichlid fishes. *Proceedings of the Royal Society of London B*, **264**, 1803-1812.

van Oppen MJH, Rico C, Deutsch JC *et al.* (1997b) Isolation and characterisation of microsatellite loci in the cichlid fish *Pseudotropheus zebra*. *Molecular Ecology*, **6,** 387-388.

van Oppen MJH, Turner GF, Rico C *et al*. (1998) Assortative mating among rock-dwelling cichlid fishes supports high estimates of species richness from Lake Malawi. *Molecular Ecology,* **7,** 991-1001.

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

Vasquez EA, Glenn EP, Brown JJ *et al.* (2005) Salt tolerance underlies the cryptic invasion of North American salt marshes by an introduced haplotype of the common reed *Phragmites australis* (Poaceae). *Marine Ecology Progress Series,* **298** 1-8.

Veen T, Borge T, Griffith SC *et al.* (2001) Hybridisation and adaptive mate choice in flycatchers. *Nature,* **411,** 45-50.

Vega U, Frey KJ (1980) Transgressive segregation in inter-and intraspecific crosses of barley. *Euphytica,* **29,** 585-694.

Vellend M, Harmon LJ, Julie L *et al.* (2007) Effects of exotic species on evolutionary diversification. *Trends in Ecology and Evolution,* **22,** 481-488.

Verheyen E, Salzburger W, Snoeks J, Meyer A (2003) Origin of the superflock of cichlid fishes from Lake Victoria, East Africa. *Science,* **300,** 325 - 329.

Viard F, Franck P, Dubois M-P *et al.* (1998) Variation of microsatellite size homoplasy across electromorphs, loci and populations in three invertebrate species. *Journal of Molecular Evolution,* **47,** 42-51.

Vigouroux Y, Matsuoka Y, Doebley J (2003) Directional Evolution for Microsatellite Size in Maize. *Molecular Biology and Evolution,* **20,** 1480–1483.

Vitousek PM, Walker LR (1989) Biological invasions by *Myriaca faya* in Hawaii: plant demography, nitrogen fixation, ecosystem effects. *Ecological Monographs,* **59,** 247-265.

Waldick RC, Kraus SS, Brown M, White BN (2002) Evaluating the effects of historic bottleneck events: an assessment of microsatellite variability in the endangered, North Atlantic right whale. *Molecular Ecology,* **11,** 2241-2250.

Ward RD, Grewe PM (1994) Appraisal of molecular genetic techniques in fisheries. *Reviews in Fish Biology and Fisheries,* **4,** 300-325.

Wattier R, Engel CR, Saumitou-Laprade P, Valero M (1998) Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Molecular Ecology,***7,** 1569-1573.

Wayne RK, Jenks SM (1991) Mitochondrial DNA analysis implying extensive hybridisation of the endangered red wolf, *Canis rufus*. *Nature*, **351,** 565-568.

Weigel DE, Peterson JT, Spruell P (2002) A model using phenotypic characteristics to detect introgressive hybridisation in wild westslope cutthroat trout and rainbow trout. *Transactions of the American Fisheries Society*, 131, 389-403.

Weinig C (2000) Plasticity versus canalization: population differences in the timing of shade-avoidance responses. *Evolution,* **54,** 441-451.

Weinreich DM, Rand DM (2000) Contrasting patterns of nonneutral evolution in proteins encoded in nuclear and mitochondrial genomes. *Genetics*, **156,** 385-399.

Welcome RL (1988) International introductions of Inland aquatic species. *FAO Fisheries Technical Papers,* **294**, 318.

Whittier JM, Limpus D (1996) Reproductive patterns of a biologically invasive species: the brown tree snake *Boiga irregularis* in eastern Australia. *Journal of Zoology,* **238,** 591-597.

Wilson AC, Cann RL, Carr M *et al.* (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of Linnean Society*, **26,** 375-400.

Wilson CC, Bernatchez L (1998) The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Molecular Ecology,* **7,** 127-132.

Witte F, Goudswaard PC (1985) Aspects of the haplochromine fishery in Southern lake Victoria. *FAO Fisheries Reports,* **335**, 81-8.

Wintero AK, Fredholm M, Thomsen PD (1992) Variable $(dG\text{-}dT)_n$ $(dC\text{-}dA)_n$ sequences in the porcine genome. *Genomics,* **12,** 28 1-288.

