UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea magistrale in Biologia Evoluzionistica LM-6

TESI DI LAUREA

Resolving the taxonomy of Malagasy cichlids with *museomics*: preliminary phylogenetic results for the genus *Paratilapia*

Relatore: Prof.ssa Chiara Papetti
Dipartimento di Ecologia Molecolare, Università di Padova

Correlatore: Prof. Michael Matschiner
Dipartimento di Bioscienze, Museo di Storia Naturale, Università di Oslo

Laureando/a: Alessia Prestanti

ANNO ACCADEMICO 2022/2023

ABSTRACT

Malagasy cichlids of the genus *Paratilapia* bear an ambiguous and unresolved taxonomy at the species level; apart from a single species, *Paratilapia polleni* Bleeker, 1868, currently recognized, two other species – commonly treated as synonyms – have been described in the past: *P. typus* (Bleeker, 1878) and *P. bleekeri* Sauvage, 1882. To resolve the taxonomy within the genus *Paratilapia,* this study aimed to analyse mitochondrial DNA sequences obtained from both modern and historical tissue samples of *Paratilapia* specimens. The molecular dataset included data of one syntype specimen of *P. polleni*, recovered by applying *museomics* protocols for historical DNA extraction from wet collection specimens. The phylogenetic analyses were performed using the sequences of the mitochondrial gene fragments 16S rRNA (16S), Cytochrome Oxidase Subunit 1 (COI), and NADH dehydrogenase 2 (NADH 2). The results of this study show the consistent phylogenetic placement of the syntype specimen of *P. polleni* Bleeker, 1868 as closely related to *Paratilapia* specimens collected from the Malagasy offshore island Nosy Be, confirming their species status.

RIASSUNTO

I Ciclidi del Madagascar appartenenti al genere *Paratilapia* presentano una tassonomia ambigua e parzialmente irrisolta a livello di specie; oltre a *Paratilapia polleni* Bleeker, 1868, unica specie attualmente riconosciuta all'intero del genere, altre due specie sono state descritte in passato, ma comunemente trattate come sinonimi: *P. typus* (Bleeker, 1878) e *P. bleekeri* Sauvage, 1882. Per risolvere le ambiguità tassonomiche interne al genere *Paratilapia*, questo studio si propone di analizzare sequenze di DNA mitocondriale ottenute da campioni tissutali sia recenti, che storici, di esemplari di *Paratilapia*. Il dataset molecolare, oltre a sequenze ottenute da esemplari recentemente campionati, include sequenze mitocondriali di un esemplare sintipo di *P. polleni*, ottenute grazie all'applicazione di protocolli museomici. Queste procedure consentono di estrarre DNA storico da campioni museali conservati in collezioni zoologiche. Le analisi filogenetiche sono state eseguite utilizzando sequenze nucleotidiche relative a frammenti dei geni mitocondriali 16S rRNA (16S), Citocromo Ossidasi Subunità 1 (COI) e NADH deidrogenasi 2 (NADH 2). I risultati di questo studio mostrano una coerente collocazione filogenetica dell'esemplare sintipo di *P. polleni* Bleeker, 1868; tale sintipo risulta strettamente imparentato con esemplari di *Paratilapia* campionati recentemente sull'isola malgascia di Nosy Be, confermandone la specie.

1. INTRODUCTION

Madagascar – the world's fourth-largest island – is separated from the African continent's south-eastern coast by a distance of 400 km. The island has been separated from other landmasses, originally forming the supercontinent Gondwana, for over 85-90 million years (Ali & Aitchison, 2008; Matschiner, 2019). It was already fully isolated when the mass extinction at the Cretaceous–Paleogene (K–Pg) boundary (Longrich et al., 2011) came to an end, leading most of the land-vertebrate fauna to perish (e.g., Krause et al., 2020). Following multiple colonization events, new founder taxa thus had access to various unoccupied niches that enabled them to diversify widely, at least until the last few million years (e.g., Samonds et al., 2013; Burbrink et al., 2019; Belluardo et al., 2022). This diversification, alongside the radiation of relict species and the long-term isolation of the island, produced the high levels of endemism and the impressive species richness of Madagascar, where some clades have proliferated dramatically (Yoder & Nowak, 2006; Vences et al., 2009; Ali & Hedges, 2023). Among the endemic land-vertebrate fauna, for instance, lemurs, birds of the family Vangidae, chameleons of the genus *Brookeria*, pseudoxyrhophine snakes, mantellid frogs, microhylid frogs, each comprise tens or even hundreds of species (Crottini et al., 2012; Reddy et al., 2012).

Since most of the endemic taxa are currently threatened, mainly by deforestation and introduction of non-native invasive species (Myers et al., 2000; Jones et al., 2009; Vences et al., 2017), Madagascar is defined as an international conservation priority (Benstead et al., 2003). The island also harbours a unique freshwater biota (Goodman & Benstead, 2003) and has been recognized as a global hotspot for freshwater biodiversity (Groombridge & Jenkins, 1998; Ganzhorn et al., 2008).

While some studies reported Cretaceous (145–66 million years ago; Ma) ages estimates for Madagascar's three particularly species-rich freshwater fish families (Cichlidae, Bedotiidae, Aplocheilidae), based on molecular time trees (e.g., Azuma et al., 2008), other analyses supported a later origin of these lineages in the Paleogene (66–23 Ma; Vences et al., 2001; Crottini et al., 2012; Matschiner et al. 2020). The earlier estimates imply that these major freshwater fish groups were well established in Madagascar before the K–Pg boundary (Gottfried et al., 1998; Sparks & Stiassny, 2003; Crottini et al., 2012), supporting a vicariance origin of these families linked to the break-up of Gondwana (Samonds et al., 2013). This view agrees with an assumed disadvantage for trans-oceanic dispersal of these fish, for it is more prohibitive than that of terrestrial taxa (Samonds et al., 2012). Contrary to freshwater fish, these terrestrial lineages, mainly from the mainland Africa, might have crossed the Mozambique Channel rafting on floating environmental structures (Yoder et al., 2003; Ali & Huber, 2010), or along temporarily corridors above the sea level (Wit, 2003; Masters, De Wit & Asher, 2006). The "dispersal-disadvantage" for freshwater fish would also be supported by an apparent rarity of these fish lineages that arrived clearly after Madagascar's isolation, despite the niche displacement that affected other groups (Samonds et al., 2012; Masters, De Wit & Asher, 2006).

In this thesis, I focus on one of the Malagasy fish families, the Cichlidae. Within vertebrates, cichlids are among the most speciose families, exhibiting a wide range of ecological, morphological, and behavioural variation (Sparks, 2004). For these

reasons, cichlid fishes have become one of the most important model groups in evolutionary biology and adaptive radiation research (Brawand et al., 2014). Despite the Gondwanan distribution of extant cichlids (Africa, Neotropic, Madagascar, Southern India and Sri Lanka), the vicariance hypothesis seems to be refuted in this case by Bayesian age estimation based on molecular sequence data and fossil ages (Matschiner et al., 2017). This timeline implies that cichlid diversification in both the Neotropics and Madagascar occurred long after the split of the Gondwanan landmasses, and thus supports trans-oceanic dispersal rather than vicariance as the explanation of their biogeographic origin (Matschiner et al., 2017, 2020; Matschiner, 2019).

1.1. Cichlids of Madagascar

According to the classification by Sparks and Smith (2004), the family Cichlidae represents a monophyletic clade and is subdivided into four subfamilies: Etroplinae (occurring in India/Sri Lanka and Madagascar), Ptychochrominae (Madagascar), Cichlinae (Neotropics) and Pseudocrenilabrinae (Africa).

While the African and Neotropical cichlids have received great scientific attention, there are still only few studies addressing the biology and evolution of endemic Malagasy cichlid fishes, even though their persistence is highly threatened by overfishing, deforestation, and the introduction of non-native species (Benstead et al., 2003; Sparks & Stiassny, 2003).

This information gap is mainly owed to the restrictive rules on collecting live organisms imposed by the government during the 1970s (de Rham & Nourissat, 2004), when Madagascar was under military rule. The more stable political situation recovered at the beginning of the 1980s allowed researchers to restart the exploration of the inland waters of Madagascar, leading them to discover many new cichlids species. However, it is likely that more species remain to be discovered (de Rham & Nourissat, 2004).

The phylogenetic relationships among several of these species are still unresolved, so that gaining a better knowledge of the Malagasy cichlid phylogeny and diversity remains important, both for developing our understanding of their diversification processes, and for conservation purposes (Benstead et al., 2003; Lovaharisoa & Ranalison, 2018).

1.1.1. Etroplinae and Ptychochrominae

The two cichlid subfamilies that live in Madagascar's freshwater systems are Ptychochrominae – comprising the endemic genera *Ptychochromis*, *Ptychochromoides*, *Katria*, *Oxylapia* and *Paratilapia* – and Etroplinae – comprising a single genus *Paretroplus*, endemic to Madagascar, and its sister taxon *Etroplus,* endemic to southern India and Sri Lanka (Stiassny & Sparks, 2006). Given that Etroplinae form the sister taxon to all other cichlids (Sparks & Smith, 2004a; Matschiner et al., 2020), the Malagasy and South Asian cichlids do not represent a monophyletic group, in contrast to results of earlier studies (Farias et al., 1999; Sparks, 2001), in which such a clade was only weakly supported. On the other hand, the placement of Etroplinae as the sister taxon of all the other cichlids (Sparks & Smith, 2004) is strongly supported by genome-wide evidence (Matschiner et al., 2020) and reinforced by morphological evidence (Sparks, 2001).

Etroplinae, with 16 species described, bear several phenotypic specializations that make this clade morphologically distinct from other cichlids (Cichocki, 1976; Sparks, 2001; Stiassny et al., 2001). Within this subfamily, the genus *Paretroplus* is endemic to Madagascar and recovered as monophyletic with strong support, by molecular data as well as its numerous synapomorphic morphological features (Cichocki, 1976; Sparks, 2001, 2004a; Stiassny et al., 2001). Morphological evidence suggests the distinction of three clades within the genus *Paretroplus*, one comprising the elongate, primarily riverine species *P. damii* Bleeker, 1868, *P. nourissati* (Allgayer, 1998), *P. tsimoly* Stiassny, Chakrabarty & Loiselle, 2001, and *P. lamenabe* Sparks, 2008 (Sparks, 2008). The second clade comprises the deep-bodied, primarily lacustrine species *P. polyactis* Bleeker, 1878, *P. petiti* Pellegrin, 1929, *P. maculatus* Kiener & Maugè 1966, *P. menarambo* Allgayer, 1996, *P. maromandia* Sparks & Reinthal, 1999, and *P. dambabe* Sparks, 2002; finally, the third clade within the genus *Paretroplus* includes the shallow-bodied and highly mottled species *P. kieneri* Arnoult 1960, and *P. gymnopreopercularis* Sparks, 2008, which occur in both lentic and lotic habitats (Sparks, 2008).

