UNIVERSITÀ DEGLI STUDI DELL'INSUBRIA
FACOLTÀ DI SCIENZE MM. FF. NN.

Ph.D. School in Biological and Medical Sciences

Ph.D. Program in Analysis, Management and Protection of the Biodiversity

XXIII Cycle

Evolutionary processes in an environmental challenging site: the soda-lake Natron (Tanzania)

Ph.D. Thesis of

Isabella Vanetti

Supervisor

Prof. Giorgio Binelli

2009 / 2010

INDEX

ABSTRACT

Lakes Natron and Magadi are neighbouring small soda localised in the Eastern Rift Valley, between Kenya and Tanzania, and are the remnants of the Paleolake Orolonga after a split that occurred around 9,000 years BP. These basins represent a really harsh environment, characterized by a complex geo-morphological structure, which may increase, by habitat fragmentation and isolation, the effects of evolutionary forces (like genetic drift and selection) on the native species and populations. These lakes harbour a small species flock of quite recent origin, the *Alcolapia* flock, endemic of this region and not found in any other place around the world. This flock is characterized by four different morphotypes: *A. alcalicus, A. latilabris*, *A. ndalalani*, endemic of Lake Natron, and *A. grahami,* found only in the Lake Magadi.

With the aim to understand the evolutionary processes that are shaping genetic diversity of this species, we have analysed by means of molecular genetics tools (mtDNA and nuclear DNA markers), 310 specimens of *Alcolapia* collected from eight populations placed around the Lake Natron and from one population located in the north-east part of Lake Magadi.

Phylogenetic analyses based on D-loop sequences of a subset of 69 *Alcolapia* have shown a monophyletic structure of the flock, as suggested by the more frequent haplotype (*2lat*) shared by all the morphotypes and corresponding to the Orolonga haplotype identified in previous studies (Seegers et al, 1999; Wilson et al., 2000 and 2004). Besides, the main starburst radiation occurred in Lake Natron *Alcolapia*, which evolved in relatively recent times,

following the Natron/Magadi separation. The successive demographic expansion of the founding population determined the sharing of sequence polymorphisms among the flock.

Population and bayesian analyses were conducted based on the genotyping at 7 heterologous SSR loci of the whole dataset. The genetic differentiation of the *Alcolapia* flock is mainly due to the separation between the Natron and Magadi populations. The genetic structure of the Natron populations, on the other hand, is really weak, due to the existence of gene flow between the *Alcolapia* populations, in spite of the chemical and physical parameters of the main lake. The subsequent evolution of the three distinct *Alcolapia* morphotypes in the Natron basin could be the beginning of their speciation process, although their current levels of genetic diversity remain low.

From these analyses, we suggest that their rapid evolution has been driven by the association of the frequent environmental perturbations in the catchments. Whilst this may have also been in association with interspecific hybridisation events where the morphotypes co-occurred, as proposed by Seehausen (2006), no hybrids were documented in the study.

Thus, future work could focus on identifying the presence of hybrid phenotypes in sympatric communities to reveal whether the hypothesis that interspecific hybridisation facilitates rapid contemporary evolution is valid. Besides, we plan to assess genetic diversity and differentiation for those loci which could be possibly involved in speciation by a candidate gene approach. To this purpose, homeogenes for the development of the mouth and gene responsible for high temperature and high saline tolerance will represent suitable candidates.

INTRODUCTION

Biodiversity

2010 is the International Year of Biodiversity as declared by the United Nations (Marton-Lefèvre, 2006). The goal of this UN initiative is to increase public awareness of biodiversity and to encourage people to protect it.

But what does biodiversity mean and why should we care?

The term "biodiversity" was introduced in the eighties of the last century by WG Rosen (1985) as the combination of two distinct terms: "biological" and "diversity". The biodiversity was defined by the UN Earth Summit that took place in Rio de Janeiro in 1992 as "the variability among living organisms from all sources, including, inter alia, terrestrial, marine, and other aquatic ecosystems, and the ecological complexes of which they are part: this includes diversity within species, between species and of ecosystems"; this definition is in complete agreement with what was formerly defined by Norse & McManus in 1980, that is that biological diversity summarizes the concepts of “genetic diversity” and “ecological diversity”. Genetic diversity regards the variability of genetic information found between species and, at the species level, within and between populations. The ecological diversity concerns the number of species in a given area, the ecological role they played, the way in which their composition varied and their grouping in a given space.

The loss of biodiversity is undoubtedly become a reality for our planet in the last years, so that in 2002 the present year was taken as dead-line to reduce this phenomenon at the global, regional and national levels by the 190 government Parties that took part to the Johannesburg Convention on Biological Diversity (CBD) (Decision VI/26). This was followed later by a commitment at the UN World Summit on Sustainable Development to achieve "by 2010 a significant reduction in the current rate of loss of biological diversity" (Mace & Baillie, 2007; Marton-Lefèvre, 2006).

Evidences of the ongoing loss of biodiversity can be found looking at The Red List of Threatened Species redacted by the International Union for Conservation of Nature (IUCN) that documents the extinction risk of 47,677 species: 17,291 are threatened, including 12% of birds, 21% of mammals, 30% of amphibians, 27% of reef-building corals, and 35% of conifers and cycads. The Living Planet Index reveals that populations of wild species have declined by 30% since 1970; mangrove forests have lost a fifth of their area since 1980, and 29% of seagrass beds are gone (Marton-Lefèvre, 2006).

The anthropic factor certainly has been the protagonist in the biodiversity level decrease. The use of fossil fuels and deforestation have increased over the last three centuries the concentration of carbon dioxide (CO_2) by 30% and about half of this increase occurred in the last 40 years. Human activities have doubled the concentration of methane and increased the concentration of other gas involved in global warming. The expansion of urban and agricultural areas has incorporated grasslands, forests and wetlands (e.g. more than half of all natural habitat on agriculturally useable

land has already been cleared for cropland or permanent pasture, and much of the rest has been altered by temporary grazing (FAOSTAT 2000)) pouring into water basins increasing amounts of water nutrients, responsible for changes at ecological level of estuaries and coasts. A third of dry land productivity and 8% of that of sea are controlled by man, as well as 54% of available water, the use of which presumably will rise up to 70% by 2050. At sea, three-quarters of harvested fish populations monitored by the FAO are already overexploited, or will become so without stringent management intervention (Balmford *et al.*, 2005;FAO 2000).

Fresh waters as a whole are a hotspot for biodiversity, with 125,000 species described that represent 9.5% of all known animal species on the planet (including 1/3 of all vertebrate species), even though fresh waters cover just 0.8% of the Earth's surface area. A few freshwater species have large geographic ranges, but the insular nature of freshwater habitats has led to the evolution of many species with small geographic ranges, often encompassing just a single lake or drainage basin, resulting in high levels of endemism and habitat fragmentation (Strayer *et al.*, 2010). Habitat isolation and dispersal limitation that have generated high freshwater fish diversity can also increase the risk of species extinction. Recent global estimates indicate that 25% of evaluated freshwater fish species are considered threatened with extinction, but this should not be surprising, given that freshwaters are subjected to a panoply of anthropogenic threats, including habitat loss and fragmentation, hydrologic alteration, climate change, overexploitation, pollution and the spread of invasive species. For example,

humans now appropriate >50% of available freshwater run-off, reservoirs trap 25% of the global sediment load before it reaches the oceans, river systems have been fragmented by about 1 million dams globally, and many inland fisheries are vulnerable to collapse. The end result is that freshwater fishes are among the most imperilled faunas worldwide (Olden *et al.*, 2010).

All these factors surely leads to an alteration of ecosystems that has already favourited the extinction of an unknown number of species and endangered many others due to the reduction in size of the populations.

Extinction by itself represents a natural step of the evolutionary process: a species persists on average for about 5-10 million years and, if the extinctions are balanced by the origin of new species (speciation), the degree of biodiversity can be maintained. Currently, however, the speciation is not able to compensate the amount of loss of species and current forecasts indicate that, with the continuous rising in human population size, the rate of extinction is expected to further increase, up to 1000 times the natural rate. Over the last 500 million years there have been five mass extinctions, including the cataclysm that led to the disappearance of the dinosaurs and eliminated much of the flora and fauna of the Cretaceous, but in none of these man has played a role. The severity of the current situation has brought the community scientific to talk about a "sixth extinction".

Human involvement to resolve the problem is essential and is of paramount importance not only to ensure the survival of endangered species but also to ensure the welfare of humankind itself. Mankind, in fact, derives from the living world many direct and indirect benefits: food, pharmaceutical drugs,

fuel, clothing fibres and building materials are just some examples of the available bioresources. Living organisms also ensure a range of "services", useful, essential biological functions that are provided free of charge: production of oxygen by plants, climate control implemented by forest, natural pest control, pollination of crop plant, etc. Even the purely aesthetic value of biotic resources must not be forgotten, as expressed in growing ornamental plants, keeping pets, visiting zoos and nature reserves, and ecotourism. The reduction of the biodiversity degree results in a reduction of all these resources and is therefore disadvantageous for humankind.

The realization that the current decline in biodiversity is largely due to human activities and that therefore we can control these factors through appropriate management policies, has placed the need to develop appropriate conservation programs at government level, global and local.

The Convention on Biological Diversity (CBD) signed in 1992 at the United Nations Conference on Environment and Development (UNCED) held in Rio de Janeiro highlighted the problem of biodiversity conservation: the objectives propose in Article 1 were "the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources, including by appropriate access to genetic resources and by appropriate transfer of relevant technologies, taking into account all rights over those resources and to technologies, and by appropriate funding". Italy has ratified and implemented the CBD with the law n ° 124, 02/14/1994, published in the Official Gazette of February 23, 1994. The CBD has also stated the

responsibility of each state in biodiversity conservation on its territory (Article 6), underlining the need to develop scientific capacity that may allow an higher, wider level of information about the biodiversity issue and has identified the major policies for the conservation "*in situ*" and "*ex situ*" (Articles 8 and 9). The conservation of biodiversity "*ex situ*" is defined (Article 2) as "The conservation of components of biological diversity outside their natural habitats".

The biodiversity conservation "*in situ*" is defined as "The conservation of the ecosystems and natural habitats and maintenance and recovery of viable populations of species in their natural surroundings and, in the case of domesticated or cultivated species, in the surroundings where they have developed their distinctive properties". The "*in situ*" strategy has therefore as its objective the conservation of natural populations directly in their original environment. To accomplish this, however, is necessary to assess the genetic resources that are really available and to investigate the evolutionary mechanisms that regulate their maintenance. A useful aid in this sense comes from Conservation Genetics, an applied science that involves the application of evolutionary and molecular genetics to biodiversity conservation, dealing with the genetic factors that affect extinction risk and genetic management regimes required to minimise these risks.

The main genetic issues in conservation biology can be resumed as:

- genetic contributions to resolving taxonomic uncertainties,
- defining evolutionary diverged units for separate management within species,

- genetic management to minimize inbreeding and loss of genetic diversity in populations, and species extinction risk in the wild,
- genetic management of captive populations to minimize inbreeding and loss of genetic diversity in captivity, and to maximize reintroduction success,
- contributions of genetics to management of invasive species and their impacts on threatened species,
- application of molecular genetics to obtain information important to species conservation (sex, population size, demographic history, mating system, population structure, gene flow, parentage, diet, diseases),
- use of molecular genetic analyses in forensics,
- integrating genetics with demographic and environmental variables, catastrophes and human impacts to predict extinction risk and compare alternative options in species recovery programs,
- landscape genetics, an interdisciplinary field that combines population genetics and landscape ecology, to retrace routes for species expansions and to predict new ones.

(Frankham *et al.*, 2002; Frankham, 2010)

Cichlids

Cichlids are teleost freshwater fishes that represent, inside the vertebrates, one of the most species-rich families. The Chilidae family, infact, is made by more than 3000 species, naturally distributed across Africa, Madagascar, India, southern and central America. The phylogenetic relationships among the major lineages of cichlids, as assessed by means of molecular analyses, suggest a basal and paraphyletic position for the Indian and Malagasy lineages, whereas placing the monophyletic African and American lineages as sister groups. The basal lineages account only for about a dozen of species, whereas the American group is represented by 400 or at least 500 species. The real hotspot of cichlid biodiversity, however, occurs in Eastern Africa, where several rivers and lakes of various sizes and ages harbour over 1800 different cichlid species, characterised by high level of diversity for morphology, behaviour and ecology (Genner *et al.*, 2007; Salzburger and Meyer, 2004).

The three major lakes of East Africa, Malawi, Tanganyika and Victoria, account collectively for over 1000 species, organised in communities with different species-richness levels, spanning from fewer than 10 species to over 100, even in the same basin and often living in complete sympatry conditions. Due to the surprising biodiversity levels of the three lakes, the natural and evolutionary history of Lakes Malawi, Tanganyika and Victoria was widely investigated during the last 110 years, starting from Boulenger

that published the first report on East African fish fauna in 1898. (Hulsey, 2009; Kornfield & Smith, 2000; Kocher, 2004; Turner *et al.*, 2001).

Lake Tanganyika, with an estimated age between 9 and 12 million years BP, is the oldest of the three basins (Cohen *et al* . 1993) and harbours at least 197 endemic species of cichlids belonging to 49 endemic genera (Poll, 1986) that can be grouped into 12 separate tribes, thought to have arisen from seven or nine distinct ancestral lineages (Meyer 1993; Koblmüller *et al.*, 2007). Lake Tanganyika's cichlids are relatively old compared to other East African cichlid lineages and older than either the Victoria or Malawi lake basins (Fryer & Iles, 1972), given that phylogenetic evidences suggest for some of the tribes a time of divergence of five million years (Sturmbauer & Meyer, 1993). Tanganyika's cichlids are morphologically and behaviourally more diverse than the cichlids of Lakes Victoria and Malawi (Fryer & Iles, 1972), however, the latter two lakes each possess a greater number of cichlid species than Lake Tanganyika.

More than 300 endemic species of cichlids inhabit the Lake Victoria, the youngest (originated 250,000–750,000 years ago) and shallowest basin (Johnson *et al.*, 1996; Seehausen,1996). It appears that all species were derived from a single common ancestor (Meyer *et al.,* 1990) even if recent molecular data suggest that Lake Victoria was colonized by at least two separate lineages, which invaded the lake and rapidly diverged within the last 12,500 years, after the completed drying of the basin during the last Ice Age (Nagl *et al.,* 2000).