Won YJ, Sivasundar A, Wang Y, Hey J (2005) On the origin of Lake Malawi cichlid species: A population genetic analysis of divergence. *Proceedings of the National Academy of Sciences USA*, **102,** 6581-6586.

Wong AK, Yee HA, van de Sande JH, Rattner JB (1990) Distribution of CT-rich tracts inversed in vertebrate chromosomes. *Chromosoma*, **99,** 344-351.

Wright S (1942) Statistical genetics and evolution. *Bulleting of American Mathematical Society*, **48,** 223-246.

Wright JM, Bentzen P (1994) Microsatellites: genetic markers for the future. *Reviews in Fish Biology and Fisheries*, **4**, 384-388.

Yina B, Yanga S, Weia W, Zhang Y (2009) Male reproductive success in plateau pikas (*Ochotona curzoniae*): A microsatellite analysis. *Mammalian Biology,* **74,** 344-350.

Young KA, Snoeks J, Seehausen O (2009a) Morphological diversity and the roles of contingency, chance and determinism in African Cichlid radiations. *PLoS One,* **4,** e4740.

Young KA, Whitman JM, Turner GF (2009b) Secondary contact during adaptive radiation: a community matrix for Lake Malawi cichlids. *Journal of Evolutionary Biology*, **22,** 882-889.

Young KA, Genner MJ, Joyce DA, Haesler MP (2009c) Hotshots, hot spots, and female preference: exploring lek formation models with a bower-building cichlid fish. *Behavioral Ecology,* **20,** 609-615.

Zachos FE, Cirovic D, Kirschning J *et al.* (2009) Genetic Variability, Differentiation, and Founder Effect in Golden Jackals (*Canis aureus*) from Serbia as Revealed by Mitochondrial DNA and Nuclear Microsatellite Loci. *Biochemical Genetics,* **47,** 241-250.

Zardus JD, Hadfield MG (2005) Multiple origins and incursions of the Atlantic barnacle *Chthamalus proteus* in the Pacific. *Molecular Ecology,* **14,** 3719-3733.

Zidana H, Turner GF, van Oosterhout C, Hänfling B (2009) Elevated mtDNA diversity in introduced populations of *Cynotilapia afra* (Günther 1894) in Lake Malawi National Park is evidence for multiple source populations and hybridisation. *Molecular Ecology,* **18,** 4380-4389.

Appendices

Appendix I: Simulation Model of founder events

Simulating the effects of founder events on the genetic variation at microsatellite loci and mtDNA control region

To further investigate how founder events that originate from multiple (genetically diverged) source populations may affect the genetic variation and genetic divergence at mtDNA and microsatellite loci in introduced populations, we used an individual based model to simulate this study.

We simulated a founder event with size (number of founders=N_f) from source populations with a known allele or haplotype frequency distribution. The source populations contribute founder individuals proportional to the observed mtDNA haplotype frequency in the introduced population (Zidana *et al.* 2009). The demographic parameters (N_f and the time of introduction, t) are based on analysis of microsatellite variation using DIY ABC software version 7.1 (Cornuet *et al*. 2008). These estimates are consistent with the introduction history of *C. afra*. The census size (N) is estimated from direct observations and counts of the number of territorial males (Young *et al.* 2009).

The population post-founder event with size N was simulated over t generations and introduced *C. afra* population at Domwe was used as a typical introduced population basing on identification of source all of its haplotypes as follows: N_f=15% from Mara rocks, 15% from Nkhata Bay and 70% from Thumbi west Island site-A (Zidana *et al.* 2009). The Domwe population was founded 20-30 generations ago by a founder size of $181 \leq N_f \leq 4124$, estimated from DIY ABC. The mtDNA sequence of 500bp has a mutation rate $\mu=5x10^{-6}$ and mutates according to an infinite allele model (IAM) (Kimura & Crow 1964), while microsatellite loci have a mutation rate $\mu=5x10^{-4}$ and mutates according to stepwise mutation model (SMM) (Ohta & Kimura 1973).

We analysed the effect of genetic drift and multiple source founder event on the change in effective number of alleles (n_e) between the sources and introduced *C. afra* populations, and compared the simulated data set with the results from observed empirical data. I used General Linear Model (GLM) to test the significance of the results in MINITAB

version 12.1 (MINITAB Inc.). We predicted that the mtDNA variation in the introduced populations will be relatively elevated, whereas the microsatellite variation will be reduced compared to the source populations.