The greatest diversity of Etroplinae is found in the north-western part of Madagascar, where 11 species occur, whereas only a single taxon, *P. polyactis*, is restricted to eastern drainages and brackish waters (Sparks & Smith, 2004; Sparks, 2008). The phylogenetic relationships at species level within the genus *Paretroplus* were investigated using nucleotide markers (Sparks, 2004a; Sparks & Smith, 2004) as well as morphological features, and have been shown to be largely resolved (Sparks, 2008).

Recent molecular analyses warrant the placement of Ptychochrominae, with 15 valid species, as the sister taxon to the African-Neotropical clade (Matschiner, 2020), in agreement with an earlier study by Sparks and Smith (2004a).

Within Ptychochrominae, the genus *Katria* was formally recognized only in 2006, in account of new morphological evidence (Stiassny & Sparks, 2006) that corroborated the phylogenetic position of the species formerly known as *Ptychochromoides katria* (Reinthal & Stiassny, 1997) as the sister taxon to the genus *Ptychochromis.* This position had also received support already from molecular data (Sparks & Smith, 2004). As a result of this acknowledgement, *P. katria* was removed from *Ptychochromoides* (a genus that includes three other species) and placed in a separate genus, *Katria* (Stiassny & Sparks, 2006).

The inclusion of *Paratilapia* within Ptychochrominae remained unclear for longer, since depending on whether morphological or molecular characters were analysed, *Paratilapia* was equivocally placed as the sister taxon to either the etropline or ptychochromine cichlids (Sparks, 2001; Sparks, 2004a), to a clade comprising *Oxylapia* and *Ptychochromoides* (Sparks, 2004a), to *Paretroplus* (Farias et al., 2001), or as the sister taxon to a clade comprising both Etroplinae and Ptychochrominae (Sparks, 2001). Thanks to simultaneous analyses of mitochondrial and nuclear marker genes, *Paratilapia* was finally recovered as the sister taxon to the ptychochromine cichlids by Sparks (2003, 2004a), allowing the phylogenetic placement of this genus within the subfamily Ptychochrominae (Sparks & Smith, 2004). Members of this subfamily are distributed both in eastern and western drainages of Madagascar, and their inter- and intra-generic phylogenetic relationships are mainly resolved, with few exceptions.

Within Ptychochrominae, the endemic Malagasy genus *Paratilapia* remains one of the most taxonomically unresolved and understudied. Despite the attempts of different authors to identify new species, *Paratilapia polleni* Bleeker, 1868 is so far the only species within the genus that is recognized as valid.

1.2. The genus *Paratilapia*

Prior to the extensive deforestation of much of the island, the genus *Paratilapia* had the largest distribution among all native Malagasy cichlids (de Rham & Nourissat, 2004) and seemed to be absent only from a few areas in the West and Southwest of Madagascar, as well as from parts of the highlands, in regions above an altitude of 1500 m (Kiener, 1963; Kiener & Maugè, 1966). Different populations have shown a capability of surviving in a wide variety of habitats, tolerating a broad range of water temperatures and water chemistry, that characterize various different Malagasy freshwater habitats (Stiassny & Gerstner, 1992). Although *Paratilapia* cichlids cannot tolerate temperatures lower than 12–15 °C (Kiener, 1963), they are able to thrive in the small alkaline lakes in southern Madagascar, where it is quite common to record temperatures of 40 °C or more (Catala, 1977).

Adults can rarely be encountered in large numbers in fish catches and exhibit a more-or-less solitary social behaviour, while juveniles might occur in flocks (Stiassny & Gerstner, 1992).

Paratilapia cichlids have a mixed diet, feeding on insects and aquatic larvae, crustaceans, tadpoles, small frogs and, occasionally, small fish (Stiassny & Gerstner, 1992). Growth is relatively slow: usually, three years are required for individuals to attain lengths of about 25 cm. Sexual maturation occurs at a small size: at 12 cm of total length on the central plateau, and at 8 cm in the nutrient poor acid waters of the coastal plain where growth rates are slower (Kiener, 1963).

Despite the weak sexual dimorphism, *Paratilapia* males often display longer dorsal and anal fins. During reproductive activity their fins can be more vividly blackened with a more strongly contrasted blue spangling than the female, whose ground colour is usually olive-brown with some golden and blue spangling on the flanks. Females also tend to be slightly deeper-bodied than males of equal size (Stiassny & Gerstener, 1992). Juveniles often exhibit a series of dark vertical bands along the flanks and have a typically ocellated “tilapia-spot” (Thys, 1968) on the dorsal fin (Stiassny & Gerstener, 1992).

Paratilapia populations began decreasing around the end of the 1990s, due to anthropogenic pressures such as habitat degradation and the introduction of non-native invasive species, for instance, the African bony-tongue, carp, goldfish, and the African ‘tilapias’ (*Oreochromis* and *Coptodon*) (Šimková et al., 2019), which caused their general rarefaction and local extinction, especially from lowland habitats (de Rham & Nourissat, 2004). According to IUCN risk assessment (IUCN 2020), the conservation status of *Paratilapia* is considered as Vulnerable, and the current population size trend is decreasing. *Paratilapia* seems to be still abundant on the Nosy Be Island and in some lakes in the Sambirano biogeographic region of north-western Madagascar, while riverine populations have become very depauperated (IUCN, 2020).

Despite the taxonomic assignment of several nominal species within the genus *Paratilapia* (Loiselle & Stiassny, 1993; de Rham & Nourissat, 2004; Loiselle, 2011; Lovaharisoa & Ranalison, 2018), currently only a single species is considered valid: *Paratilapia polleni* Bleeker, 1868. According to de Rham and Nourissat (2004), the type locality of this species is the Ambazoana River, in north-western Madagascar, where Francois Pollen and Douwe Casparus van Dam collected the type material that led the Dutch ichthyologist Pieter Bleeker to describe *P. polleni* in 1868 (Loiselle & Stiassny, 1993).

This species is phenotypically characterized by a jet-black or deep dark blackish-blue body, covered with iridescent clear spots, which are usually white, but occasionally light blue, greenish, pinkish or yellow (de Rham & Nourissat, 2004). According to Bleeker's original description (Bleeker, 1868), the body of this species is oblong and compressed, and exhibits a blackish-violet pigmentation. The intensity of this colouration decreases in the abdominal region, which is characterized by violet-brown ventral spots usually scattered with blue ones. The scales on both the head and most of the body are described as having a small irregular spot shining blue, and the colour of the fins is reported to be blackish-purple or brownish-purple (Bleeker, 1868).

Different populations of *Paratilapia* across the island exhibit a range of variation in both morphology and colour patterns, but the most distinctive feature is the "size-of-spots", a single character which seems to be enough by itself to separate the *Paratilapia* populations into two main strains: the large spotted, and small spotted *Paratilapia* (de Rham & Nourissat, 2004). However, additional evidence suggests that the systematics of *Paratilapia* are more complex than a separation into two species according to spot size.

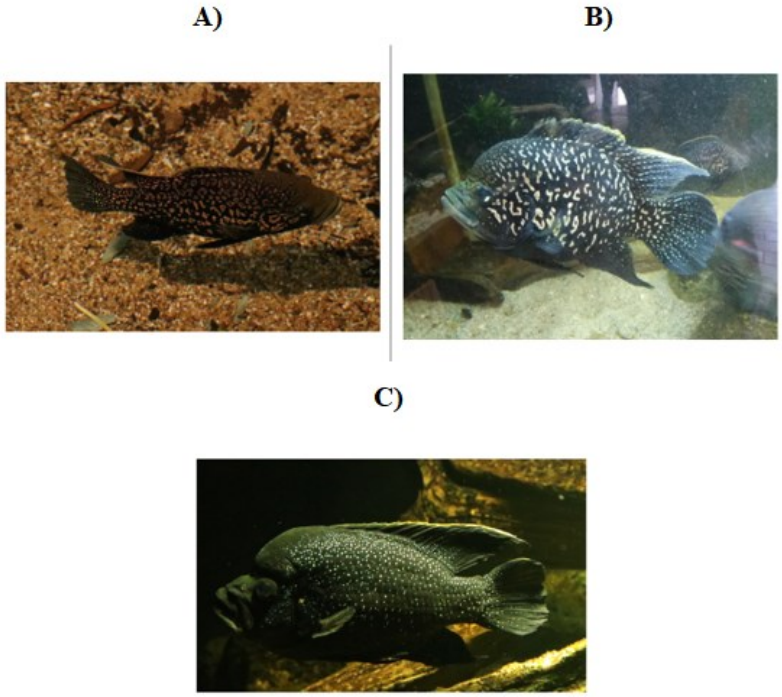

Figure 1: In A and B, large-spotted Paratilapia *specimens from the breeding stock of Patrick de Rham, referred to as* Paratilapia *cf.* polleni "Andapa". *The individual in B is from the Zoo in Cologne. In C, a small-spotted* Paratilapia *specimen with no precise locality information from the Zoo in Berlin. Photos by Thore Koppetsch.*

1.2.1. Taxonomic uncertainties*: Paratilapia typus* & *Paratilapia bleekeri*

In 1878, Bleeker described the species *Paracara typus* based on material collected from the Mananara River, near the town of Mananara Nord, also called Mananara Avaratra, situated on the northeast coast of Madagascar. Bleeker depicted *Paracara typus* as a freshwater fish with a compressed and oblong body, with a dorsal colouration being greenish yellow or silvery, with a violet-green iridescence. He also mentioned the presence of purplish or deep brown spots at the back of the gills and irregular brown spots on the flanks (Bleeker, 1878).

Being placed in the genus *Paratilapia* subsequently, the species *P. typus* was soon put into synonymy with *P. polleni (*Pellegrin, 1904; Maréchal & Poll, 1991). However, the name *Paratilapia typus* (Bleeker, 1878) can still be found in scientific literature when referring to the *Paratilapia* population living in the Mananara area (de Rham & Nourissat, 2004; Loiselle, 2011).

In 2006, Paul Loiselle (2011) led a sampling trip in Madagascar that enabled the comparison between the type specimens of *Paratilapia typus* (Bleeker, 1878), which are juveniles in poor conditions, with more recent materials. During this expedition, Paul Loiselle managed to reach the town of Mananara Nord and succeeded in collecting topotypical material of *P. typus* (Bleeker, 1878), that proved to be large-spotted (Loiselle, 2011).