Cichlid fishes are thought to have invaded Lake Malawi approximately 700,000 years ago, and their morphological diversity is considerably greater than the much younger species flock of Lake Victoria. In addition, the Malawian radiation has produced the greatest number of endemic species of the three species flocks (well over 400 endemic species distributed among 49 endemic genera) and appears to be monophyletic in origin (Meyer *et al.*, 1990; Kocher *et al.*, 1993; Moran *et al.,* 1994). The phylogenetic history of the Lake Malawi flock suggests a pattern of radiation in three stages: occupation of different macrohabitats (e.g., rocky versus sandy substrates), followed by a radiation in feeding morphology (that produced most of the currently recognized genera) and finally, presumably under the action of sexual selection, diversification in male breeding colours for certain lineages (Albertson *et al.*, 2003; Danley & Kocher, 2001).

This astonishing biodiversity richness has been shaped in the last several thousand to a few million years by explosive speciation and adaptive radiations phenomena, not yet observed in any other vertebrate group. East African cichlids, for this reason, have become a really great point of attraction for evolutionary biologists that have recognised the importance of these fishes as a strong biological and ecological model to investigate evolutionary processes and test theory on speciation and differentiation (Hulsey, 2009; Kornfield & Smith, 2000; Kocher, 2004; Schwarzer *et al.*, 2009; Turner *et al.*, 2001).

Haplochromines and tilapiines are the two major lineages of cichlid fishes in East Africa (Regan, 1920). The haplochromines are widely dispersed in the

East African Great Lakes with hundreds of species that have been studied extensively (Brandstätter *et al.,* 2005; Elmer *et al.,* 2009; Salzburger *et al.,* 2002; Salzburger *et al.,* 2005; Verheyen *et al.,* 2003). The tilapiines, on the other hand, comprise smaller flocks of typically less than 10 species. Although generally confined to river systems and unable to colonize lakes occupied by the more specialised haplochromine cichlids, their ability to adapt to extreme environmental conditions, including water temperature, alkalinity, and salinity, has increasingly been receiving attention (Beveridge & McAndrew, 2000; Klett & Meyer, 2002; Nagl *et al.,* 2001; Stiassny *et al.,* 1992).

Based on the breeding habits the tilapiines have been divided into three major genera: *Sarotherodon,* the biparental and paternal mouthbrooders (9 species); *Tilapia,* the substrate spawners (Trewavas, 1981, 1982, 1983) with 70 or so species commonly referred to as the "tilapias" (Trewavas, 1966a,b) and *Oreochromis,* the maternal mouthbrooders (31 species), divided by Trewavas (1983) into five subgenera: *Oreochromis, Alcolapia, Vallicolla, Nyasalapia,* and *Neotilapia* (Nagl *et al.*, 2001).

The *Alcolapia* flock

The subgenus *Alcolapia*, recently advanced at genus rank (Seegers *et al.*, 1999) and identified as single cluster by Nagl *et al.* (2001), represents a small and young flock of quite recent origin with a clear monophyletic structure.

The natural distribution of this flock encompass two of the most harsh environments in the Eastern African Rift Valley: the Soda Lakes Magadi and Natron, located in Kenya and Tanzania, respectively. Lake Magadi is a closed lake at the southern end of the Eastern (Gregory) Rift that displays particularly harsh conditions including water of pH 9.8-10.5, titration alkalinity >300 mM, osmolality 525 mOsm, temperatures as high as 42°C, O_2 levels fluctuating diurnally between extreme hyperoxia and virtual anoxia, and intense levels of avian predation and UV radiation (Narahara *et al.*, 1996; Pörtner *et al.*, 2010; Wilson *et al*, 2004). Lake Natron is wider (600 km^2 vs. 100km^2), placed southern and, unlike Magadi, receives substantial surface water inflows, primarily from the Ewaso-Ngiro river. In Lake Natron, measurements recorded in April 2009 revealed water temperatures to 42°C, conductivity values to 107,000 µS cm and pH levels up to 10. Where conductivity salinity was high, dissolved oxygen was low (< 3 mg/L).

The geological record indicates that Lakes Magadi and Lake Natron were joined in the Paleolake Orolonga at the end of the Pleistocene, approximately between 12,500 and 9,600 yr BP (Butzer *et al.*, 1972; Goetz & Hillaire-Marcel, 1992). The lake level was 56 m above the current surface of Lake Magadi (600 m. alt.) and 48 m above that of Lake Natron and the water

chemistry was much less alkaline and less salty respect the present time. The Younger Dryas, a period of intense climate change that outcame in progressive drying in this part of Africa as a result of a reduction in ocean to land moisture flux (Roberts *et al.* 1993), probably favourited between 9,000 and 8,000 years BP the separation of Paleolake Orolonga into smaller Lakes Natron and Magadi, and even smaller Little Magadi (0.5 km^2, 620 m alt., north of the Magadi basin); it seems that the split-up was complete by 7,000 years BP (Butzer *et al.*, 1972; Tichy & Seegers 1999). Increasing contraction of the lakes has occurred since then, shaping the common geomorphological structure of the basins, characterised by floating precipitate of sodium bicarbonate (trona), underlain by anoxic lake water, which occupies large parts of the lakes bed and defines separated discrete shallow water bodies, or lagoons (the maximum depth for Natron is around 3 meters) (Pörtner *et al.*, 2010; Seegers *et al.*, 1999). *Alcolapia* populations that inhabit the lakes persist only in these relatively small freshwater areas including swamp, creeks, brooks and springs that flow into the lagoons, (Coe, 1966; Coe, 1969). The alkalinity and the temperature of the water seem to prevent fish dispersion within the littoral zone, reducing their dispersal capability during short seasonal spots of intense rainfall, which provide temporary layer of fresh water across the lakes (Pörtner *et al.*, 2010; Vincens & Casanova, 1987; Wilson *et al.*, 2004).

Figure1 Aerial view of Lake Natron.

Despite the severe habitat conditions, fishes of the *Alcolapia* genus have shown an astonishing ecological adaptability, developing a wide spectrum of physiological adaptations and tolerances to pH, temperature and salinity, well investigated by scientists, which allow them the survival in such hostile environment:

- expression of the ornithine-urea cycle in white muscle and liver that permits a condition of 100% ureotely, unique within teleost fishes;

- extremely thin blood-water diffusion barrier, with very high blood O_2 affinity and an exceptionally high diffusing capacity of the gills for O_2;
- use of the swim-bladder as a primitive air-breathing organ capable of supplementing air-breathing during hypoxic conditions;
- a unique pyloric bypass system to allow drinking of alkaline lake water despite the presence of an acidic stomach;
- an exceptionally high metabolic rate that reflects the high environmental temperatures and activity levels of these remarkable fishes

(see Bergman *et al*, 2003; Coe, 1965; Laurent *et al.*, 1995; Lindley *et al.*, 1999; Maina, 1990; Maina *et al.*, 1995; Maina *et al.*, 1996; Narahara *et al.*, 1996; Randall *et al.*, 1989; Wood *et al.*, 1994; Wilkie and Wood, 1996; Wilson *et al.*, 2000).

The peculiarity of these extreme freshwater fishes, undoubtedly, has represented and still represents an interesting challenge for evolutionary biologists, given that:

- the surprising range of morphological and physiological adaptations observed was developed in a few thousand years (i.e. from the split up of Lake Orolonga that has changed water chemistry), thus representing an exceptional rapid adaptive radiation;
- the habitat fragmentation and isolation could have reduced the population of the species that live in these lakes to the condition of "small populations", making them potentially more sensitive to the

effects of genetic erosion due to both genetic drift and the potential selective effects of their peculiar environmental conditions.

Classic taxonomic methods of investigation jointed with more actual genetic molecular analysis have been so applied by several scientists to elucidate the origin and the speciation dynamics of this fascinating fishes flock.

The morphological characterisation of several populations of the *Alcolapia* flock from Lakes Natron and Magadi has revealed the presence of only four morphologically quite distinct species: *A. alcalicus* (Hilgendorf, 1905), *A. ndalalani* and *A. latilabris*, endemic of the lake Natron and *A. grahami* (Boulenger, 1912) located exclusively in Magadi and Little Magadi. The distinctiveness of these four alternative morphotypes is due primarily to head shape and oral dentition, as yet observed in other cichlid species from the Great African Lakes (Seegers & Tichy, 1999; Seegers *et al.*, 1999; Tichy & Seegers, 1999; Trewavas, 1983).

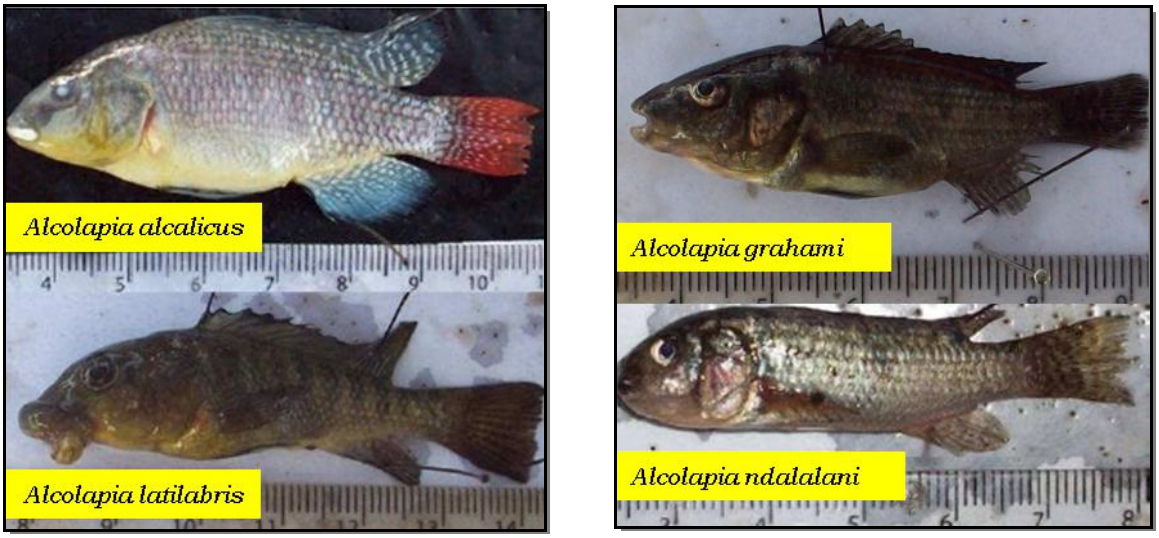


Figure 2 Morphotypes of the Alcolapia flock fishes.

Analyses of mitochondrial DNA sequence have identified 18 haplotypes in a collection of fishes sampled mainly from Lake Natron (Seegers and Tichy 1999; Seegers *et al.* 1999) and 13 haplotypes in a subset of animals collected primarily from Lakes Magadi and Little Magadi (Wilson *et al.* 2000, 2004). A perfect match for some of the haplotypes from the two studies has been identified: a single haplotype common to all populations, termed *A1* by Seegers *et al.* (1999) and *B* by Wilson *et al.* (2000, 2004), and a dominant Magadi specific haplotype, termed *A17* (most frequent of five) by Seegers *et al.* (1999) and *A* (most frequent of seven) by Wilson *et al.* (2000, 2004). Both studies have also identified several other lake-specific haplotypes evolved by different mutation steps from this two haplotypes.

The authors of these studies suggest that the haplotype *B (A1)* shared by all the lakes and by all the species represents an ancestral haplotype retained from the fish resident in Paleolake Orolonga, likely a cichlid with basilar similarity to present day *Alcolapia spp.* This finding seems to be supported by fossil tilapia, very similar in morphology to but much larger in size than present-day *A. grahami*, found in deposits approximately 20–40 m above the current surface of the basins (Coe 1966; Pörtner *et al.*, 2010; Tichy and Seegers 1999).

The haplotype *B,* so, must have an age of at least of 8000-9000 years, in agreement with the timing of the split up of Paleolake Orolonga and then the radiation from which have arisen all the others haplotypes has occurred in a really short evolutionary period.

Why *Alcolapia*? Evolution and conservation considerations

The aim of the present work is to investigate, by means of molecular genetics tools (mtDNA and nuclear DNA genetic markers), the evolutionary processes that are shaping the genetic diversity of the *Alcolapia* flock from lake Natron and Magadi.

The real question we have to propose, however, is "Why *Alcolapia*?".

First this flock is important for conservation.

This species complex, endemic of lakes Natron and Magadi, and artificially introduced in lakes Elementeita and Kivu between the fifties and the beginning of the sixties of the last century, cannot be found in any other part of the world. At the present time *Alcolapia spp.* are classified as "vulnerable or endangered" by the Red List redacted by The International Union for Conservation of Nature (IUCN), due to their restricted distribution. The peculiar environment of soda lakes, moreover, results particularly sensitive to the effects of climate change that tends to reduce by evaporation the water bodies; the resulting exasperation of water chemistry and temperature parameters can reach toxic or lethal level even for fishes used to live in extreme conditions.

Besides, Lake Natron is threatened by the proposals to establish an hydro-electric plant and a facility to extract and process soda ash; these solutions could surely led to an alteration of the environment with potential extinction risk for the *Alcolapia* species.

The possibility to lose definitively a such interesting and unique biological reality suggest the need to establish adequate conservation plans.

Conservation genetics provide us with the right tools to do it, starting from the assessment of genetic variability levels. An appropriate degree of genetic variability infact guarantees the populations the ability to evolve and adapt to environmental changes and to maintain reproductive fitness.

But what is the object of conservation genetics?

Genetic variability, or genetic diversity, is the variability of genetic information present at the individual level in all living organisms and is manifested by the presence of multiple allelic variants and/or genotype; it is present both between species and between populations of the same species. At the individual level events of mutations in the genome contribute to the formation of new allelic variants, but the occurrence of such events is extremely low to create by themselves genetic variability. The real forces that can create genetic diversity at the population level, mobilizing between populations the allelic variants created by mutation, are represented by the migration of individuals and the recombination. IUCN recognizes the need to preserve genetic diversity as one of three global conservation priorities. Genetic variability, in fact, represents an essential condition for evolution, understood as changes in allele frequencies over time, and it is the starting point for the evolutionary forces of selection and genetic drift. Selection changes allele frequencies in the direction of better adaptation of individuals of a population to environmental conditions. Genetic drift changes allele

frequencies at random and can be understood as the random sampling of alleles in each generation.