Appendix II

Table II.1: The observed and simulated microsatellite effective number of alleles (n_e) from native and introduced *C. afra* populations in Lake Malawi

Population	Pzeb3	Ppun32	Ppun5	Ppun21	UNH154	Ppun7	Mean	SE
Introduced (Observed)								
Otter point	1.535	1.581	10.683	8.372	11.520	6.180	6.645	0.727
Thumbi Site A	1.812	1.858	17.143	13.900	17.266	17.225	11.534	1.270
Thumbi Site B	1.185	2.194	8.136	9.690	18.136	12.478	8.637	1.063
Domwe	1.471	1.703	14.784	9.149	11.374	9.549	8.005	0.893
Mean	1.501	1.834	12.687	10.278	14.574	11.358	8.705	0.988
Introduced (Simulated)								
Domwe	1.827	1.877	11.551	10.304	12.658	13.078	8.549	4.816
Native (Observed)								
Ntekete	2.010	1.821	11.669	14.845	8.696	14.090	8.855	0.965
Chiofu	2.314	2.177	11.357	16.000	11.483	18.462	10.299	1.134
Mbenji	1.858	2.507	13.820	15.965	26.866	18.799	13.302	1.614
Nkhata Bay	2.335	1.817	15.721	14.286	22.360	31.304	14.637	1.907
Mara	1.349	1.934	13.873	10.909	17.433	15.859	10.226	1.167
Likoma	1.209	1.586	17.822	12.857	24.242	16.216	12.322	1.539
Mean	1.846	1.974	14.044	14.144	18.513	19.122	11.607	1.388

Appendix III

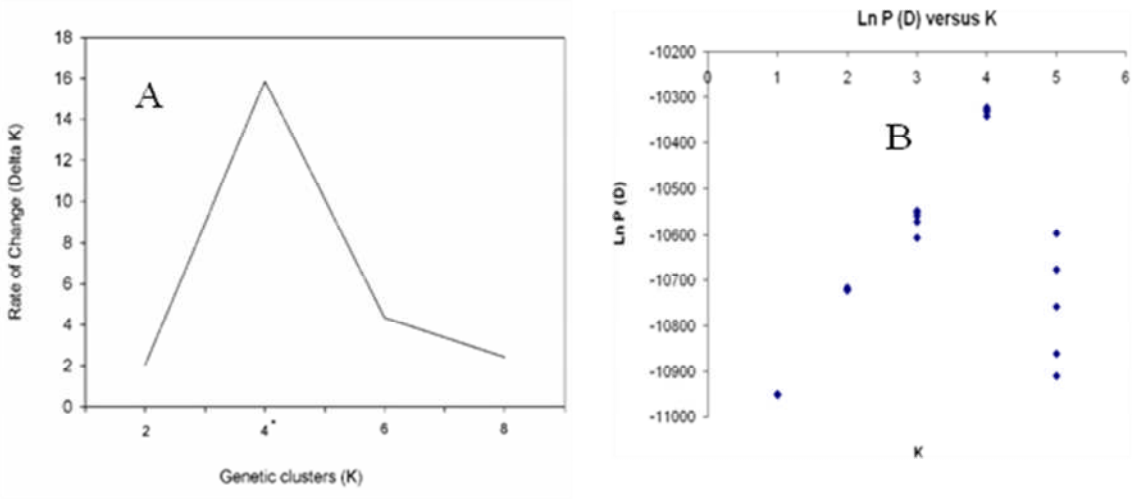

Fig. III.1. The modal value of this distribution is the true K (*) or the uppermost level of genetic structure among *C. afra* populations, here four clusters (A=Evanno *et al.* 2005 procedure, B=Pritchard *et al.* 2000). Ln P(D)=Log-likelihood for number of K clusters.