As yet another member of the genus *Paratilapia*, the French ichthyologist Henri Émile Sauvage described a new Malagasy cichlid species in 1882 and named it in honour of the person that erected the genus: *Paratilapia bleekeri* Sauvage, 1882. The description of this species, published in 1891, was based on specimens collected by Alfred Grandidier from rice paddies and marshes to the east of Antananarivo, the capital of Madagascar (Loiselle & Stiassny, 1993). Sauvage (1891) emphasized the more truncated dorsal and anal fins of the type specimens compared to *P. polleni*, suggesting these traits to be diagnostic characters of this new species. According to his original description, *P. bleekeri* could be easily distinguished from *P. polleni* by the morphology of the fins, since the caudal fin is depicted as much shorter and more rounded than in *P. polleni*, and the dorsal and anal fins are described as less elongated (Sauvage, 1891).

The colouration is the second character which was supposed to allow the distinction between the two species. According to Sauvage's description (1891) of *P. bleekeri*, the species' scales were blue in the centre and a few blue spots could be found below the eyes. Additionally, Sauvage stressed the presence of elongated blue spots at the base of each dorsal and anal spine as well as on the anal, dorsal, and caudal fins (Sauvage, 1891).

However, Jacques Pellegrin synonymized *P. polleni* and *P. bleekeri* in his revision of cichlid taxonomy published in 1904. He concluded that the differences in colour pattern and buccal dentition presented in Sauvage's description could not justify recognizing the latter as a distinct taxon (Pellegrin, 1904). Despite the formal rejection of *P. bleekeri*'s species status over a hundred years ago, the large-spotted specimens living around Antananarivo still continue to be referred to as *P. bleekeri* in the literature (Loiselle & Stiassny, 1993; Loiselle, 2011). Topotypical specimens of *P. bleekeri* have not been collected since 1979 (Loiselle, 2011).

Although the two nomina *Paratilapia typus* (Bleeker, 1878), and *Paratilapia bleekeri* Sauvage, 1882 were synonymized with *Paratilapia polleni* Bleeker, 1868, the paucity of information derived from molecular analyses has so far not allowed to exclude a different interpretation of the taxonomy within the genus *Paratilapia*.

1.2.2. "Large-spotted" & "small-spotted" *Paratilapia*

In 1988, Peter Reinthal and Melany Stiassny collected a large series of juvenile *P. polleni* from a lake in the Pangalanes canal system near the village of Sahely on the east coast of Madagascar (Reinthal & Stiassny, 1991). Some of these living individuals were brought to the American Museum of Natural History (AMNH), in New York, USA, where a pair bred successfully in 1991 (Stiassny & Gerster, 1992). Offsprings of this single pair showed the small-spotted phenotype as adults (Loiselle & Stiassny, 1993).

Also in 1991, the French aquarist Jean-Claude Nourissat led a collecting trip to Madagascar with the purpose of bringing back to Europe as many native cichlids as possible. He managed to capture some specimens from the north-western basin of the Sofia River, which afterwards bred successfully in captivity (de Rham & Nourissat, 2004). Fry were distributed through commercial channels to breeders on both sides of the Atlantic; these individuals showed a large-spotted phenotype as adults (Loiselle & Stiassny, 1993; de Rham & Nourissat, 2004; Loiselle, 2011).

Further differences between Reinthal and Stiassny's small-spotted individuals and Nourissat's large-spotted strain are found in the size and disposition of their scales, fright colouration, body proportions and the expression of aggressive behaviour. These differences were considered pronounced enough to warrant a re-examination of preserved materials in the AMNH collections and those in the Museum National d'Histoire Naturelle (MNHN) in Paris, to determine whether these two phenotypes might represent two distinct species and, eventually, which of them would have matched the original description of *P. polleni* (Loiselle & Stiassny, 1993). This effort revealed a range of suitable diagnostic traits, which allowed preserved materials to be grouped into three distinct groups without any reference to colour pattern. Consequently, one of these clusters was assumed to correspond to *P. polleni*, while the taxonomic status of the other two groups remained uncertain (Loiselle & Stiassny, 1993). Taking into account Bleeker's illustration of the type specimen, which depicts a black fish with small lighter spots on the body and fins, it was intuitive to assign *P. polleni* to the small-spotted phenotype strain. *P. bleekeri* was considered as the proper name for the large-spotted strain, based on morphological features found for both the type specimens and more recently preserved materials, as well as Sauvage's description (1891) of the species' colours in life (de Rham & Nourissat, 2004). The third putative species appeared to be restricted to the extreme south-western region of Madagascar, so that it could not be matched with the third available nomen, P. *typus,* originating from the northeast, and therefore wasn't assigned to any published name (Loiselle & Stiassny, 1993).

However, the assignment of *P. polleni* and *P. bleekeri* to the small- and large-spotted phenotypes turned out to be unreliable, since it was based on Bleeker's illustration only, which was later interpreted as an artifact of preservation. In the early years of the nineteenth century, Bleeker was obliged to preserve the type materials by placing freshly caught specimens of *Paratilapia* directly into alcohol,

causing a quite dramatic contraction of the metallic spots (Loiselle, 2011). The identification of this alcohol distortion effect, which led Bleeker to depict a misleading illustration of the colour pattern, was fundamental for the realization that *P. polleni* was in fact large-spotted, so that the primary diagnostic character that allowed to distinguish this species from *P. bleekeri* was untenable and the two species fell back into synonymy (de Rham & Nourissat, 2004; Loiselle, 2011).

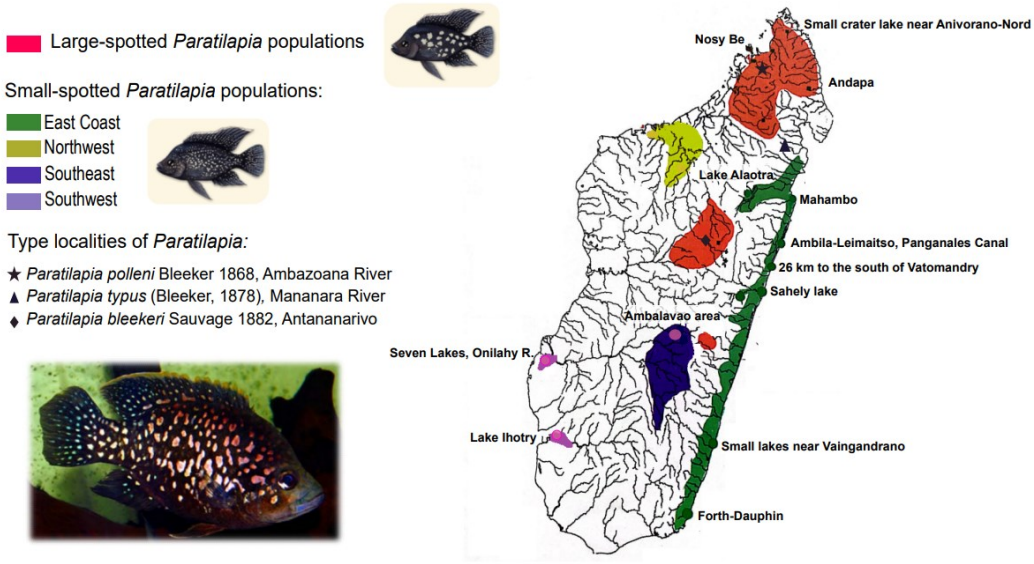

Figure 2: Map of Madagascar with small-spotted and large-spotted Paratilapia *populations marked in different colours. The type localities of* P. polleni, P. typus *and* P. bleekeri *are reported on the map in different shapes. Coloured dots and squares highlight collection localities of* Paratilapia *specimens. The map is based on that reported by de Rham and Nourissat (2004). The picture of the living specimen is a female of* Paratilapia *sp. from the Andapa region (Loiselle, 2011).*

Considering the localities in which the different morphs have been recorded, nevertheless a general pattern could be noticed (Fig. 2): the large-spotted populations seem to be distributed in northern Madagascar, occurring in the Ifasy and Sambirano drainages, and in the crater lakes of the island Nosy Be (Loiselle, 2011). Additionally, this northern population is characterized by a massive body, and males, that grow larger than females, and that do not develop a pronounced cephalic hump. The large-spotted phenotype is also present in the central highlands close to the capital Antananarivo, but this population now seems to be very depauperated (Loiselle, 2011; de Rham & Nourissat, 2004). Small-spotted populations, on the other hand, inhabit mainly the weakly acidic black water habitats close to the east coast and demonstrate a broad geographical distribution, from Fort Dauphin (=Tolagnaro) in the South, up north to Soaniera-Ivongo (de Rham & Nourissat, 2004). Small-spotted specimens were also found close to the city Toliara (=Tulear), along the south-west coast, and in the lower reaches of the Betsiboka drainage (Loiselle, 2011). All the *Paratilapia* specimens from the east coast share the same general appearance, with the external edging of the dorsal fin either yellow or orange, and males are acquiring a pronounced frontal hump as they get older. Small-spotted populations of the southern half of the east coast show a colouration gradient, since the variation in size between the spots of females and males increases southwards in Madagascar (de Rham & Nourissat, 2004).

1.2.3. Synonymized species of *Paratilapia*

Different authors claimed to have discovered so far unrecognized diversity within *Paratilapia*, based on morphological and genetic differences between populations, indicating a diversification of this genus into more than a single taxonomically recognized species (Loiselle & Stiassny, 1993; de Rham & Nourissat, 2004; Loiselle, 2011; Lovaharisoa & Ranalison, 2018). The two main interpretations of the diversity within the genus *Paratilpia* were originally proposed by de Rham and Nourissat (2004) and Loiselle (2011), whose papers provided evidence for the existence of more than one *Paratilpia* species, showing an intricated taxonomy.