Endangered species are often composed of small populations, in which genetic drift is particularly important, contributing most strongly to the loss of genetic variability for loss and fixation of alleles. In small populations, besides, the rate of inbreeding increases with further loss of genetic variability due to the heterozygosity reduction caused by this coupling system. These considerations can be certainly applied to the *Alcolapia* species we are studying, due to the habitat fragmentation and isolation they are experimenting and to the special environmental conditions that can play selective effect on populations.

Second, the *Alcolapia* flock of Lakes Natron and Magadi represents also an original and contemporary evolutionary radiation model, useful for improving our general understanding of speciation processes.

Speciation mechanisms have been classified in many ways, depending on the importance that is ascribed to different factors.

The traditional, geographic, classification proposed by Mayr (1963) recognises three main speciation pattern:

- **Allopatric speciation:** the gradual divergence of populations with completely separate geographic ranges;
- **Parapatric speciation**: the divergence of populations with adjacent geographic ranges, and which occasionally exchange genes;
- **Sympatric speciation**: the divergence of populations in close physical contact in the same geographic area.

For most of the last century allopatric speciation was considered the main mechanism for new species arising: the gradual development of reproductive isolation among populations, achieved by incompatibilities accumulation, led populations fail to produce offspring when they cross each other, and they would be considered separate species. Sympatric speciation, moreover, was thought to be unlikely because high levels of gene flow would prevent differentiation (Chinsembu, 2009; Kocher, 2004).

Allopatric speciation has been surely important in the radiation of many groups, as assessed by Mayr (1963, 1984) for African cichlids; however it would be a mistake to conclude that only this mechanism explains the origin of all new species. Lake Victoria cichlid radiation, e.g., has taken about 12,500 years, a really short evolutionary period to be compatible with an allopatric speciation pattern (Meyer *et al.*, 1990; Seehausen, 1996).

Today molecular evidences suggest the presence of parapatric and sympatric speciation among cichlids, e.g., in lake Malawi (Danley *et al.*, 2000; Danley & Kocher 2001), in a tropical lake of Mexico (Strecker , 2006) and among some crater lakes in Nicaragua (Barulenga *et al.*, 2006; Elmer *et al.*, 2010).

Recent classifications of speciation mechanisms, however, have focused on the selective forces that are responsible for the differentiation of populations, regardless of the levels of gene flow among incipient species: selection on ecological traits, sexual selection and genetic conflicts (Schluter, 2001).

- **Ecological selection:** The role for ecological selection in the radiation of jaw shape has been suggested by the presence of similar feeding morphologies evolved repeatedly within different cichlid assemblages

and lakes. Cichlids have developed, in a remarkably short evolutionary period, a wide array of jaw and tooth morphologies, adapted for the three main different modes of feeding: biting, sucking, and ram feeding. The functional design of the feeding apparatus can predict the feeding modality. Ram feeders are characterized by a long, streamlined head optimized for pursuit and overtaking of prey. Both sucking and biting species have a short, cone-shaped head, but differ at discrete anatomical points on the upper and lower jaw . Further specialization of jaws and teeth is correlated with occupation of specific foraging niches (i.e., biting fins or biting algae) (Albertson *et al.*, 2003; Albertson & Kocher, 2006).

- **Sexual selection:** acts on an organism's ability to successfully copulate with a mate. Female cichlids lay a small number of relatively large eggs, and care for their young for several weeks until the yolk is absorbed. Most of the African lake cichlids are maternal mouthbrooders, i.e. the females pick up the eggs immediately after laying and incubate them in their mouths for several weeks. In some species, maternal care continues for several weeks after the free-swimming juveniles are released. By contrast, males of most species contribute nothing but genes to their offspring. This asymmetric parental investment leads to strong sexual selection, sexual dimorphism and has further consequences for cichlid life history and dispersal. In Lakes Malawi and Victoria cichlids, females select their mates based on male nuptial colouration (Seehausen & Van Alphen,

1998). Thus there is a possibility that female choice of male nuptial colours is a special driving force for speciation (Seehausen *et al.*, 1999). It has been postulated that female cichlids have preferences for different male colours and that these female preferences cause reproductive isolation between incipient species (Chinsembu, 2009; Kocher, 2004).

- **Genetic conflicts:** the powerful evolutionary force represented by conflicts between different elements of the genome (Hurst *et al.*, 1996). The most frequent conflicts arise between genes that reside in different cellular compartments (e.g. cytoplasmic versus nuclear genes) or which are inherited asymmetrically (e.g. autosomes versus sex chromosomes). Conflicts can also arise between genomes (e.g. maternal versus paternal or parental versus zygotic). Intersexual conflict, or sexual antagonism, is thought to be responsible for the rapid evolution of proteins in the male and female reproductive tracts, and to thereby cause both pre- and post-zygotic hybrid infertility among closely related species (Chinsembu, 2009; Kocher, 2004; Seehausen & Van Alphen, 1998).

The *Alcolapia* flock of Lake Natron thus represents a very good living lab to study the mechanisms of sympatric speciation, because all different hypotheses can be tested by using genetic tools. In the last decades, a new, powerful tool has become available to this purpose, being represented by molecular genetic markers.

Molecular markers

The use of molecular markers has allowed in recent decades to monitor wide areas of the genome in a large number of individuals, giving a strong push forward to the studies of population genetics. A molecular genetic marker is defined as any locus that is able to uniquely and constantly identify a region of DNA. At the present time several types of molecular genetic markers, including mitochondrial DNA (mtDNA) and nuclear DNA (ncDNA) markers, are available but none of them can be regarded as optimal for all applications, thereby, often the choice of marker class is based on the context in which they are used and on the objectives proposed (Arif & Khan, 2009; Frankham *et al.*, 2002; Sunnucks, 2000).

In this work we have used two different molecular marker: mtDNA D-loop and nuclear SSRs.

Mitochondrial DNA is present in multiple copies in the cell, shows a gene content strongly conserved across animals, with very few duplications, no intron, and very short intergenic regions and displays some highly conserved regions (e.g. ribosomal DNA), on which PCR primers can be designed (Gissi *et al.* 2008). Due to these characteristics mtDNA yields good amplifications and thereby represents an easy and economic molecular marker to use. mtDNA, moreover, shows an high mutation rate rates that confers on average high information content per base pair sequenced and it is generally transmitted exclusively through the maternal line, with almost no recombination at all. These features considerably simplifies the

representation and analysis of within-species variation data and so, with the widespread use of direct sequencing, mtDNA has become the marker of choice for phylogenetic and phylogeographic studies (Brito & Edwards,, 2009; Gissi *et al.* 2008) .

The control region of mtDNA or D-loop (Displacement Loop) present in all vertebrates is a non–coding region of about 1 kb, placed between the phenylalanine tRNA ($tRNA^{Phe}$) and the proline tRNA ($tRNA^{Pro}$) that exhibits a comparatively higher level of variation than protein–coding sequences, due to reduced functional constraints and relaxed selection pressure. The high informative level that results make it a good maker of choice to investigate population history of species, especially when we lack preliminary information about it.

SSRs (Simple Sequence Repeats) or microsatellites, due to the high degree of polymorphism that are able to highlight, are an important and widely used tool for population genetics analysis.

They consist of short sequences (1-4bp) highly repeated in tandem in the genome and are present in all eukaryotes both in the nuclear DNA and in the organelles (mitochondria and chloroplasts).

It is possible to distinguish three classes of SSRs (Jarne & Lagoda, 1996):

- Pure: formed by a single repeat sequence (for example CACACACACACACACACACA)
- Compounds: formed by the repetition of multiple sequences (for example CACACACACAGAGAGAGAGA)

- Broken: consist of repeated sequences interspersed with non-repeated bases (for example CACATTCACACATTCATTCA)

SSRs are codominant markers, thus allowing us to discriminate homozygous from heterozygous individuals at a given locus. The polymorphisms identified is given by variations in the number of repetitions due to a particular replicative phenomenon called "Slipped strand mispairing" in which the DNA polymerase adds one or more repetitions of the sequence that is replicating, or due to phenomena of unequal crossing over. The alleles at a given locus are detected by the classical technique of PCR in which primers are designed according to sequences flanking the microsatellites. The technique is simple, however, very laborious in the initial phase, when dealing with a new species, for which no sequence information is known.

In this case preliminary work must be done, isolating and sequencing microsatellite sequences from a specifically designed library. However, in the vast majority of cases, it is possible to amplify SSR markers by means of primers already designed in an evolutionary close species. Infact, it appears that the sequences flanking the SSRs have been preserved in the different taxa.

In this work, to this purpose, we used primer information derived from the related species *O. niloticus*.

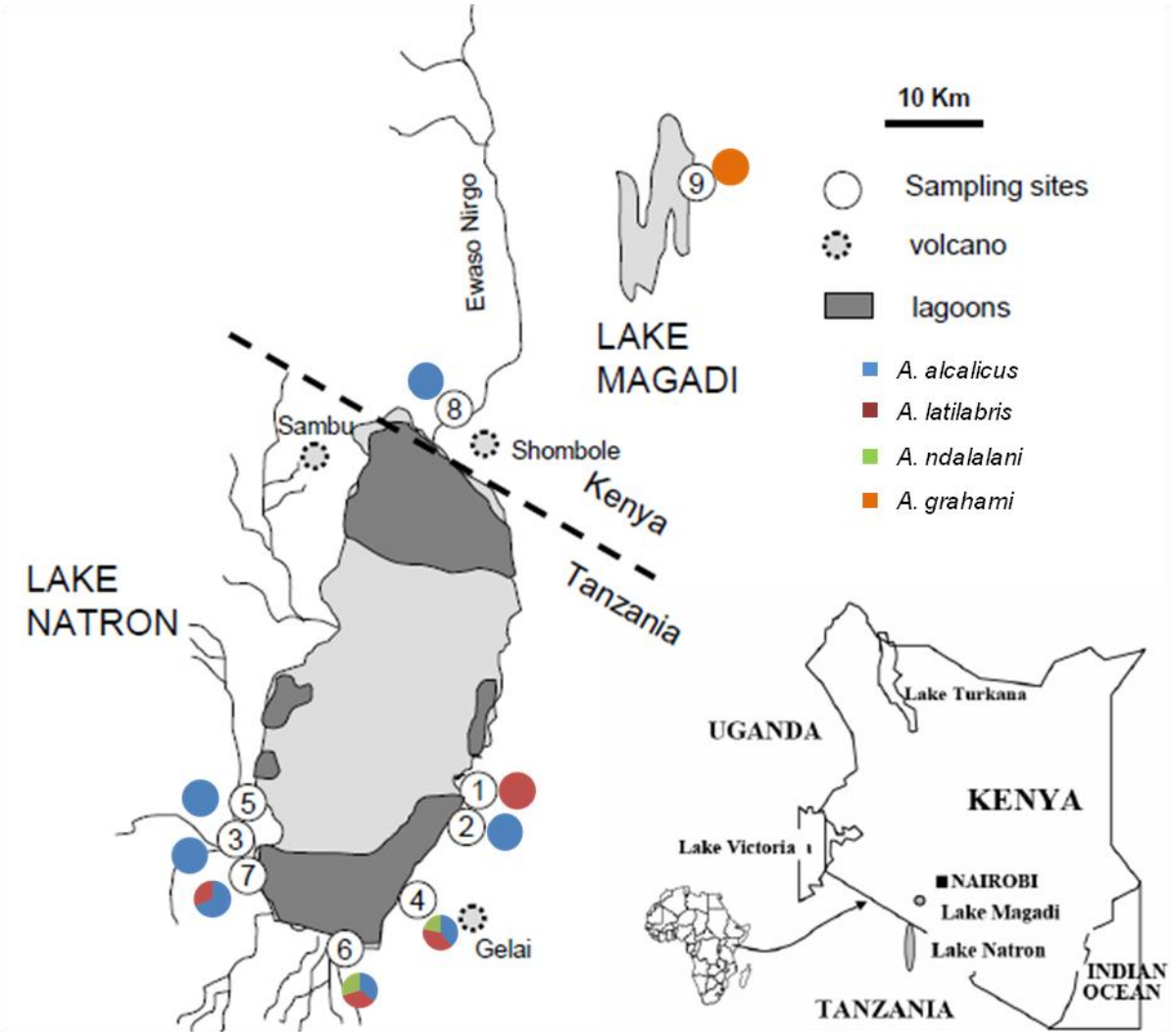
MATERIALS AND METHODS

Fish sampling and DNA extraction

310 samples of *Alcolapia* fishes were collected between April and December 2009 from eight populations placed around the Lake Natron and from one population located in the north-east part of Lake Magadi (Fig.3, Tab.1), using a combination of seine and cast netting (method dependent upon habitat). Morphological identification was carried out in the field and the fishes were assigned to the variously designated, species, subspecies or morphotypes *A. alcalicus*, *A. grahami*, *A. ndalalani* or *A. latilabris*, according to the descriptions of Seegers & Tichy (1999). The Lake Magadi population was constituted only by *A. grahami* whilst the Lake Natron populations were composed by *A. alcalicus*, *A. ndalalani* and *A. latilabris*

Fishes were anesthetized and a tissue sample (pelvic fin) was taken from each one and stored in ethanol (98%), then the fishes were released.

Total DNA was obtained by sodium chloride extraction and ethanol precipitation after initial proteinase K digestion (extraction protocol adapted from Aljanabi & Martinez, 1997).