Appendix IV

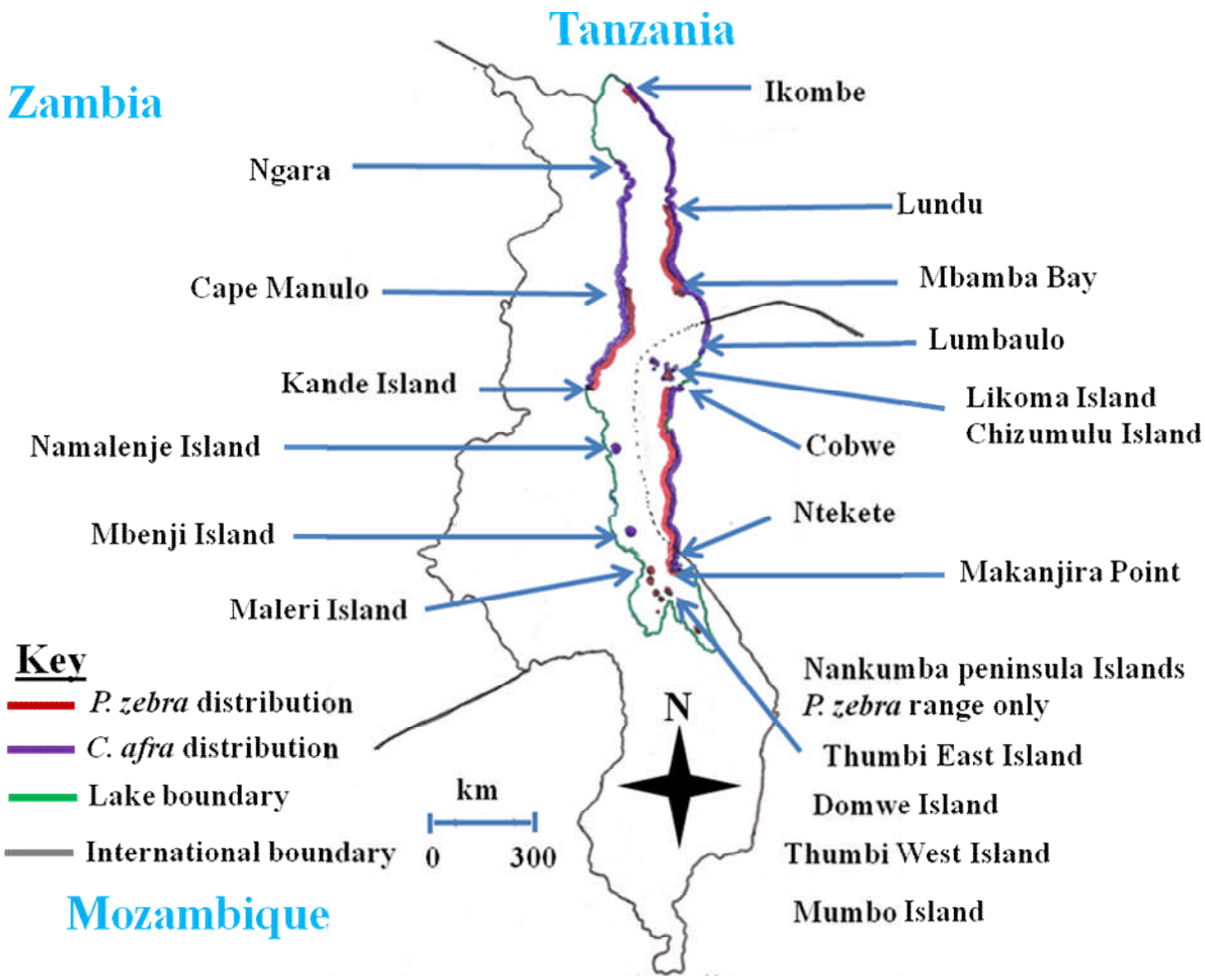

Fig. IV. 1. The map of Malawi with its neighbouring boundaries showing Lake Malawi and the natural distribution range for *C. afra* (purple line) and introduced range for *P. zebra* (red line) and dotted points for Island distribution.

Appendix V

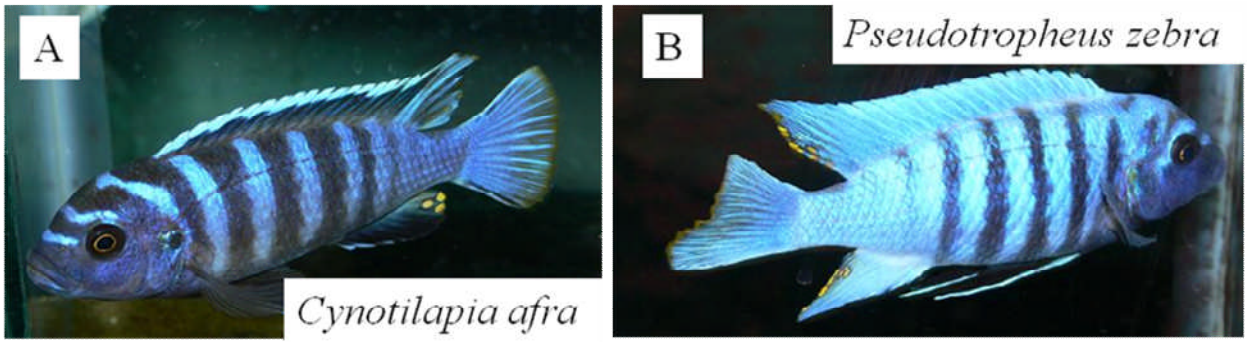

Fig. V.1. The sample species collected from Thumbi west Island (A) *Cynotilapia afra* and (B) *Pseudotropheus zebra* in Lake Malawi (Pictures taken by Alan Smith).