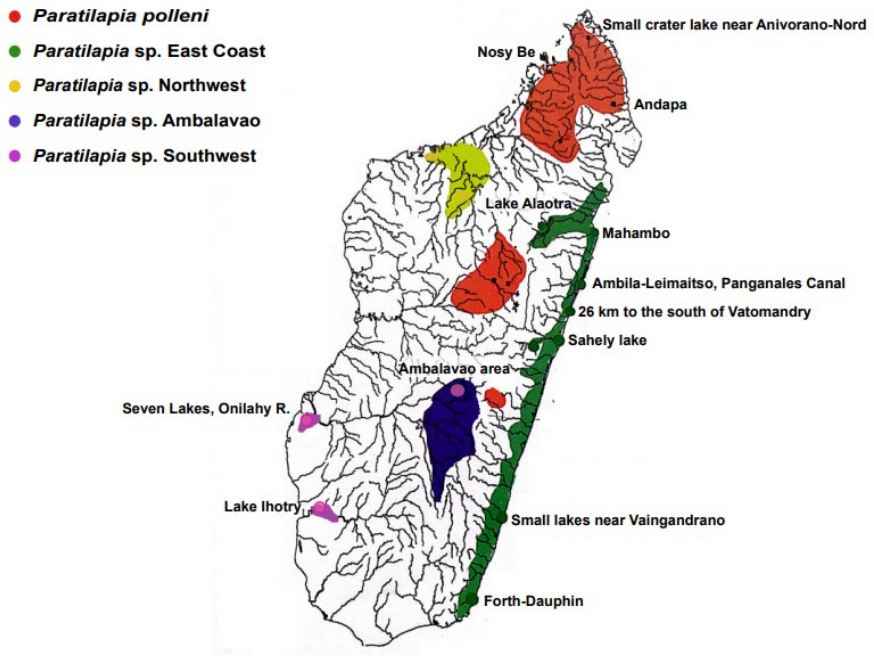

Figure 3: Distribution of Paratilapia *taxa according to de Rham and Nourissat (2004).*

De Rham and Nourissat (2004) characterized the species diversity within *Paratilapia* based on morphological traits, which suggests the presence of at least four new taxa still undescribed (Fig. 3). The authors assigned all the large-spotted populations from northern Madagascar to *P. polleni sensu stricto*, including not only those occurring in the Ifasy and Sambirano drainages, close to the type locality, but also the isolated population that can be found in the Andapa basin (Lokoho River drainage, Northeast) and populations recorded from Nosy Be and in small crater lakes near Anivorano-Nord. Even though the *Paratilapia* population living in swamps near Antananarivo, the type locality of *P. bleekeri,* seems to be near extinction, a relict large-spotted population has been reported to still live in several lakes around the capital (Loiselle, 2011). This relict population highlights the importance of comparisons between wild specimens from near Antananarivo and the large-spotted fish from the north of Madagascar. The lack of evidence of a clear distinction between these two lineages led de Rham and Nourissat (2004) to treat this population of the central highlands as *P. polleni*.

In their overview, de Rham and Nourissat (2004) further suggest that all the small-spotted *Paratilapia* populations should be split into four different species: *P.* sp. East Coast*, P.* sp. Ambalavao, *P.* sp. Southwest, and *P.* sp. Northwest. Of these four

putative species, *P.* sp. East Coast would have the widest distribution, since it's known from several localities along the east coast between the most southerly sampling site, the river Henandrano, 12 km to the north of Fort Dauphin, and the most northern site, a small lake near Mahambo, 150 km to the north of the port of Tamatave (=Toamasina). This species would include also the small-spotted *Paratilapia* which inhabits lake Alaotra, located northeast of Antananarivo (de Rham & Nourissat, 2004).

In 1993, de Rham and Nourissat collected small-spotted specimens in the Manantanana river, situated in the vicinity of the municipality of Ambalavao, in the southern highlands of the Betsileo area (de Rham & Nourissat, 2004). These specimens belonged to the *Paratilapia* population associated with the local name "Fiamanga", characterized by a dark-blue body spangled with many fluorescent, silver, blue, green or gold small spots, that are equally abundant in males and females (de Rham & Nourissat, 2004). De Rham and Nourissat (2004) suggested that these southern highlands *Paratilapia* might belong to an undescribed species that they called *Paratilapia* sp. Ambalavao (also *Paratilapia* sp. Betsileo highlands), and that this species would be identical with the "Fiamanga" population. *Paratilapia* sp. "Fiamanga" is also recognized in Loiselle's classification (2011), because of strong phenotypical differences to the east-coast population, which is less brightly coloured and has a different general appearance. Curiously, even though the specimens that were brought back from Ambalavo to France were lost or interbred with other stocks (de Rham & Nourissat, 2004), this species still seems to be available in commercial pet trade in the USA under the name *"Paratilapia* Fianarantosa" (Albering, 2006).

A new undescribed *Paratilapia* species from the southwest of Madagascar was first recognized by Loiselle and Stiassny (1993), after the morphological examination of specimens preserved in the collections of the AMNH; these specimens were labeled as originating from Lake Ihotry, Madagascar's second-largest lake south of the river Mangoky (Loiselle & Stiassny, 1993). Even though this population has apparently become extinct, de Rham and Nourissat subsequently managed to capture a few small-spotted specimens in one of the lakes which are part of the Onilahy River drainage, situated around 200 km to the southeast of Lake Ihotry (de Rham & Nourissat, 2004). They decided to assign the name *P.* sp. Southwest to this putative undescribed species, since the relationship with the *Paratilapia* population found in Lake Ihotry has never been investigated (de Rham & Nourissat, 2004).

A fourth so far undescribed small-spotted species identified by de Rham and Nourissat is *Paratilapia* sp. Northwest, whose populations are assumed to inhabit the lower Ikopa-Betsiboka drainage system, even though few localities have been sampled. Contrary to Loiselle (2011), de Rham and Nourissat (2004) strongly believed that the large-spotted *Paratilapia* living in swamps (Ikopa rivers watershed) near Antananarivo until the 1980's (Raminosa 1987; Loiselle & Stiassny, 1993) did not mix with the small-spotted species which inhabits the lower part of the same drainage.

According to Loiselle (2011), both this small-spotted phenotype population from the lower reaches of the Bestiboka drainage and large-spotted individuals still living in central highlands near the capital city should be referred to as *P. bleekeri,* given

the absence of physical barriers to the downstream movements of the fishes, and to gene flow across the Bestiboka drainage.

The picture of the species diversity within the genus *Paratilapia* suggested by Loiselle (2011) is much more complex than that drawn by de Rham and Nourissat (2004), as it's shown in Figure 4. It implies the recognition of *P. polleni*, both the two previously described, but not formally recognized species *P. bleekeri* and *P. typus,* as well as eight undescribed species: *P.* sp. "Andapa", *P.* sp. "Fony", *P.* sp. "Vondrozo", *P.* sp. "Anosy", *P.* sp. "Fiamanga", *P.* sp. "Onilahy", *P.* sp "Ihotry" and *P.* sp. "Ilempo". Although the main points of this taxonomic review have apparently emerged from genetic analyses carried out by Sparks and colleagues (Loiselle, 2011; also see Sparks, 2023), details of these analyses have so far remained unpublished and were only indirectly referred to.

Among these undescribed species, *P.* sp. "Andapa", known to inhabit the Lokoho River in the Sava region of north-eastern Madagascar (Loiselle, 2011), is the only one that exhibits a large-spotted phenotype. The species was first reproduced in captivity by the Malagasy fish breeder Guy Tam Hyok and subsequently brought to France by Nourissat, who bred it and gave a pair to Loiselle. According to Sparks' genetic analyses, fish from Andapa were not only distinct from *P. polleni,* but, despite the differences in colour pattern, they would be more closely related to the small-spotted *Paratilapia* of the east coast (Loiselle, 2011).

P. sp. "Fony" has the widest distribution along the east coast, since its range extends from the Maningory River in the North, southward to the Mananara du Sud River, inclusive of the basin of Lake Alaotra (Loiselle, 2011).

P. sp. "Vondrozo" is found in the upper parts of the Manampatrana River near the town of Vondrozo. It is distinct from other *Paratilapia* by its almost uniform black colouration of older males, which, when sexually active, develop a massive nuchal hump; the sparse blue spotting is restricted to the head and fins (Loiselle, 2011).

The coastal rivers of the Anosy region, immediately to the North of Fort Dauphin (=Tolanaro), harbour *P.* sp. "Anosy", whose spot pattern is sparser than that of *Paratilapia* sp. "Fony". Additionally, the melanophore pattern of stressed individuals also differs from that of other east coast populations (Loiselle, 2011).

Paratilapia sp. "Fiamanga" corresponds to the species recognized by de Rham and Nourissat (2004) as *Paratilapia* sp. Ambalavao, found in the headwaters of the river Mangoky.

Paratilapia sp. "Onilahy" is native to the Onilahy River, draining to the southwest of Madagascar near the town of Toliara (=Tulear).

Paratilapia sp. "Ihotry" may represent a species restricted to Lake Ihotry, south of the Mangoky River. It was brought near extinction in the first years of the 2000s, possibly in connection to a dramatic reduction of the lake's area since the 1930s. According to Loiselle (2011), it remains unclear whether Lake Ihotry harbours its own endemic species or either a population of the "Fiamanga" species or the "Onilahy" species.

Finally, the last undescribed species presented by the author, considering the results of Sparks' unpublished genetic analyses, is *Paratilapia* sp. "Ilempo". A population

assigned to this species was found in Lake Ilempo, a small lake adjacent to the town of Itasy (Loiselle, 2011).

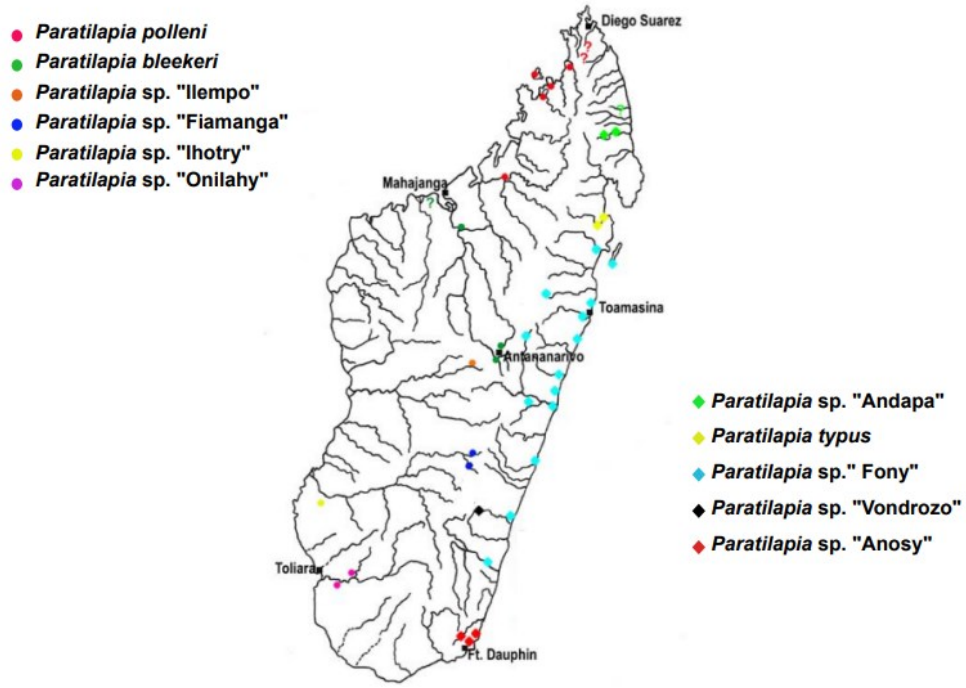

Figure 4: Distribution of Paratilapia *taxa according to Loiselle (2011).*

1.3. Aims of the thesis

To resolve the taxonomy within the genus *Paratilapia*, this master's project aims to analyse mitochondrial DNA sequences obtained from both modern and historical tissues of the alleged species' specimens, including type material of *Paratilapia polleni* Bleeker, 1868 and *P. typus* (Bleeker, 1878).