Code	Sampling localities	Basin	Fish community	S	E
1	Top Eastern Spring	Natron	*Al*	02 27.366	036 05.287
2	Second Esatern Spring	Natron	*Aa*	02 26.081	036 05.873
3	Pool and Brook on river	Natron	*Aa*	02 30.653	035 53.144
4	Spring South-East	Natron	*Aa, Al, An*	02 31.629	036 02.762
5	Lagoon South-West	Natron	*Aa*	02 27.353	035 53.478
6	Matt's Brook and spring	Natron	*Aa, Al, An*	02 35.415	036 00.427
7	Olomotony	Natron	*Aa, Al*	02 35.400	035 55.117
8	Shombole	Natron	*Aa*	02 06.495	036 00.379
9	Magadi	Magadi	*Ag*	01 53.270	036 18.494

Figure 3 and Table 1 Sampling locations in the Lake Natron and Magadi catchment, and composition of the fish community, where *Aa* = *Alcolapia alcalicus*; *Al* = *Alcolapia latilabris*; *An* = *Alcolapia ndalalani*; *Ag* = *Alcolapia grahami*. Geographic coordinates are given. The map is modified by Seegers et al., 1999.

mtDNA amplification and sequencing

Primers L-15926 (5'- TCAAAGCTTACACCAGTCTTGTAAACC-3') (Seegers *et al.* 1999) and Ormt-917LP (5'- CTTATGCAAGCGTCGATGAAA-3') (Nagl *et al.*, 2001) were used to amplify by polymerase chain reaction (PCR) a fragment of the mitochondrial control region. PCR amplifications were performed using a Mastercycler EPgradient thermal cycler (Eppendorf®).

PCR reactions were carried out in 25 µl reaction volume containing: 10 ng of DNA, 1 unit of Taq DNA Polymerase (Sigma-Aldrich, Germany), 1X reaction buffer, 2.5 mM $MgCl_2$, 200µM of each dNTP (Promega Corporation USA) and 12.5 picomoles of each primer, under the following conditions: 94°C (2 min), then 35 cycles at 94°C (45 s)/ 62°C (60 s)/ 72°C (60 s) and final extension at 72°C for 10 min. Amplification products were purified using the Montage PCR_{96}Millipore Cleanup Kit (Millipore) and two microliters of purified PCR product were used as template in the cycle sequencing reaction, with L-15926 primer and DYEnamic ET Dye Terminator Cycle Sequencing Kit (GE Healthcare). The cycle-sequenced products were purified using the Montage SEQ_{96} Sequencing Reaction Cleanup (Millipore) and analysed on an automated capillary DNA sequencer (MegaBace ® 500, GE Healthcare).

Microsatellites amplification and genotyping

Seven heterologous microsatellite loci, developed in *Oreochromis niloticus* (UNH843, UNH851, UNH874, UNH891, UNH958, UNH989 and UNH915; Carleton *et al.*, 2002) and selected as they yielded a simple amplification pattern when used to genotype the samples being studied, were chosen for population analysis (Tab.2).

SSR marker	5' – 3' sequence	Ta	LG	Range size expected	Genbank accession n°
UNH843	**F:** CGTTCTACTCTGAAGAAAGACATGA **R:** CCACTCGACGGACGTTTTAG	55°C	9	108÷124	G68182
UNH851	**F:** TGGCTGTCACATAATCTGGTG **R:** CCTATCGTCGGTGATTGGTC	55°C	7	110÷126	G68188
UNH874	**F:** AGTAAAATGGGCGAACGTGT **R:** TGAAGCTGGGAGTTTCCTGT	59°C	12	202÷236	G68202
UNH891	**F:** GCAGCCCAAATTGTTGAGTC **R:** TGCTGCTGAAGGTTCTAAAGG	55°C	11	158÷184	G68213
UNH915	**F:** CAAGGGCAAGGGTAATTTTG **R:** CGCAGATTAGTAACATCAAGCA	55°C	10	145÷159	G68227
UNH958	**F:** TGCTTGTCTGACAGGGAAAA **R:** TTCCAGCGAGTCACACATTC	55°C	15	143÷155	G68252
UNH989	**F:** CCACCTCATACACGCAAATG **R:** CAGCAGCAGCTGTCCACTAA	57°C	14	140÷152	G68269

Table 2. Characteristics of microsatellite marker used in this study.

Fluorescent labelling of PCR fragments was carried out as suggested by Schuelke (2000), using 6-FAM, PET, VIC and NED dyes. PCR amplifications were performed in a final volume of 20 µL containing 10 ng of genomic DNA, 0.5 units of GoTAQ® Flexi DNA Polymerase (Promega Corporation USA), 1x reaction buffer, 1.5–2mM, depending on the primer used (1.5mM for

UNH989, 2mM for all the others), of $MgCl_2$, 200 µM of each dNTP (Promega Corporation USA), 3.2 picomoles of each reverse and labelled-M13(-21) primer and 0.8 picomoles of the forward primer. Conditions of the PCR amplification were as follows: 94°C (5 min), then 30 cycles at 94°C (30 s) / Ta°C (45 s) /72°C (45 s), followed by 8 cycles 94°C (30 s) /53°C (45 s) / 72°C (45 s), and a final extension at 72°C for 10 min. The amplification products were run on a capillary ABI3730XL sequencer (Applied Biosystems) and the raw data were analysed using the Peak Scanner™ Software Version 1.0 (Applied Biosystems) to score the single-fish genotypes.

Data analysis

mtDNA D-loop analysis

The DNA sequences were edited and aligned using ClustalX 1.81 software (Thompson *et al.*, 1997) and Bioedit (Hall, 2003). The sequences have been deposited in the GeneBank database (Accession Nos. HQ637440-HQ637456).

In the first step, we reconstructed the *Alcolapia* phylogeny by adding to our dataset 36 mtDNA D-loop haplotypes of *Alcolapia* (20 from Seegers *et al.*, 1999; 6 from Wilson *et al.*, 2000; 10 from Nagl *et al.*, 2001) and five haplotypes from two African tilapiines species: three haplotypes of *Oreochromis amphimelas* (HT9102, HT9114 and HT9325) from Lake Manyara and Lake Singida (Tanzania) and two haplotypes of *O. niloticus vulcani* (HT7515 and BJMVUL1F) (Nagl *et al.*, 2001), a variant of *O. niloticus* endemic to crater lakes on the Central Island of Lake Turkana (Kenya) and adapted to life in temperature of 37-40°C. The tree was rooted by a haplochromine sequence of *Astatoreochromis alluaudi* (Asal6744, Nagl *et al.*, 2000). The phylogenetic tree was built by Maximum Likelihood (ML), as implemented by the PhyML program package (Guindon & Gascuel, 2003), in which HKY85 (gamma distribution = 0.2516) (Hasegawa *et al.*, 1985) was used as the best-fit evolutionary model determined within a Hierarchical Likelihood Ratio Tests (hLRTs) framework by ModelTest 3.06 software (Posada & Crandall, 1998). The Maximum Parsimony Consensus tree (MP) was reconstructed

using PAUP* 4.0b10 program package (Swofford, 2002). Confidence in the nodes was assessed using bootstrap procedures with 1,000 replications (Felsenstein 1985). The Bayesian approach (Huelsenbeck *et al.*, 2001) within the MrBayes program ver.3.1 (Ronquist & Huelsenbeck, 2003) was also performed. Four Markov chains from random trees were started and run for 500,000 generations, with the first 125,000 generations (1,250 trees) discarded as the burn-in. The analysis was run independently four times. The demographic history was explored by constructing the haplotype network and the 'mismatch' distribution analyses of genetic pairwise differences between sequences were performed, testing the sudden expansion model with a non-linear least square regression (Schneider & Excoffier, 1999). The significance of the sum of square deviations (SSD) was assessed with bootstrap approach (5,000 simulations), as implemented in Arlequin 3.1 (Excoffier *et al.*, 2005). The nucleotide (π) and haplotype (h) diversity were calculated. To infer the amount of genetic differentiation between morphotypes pairwise mtDNA F_{ST} were also estimated (Reynolds *et al.*, 1983).

Microsatellite markers analysis

Allele frequencies and observed (H_o) and expected (H_e) heterozygosities were estimated at each locus for all populations. Departure from Hardy-Weinberg equilibrium for each population and each locus was assessed by a Fisher's exact test using the Markov Chain algorithm (Guo & Thompson, 1992). The

non-random association of the alleles between pairs of loci, or linkage disequilibrium (LD), was tested at the single population level and across all populations with Fisher's exact test. Genetic differentiation between populations was estimated by both Wright's F-statistic (F_{ST}) (Weir & Cockerham, 1984), based on differences in allele frequencies, and R-statistic (R_{ST}) (Slatkin, 1995), the F_{ST} analogue for microsatellites, to include the molecular information relative to the size of differences between the alleles. R_{ST}, infact, is based on the stepwise mutation model and so is able to detect differentiation events older than those detected by F_{ST}. Their significance was tested by a permutation procedure based on 1,000 permutations of the data.

Mantel test (1967) estimates the correlation between two distance matrices, typically a genetic distance matrix and a geographic distance matrix, with tests of significance by permutation. When spatial genetic structure is strong and extends over the full geographic range, significant positive correlation can be detected by this test. A Mantel test was applied to the matrices of pairwise (F_{ST} /1- F_{ST}) and log-transformed geographical distance between Lake Natron populations to assess isolation-by-distance, namely the model under which genetic differentiation between populations is the result of drift. The indirect estimate of the number of migrants (Nm) was performed by the private allele method (Barton & Slatkin, 1986).

In order to identify the genetically homogeneous groups in our population sample, the Bayesian clustering method proposed by Pritchard and colleagues (2000), implemented by a Markov Chain Monte Carlo (MCMC) algorithm, was applied The estimate of K (the genetically homogeneous

inferred populations) was performed based on ΔK, the second order rate of change of the likelihood function with respect to K, as suggested by Evanno *et al.* (2005); the modal value of the distribution of ΔK is located at the real K. Ten runs were carried out for each K value from 1 to 12 (the number of real populations plus 3) tested. The analysis (admixture model, correlated allele frequencies between populations) was run with a length of burn-in period of 10,000 and with 50,000 MCMC replications after burn-in.

Analysis of molecular variance (AMOVA) is a method of estimating population differentiation directly from molecular data and testing different hypotheses about such differentiation; this statistical procedure allows the hierarchical partitioning of genetic variation (total F_{ST}) among populations and regions. AMOVA was performed to partition the total genetic variation among regions suggested by the Bayesian analysis and between populations within regions (Excoffier *et al.*, 1992; Huff *et al.*, 1993).

The test of significance for the AMOVA was carried out on 1,000 permutations of the data.

Multivariate analysis was employed, in order to get an easy-to-read visualization of the relationship between the fishes of the flock, applying the Factorial Correspondence Analysis (FCA) implemented by software Genetix (Belkhir *et al.*, 1996). This method of analysis permits to represent individuals in a multivariate Euclidean space by transforming their genotypes trough a principal component approach.

Spatial Autocorrelation Analysis (ACS), as implemented by GenAlEx v.6 software (Peakall & Smouse, 2006), was carried out only for Natron populations, in order to detect a possible genetic structure at basin level.

ACS technique measures the degree of similarity between samples for a given variable as a function of spatial distance, if the variable is the genotype, then what we get is a description of spatial pattern of genetic variation.

The autocorrelation coefficient generated (r) by GenAlEx software is a proper correlation coefficient, bounded by [–1, +1] and in the absence of autocorrelation takes the value 0; r provides a measure of the genetic similarity between pairs of individuals whose geographic separation falls within the specified distance class.

The autocorrelation coefficient r is estimated through a series of permutations and bootstrap procedures performed on the dataset under consideration. Confidence interval limits at 95% are calculated by 1,000 random permutations of the data (random shuffling of all individuals among the geographic locations), under the null hypothesis of no spatial structure; values of the 25th and 975th ranked r values are taken to define respectively the upper and lower bounds of the confidence interval. To reject the null hypothesis, and then to say that there is a spatial genetic structure, are considered statistically significant those r values that fall outside the confidence interval with a probability $P < 0.05$.

The software also estimates a 95% confidence interval around the estimated value of r by drawing with replacement from a set of pairwise comparisons for each distance class. For each of 1,000 bootstrap cycles, the value of r is

calculated for each distance class. The 25th and 975th ranked *r values* are then taken to define 95% confidence interval. When the bootstrap confidence interval contains the value 0, you can accept the null hypothesis of no spatial structure. The bootstrap test is less powerful than the permutation test, since the number of samples per distance class is much less than the number of comparisons n (n – 1) / 2, used during the permutations.

The results of spatial autocorrelation are graphically presented by autocorrelogram, a graph showing the genetic correlation between individuals as a function of distance and that is a valuable tool to represent the spatial genetic structure. By means of ACS, infact, it is possible to detect the genetic patch diameter, i.e. the distance at which the samples become genetically independent, a parameter useful to define geographic range of independent samples, and, consequently, to define small units for conservation and management (Diniz-Filho & Telles, 2002).

The software packages used to analyse the genetic data were Genepop (Raymond & Rousset, 1995), Genetix, GenAlEx v.6, Rst calc (Goodman, 1997) and Structure v.2.1 (Pritchard *et al.*, 2000).

RESULTS

mtDNA

The length of D-loop mtDNA sequences resulted in 350 bp. The sequences differed from each other by 1 to 18 substitutions, with 10 parsimony informative polymorphic sites. The mean G+C content was 32 % and the overall transition/transversion ratio was 1.17. A total of 17 haplotypes were identified from the 69 *Alcolapia* fishes analysed; a detailed list of sampling localities, haplotype distribution and haplotype morphologic assignment as *A. alcalicus*, *A. ndalalani*, *A. latilabris* and *A. grahami* are given (Table 3). The sequences have been deposited in the GeneBank database (Accession Nos. HQ637440-HQ637456).