Appendix VI

Table VI. 1: The microsatellite mean number of alleles (*A*), observed heterozygosity (H_O), expected heterozygosity (H_E) and standardised (G'_{ST}) from sympatric populations of *C. afra* and *P. zebra* populations in Lake Malawi National Park

Otter point	*C. afra*			*P. zebra*			
Locus	*A*	H_O	H_E	*A*	H_O	H_E	G'_{ST}
PPUN32	5	0.27	0.37	4	0.53	0.60	0.20
PPUN5	18	0.90	0.91	19	0.88	0.93	0.86
PPUN21	18	0.90	0.89	18	0.80	0.91	1.09
UNH154	25	0.95	0.92	28	0.92	0.92	1.29
PPUN7	14	0.62	0.85	23	0.47	0.91	1.27
PZEB3	6	0.33	0.35	7	0.43	0.74	0.92
Mean	14.33	0.66	0.72	16.50	0.67	0.84	0.94
SE	1.70	0.05	0.05	1.95	0.02	0.05	0.18
Thumbi west A							
PPUN32	7	0.42	0.47	8	0.68	0.71	0.65
PPUN5	24	0.95	0.95	21	0.88	0.94	0.34
PPUN21	23	0.93	0.94	23	0.75	0.94	0.49
UNH154	31	0.95	0.95	36	0.98	0.97	0.56
PPUN7	25	0.65	0.95	27	0.87	0.96	0.13
PZEB3	9	0.48	0.45	8	0.50	0.47	0.00
Mean	19.83	0.73	0.78	20.50	0.78	0.83	0.36
SE	1.93	0.03	0.05	2.51	0.05	0.08	0.05
Domwe Island							
PPUN32	4	0.45	0.42	5	0.60	0.68	0.43
PPUN5	20	0.83	0.94	20	0.93	0.94	0.75
PPUN21	16	0.87	0.90	19	0.85	0.93	0.42
UNH154	24	0.83	0.92	30	0.95	0.95	0.94
PPUN7	17	0.50	0.90	23	0.73	0.93	0.66
PZEB3	5	0.32	0.32	7	0.38	0.44	0.07
Mean	14.33	0.63	0.73	17.33	0.74	0.81	0.55
SE	1.57	0.03	0.06	2.08	0.07	0.08	0.09

Appendix VII

Table VII. 1: The microsatellite genetic differentiation (F_{ST}) among *C. afra* and *P. zebra* populations in three localities of Lake Malawi National Park

	Otter *C. afra*	Otter *P. zebra*	Thumbi *C. afra*	Thumbi *P. zebra*	Domwe *C. afra*
Otter *P. zebra*	0.090				
Thumbi *C. afra*	0.033	0.048			
Thumbi *P. zebra*	0.063	0.045	0.032		
Domwe *C. afra*	0.053	0.070	0.020	0.048	
Domwe *P. zebra*	0.058	0.053	0.025	0.012	0.040

Appendix VIII

Table VIII.1: The microsatellite genetic differentiation (F_{ST}) among *C. afra* and *P. zebra* populations in all sampled localities of Lake Malawi

	OP	OP z	TH-A	TH-A z	TH-B	DO	DO z	NT	CH	MB	NK	LK	MA
OP													
OP z	0.090												
TH-A	0.033	0.048											
TH-A z	0.063	0.045	0.032										
TH-B	0.049	0.084	0.030	0.036									
DO	0.053	0.070	0.020	0.048	0.030								
DO z	0.058	0.053	0.025	0.012	0.036	0.040							
NT	0.051	0.044	0.018	0.038	0.041	0.041	0.034						
CH	0.074	0.045	0.038	0.022	0.053	0.063	0.031	0.043					
MB	0.047	0.043	0.018	0.007	0.028	0.038	0.014	0.023	0.012				
NK	0.040	0.040	0.012	0.030	0.034	0.029	0.024	0.023	0.023	0.014			
LK	0.029	0.074	0.014	0.048	0.032	0.025	0.039	0.027	0.058	0.030	0.020		
MA	0.035	0.071	0.014	0.037	0.031	0.030	0.030	0.029	0.050	0.028	0.019	0.017	