Type specimens stored in biological collections and natural history museums worldwide are extremely valuable, since these are key references to species description, and often the only available reliable sources for the reconstruction of the tree of life (Straube et al., 2021). Despite the relevance of type material for biological studies, type specimens could be included in routine genetic analyses only in recent years, thanks to advances in DNA extraction protocols (e.g., Rohland et al., 2004; Dabney et al., 2013; Straube et al., 2021) and next-generation sequencing (NGS) methods (e.g., Gansauge, 2013; Gansauge et al., 2017; Straube et al., 2021). The *museomics* protocols applied in the present study allow to recover nucleotide sequences from old tissues, despite the high degree of DNA fragmentation expected for both the effect of time and of preservation methods of museum specimens, which negatively affect the integrity of DNA molecules.

DNA preserved in museum specimens that are archived in collections as voucher specimens is called historical DNA (hDNA) and is usually less than 200 years old (Raxworthy & Smith, 2021).

2. MATERIALS AND METHODS

The molecular dataset compiled for this study (Table 1) includes three different mitochondrial genes, represented by 21 sequences of the Cytochrome Oxidase Subunit 1 (COI), 20 sequences of the 16S rRNA (16S), and 11 sequences of the NADH dehydrogenase 2 (ND2). All the ND2 sequences and nine of the 16S sequences have been newly obtained from nine recent tissue samples, alongside the DNA sequences obtained from one syntype specimen of *P. polleni* Bleeker, 1868. The dataset also includes sequences of these marker genes available from NCBI's GenBank for *P. polleni*, as shown in Table 1.

The geographical distribution of the specimens investigated in this study covers the North of Madagascar, including twelve specimens from two crater lakes originating from the offshore island of Nosy Be (Lake Amparihibe and Lake Amparihimirahavavy), and two specimens from the breeding stocks of the Zoological Institute in Basel and the Zoo of Cologne, which originally derive from individuals collected in the Andapa region, in northeast Madagascar (Indermaur, pers. comm.). The dataset also includes four specimens from eastern Madagascar and two syntype specimens of *P. polleni* Bleeker, 1868, which were collected on the north-western side of the island (de Rham & Nourissat, 2004; Stiassny & Loiselle, 1993). Additionally, the type specimen of *P. typus* (Bleeker, 1878) was included in Table 1.

Recent tissue samples have been preserved in absolute ethanol, while the conservation technique originally adopted for historical type specimens is not precisely known. However, given that formalin – a formaldehyde solution frequently used for the fixation of specimens – became commercially available only around the end of the nineteenth century, preservation in ethanol is assumed for the type specimens collected in 1868 and 1878.

Taxon	*Voucher ID*	*Locality*	*GeneBank Accession #*			*Sequenced Sample*		
			COI	16S	ND2	16S	ND2	COI
P. polleni	RDR_0789	Nosy Be, Lake Amparihibe	ON604401	-	-	*Yes*	*Yes*	*No*
P. polleni	RDR_0791	Nosy Be, Lake Amparihibe	ON604402	-	-	*Yes*	*Yes*	*No*
P. polleni	RDR_0792	Nosy Be, Lake Amparihibe	ON604403	-	-	*Yes*	*No*	*No*
P. polleni	RDR_0793	Nosy Be, Lake Amparihibe	ON604398	-	-	*Yes*	*Yes*	*No*
P. polleni	RDR_0794	Nosy Be, Lake Amparihibe	ON604399	-	-	*Yes*	*Yes*	*No*
P. polleni	RDR_0795	Nosy Be, Lake Amparihibe	ON604400	-	-	*Yes*	*Yes*	*No*
P. polleni	RDR_0803	Nosy Be, Lake Amparihibe	ON604404	-	-	*Yes*	*Yes*	*No*
P. polleni	TZTIS35	Zoo Cologne	ON604407	MW217523	-	*Yes*	*Yes*	*No*
P. polleni	PP	Zoological Institute Basel	-	-	-	*Yes*	*Yes*	*No*
P. polleni	FGMV 2002.F-B25	Nosy Be, Lake Amparihimirahavavy	ON604405	ON611972	-	*No*	*No*	*No*
P. polleni	FGMV 2002.F-B35	Nosy Be, Lake Amparihimirahavavy	ON604406	-	-	*No*	*No*	*No*
P. polleni	FGMV 2002.F-B36	Nosy Be, Lake Amparihimirahavavy	ON604409	ON611973	-	*No*	*No*	*No*
P. polleni	FGMV 2002.F-B37	Nosy Be, Lake Amparihimirahavavy	ON604408	ON611970	-	*No*	*No*	*No*
P. polleni	FGMV 2002.F-B42	Nosy Be, Lake Amparihimirahavavy	-	ON611971	-	*No*	*No*	*No*
P. polleni	ZCMV408	Madagascar	ON604411	-	-	*No*	*No*	*No*
P. polleni	ZCMV419	Manombo Camp	ON604410	ON611969	-	*No*	*No*	*No*
P. polleni sp. "East"	UMMZ JSS 94-26	Mananjary	AY263884	AY263818	-	*No*	*No*	*No*
P. polleni sp. "Manombo"	UMMZ PNR 96-3-4	Southeast coast: Manombo	AY263885	AY263819	-	*No*	*No*	*No*
P. polleni sp. "Northwest"	AY263887	Northwest Madagascar	AY263887		-	*No*	*No*	*No*
P. polleni	UMMZ_243192	Madagascar	DQ119222	DQ119193	-	*No*	*No*	*No*
P. polleni	AP009508	Madagascar	AP009508.1:5474-7024	AP009508.1:1092-2782	AP009508.1:4044-5090	*No*	*No*	*No*
***P. polleni* Bleeker, 1868 - syntype**	**RMNH.PISC.3934**	**Ambazoana River**	-	-	-	***Yes***	***Yes***	***Yes***
P. polleni Bleeker, 1868 - syntype	RMNH.PISC.4483	Ambazoana River	-	-	-	*failed*	*failed*	*failed*
P. typus (Bleeker, 1878) - syntype	RMNH.PISC.6692	Mananara River	-	-	-	*failed*	*failed*	*failed*
Katria katria - outgroup	AP009507	-	AP009507.1:5484-7073	AP009507.1:1094-2786	AP009507.1:4048-5094	*No*	*No*	*No*

Table 1: Vaucher IDs and GenBank accession number of specimens used in this study. The line in bold represents the type specimen of P. polleni *included in the phylogenetic analysis.*

2.1. Obtaining DNA sequences from recent tissues

The recent tissue samples of *P. polleni* included in this study were first used for DNA extraction by Vences et al. (2022), who also sequenced and published the mitochondrial gene fragment COI from these samples. The DNA was re-extracted from the same samples at the DNA lab of the Natural History Museum of Oslo using a column-based approach (E.Z.N.A. Tissue DNA Kit, Omega BIO-TEK). Additionally, a DNA extraction from the tissue sample of specimens originating from the breeding stock in Basel (PP) was performed using Nanobind HMW DNA Extraction-Standard TissueRuptor Tissue Protocol (Circulomics).

Using the DNA extracts of these recent samples, the 16S rRNA and ND2 mitochondrial gene fragments were amplified with polymerase chain reactions (PCRs) and sequenced using a Sanger sequencing approach. The amplification of the 16S gene was performed using the primers 16SA-L (5'-CGCCTGTTTATCAAAAACAT-3') and 16Sbr-H (5'-CCGGTCTGAACTCAGATCACGT-3') (Kocher et al., 1989; Palumbi, 1996). The touchdown PCR thermal profile included 3 minutes of denaturation at 95 °C, followed by 30 cycles of 30 seconds denaturation at 95 °C, 45 seconds annealing, starting from 65 °C for the first 14 cycles, decreasing temperature by 0.5 °C per cycle up to 58 °C, 45 seconds elongation at 72 °C, and final elongation at 72 °C for 5 minutes.

The mitochondrial gene ND2 was PCR-amplified using the primers ND2Met (5'-CATACCCCAAACATGTTGGT-3') and ND2Trp (5'-GTSGSTTTTCACTCCCGCTTA-3') (Kocher et al., 1995). The PCR thermal profile consisted of 3 minutes at 95 °C, followed by 34 cycles of 30 seconds denaturation at 95 °C, 30 seconds annealing at 49 °C, 46 seconds extension at 72 °C, and final elongation at 72 °C for 5 minutes.

Following the PCR-amplification, generated amplicons were sequenced using the Sanger sequencing approach on the Applied Biosystem 3730xl DNA Analyzer machine provided by Macrogen Europe.

2.2. Obtaining DNA sequences from historic type material

Tissue samples of two syntype specimens of *P. polleni* Bleeker, 1868 and one type specimen of *P. typus* (Bleeker, 1878) were requested from the Naturalis Museum in the Netherlands.

The initial weight of the tissue pieces obtained from the syntype specimens of *P. polleni* Bleeker, 1868 with the voucher numbers RMNH.PISC.3934 and RMNH.PISC.4483 was, respectively, 54 mg and 20 mg. The tissue piece obtained from the syntype specimen of *P. typus* (Bleeker, 1878) with the voucher number RMNH.PISC.6692 had a weight of 12.8 mg, but, differently from the other two tissue samples, it was completely dried up due to evaporation of its preservation liquid over time. This type material was then newly analysed for this study in collaboration with Prof. Nicolas Straube from the University of Bergen, who optimized *museomics* protocols for hDNA extraction and single-stranded library preparation (Straube et al., 2021).

2.2.1. Historical DNA extraction

Prior to DNA extraction, the three tissue samples from the type specimens were weighted and washed with 1 ml PBS buffer twice to decrease the amount of ethanol and other potential inhibitors. The procedure followed was based on the widely adopted method for ancient DNA extraction developed by Dabney et al. (2013), in which DNA is bound to a silica membrane in the presence of a chaotropic salt (guanidine hydrochloride) buffer.

Tissue samples and two negative controls were incubated in 1 ml of extraction buffer (5 M GuSCN, 50 mM Tris pH 8.0, 25 mM NaCl, 20 mM EDTA, 1% Tween 20, 1% 2-mercaptoethanol) (Rohland et al., 2004) for ~18 hr rotating at room temperature. After centrifugation, 1 ml of the supernatant was added to 13 ml of binding buffer (5 M guanidine hydrochloride, 40% isopropanol, 0.05% Tween 20, 90 mM sodium acetate) as described in Dabney et al. (2013). DNA was purified using the MinElute silica spin column (QIAGEN) and eluted twice in TET buffer (10 mM Tris-HCL, 1 mM EDTA, 0.05% Tween 20) for a total of 25 µl. Afterwards, the DNA concentration of each extract was measured with a Qubit fluorometer (Thermo Fisher) with high sensitivity reagents, using 1 µl of DNA extract.