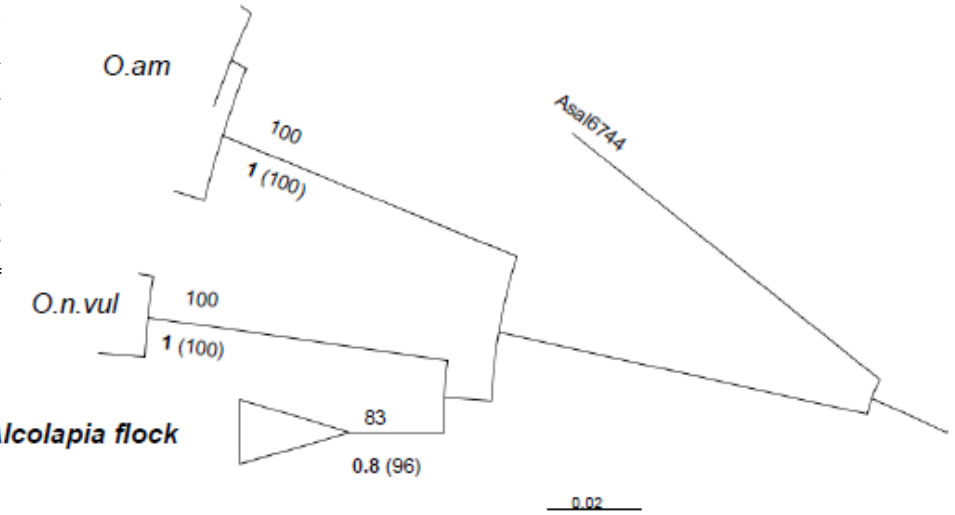
The different *Alcolapia* phylogenetic trees, constructed using the enlarged dataset and applying the different methods as detailed above, showed a congruent and well supported topology (Fig. 4A). The *Alcolapia* genus resulted in monophyletic structure with similar values of uncorrected genetic distance from *O. amphimelas* and *O. niloticus vulcani*, 9.7% (SD 0.3) and 9.1% (SD 0.4), respectively. By contrast, the *Oreochromis* genus, represented by the two selected species and distributed only in close alkaline lakes, evidenced a divergence distance of 12% (SD 0.3). The phylogenetic relationships within *Alcolapia* genus remained unresolved due to the low values of intraspecific genetic distance ranging from 0.3% to 1.7%.

In our haplotype network, all *Alcolapia* haplotypes from Lake Natron and Lake Magadi were placed in the same network, depicting a radial structure from a central haplotype (*2lat*) (Fig. 4B) which showed the highest frequency within our dataset (0.6) and was widespread in both lakes and shared by all morphotypes (Table 3). Moreover, the haplotypes differed for a maximum number of 11 mutation steps and few apomorphic positions created haplotype in single copy, contributing to a general star-phylogeny structure (Fig. 4B), in which *2lat* was the most probable *Alcolapia* flock ancestor. Only three haplotypes (*03, 01* and *34*), linked to the *2lat* by 4 to 7 mutation steps, generated a weak independent branch (Fig. 4B), prevalently distributed in Lake Magadi. For 5 out of 17 halpotypes identified in this work we have also found a complete correspondence with haplotypes identified in previous studies by Seegers *et al.* (1999) and Wilson *et al.* (2000; 2004).

Pop	N	*50*	*3lat*	*68*	*53*	*70*	*305*	*7*	*42*	*2lat*	*241*	*243*	*398*	*238*	*125*	*34*	*01*	*3lat*
1	4									4 l								
3	13		7 a	1 a		1 a				4 a								
4	6						1 a			3 a 2 l 1 n								
5	5									4 a					1 a			
6	18		1 n							4 a 3 n 6 l	1 a	1 a	1 l	1 a				
7	12	2 a	2 l 1 a	1 a			1 l	1 a	2 a 2 l									
8	5		2 a							3 a								
9	6									3 g						1 g	1 g	1 g
total	**69**	2	12	1	1	1	1	1	1	41	1	1	1	1	1	1	1	1

Table 3 Haplotype distribution in the *Alcolapia* based on 350bp length of mtDNA D-loop. Number of sequenced individuals per sampling site and morphological attribution are reported (*a* = *A. alcalicus*; *l* = *A. latilabris*; *n* = *A. ndalalani*; *g* = *A. grahami*).

Figure 4A D-loop mitochondrial DNA analyses. (A) Maximum Likelihood analysis (with HKY85 and gamma distribution = 0.2516). Number above the branches represent values from ML method obtained with PhyML; numbers below represent Bayesian inference (in bold) obtained with MrBayes program and the Maximum Parsimony values (in brackets) obtained with PAUP. This analysis combines sequence data from *Oreochromis* species from East African alkaline lakes (*O. amphimelas*= *O. am* and *O. niloticus volcani*= *O. n.vul*).



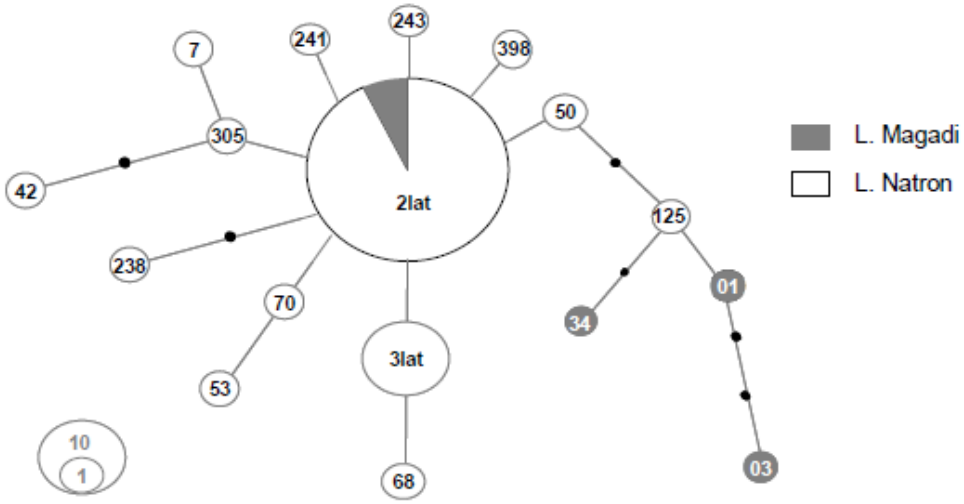


Figure 4B D-loop mitochondrial DNA analyses. Unrooted haplotype network of the *Alcalicus* flock. Haplotypes from Lake Natron are white and haplotypes from Lake Magadi are grey. The sizes of the haplotypes reflect the number of specimens sharing the same haplotype (see scale in the lower left corner). Interior haplotypes not detected in the sample are represented by small black circles.

The demographic pattern, characterised by low levels of genetic diversity, was explored through mismatch distribution analysis. The pairwise difference wave was unimodal and compatible with the "sudden expansion" model (SSD = 0.004, p = 0.49; r = 0.057 P = 0.82; τ = 0.467, θ_0= 0.631 for $\theta_1 \rightarrow \infty$) Therefore, low values of nucleotide diversity (π = 0.004 SD 0.003) and high value of haplotype diversity (H = 0.63 SD 0.06) support the starburst radiation of the *Alcolapia* flock from a low number of founders.

The molecular indices and the mismatch distribution analyses were also calculated for each morphotype, according to the morphological discrimination. The pairwise difference waves of the mismatch distribution analyses showed similar unimodal trends (Fig. 5), evidencing rapid range expansion traits in *A. alcalicus*, *A. ndalalani* and *A. latilabris*, also supported by high values of H (0.4 ± 0.14 to 0.7 ± 0.01), and low values of π (0.001 to 0.003). Conversely, *A. grahami* showed the highest value for nucleotide diversity (π = 0.013 SD 0.009) with more fragmented pairwise distribution wave (Fig. 5), revealing a more stable demographic equilibrium in comparison to the Lake Natron morphotypes.

The genetic differentiation between morphotypes pairwise, assessed by mtDNA F_{ST}, has revealed no differentiation between the three morphotypes endemic of lake Natron (see Tab. 6), as suggested by non significant F_{ST} values found. High and significant values of mtDNA F_{ST}, however, have been reported for comparisons between *A. grahami* and *A. alcalicus* (0.329) and between *A. grahami* and *A. latilabris* (0.405).

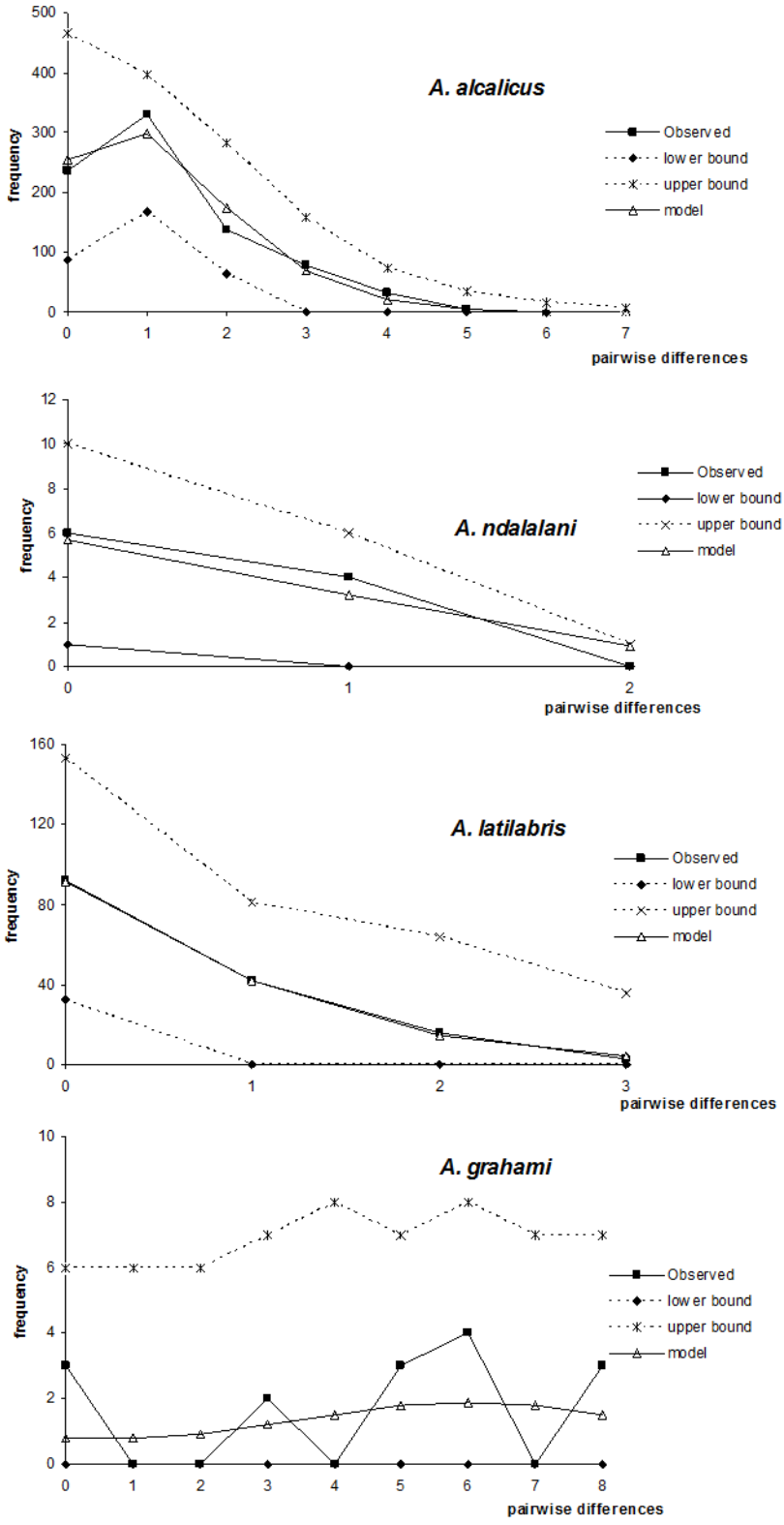


Figure 5 Results of mismatch distribution analyses using ARLEQUIN. Trends in *A. alcalicus*, *A. ndalalani*, *A. latilabris* and *A. grahami* morphotypes. The upper and lower bound curves are 5 and 95 percentile values of 5,000 simulations.

Microsatellites

Genetic variability

A total of 310 individuals of the *Alcolapia* flock were analysed using seven heterologous microsatellite markers, and 147 distinct alleles were identified. All the loci studied are polymorphic: the number of detected alleles per locus ranges from 3 (locus UNH851) to 48 (locus UNH989) (Table 4). Genetic diversity, as measured by Nei's heterozygosity *(He*), ranges from 0.068 (locus UNH891) to 0.920 (locus UNH989) (Table 4). Average *He* estimated for each population maintains a medium level with a more restricted range (0.505< *He* < 0.626).

Microsatellite locus	N_A	*Ho*	*He*	Size range (bp)
UNH843	8	0.553	0.529	116 ÷ 136
UNH851	3	0.337	0.338	122 ÷ 128
UNH874	26	0.412	0.447	210 ÷ 274
UNH891	5	0.029	0.068	162 ÷ 172
UNH915	33	0.834	0.885	139 ÷ 209
UNH958	24	0.740	0.886	161 ÷ 211
UNH989	48	0.651	0.920	152 ÷ 256

Table 4 Number of alleles (N_A), observed (*Ho*) and expected (*He*) heterozygosity, and allele size range detected for each locus.

The Hardy-Weinberg equilibrium was verified for all the loci and populations by testing the departure of F_{IS} from zero under the null hypothesis. Locus UNH989 shows significant deviations from Hardy-Weinberg equilibrium,

associated with positive values of F_{IS} in all the populations sampled, with the exception of population 9 sampled in the Lake Magadi. Population 9, on the other hand, is the only one at H-W equilibrium for all the loci analysed.

Genotypic disequilibrium

Genotypic disequilibrium was analysed for all pairs of SSR markers for each population and across all populations. A significant departure from equilibrium at the 5 % level was found for almost all pairs of loci within population. An exception to this trend is represented by population 9, for which three loci only (UNH874, UNH958 and UNH915) show LD in all three comparisons.

Isolation by distance

A Mantel test was applied on the matrix of pairwise F_{ST} /(1-F_{ST}) distances between populations and the matrix of log-transformed geographic distance between populations to assess the presence isolation-by-distance. The test, however, was not significant (z = 2.77, g = -0.059, p = 0.515) .

The indirect estimate of the number of migrants (*N*m), performed by the private allele method (Barton & Slatkin 1986), resulted in *N*m=3.62.

Genetic differentiation among populations and population structure

The genetic divergence between populations was measured using both F_{ST} and $R_{ST,}$ so as to include in the differentiation estimates the molecular information relative to the size of differences between the alleles (Tab. 5), and their significance was tested by a permutation procedure based on 1000 permutations of the data.