All pairwise differentiations were significant, $p < 0.05$. OP=Otter Point *C. afra*, OPz=Otter Point *P. zebra*, TH-A=Thumbi West C. afra, TH-A z=Thumbi West *P. zebra*, TH-B=Thumbi West site B *C. afra*, DO=Domwe Island *C. afra*, DOz=Domwe Island *P. zebra*, NT=Ntekete *C afra*, CH=Chiofu *C afra*, MB=Mbenji *C afra*, NK=Nkhata Bay *C afra*, LK=Likoma *C afra* and MA=Mara *C afra*.

Appendix IX

Table IX.1: Within population genetic diversity mean number of microsatellite alleles (*A*) in *C. afra* and *P. zebra* populations sampled in localities of Lake Malawi

Locus	OP	OP z	TH-A	TH-A z	TH-B	DO	DO z	NT	CH	MB	NK	LK	MA
PPUN32	5	4	7	8	7	4	5	4	4	5	4	6	8
PPUN5	18	19	24	21	19	20	20	21	17	22	22	24	22
PPUN21	18	18	23	23	20	16	19	20	26	23	22	19	18
UNH154	25	28	31	36	30	24	30	27	28	38	33	35	29
PPUN7	14	23	25	27	25	17	23	25	29	30	44	28	27
PZEB3	6	7	9	8	4	5	7	6	6	6	10	5	6
Mean	14.33	16.50	19.83	20.50	17.50	14.33	17.33	17.17	18.33	20.67	22.50	19.50	18.33
SD	7.04	8.46	8.76	10.01	9.25	7.41	8.77	8.93	10.21	11.94	13.36	11.00	8.77

OP=Otter Point C. afra, OPz=Otter Point P. zebra, TH-A=Thumbi West C. afra, TH-A z=Thumbi West P. zebra, TH-B=Thumbi West site B C. afra, DO=Domwe Island C. afra, DOz=Domwe Island P. zebra, NT=Ntekete *C afra*, CH=Chiofu *C afra*, MB=Mbenji *C afra*, NK=Nkhata Bay *C afra*, LK=Likoma *C afra* and MA=Mara *C afra*.

Appendix X

Table X.1: Within population genetic diversity expected heterozygosity (H_E) at six microsatellite alleles in *C. afra* and *P. zebra* populations sampled in localities of Lake Malawi

Locus	OP	OP z	TH-A	TH-A z	TH-B	DO	DO z	NT	CH	MB	NK	LK	MA
PPUN32	0.37	0.60	0.47	0.71	0.55	0.42	0.68	0.45	0.55	0.61	0.45	0.37	0.49
PPUN5	0.91	0.93	0.95	0.94	0.88	0.94	0.94	0.92	0.92	0.94	0.94	0.95	0.94
PPUN21	0.89	0.91	0.94	0.94	0.90	0.90	0.93	0.94	0.95	0.95	0.94	0.93	0.92
UNH154	0.92	0.92	0.95	0.97	0.95	0.92	0.95	0.89	0.92	0.97	0.96	0.97	0.95
PPUN7	0.85	0.91	0.95	0.96	0.93	0.90	0.93	0.94	0.95	0.95	0.98	0.95	0.94
PZEB3	0.35	0.74	0.45	0.47	0.16	0.32	0.44	0.51	0.57	0.47	0.58	0.17	0.26
Mean	0.72	0.84	0.78	0.83	0.73	0.73	0.81	0.78	0.81	0.81	0.81	0.72	0.75
SD	0.25	0.12	0.23	0.19	0.29	0.26	0.19	0.21	0.18	0.20	0.21	0.32	0.27

OP=Otter Point *C. afra*, OPz=Otter Point *P. zebra*, TH-A=Thumbi West *C. afra*, TH-A z=Thumbi West *P. zebra*, TH-B=Thumbi West site B *C. afra*, DO=Domwe Island *C. afra*, DOz=Domwe Island *P. zebra*, NT=Ntekete *C afra*, CH=Chiofu *C afra*, MB=Mbenji *C afra*, NK=Nkhata Bay *C afra*, LK=Likoma *C afra* and MA=Mara *C afra*.