2.2.2. Single-stranded library preparation

Dual-indexed single-stranded libraries were prepared from each DNA extract following the procedure implemented by Prof. Straube, based on Gansauge and Meyer (2013, 2017). Alongside the DNA extracts of type specimens, a library blank was also processed to serve as a negative control.

Extracts were initially treated with uracil-DNA glycosylase to remove deoxyuracils resulting from cytosine deamination, and Endonuclease VIII, to cleave abasic sites. Following heat denaturation, *T4* DNA ligase promotes ligation of biotinylated adapter oligos to the 3' ends of the single stranded DNA fragments.

Ligation products are immobilized on streptavidin-coated magnetic beads and, in subsequent steps, the strand complementary to the original template is filled in. Synthesis of the second strand is carried out using the Klenow fragment of *Escherichia coli* DNA polymerase I, incorporating the proximal sections of 5' and 3' Illumina adapter sequences, which serve as priming sites for dual indexing PCR.

The optimal number of dual-indexing PCR cycles was individually determined for each library by qPCR, carried out with three replicates per library and an additional negative control (qPCR blank) prior to amplification. Libraries were amplified using P7 and P5 index primers individually assigned to each sample. The amplified libraries were pooled, purified using the QIAGEN MinElute PCR Purification kit, and then quantified using a DNA1000 chip on the Bioanalyzer 2100 (Agilent Technologies). Finally, libraries were sequenced using a next-generation sequencing approach (NGS), on Illumina's Miniseq instrument.

2.3. Bioinformatic analyses for phylogenetic inference

Sequences of the 16S, COI, and ND2 markers, obtained with the Sanger sequencing technique from the nine recent tissue samples, were trimmed and edited using Geneious Prime 2023.1.2 (https://www.geneious.com).

Sequencing of the amplified libraries for type specimens on Illumina's Miniseq instrument generated per-cycle BCL files as primary output; these files needed to be converted into FASTQ format for downstream analyses. The software bcl2fastq v. 2.0 (Ewel et al., 2016) was run both to translate BCL files into FASTQ files and to separate multiplexed samples (demultiplexing), identified by index sequences.

The multiplexed reads were assigned to each sample and stored in corresponding FASTQ files. Before mapping, reads were processed with the software Fastp (Chen et al., 2018; https://academic.oup.com/bioinformatics/article/34/17/i884/5093234) for quality filtering. Reads had to have a minimum length of 15 bp to pass the quality filtering, reads with more than 10 unresolved nucleotides (coded as "N") were discarded, reads with a poly-G tail of more than 10 "G" at the end were excluded, adapter sequences were automatically detected and trimmed, and duplicated reads were discarded.

Following quality filtering, each read was matched to the NCBI non-redundant (NR) sequence database (downloaded on 16 May 2023) with the BLASTX algorithm as implemented in Diamond v.2.1.6 (Altschul et al. 1990; Buchfink et al. 2015), to test for the presence of endogenous DNA. The resulting taxon assignments were plotted with MEGAN v.6.21.4 (Huson et al., 2016).

The reads were then mapped to a reference mitochondrial genome of *P. polleni* (NCBI accession number AP009508; Azuma et al., 2008) using the software BWA (Li & Durbin, 2009). Overlapping reads allowed to extract sequences for the three marker genes under investigation (16S, COI, and ND2) using the software IGV v. 2.16.1 (Robinson et al., 2011).

Ambiguous bases were replaced with letter code according to the Nucleotide Ambiguity Code (IUPAC). Subsequently, DNA sequences of each marker were aligned using the software MAFFT v. 7.511 (Katoh & Toh, 2008). Multiple sequence alignments were used as input files for the software IQ-TREE2 v. 2.0 (Minh et al., 2020) to build gene trees, and thus to estimate the phylogenetic relationships among the samples based on the maximum likelihood method. As outgroup in the phylogenetic analyses, mitochondrial sequences of *K. katria* (NCBI accession number AP009507; Azuma et al., 2008) were included and represented the sister group of *Paratilapia*, the clade combining all other genera within Ptychochrominae.

Besides analysing each gene tree separately, the multiple sequence alignments for 16S and COI were also concatenated to generate a tree for a larger fraction of the mitochondrial genome that includes all specimens of the dataset. The multiple sequence alignment for ND2, however, was not concatenated with the other two alignments due to the large amount of missing data in the ND2 alignment. Bootstraps values were calculated for all trees to assess node support.

3. RESULTS

3.1. PCR amplification and Sanger sequencing applied to recent tissues

Initial attempts to PCR-amplify 16S failed, suggesting that the used 16S amplification primers might not bind to the *Paratilapia* mitochondrial genome. The results of a similarity search executed with the tool BLAST (Altschul et al., 1990) proved that the amplified sequences did not represent the mitochondrial genome of *P. polleni*. This was due to the use of the 16S H-new reverse primer, which was listed in Vences et al. (2022) as a primer suitable for cichlids, while it was originally designed to amplify frog DNA. The PCR was repeated using the 16Sbr-H reverse primer (Kocher et al., 1989; Palumbi, 1996), which led to the successful amplification of the 16S gene for all specimens with fresh extracts available.

Even though amplified fragments were Sanger-sequenced in both forward and reverse directions, only the sequences obtained with the forward primer, 16SA-L were retained and used in sequence alignments, owing to a very low base call quality in the sequences with the reverse primer, 16Sbr-H.

The gene ND2 could be amplified and sequenced successfully for all samples, except for RDR_0792, that showed unclear chromatograms for sequences amplified with both the forward and the reverse primer. This sample was therefore excluded from the ND2 alignment. Among the recent samples, no COI sequence was available for the specimen PP.

3.2. DNA extraction and NGS applied to type specimens

The DNA extractions from the syntype specimens of *P. polleni* were successful, although the Qubit measurements showed evidence of a less effective extraction from the tissue sample of *P. polleni* Bleeker, 1868 with the voucher number RMNH.PISC.4483. The DNA extract of this sample had a lower DNA concentration than those of the *P. typus* sample and the *P. polleni* sample with the voucher number RMNH.PISC.3934 (Table 2).

Results of Qubit measurements for each indexed library reported in Table 2 show contrasting values compared to previous Qubit results of DNA extracts. The sample of *P. polleni* (RMNH.PISC.4483) that had the lowest Qubit value for the DNA extract had the highest value when measuring the DNA concentration of corresponding indexed library. In Table 2 are also reported the results of the size distribution's analysis of DNA fragments performed with the Agilent 2100 Bioanalyzer system. The peak size of each library – ranging between 151 bp and 162 bp – is a proxy of the DNA fragments' size.

In combination with the Qubit measurements, the DNA concentration associated with each library's fragment length peak was used to compute the molarity, and ultimately to pool samples in equimolar ratios for the sequencing run.

Syntype	Sample ID	Qubit measurements		Bioanalyzer	
		DNA extracts (ng/µl)	Indexed libraries (ng/µl)	Peak Size (bp)	Concentration (pg/µl)
P. typus (Bleeker, 1878)	RMNH.PISC.6692	1.17	3.53	158	1371.31
P. polleni Bleeker, 1868	RMNH.PISC.3934	1.50	2.87	162	1297.53
P. polleni Bleeker, 1868	RMNH.PISC.4483	0.87	17.6	151	717.11
-	(-)E	-	30.4	151	9 742.94
-	(-)L	-	25.1	151	9 145.16

Table 2: Qubit results for each sample sequenced with Illumina platform, after both the DNA extraction and the PCR amplification of indexed libraries. Qubit results of indexed libraries are reported also for extraction negative (-) E and library negative (-) L. The fragment length peak size estimated in the Bioanalyzer analysis and the corresponding DNA concentrations are also included.

Next-generation sequencing (NGS) of the samples taken from type specimens (*Paratilpia polleni* Bleeker, 1868 and *P. typus* (Bleeker, 1878)) led to recover a total of 7,482,886 reads, including those obtained from library negative, extraction negative, and undetermined reads, that could not be assigned to any sample after demultiplexing. The numbers of reads before and after quality filtering and the percentages of reads passing filtering for each library are reported in Table 3.

Syntype	Library	Number of reads	Number of reads after filtering	Percentage of reads passing filtering
P. typus (Bleeker, 1878)	RMNH.PISC.6692	1,244,644	385,877	31.0
P. polleni Bleeker, 1868	RMNH.PISC.3934	576,005	531,056	92.2
P. polleni Bleeker, 1868	RMNH.PISC.4483	1,255,284	283,663	22.6
	undetermined	420,614	331,169	78.7
-	(-)E	2,214,739	403,809	18.2
-	(-)L	1,771,600	210,171	11.9
	Sum	7,482,886	2,145,745	28.7

Table 3: Numbers of reads before and after quality filtering, including those reads obtained from extraction negative (-) E, library negative (-) L, and undetermined reads. The underscored line represents the type specimen included in the phylogenetic analysis.

Plots representing the taxonomic assignments for each library show the prevalence of reads derived from bacterial contamination in both the sample of *P. typus* (Fig. 5) and the *P. polleni* sample with the voucher number RMNH.PISC.4483 (Fig. 6). In contrast, the taxonomic assignments of reads obtained from the sample of *P. polleni* with the voucher number RMNH.PISC.3934 reveals frequent matches with cichlid sequences, indicating the presence of endogenous DNA alongside prokaryotic sequences (Fig. 7).

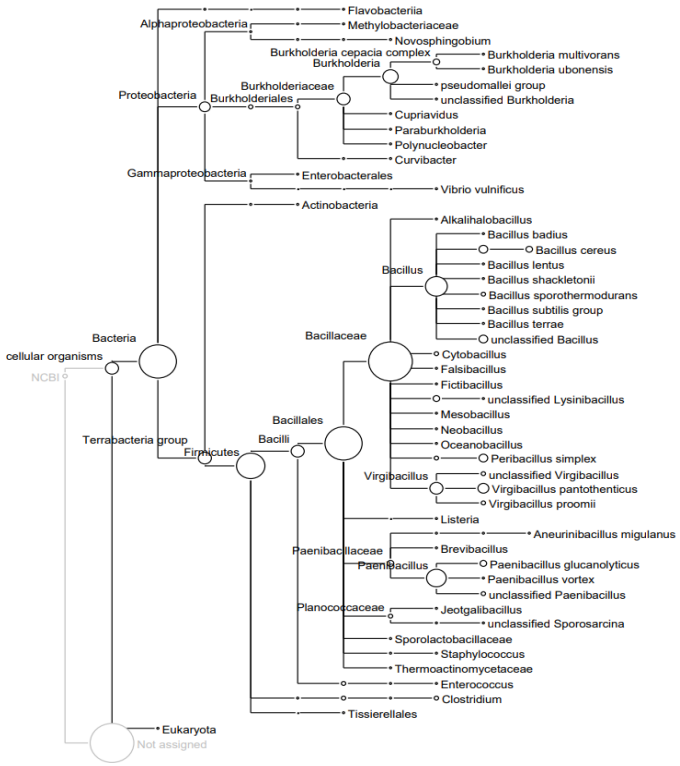

Figure 5: Taxonomic assignments for reads extracted from the type specimen of Paratilapia typus *(Bleeker, 1878). Plot generated with MEGAN v.6.21.4.*