Pop.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1		0.044	0.055	0.014	0.035	0.028	0.050	0.042	0.166
2	0.058		*0.008*	0.050	*0*	0.024	0.040	*0.007*	0.061
3	0.087	*0*		0.064	0.009	0.036	0.040	0.020	0.101
4	0.061	0.072	0.098		0.052	0.022	0.061	0.051	0.153
5	0.058	*0.017*	*0.012*	0.081		0.029	0.031	0.011	0.076
6	0.038	0.036	0.043	0.125	0.05		0.040	0.029	0.112
7	0.041	*0.020*	0.054	0.129	0.024	0.059		0.040	0.149
8	0.081	*0.043*	0.056	0.094	0.048	0.080	0.060		0.082
9	0.210	0.034	0.101	0.205	0.077	0.159	0.153	0.122	

Table 5 F_{ST} (above diagonal) and R_{ST} (below diagonal) values for each population pair. Values in italics does not significantly differ from zero, according to the results of a permutation test.

Only 3 F_{ST} values out of 36 were not significantly different from zero: 2 and 3, 2 and 5, 2 and 8. For R_{ST} values 6 comparisons out of 36 were not significantly different from zero: the same three of F_{ST} and, in addition, 2 and 7, 2 and 9 and 3 and 5. The maximum F_{ST} and R_{ST} values were found between population 9, the Magadi population, and all the other populations, the Natron populations, whereas differentiations values measured inside Lake Natron were very low.

To obtain a graphically clear representation of the relationship between the populations studied, we used multivariate analysis in the form of Factorial Correspondence Analysis. Each fish was represented in the multivariate space by its genotype at the seven loci tested and the principal coordinate method was applied to visualize into the three-dimensional space the Euclidean orthogonal distances between individuals. The results are shown in Figure 6, the three axis explaining 19%, 16% and 13% of the total variability, respectively. The main feature of this graph is the clear separation of the Magadi population from all others, which result in a dense cloud in the centre of the graph. It is worth noting that the Magadi population is also the one showing the highest pairwise F_{ST} values.

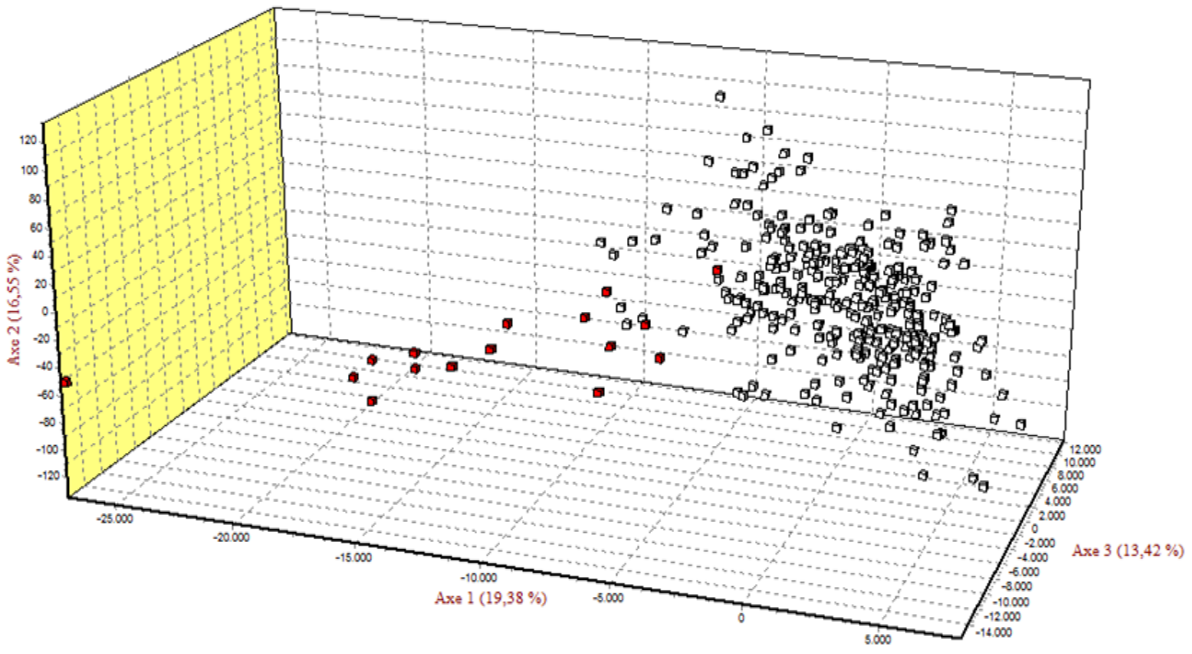


Figure 6 Factorial Correspondence Analysis based upon the multilocus genotype of the single fhises (squares). Fishes from the Magadi population are in red.

The overall genetic differentiation of the flock was low, as indicated by F_{ST} = 0.044 and R_{ST} = 0.064, but significant (P<0.001). A confidence interval at the 95% level for F_{ST} was estimated by running 1000 bootstrap replicates, resulting in 0.023< F_{ST} < 0.079; the same estimate for R_{ST} gave a confidence interval of 0.060< R_{ST} <0.10. The higher and significant values for R_{ST} indicates that differentiation of actual population has started in a non recent period.

In order to identify the genetically homogeneous groups in the *Alcolapia* flock a Bayesian clustering method as implemented by STRUCTURE (Pritchard *et al.*, 2000) was applied. The rationale for this was to estimate the most probable number K of homogeneous gene pools from which the fishes of the actual *Alcolapia* flock could have arisen. We found a signal at K = 2 (average LnP(D)= -7662.33, ΔK=17.7), suggesting the presence of two distinct gene pools. The level of admixture in each of the populations is however very high, thus suggesting a weak genetic structure of the *Alcolapia* flock. The less admixed population is population 9, made up of *A. grahami* only and located in a different basin than the others.

Based upon analysis of the population structure, the hypothesis was tested by AMOVA that a major differentiation factor for the populations of the *Alcolapia* flock is the division into two geographic regions, corresponding to Natron and Magadi Lakes. The results of AMOVA show that the among regions component is statistically significant (P=0.001) and accounts for 8% of the total variance.

Since population 9 is composed exclusively by *A. grahami* individuals, we tested the genetic structure of the Natron populations only. Bayesian analysis yielded in this case K=3 (average LnP(D)= -6881.05, ΔK=167.84) as the most probable value, but the level of admixture in the eight populations is so high not to define a clear structure. Besides the AMOVA based on the partition of the populations based on K=3 identified an among regions component explaining 1% only of the total genetic differentiation.

When Spatial Autocorrelation Analysis (ACS) is carried out only for Lake Natron, detects the presence of a spatial pattern with a genetic patch around 12 km, value which is compatible with the dimension of the lake. The shape of the graph (Fig.7) for *r* (autocorrelogram), a positive and significant correlation coefficient in the first class of distance but not significant in the last class, is typical of a stabilizing profile. The intercept in this case is useful to determine the best sampling strategy for conservation programs, because it give us an indication of the minimum distance between samples needed to conserve and assess genetic diversity with maximum efficiency and lower cost. Collecting sample situated around this minimum distances, infact, maximizes the possibility of assessing more genetic heterogeneity and thus avoiding pseudoreplication of the variability. All samples situated at a geographic distance lower than the intercept, moreover, could be considered a single genetic unit for conservation and management (Diniz-Filho & Telles, 2002).

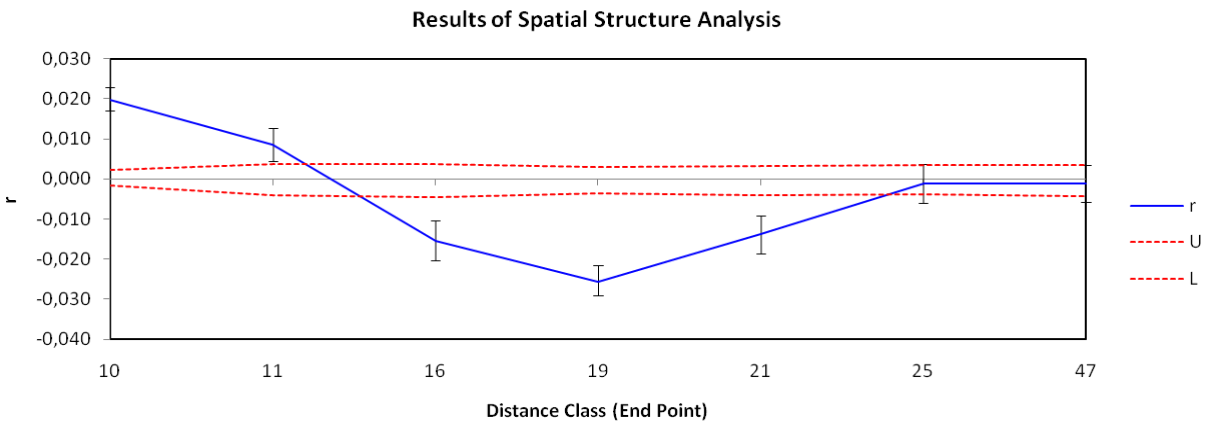


Figure 7. Autocorrelogram relative to the spatial genetic structure of the Lake Natron populations. A genetic "patch" of approximately 12 kilometres is revealed.

Thus it is possible that the presence of the Magadi population in the dataset is the major factor shaping the genetic differentiation of the *Alcolapia* flock.

To shed more light on this issue, we finally analysed by STRUCTURE the flock divided according to the prevalent morphotype of each population.

In this case we have found a signal for K=3 (average LnP(D)= -7371.07, ΔK=468.21) even though with a general high admixture level, especially for the three morphotype endemic of Lake Natron.

Pairwise differentiation by F_{ST} and R_{ST} indexes was carried out at morphotype level too (Tab 6). All the results were statistically significant. Comparisons between *A. grahami*, the morphotype endemic of the Lake, and the other three morphotypes shows the highest values both of F_{ST} and R_{ST} indexes.

basin	morphotype comparison	mtDNA F_{ST}	Microsatellite F_{ST}	Microsatellite R_{ST}
Within Natron	*alcalicus – ndalalani*	-0.103	0.014*	0.032*
	alcalicus – latilabris	0.002	0.028*	0.044*
	ndalalani - latilabris	-0.104	0.028*	0.050*
Magadi vs. Natron	*grahami – alcalicus*	0.329**	0.089*	0.097*
	grahami - ndalalani	0.177	0.091*	0.108**
	grahami – latilabris	0.405*	0.162*	0.218*

Table 6 Morphotype pairwise F_{ST} (mtDNA sequences and microsatellite loci) and R_{ST} (microsatellite loci) values.
*P<0.001; **P<0.05; all others were not significant

DISCUSSION

In this work we analysed 310 samples of *Alcolapia* fishes collected from eight populations placed around the Lake Natron (Tanzania) and from one population located in the north-east part of Lake Magadi (Kenya). A fragment of 350bp of mtDNA D-loop was sequenced and used for phylogenetic analyses. The sample were also genotyped at seven heterologous SSR loci and the resulting dataset was used for population analyses.

The phylogenetic reconstruction based on mtDNA D-loop sequences revealed monophyly in the *Alcolapia* flock of Lakes Natron and Magadi. This result is in complete agreement with previous molecular phylogenetic studies conducted both at local scale on the same *Alcolapia* flock (Seegers *et al.*, 1999) and on a wider collection comprising Tilapiine from several African Lakes (Nagl *et al.*, 2001). The monophyletic pattern detected suggests that the evolution of this flock was driven by the geographic and geological activity of the region including frequent climatic upheavals impacting lake and river levels (Johnson *et al.*, 1996; Verheyen *et al.*, 2003; Stager & Johnson 2008; Elmer *et al.* . 2009) and changes in hydrology due to volcanic activity (Roberts *et al.*, 1993; Schwarzer *et al.*, 2009). These changes then influenced the evolution of cichlid species in the region through forming isolated lakes in closed basins (e.g. Lake Magadi, Natron and Manyara). For example, whilst the *Alcolapia* and *O. amphimelas* in Lake Manyara are separated by only a short geographic distance (approximately 80 km), their

evolution has been independent following the historical separation of the fish populations (Trewavas, 1983). The volcanic activity would have also favoured the development of water of high alkalinity in these lakes (e.g. increased sodium concentrations - Trewavas, 1983), an environmental process likely to have driven the evolutionary adaptation of the cichlid species towards similar ecological strategies in being tolerant to conditions of high alkalinity and salt concentration, and high water temperatures (Tichy & Seegers, 1999; Nagl *et al.*, 2001).

Geological and molecular data available from previous studies (Coe, 1966; Pörtner *et al.*, 2010; Tichy & Seegers 1999) suggest that the *Alcolapia* spp. from lake Natron and Magadi originated from an ancestral population inhabiting the paleolake Orolonga before its split. Infact fossil tilapia, very similar in morphology to present-day *A. grahami*, has been found in deposits approximately 20–40 m above the current surface of the two lakes.

Analyses of mitochondrial DNA sequences, besides, have identified an haplotype, termed *A1* by Seegers *et al.* (1999) and *B* by Wilson *et al.* (2000, 2004), that is shared by all the populations sampled in the two lakes. The authors of these studies suggest this haplotype, shared by all the *Alcolapia* species, represents the ancestral haplotype of the fishes resident in Paleolake Orolonga.

Our molecular data support these findings.

The haplotype network obtained by TCS analysis of the 17 mtDNA D-loop haplotypes, infact, shows a star-like structure with the central position occupied by the haplotype *2lat.* This haplotype is the most frequent in our

dataset, is widespread in both lakes and shared by all morphotypes; *2lat* moreover corresponds to the Orolonga haplotype (named *B* or *A1*) previous identified as the ancestor. The radial structure comprises haplotypes that differed for a maximum number of 11 mutation steps and few apomorphic positions from the central one, as reflected by low values of nucleotide diversity ($\pi = 0.004$) and high value of haplotype diversity ($H = 0.63$). These findings should not be surprising, given that Strecker (2006) observed similar haplotype diversity and distribution in a *Cyprinodon* flock endemic of a tropical lake from Mexico (characterised by harsh environmental conditions and a recent origin) and proposed an origin from a small founding population. The unimodal trend in the mismatch distribution analysis of the pairwise differences, on the other hand, strengthened the hypothesis of sudden demographic expansion occurred from a limited number of founders. The presence in the network of the weak independent branch composed by haplotypes prevalently distributed in Lake Magadi, otherwise, suggests that the starburst radiation of the *Alcolapia* flock from the Orolonga haplotype must then have evolved independently in the two basins since their separation (between 8,000 and 9,000 years ago; Seegers *et al.*, 1999; Nagl *et al.*, 2001; Wilson *et al.*, 2004). This hypothesis is supported by the molecular indices and the mismatch distribution analyses calculated for each morphotype. The pairwise difference waves of the mismatch distribution analyses showed, infact, similar unimodal trends, supported by high values of H and low values of π, for *A. alcalicus*, *A. ndalalani* and *A. latilabris*, evidencing rapid range expansion traits in the morphotypes endemic of Lake

Natron. Conversely, *A. grahami*, the morphotype endemic of Lake Magadi, showed the highest value for nucleotide diversity (π = 0.013) with more fragmented pairwise distribution wave (Fig. 5), revealing a more stable demographic equilibrium in comparison to the Lake Natron morphotypes.