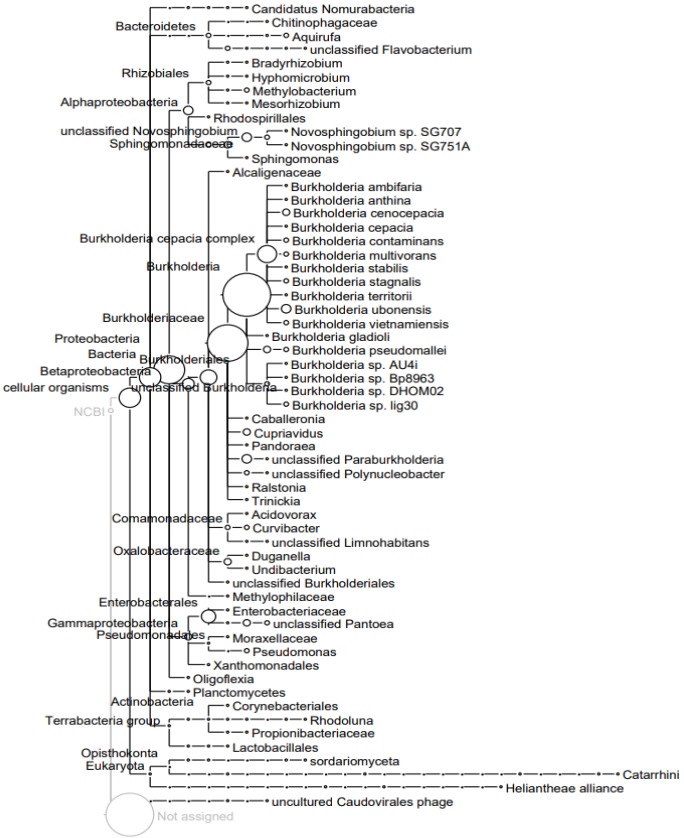

Figure 6: Taxonomic assignments for reads extracted from the syntype specimen of Paratilapia polleni *Bleeker, 1868 with the voucher number RMNH.PISC.4483. Plot generated with MEGAN v.6.21.4.*

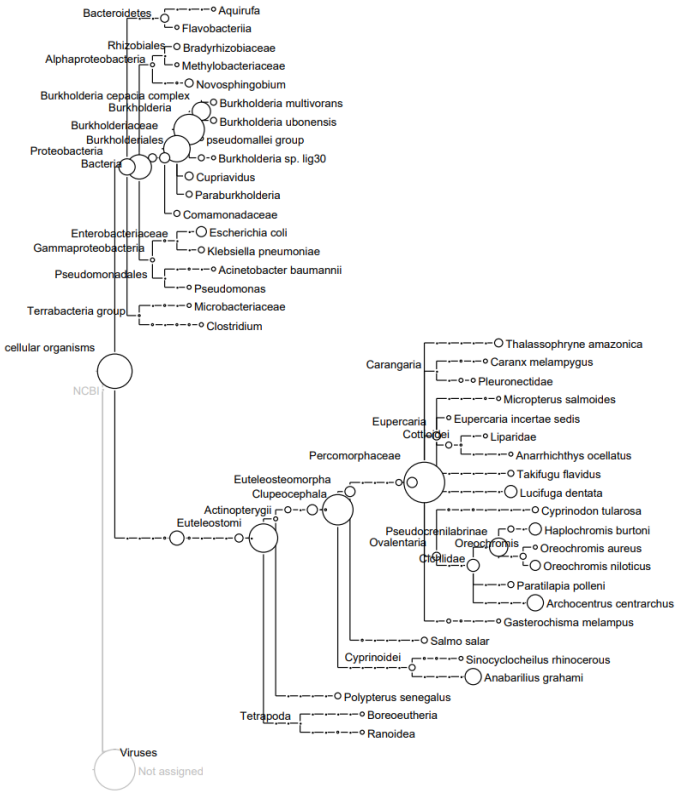

Figure 7: Taxonomic assignments for reads extracted from the type specimen of Paratilapia polleni *Bleeker, 1868 with the voucher number RMNH.PISC.3934. Plot generated with MEGAN v.6.21.4.*

Read mapping allowed to find 48,564 undetermined reads which mapped to the *P. polleni* nuclear genome assembly (NCBI accession number VUQI00000000), and 124 undetermined reads mapping to the mitochondrial genome of the same species (NCBI accession number NC_011170.1). As these could not be assigned to a sample, they were not further considered.

For the syntype specimen of *P. polleni* with the voucher number RMNH.PISC.3934, 278,596 out of 531,056 reads mapped to the *P. polleni* genome assembly, including 2,315 that mapped to its mitochondrial genome. This read concentration proved sufficient to extract sequences of all the marker genes under investigation, which were included in the phylogenetic analysis.

On the contrary, the second syntype specimen of *P. polleni* with the voucher number RMNH.PISC.4483, had only 2,157 out of 283,663 reads mapping to the *P. polleni* genome assembly, of which only 5 mapped to the mitochondrial genome. Thus, this specimen was not included in the final dataset.

For the type specimen of *P. typus*, only 148 out of 385,877 reads mapped to the *P. polleni* genome assembly. Of these, only one read (with a length of 37 bp) mapped to the mitochondrial genome of the target species, so that this sample also could not be included in the phylogenetic analysis. This failure to obtain a good amount of data from this type specimen might be a result of the dry condition that the tissue had before the DNA extraction.

The multiple sequence alignment for the 16S gene included 20 sequences, had a total length of 282 bp, and a proportion of missing and ambiguous sites of 2.5%.

The COI multiple sequences alignment included 21 sequences and had a total length of 498 bp, with a proportion of missing sites (no ambiguous sites) of 0.9%. The multiple sequence alignment included 11 sequences in the case of the gene ND2 and had a total length of 932 pb with a proportion of missing and ambiguous base of 5.8%.

3.3. Phylogenetic inference

The phylogenetic analysis of the 16S, ND2, and COI mitochondrial genes led to three corresponding gene trees, estimated with the maximum likelihood method.

The 16S gene tree (Fig. 8) shows the syntype specimen of *P. polleni* (RMNH.PISC.3934) clustering with eleven samples from two crater lakes of the island of Nosy Be (Lake Amparihibe and Lake Amparihimirahavavy). Within this group, the 16S sequence of the type is identical to four sequences of individuals collected in Lake Amparihimirahavavy. This gene tree additionally includes another cluster of five identical sequences, of which those associated with a precise collection locality are from the southeast coast of Madagascar (Mananjary and Manombo). The samples TZTIS35 and PP, from two breeding stocks in the Zoo of Cologne and in Basel (Zoological Institute), respectively, are also clustering together in 16S gene tree.

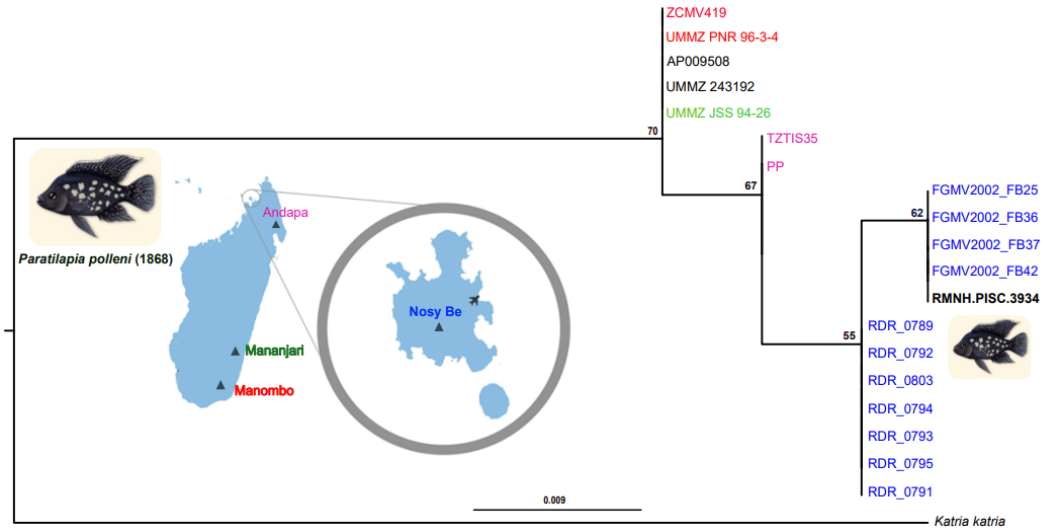

Figure 8: 16S maximum likelihood gene tree of Paratilapia *specimens included in the study. Bootstrap values above 50 are reported at each node. The specimen highlighted in bold is the syntype of* P. polleni *included in the analysis.*

In the COI gene tree (Fig. 9), the sequence of the syntype specimen is identical to seven sequences form the Lake Amparihibe, which do not show any genetic variation. In this phylogeny, the cluster comprising sequences from the southeast coast of Madagascar is nested inside the cluster of sequences from Nosy Be.

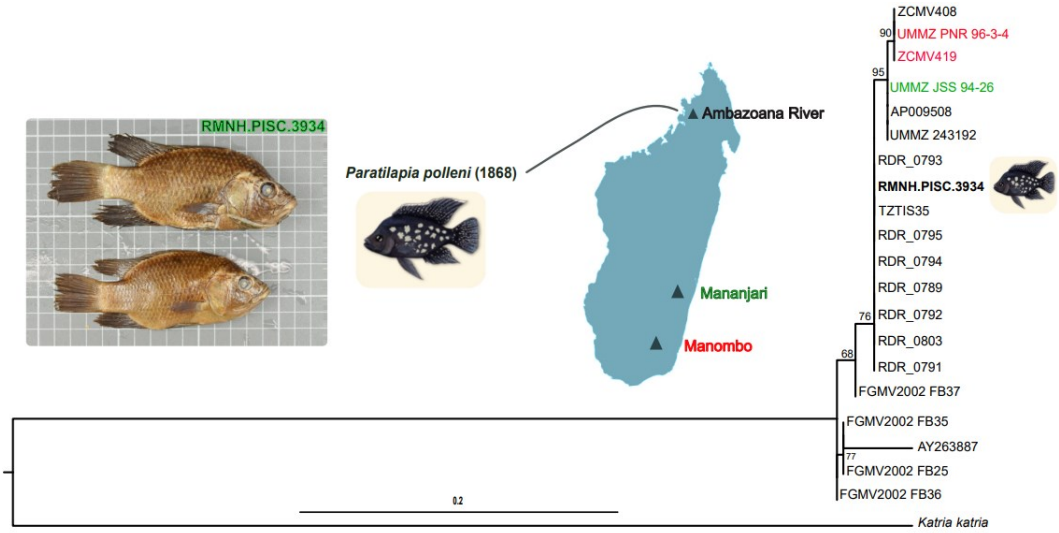

Figure 9: COI maximum likelihood gene tree of Paratilapia *specimens included in the study. Bootstrap values above 50 are reported at each node. The specimen highlighted in bold is the syntype of* P. polleni *included in the analysis. The picture of two syntypes of* P. polleni *is from the Zoology and Geology catalogues of the Naturalis Museum in Leiden.*

The ND2 gene tree (Fig. 10) is the least informative one because of the limited dataset comprising only ten sequences, six of which are from individuals collected in Lake Amparihibe. Despite the small amount of available data for the ND2, the phylogenetic position of *P. polleni* is in agree with the other gene trees, showing the sequence of the syntype specimen clustering with a sequence from Nosy Be.