Population analyses conducted using the more informative nuclear DNA markers (SSRs) have, furthermore, depicted a more detailed picture of the independent evolutionary dynamics of the two basins.

Population 9, sampled from Lake Magadi and constituted exclusively by *A. grahami* fishes, seems to be well differentiated from the Natron populations, as suggested by the highest F_{ST} and R_{ST} values scored in the pairwise comparisons and by the AFC pattern. This population, moreover, is probably experimenting a stable phase of its evolution, since shows no deviation from Hardy-Weinberg equilibrium and presents lower Linkage Disequilibrium values, suggesting that effects of the ongoing evolutionary forces are really weak. On the other hand, the populations from Lake Natron are not in HW equilibrium and display higher LD values. The detected deviation from the Hardy-Weinberg equilibrium is almost exclusively associated with a significant deficit of heterozigotes. It is possible to explain the heterozygote deficiency observed in the present study by a selection against heterozigotes or by presence of null alleles (possibly arising from mutations in the region of homology to the SSR primers), which could have led to an underestimation of *H*e.

The peculiarity of Magadi population respect the Natron ones is highlighted even by the Bayesian analysis results. The Bayesian clustering method, as

implemented by STRUCTURE, was applied in order to identify the genetically homogeneous groups in the *Alcolapia* flock and has revealed the presence of two putative gene pools (K=2) from which the fishes of the actual *Alcolapia* flock could have arisen. Natron population has shown high level of admixture, whereas Magadi population results the less admixed one.

AMOVA results has confirmed that the division into two geographic regions, corresponding to Natron and Magadi Lakes, represents a major differentiation factor for the populations of the *Alcolapia* flock .

Due to the high admixture levels of populations and to the low amounts of total variance (8%) explained by the among regions component of AMOVA analysis, we must agree that the genetic structure observed is undoubtedly weak, but however is present.

To further investigate the possible phenomena responsible for the detected genetic structure we have also applied the same analyses used for the whole flock to the Natron population only.

The peculiar geomorphological structure of Lake Natron, infact, could prevent movement of *Alcolapia* between the catchment's freshwater habitats, inhibiting gene flow between these isolated populations; the resulting differentiation and isolation of the populations can so favour the establishment of a genetic structure within the lake.

Natron population, however, resulted low differentiated, as suggested by F_{ST} and R_{ST} values, without evidences of isolation by distance as assessed by Mantel test. Even though Bayesian analysis yielded in this case K=3 as the most probable value, however the level of admixture in the eight populations

is very high, not supporting define a clear structure. Besides, the AMOVA based on the partition of the populations based on K=3 identified an among regions component explaining 1% only of the total genetic differentiation.

The genetic patch around 12 km, identified by Spatial Autocorrelation Analysis (ACS), is compatible with the dimension of the lake and in particular with the shoreline distances that connect the lagoons.

In fact, the connection between populations across the lake is forbidden by the harsh parameters of the water, namely high T and high alkalinity. Fishes cannot swim long in the hot water of Lake Natron without dying (http://www.youtube.com/watch?v=LdZg6sIYOwU&feature=related).

A genetic patch of 12 km is however compatible with a limited gene flow between adjacent populations during the rainy season, when it is possible that "freshwater highways" open in the Lake. Therefore "all sample situated at a geographic distance lower than the genetic patch could be considered a single genetic unit for conservation and management (Diniz-Filho & Telles, 2002)". Weak genetic structure, low differentiation also between distant populations and a 12 km genetic patch are compatible with considering all Alcolapia populations of Lake Natron as a single genetic unit for conservation purposes.

These findings suggest that the aforementioned environmental characteristics of the Lake Natron have not defined a strong population genetic structure within the *Alcolapia* populations. Although the fishes are currently ecologically segregated in their restricted freshwater areas as they inhabit isolated swamp, creeks, brooks and springs in the surroundings of

the basin, their moderate levels of genetic differentiation suggest they have experienced mixing events. The presence of a moderate gene flow is also confirmed by the detection of a small number of migrants/generation. Such fish dispersion and mixing events highlight both their surprisingly physiological resistance to the high pH, carbonate alkalinity and temperature of the main lake (Reite *et al.*, 1974), and the possibility of the occasional presence of higher freshwater layers due to intense rains enabling fish dispersal (Wilson *et al.* . 2003).

Consequently, the hypothesis on ecologically isolated populations can be rejected as it appears there has been mixing among populations in brief periods sufficient to generate only a general weak population structuring.

Thus it seems that the Magadi population is the major factor shaping the genetic differentiation of the *Alcolapia* flock.

To shed more light on this issue, and to investigate if the morphological discrimination of the Lake Natron morphotypes and *A. grahami* in Magadi would be reflected in their genetic diversity, we finally analysed the flock divided according to the prevalent morphotype of each population.

The population pairwise *F*-statistics of the mtDNA sequences revealed non-significant levels of genetic structuring between three of the four *Alcolapia* morphotypes, significant only for *latilabris-grahami* and *alcalicus-grahami* morphotypes. Conversely, through the more extensive microsatellite dataset, the morphotype pairwise comparisons with *F*-statistics and *R*-statistics were significant, confirming major levels of differentiation between *A. grahami* and Natron morphotypes. The genetic analyses depict two vicariant *Alcolapia*

pools in Lake Magadi and in Lake Natron, again supporting the theory that the two morphotypes, *A. grahami* widespread in Magadi and *A. alcalicus* widespread in Natron, were originated by the same founders and successively evolved independently in the two basins. In Lake Natron, the *A. alcalicus* morphotype then radiated into two further morphotypes (*A. ndalalani* and *A. latilabris*), as supported by the weak *F*-statistics among Lake Natron morphotypes and the lack of attribution of specific gene pools observed with Bayesian analysis. The co-occurrence in the same sampling sites, especially identified at Olomotony and in some smaller southern and southeaster waters flowing into the Southern Lagoon (Table 1; Fig. 3) suggest that the genetic pattern of the three morphotypes represent only a beginning of speciation process despite their significant morphological distinctions. Integrating these morphological and genetic results, the lack of congruence between the *Alcolapia* morphotypes of Lake Natron was clearly evidenced, generating an unusual pattern in which the discrimination based on morphological traits was not supported by the levels of molecular differentiation, especially in the mitochondrial DNA.

How to reconcile these two different outcomes? On the same specimens analysed in this work, a morphological analysis has also been carried out. Out of thirteen morphological characters analysed, only three resulted significant in the sense that they were able to discriminate between morphotypes after principal components analyses. The three characters were: pre-orbital distance, maxillary length and pre-pelvic distance. Of these the most relevant appears to be, from an evolutionary point of view, the

maxillary length. But a single mutation could probably be enough to differentiate for this character. Infact recent studies have highlighted that head morphology of the cichlids, and in particular the feeding apparatus morphology, is at least controlled by relatively few genes, often homeogenes that display pleiotropic effects. It has been demonstrated that polymorphisms in *bmp4* (bone morphology protein 4 gene), a gene encoding for a growth factor involved in craniofacial development of mice and Darwin's finches (*Geospiza*), contributes to differences in the shape of both upper and lower jaw elements in cichlid (Albertson *et al.*, 2003; Albertson & Kocher, 2006; Chinsembu, 2009; Kocher, 2004; Parsons, 2009). Several evidences suggest that functional divergence in feeding morphology has contributed to the radiation and maintenance of cichlid species diversity, e.g. Lake Malawi radiation in feeding morphology has produced most of the currently recognized genera (Albertson *et al.*, 2003; Danley & Kocher, 2001). Thus, adaptive variation in feeding apparatus shape is critical to the evolutionary success of cichlids. Genetic investigation of adaptive traits, as jaw and teeth shape, could be a good strategy to improve in the future our understanding of the speciation processes ongoing in Lake Natron. We must consider, infact, that if a speciation process is acting on key genes, an analysis based only on neutral genetic markers is not able to detect any difference between population, even though preliminary knowledges about neutral genetic variability are essential.

Conclusions

The outputs of our mitochondrial and nuclear DNA dataset suggest that this small, polymorphic genus of cichlids in Lakes Natron and Magadi originated from a common ancestor in Natron/Magadi paleolake and that the main starburst radiation occurred in Lake Natron *Alcolapia* that evolved relatively recently following the Natron/Magadi separation. The successive demographic expansion of the founding population determined the sharing of sequence polymorphisms among the flock, patterns also observed in the most studied haplochromine cichlids from Lake Malawi and Tanganyika (Salzburger *et al.* . 2002; Seehausen 2006). In Lake Natron, the physical conditions of the main lake have not prevented some gene flow between the *Alcolapia* populations as population structuring was not evident. The subsequent evolution of the three distinct *Alcolapia* morphotypes in the Natron basin could be the beginning of their speciation process although their current levels of genetic diversity remains low. From these analyses, we suggest their rapid evolution has been driven by the association of the frequent environmental perturbations in the catchments. Whilst this may have also been in association with interspecific hybridisation events where the morphotypes co-occurred, as proposed by Seehausen (2006), no hybrids were documented in the study.

Thus, future work could focus on identifying the presence of hybrid phenotypes in sympatric communities to reveal if the hypothesis that interspecific hybridisation facilitates rapid contemporary evolution is valid. Besides, we plan to assess genetic diversity and differentiation for those loci

which could be possibly involved in speciation by a candidate gene approach. To this purpose, homeogenes for the development of the mouth and gene responsible for high temperature and high saline tolerance will represent suitable candidates.

REFERENCES

Albertson RC, Streelman JT, Kocher TD (2003) Directional selection has shaped the oral jaws of Lake Malawi cichlid fishes. *PNAS*, **100 (9)**, 5252-5257

Albertson RC, Kocher TD (2006) Genetic and developmental basis of cichlid trophic diversity. *Heredity*, **97**, 211-221

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR based techniques. *Nucleic Acids Res*, **25**, 4692-4693

Arif IA, Khan HA (2009) Molecular markers for biodiversity analysis of wildlife animals: a brief review. *Animal Biodiversity and Conservation*, **32 (1)**, 9–17.

Balmford A, Crane P, Dobson A, Green RE, GM Mace (2005) The 2010 challenge: data availability, information needs and extraterrestrial insights. *Phil. Trans. R. Soc.* B, **360**, 221–228

Barluenga M, Stölting KN, Salzburger W, Muschick, Meyer A (2006) Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature, **439 (9)**, 719-723

Barton NH, Slatkin M (1986) A Quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity* **56**, 409-415

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.02, logiciel sous Windows™ pour la génétique des populations.

Laboratoire Génome, Populations, Interactions, CNRS, Université de Montpellier II, Montpellier, France

Bergman AN, Laurent P, Otiang'a-Owiti G, Bergman HL, Walsh PJ, Wilson P, Wood CM (2003) Physiological adaptations of the gut in the Lake Magadi tilapia, *Alcolapia grahami*, an alkaline- and saline-adapted teleost fish. *Comp. Bioch. And Physiol. Part A*, **136**, 701-715

Beveridge MCM, McAndrew BJ (2000) *Tilapias: biology & Exploitation*. Kluwer Academic, Dordrecht.

Brandstätter A, Salzburger W, Sturmbauer C (2005) Mitochondrial phylogeny of the Cyprichromini, a lineage of open-water cichlid fishes endemic to Lake Tanganyika, East Africa. *Mol. Phylogenet. Evol.*, **34**, 382-391.

Brito PH, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica*, **135**, 439-455

Butzer KW, Isaac GL, Richardson JL, Washbourn-Kamau C (1972) Radiocarbon dating of East African lake levels. *Science*, **175**,1069–1076.

Carleton KL, Streelman JT, Lee BY, Garnhart N, Kidd M, Kocher TD (2002) Rapid isolation of CA microsatellites from the tilapia genome. *Animal Genet,* **33 (2)**, 140-144.

Chinsembu KC (2009) Mechanisms and molecular genetic bases of rapid speciation in African cichlids. *Biotechnology and Molecular Biology Reviews* , **3 (4)**, 81-91

Coe MJ (1965) *Tilapia grahami* Boulenger – a study in environmental extremes. In "The Application of Biological Research to the Development of East Africa" *Journal of Applied Ecology*, **2 (2)**, 403-417 Published by: British Ecological Society

Coe MJ (1966) The biology of *Tilapia grahami* Boulenger in Lake Magadi, Kenya. *Acta Trop.*, **23**, 146-177

Coe MJ (1969) Observations on Tilapia alcalicus in Lake Natron. *Rev. Zool. Bot. Afr.*, **80**, 1-14

Cohen AS, Soreghan MJ, Scholz CA (1993) Estimating the age of formation of lakes, an example from Lake Tanganyika, East African Rift system. *Geology*, **21** , 511–514.

Danley PD, Markert JA, Arnegard ME, Kocher TD (2000) Divergence with gene flow in the rock-dwelling cichlids of lake Malawi. *Evolution*, **54 (5)**, 1725-1737

Danley PD, Kocher TD (2001) Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol. Ecol.*, **10**, 1075-1086

Diniz-Filho JAF, Telles MPDC (2002) Spatial autocorrelation analysis of operational units for conservation in continuous populations. *Conservation Biology*, **16 (4)**, 924-935

Elmer KR, Reggio C, Wirth T, Verheyen E, Salzburger W & Meyer A (2009) Pleistocene desiccation in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes. *Proc. Nat. Acad. Sci. USA*, **106**, 13404-13409.

Elmer KR, Kusche H, Lehtonen T, Meyer A (2010) Local variation and parallel evolution: morphological and genetic diversity across a species complex of neotropical crater lake cichlid fishes. *Phil. Trans. R. Soc.* B, **365**, 1763-1782

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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitocondrial DNA restriction sites. *Genetics*, **131**, 479-49

Excoffier L, Laval G, Schneider S (2005) Arlequin (v. 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* **1**, 47-50.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39,** 783–791.

FAO 2000 The state of the world fisheries and aquaculture 2000. Rome: Food and Agricultural Organisation of the United Nations.

FAOSTAT 2000 FAO Statistical Databases. Rome: Food and Agricultural Organisation of the United Nations

Frankham R, Ballou D, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biol. Cons.*, **143**, 1919–1927

Fryer G, Iles TD (1972) *The Cichlid Fishes of the Great Lakes of Africa, their Biology and Evolution*. Oliver and Boyd, Edinburgh.

Genner JM, Seehausen Ole, Lunt DH, Joyce DA, Shaw PW, Carvalho GR, Turner GF (2007) Age of cichlid: new dates for ancient lake radiations. *Mol. Biol. Evol.*, **25 (5)**,1269-1282

Gissi C, Iannelli F, and Pesole G (2008) Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity*, **101**, 301-320

Goetz C, Hillaire-Marcel C (1992) U-series disequilibria in early diagenetic minerals from Lake Magadi sediments, Kenya: dating potential. *Geochem Cosmochim Acta.*, **56**, 1331–1341

Goodman SJ (1997) RST CALC: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Mol. Ecol.*, **6**, 881-885

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *System. Biol.*, **52,** 696-704

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics* **48**, 361–372

Hall T (2003) *BioEdit — Biological Sequence Alignment Editor for Windows*. North Carolina State University, Raleigh

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting 546 by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, **22,** 160-174

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P. 2001. Bayesian inference of phylogeny and its impact on evolutionary biology**.** *Science*, **294,** 2310–2314

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss *Buchloe dactyloides* (Nutt) Engelm. *Theoretical and Applied Genetics*, **86**, 927-934

Hulsey CD (2009) Cichlid genomics and phenotypic diversity in a comparative context. *Integrative and Comparative Biology*, 1-12

Hurst LD, Atlan A, Bengtsson BO (1996) Genetic conflicts. *Q. Rev. Biol.,* **71**, 317–364

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *TREE* , **11,** 10

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

Klett V, Meyer A (2002) What, if anything, is a Tilapia? – Mitochondrial ND2 phylogeny of tilapiines and the evolution of parental care systems in the African cichlid fishes. *Mol. Biol. Evol.*, **19**, 865-883.

Koblmüller S, Egger B, Sturmbauer C, Sefc KM (2007) Evolutionary history of Lake Tanganyika's scale eating cichlid fishes. *Molecular Phylogenetics and Evolution*, **44**, 1295-1305

Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Rev. Gent.*, **5**, 288-298

Kocher TD, Conroy JA, McKaye KR, Stauffer JR (1993) Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. *Molecular Phylogenetics and Evolution*, **2** (2), 158–165

Kornfield I, Smith PF (2000) African cichlid fishes: model systems for evolutionary biology. *Ann. Rev. Ecol. Syst.*, **31**, 163-196

Laurent P, Maina JN, Bergman HL, Narahara A, Walsh PJ, Wood CM (1995) Gill structure of a fish from an alkaline lake: effect of short-term exposure to neutral conditions. *Canadian Journal of Zoology*, **73,** 1170–1181

Lindley TE, Scheiderer CL, Walsh PJ, Wood CM, Bergmani HL, Bergmani AL, Laurent P, Wilson P, Anderson PM (1999) Muscle as the Primary Site of Urea Cycle Enzyme Activity in an Alkaline Lake-adapted Tilapia, *Oreochromis alcalicus grahami**. *J. of. Biol. Chem*, **274 (42)**, 29858-29861

Mace GM, Baillie JEM (2007) The 2010 Biodiversity Indicators: Challenges for Science and Policy. *Cons. Biol.*, 21 (6), 1406–1413

Maina JN (1990) A study of the morphology of the gills of an extreme alkalinity and hyperosmotic adapted teleost *Oreochromis alcalicus grahami* (Boulenger) with particular emphasis on the ultrastructure of

the chloride cells and their modifications with water dilution: A SEM and TEM study. *Anatomy and Embryology* , **181**, 83–98

Maina JN, Wood CM, Narahara AB, Bergman HL, Laurent P, Walsh P (1995) Morphology of the swim bladder of a cichlid teleost: *Oreochromis alcalicus grahami* (Trewavas 1983), a fish adapted to a hyperosomatic, alkaline and hypoxic environment: a brief outline of the structure and function of the swim bladder. In *Horizons of New Research in Fish Morphology in the 20th Century* (Munshi JS & Dutta HM, eds.) pp 179–192 New Delhi: Oxford and IBH

Maina JN, Kisia SM, Wood CM, Narahara A, Bergman HL, Laurent P, Walsh P J (1996) A comparative allometry study of the morphology of the gills of an alkaline adapted cichlid fish *Oreochromis alcalicus grahami* of Lake Magadi Kenya. *International Journal of Salt Lake Research*, **5**, 131–156

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

Marton-Lefèvre J (2006) Biodiversity Is Our Life. *Science*, **327**, 1179

Mayr E (1963) Animal Species and Evolution. *Belknap Press, Cambridge, Massachusetts*

Mayr E (1984). Evolution of fish species flocks: a commentary. In: Echelle AA, Kornfield I (Eds). Evolution of fish species flocks. University of Maine: *Orono Press*, 3-11

Meyer A (1993) Phylogenetic relationships and evolutionary processes in East African cichlid fishes. *Trends in Ecology and Evolution*, **8**, 279–284

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature*, **347**, 550–553

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

Nagl S, Tichy H, Mayer WE, Takezaki N, Takahata N, Klein J (2000) The origin and age of haplochromine fishes in Lake Victoria, East Africa. *Proc. R. Soc. Lon.* B, **267**, 1049-1061

Nagl S, Tichy H, Mayer WE, Samonte IE, McAndrew BJ, Klein J (2001) Classification and phylogenetic relationships of African tilapiine fishes inferred from mitochondrial DNA sequences. *Mol. Phylogenet. Evol.*, **20 (3)**, 361-74

Narahara A , Bergman HL, Laurent P, Maina JN, Walsh PJ, and Wood CM (1996) Respiratory physiology of the Lake Magadi tilapia (*Oreochromis alcalicus grahami*), a fish adapted to a hot, alkaline, and frequently hypoxic environment. *Physiol Zool*, **69**,1114-1136.

Norse EA, McManus RE (1980) Ecology and living resources biological diversity. In: *Environmental Quality 1980: The eleventh annual report of the Council on Environmental Quality,* 31-80 Council on Environmental Quality, Washington DC.

Olden JD, Kennard MJ, Leprieur F, Tedesco PA, Winemiller KO, García-Berthou E (2010) Conservation biogeography of freshwater fishes: recent progress and future challenges. *Diversity and Distributions*, **16**, 496–513

Parsons KJ, Albertson RC (2009) Roles for Bmp4 and CaM1in Shaping the Jaw: Evo-Devo and Beyond. *Annu. Rev. Genet.*, **43**, 369-88

Peakall R, Smouse PE (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* **6**, 288-295. The Australian National University; Canberra, Australia. http://www.anu.edu.au/BoZo/GenAlEx/

Poll M (1986) Classification des cichlidae du lac Tangyanika, tribus, genres et espèces. *Académie Royale de Belgique Mémoires de la Classes Des Sciences, Collection in - 8°-2 Serie, T. XLV — Fascicule 2*, pp. 1–163. Académie Royale de Belgique, Bruxelles.

Pörtner HO, Schulte PM, Wood CM, F. Schiemer (2010) Niche Dimensions in Fishes: An Integrative View. *Physiological and Biochemical Zoology*, **83(5)**, 808–826

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14,** 817–818

Pritchard JK, Stephens M, Donnelly O (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959. software available on http//pritch.bsd.uchicago.edu

Randall DJ, Wood CM, Perry SF , Bergman HL, Maloiy GMO, Mommsen TP, Wright PA (1989) Urea excretion as a strategy for survival in a fish living in a very alkaline environment. *Nature* **337,** 165–166

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

Regan, CT. (1920). The classification of the fishes of the family Cichlidae. I. The Tanganyikan genera. *Ann. Mag. Nat. Hist.*, **9**, 33–53.

Reite OB, Maloiy GMO, Aasenhaug B (1974). pH, salinity and temperature tolerance of Lake Magadi Tilapia. *Nature*, **247**, 315

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestry coeffient: basis for a short-term genetic distance. *Genetics* **105**, 767–779.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic 595 inference under mixed models. *Bioinformatics,* **19,** 1572–1574.

Roberts N, Taieb M, Barker P, Damnati B, Icole M, Williamson D (1993) Timing of the younger Dryas event in East Africa from lake-level changes. *Nature* **366**,146–148.

Salzburger W, Baric S, Sturmbauer C (2002) Speciation via introgressive hybridization in East African cichlids? *Mol. Ecol.*, **11**, 619–625.

Salzburger W, Meyer A (2004) The species flocks of East African cichlid fishes: recent advances in molecular phylogenetics and population genetics. *Naturwissenschaften*, **91**, 277-290

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

Schluter D (2001) Ecology and the origin of species. *Trends Ecol. Evol.*, **16**, 372–380

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

Shuelke M (2000) An economic method for the fluorescent labelling of PCR fragments. *Nature Biotechnology.*, **18**, 233-234

Schwarzer J, Misof B, Tautz D, Schiewen U (2009) The root of East African cichlid radiations. *BMC Evol. Biol.*,**9**, 186

Seegers LR, Tichy H (1999) The *Oreochromis alcalicus* flock (Teleostei: Cichlidae) from lakes Natron and Magadi, Tanzania, and Kenya, with descriptions of two new species. *Ichthyol. Explor. Freshw.*, **10**, 97-146.

Seegers L, Sonnenberg R, Yamamoto R (1999) Molecular analysis of the *Alcolapia* flock from lakes Natron and Magadi, Tanzania and Kenya (Teleostei: Cichlidae), and implications for their systematics and evolution. *Ichthyol. Explor. Freshw.*, **10**, 175-199.

Seehausen O (1996) *Lake Victoria Rock Cichlids*. Verduijn Cichlids,Germany.

Seehausen O, Van Alphen JJM (1998). The effect of male colouration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). *Behav. Ecol. Sociobiol.*, **42**, 1-8

Seehausen O, Van Alphen JJM (1999). Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.*, **2**, 262-271

Seehausen O, Van Alphen JJM, Lande R (1999). Colour polymorphism and sex ratio distortion in a cichlid fish as an incipient stage in sympatric speciation by sexual selection. *Ecol. Lett.*, **2**, 367-378

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

Stager JC, Johnson TC (2008) The late Pleistocene desiccation of Lake Victoria and the origin of its endemic biota. *Hydrobiologia*, **596**, 5-16

Stiassny MLJ, Schliewen UK, Dominey WJ (1992) A new species flock of cichlid fishes from Lake Bermin, Cameroon with a description of eight new species of Tilapia (Labroidei: Cichlidae). *Ichthyol. Explor. Freshw.,* **3**, 311–346

Strayer DL, Dudgeon D (2010) Freshwater biodiversity conservation: recent progress and future challenges. J. *N. Am. Benthol. Soc.*, **29 (1)**, 344-358

Strecker U (2006) Genetic differentiation and reproductive isolation in a *Cyprinodon* fish species flock from Laguna Chichancanab, Mexico. *Molecular Phylogenetics and Evolution*, **39**, 865–872

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

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

Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony* (*and Other Methods). Sinauer Assoc.; Sunderland

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin FF, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids. Res.*, **25**, 4876–4882

Tichy H, Seegers L (1999) The *Oreochromis alcalicus* flock (Teleostei: Cichlidae) from lakes Natron and Magadi, Tanzania, and Kenya: a model for the evolution of "new" species flocks in historical times? *Ichthyol. Explor. Freshw.,* **10**, 147-174

Trewavas E (1966a) Fishes of the genus *Tilapia* with four anal spines in Malawi, Rhodesia, Mozambique and Southern Tanzania. *Rev. Zool. Bot. Afr.*, **74**, 50-62

Trewavas E (1966b) A preliminary review of the genus *Tilapia* in the eastward-flowing rivers of Africa, with proposals of two new specific names. *Rev. Zool. Bot. Afr.* **74**, 394-424

Trewavas E (1981) Nomenclature of the tilapia of Southern Africa. *J. Limnol. Soc. S. Afr.*, **7**, 42

Trewavas E (1982) Generic groupings of Tilapiini used in aquaculture. *Aquaculture*, **27**, 79–81.

Trewavas E (1983) *Tilapiine Fishes of the Genera Sarotherodon, Oreochromis and Danakilia.* Published by the British Museum (Natural History), London.

Turner GF, Seehausen O, Knight ME, Allender CJ, Robinson RL (2001) How many species of cichlid fishes are there in African Lakes? *Mol. Ecol.*, **10**, 793-806

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

Vincens A, Casanova J (1987) Modern background of Natron-Magadi Basin (Tanzania-Kenya): physiography, climate, hydrology and vegetation. *Sci. Geol. Bull. Strasbourg*, **40**, 9-21

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

Wilkie MP, Wood CM (1996) The Adaptations of Fish to Extremely Alkaline Environments. *Comp. Biochem. Physiol.*, **113B (4)**, 665-673

Wilson PJ, Wood CM, Maina JN, White BN (2000) Genetic structure of Lake Magadi tilapia populations. *J. Fish. Biol.*,**56**, 590-603.

Wilson PJ, Wood CM, Walsh PJ., Bergman AN, Bergman HL, Laurent P, White BN (2004) Discordance between Genetic Structure and Morphological, Ecological, and Physiological Adaptation in Lake

Magadi Tilapia. *Physiological and Biochemical Zoology*, **77 (4)**, 537-555 Published by: The University of Chicago Press

Wood CM, Bergman HL, Laurent P, Maina JN, Narahara A, Walsh PJ (1994) Urea production, acid-base regulation and their interactions in the Lake Magadi tilapia, a unique teleost adapted to a highly alkaline environment. *Journal of Experimental Biology*, **189,** 13–36.