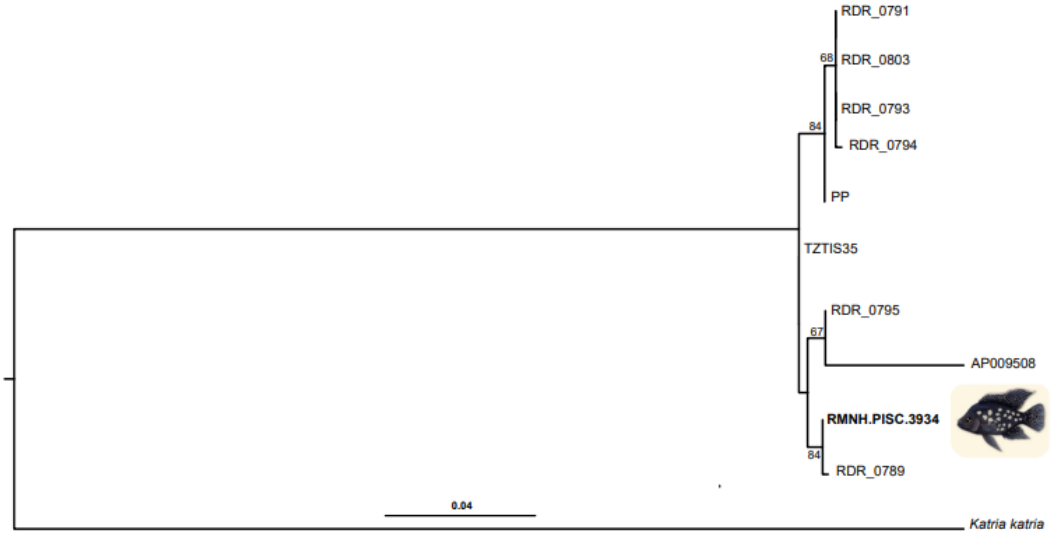

Figure 10: ND2 maximum likelihood gene tree of Paratilapia *specimens included in the study. Bootstrap values above 50 are reported at each node. The specimen highlighted in bold is the syntype of* P. polleni *included in the analysis.*

As expected from the gene trees of 16S and COI, the tree based on the concatenated alignment (Fig. 11) also exhibits a phylogenetic pattern in which the sequence of the syntype clusters with sequences from the north-western island of Nosy Be.

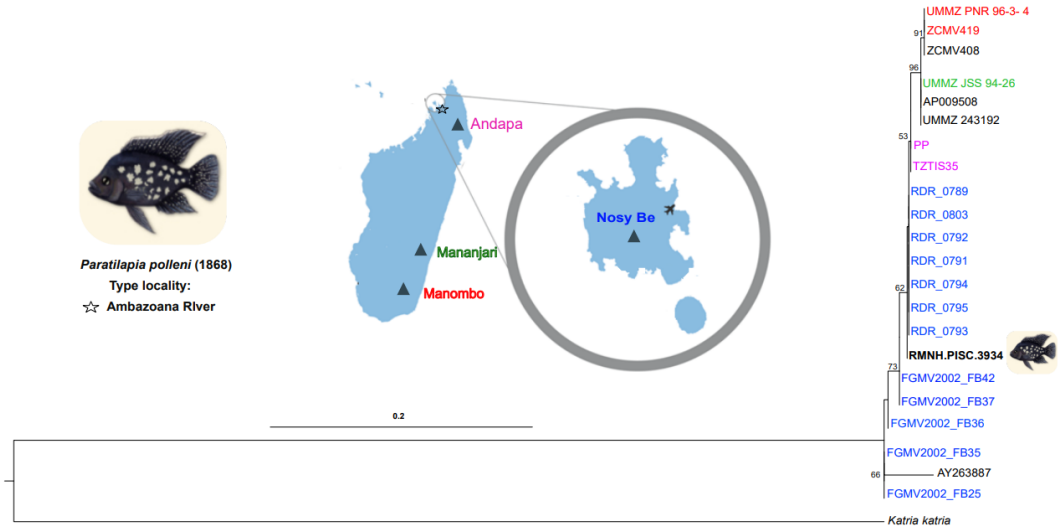

Figure 11: Maximum likelihood tree generated using the concatenated alignment for the 16S and COI markers. Bootstrap values above 50 are reported at each node. The specimen highlighted in bold is the syntype of P. polleni *included in the analysis.*

4. DISCUSSION

The phylogenetic analysis described herein clearly indicates that individuals living in freshwaters of the offshore island of Nosy Be should be assigned to the species described by Bleeker as *Paratilapia polleni* in 1868. Although the phylograms built for the three marker genes show slightly different topologies, these trees are all consistent with the phylogenetic placement of *P. polleni* as closely related to individuals collected in Nosy Be. This view is supported by the sequence identity, recovered in both 16S and COI gene trees, between the mitochondrial sequences of the type specimen and of individuals collected in the crater lakes of Amparihibe (13°18'54.0"S 048°12'35.2"E) and Amparihimirahavavy (13°18'08.3"S 48°14'08.9"E), situated on Nosy Be Island. This perfect match of *P. polleni* with Nosy Be specimens at the mitochondrial level, supports this island as a more likely type locality than the Ambazoana river (de Rham & Nourissat, 2004), as assumed by Stiassny and Loiselle (1993).

Lake Amparihibe is the largest of the nine lakes surrounding the peak of Mont Passot, situated in the western side of Nosy Be (Reinhardt et al., 2022), but its size is still smaller than most lakes at the mainland (surface of 160 hectares; volume of 46,000,000 m^3; depth between 29 and 43 m) (Razafitsiferana, 2015). In this study, the specimens collected in the Lake Amparihibe are clustering together in both the 16S and COI gene tree, showing no genetic variation in these markers' sequences. On the contrary, the phylogeny proposed by Lovaharisoa and Ranalison (2018) suggests the separation of two *Paratilapia* populations living syntopically in Lake Amparihibe. However, the coexistence of two phylogenetically distinct populations persisting within this lake seems unlikely, since the specimens, even if originating from two independent introductions, are not facing any barriers to interbreed in that comparatively small and homogeneous crater lake. Additionally, dating back any possible introductions of cichlids in Lake Amparihibe might be difficult, since several studies showed that human settlements and human-mediated introductions of livestock from the mainland existed for centuries in Madagascar, starting about two millennia ago (e.g., Burney et al., 2004; Reinhardt et al., 2022).

The 16S gene tree obtained in this study shows the strongest phylogeographic signal, for its topology identifies at least three clusters of sequences coherent with the geographic distribution of the samples. The two main clusters include sequences from the northwest and from the southeast of Madagascar, respectively, while the third cluster groups together the specimens PP and TZTIS35, captive bred descendants of individuals collected in the Andapa region, on the northeast coast of Madagascar (Indermaur pers. comm.).

The cluster of 16S sequences from the southeast of the island is consistent with the conclusions drawn by Lovaharisoa and Ranalison (2018), based on their phylogenetic analysis. These authors also assume the identification of two different species living in the North − *P. polleni sensu strictu* and *P.* sp. Nosy Be (Amparihibe) − and one small-spotted species in the East, namely *P.* sp. smallspots Est (Lovaharisoa & Ranalison, 2018). However, to verify the existence of these separate species, further results showing the phylogenetic position of all type specimens will be required.

The lack of information about the colouration patterns of the specimens included in the phylogeny prevents this research from solving the ambiguous dichotomy between large-spotted and small-spotted strains of *P. polleni.* Despite the genetic differences between south-eastern individuals and north-western individuals recovered in the 16S gene tree, the character "size-of-spots" is not recorded for these specimens, and thus the distinction between the large-spotted species *P. polleni* and one or more small-spotted species living in the East and South of Madagascar (de Rham & Nourissat, 2004; Loiselle, 2011) cannot be supported by this study. If the cluster of 16S sequences from the South-east had been associated with small-spotted specimens, this would have been a good indication of the genetic distance between the two strains of *P. polleni.*

To resolve the taxonomy of this genus, our dataset should be expanded, both by including the type specimens of *P. typus* and *P. bleekeri* and by increasing the sampling efforts of fresh tissue samples obtained from recently collected specimens, especially from the south-western Madagascar.

In this study, the possibility to obtain DNA sequences from type specimens was owed to the development of specific *museomics* protocols, allowing historical DNA extraction and sequencing from museum specimens (Gilbert et al., 2007; McGuire et al., 2018; Ruane & Austin, 2017; Tang, 2006). The DNA extraction from wet collection specimens requires more effort than from dry tissues (dried skin, pinned insect specimens, bird feathers, etc.), mainly because of the DNA damage induced by chemicals used for preservation and fixation (Straube et al., 2021). For specimens stored in alcohol (typically 70-75% ethanol), hydrolytic damage may occur if the alcohol concentration drops due to evaporation (Raxworthy & Smith, 2021). Additionally, from 1883 a solution of diluted formaldehyde started to be used for specimen fixation. Commonly called formalin, this solution hardens the tissue and thus helps to preserve its original shape. However, formaldehyde also induces DNA alteration such as base-pairing, promotion of denaturation, and cross-linking between DNA and proteins (Hoffman et al., 2015). These processes degrade the DNA of museum specimens and thus limit the outcome of even advanced *museomics* protocols for DNA extraction.

Despite the failure to extract sufficient endogenous DNA from one of the two *P. polleni* syntypes (RMNH.PISC.4483) and the syntype of *P. typus*, these results might be useful to develop further improvements of the already adopted procedures to extract DNA from collection specimens, currently based on those applied to ancient DNA (aDNA). All DNA preserved under natural conditions in biological material for more than 200 years is known as aDNA, and is mainly recovered from bones, plant matter, environmental samples, and subfossils (Raxworthy & Smith, 2021). The aDNA is in trace amounts, heavily degraded and, unlike hDNA, usually between thousands of years to a million years old (van der Valk et al., 2021). Despite this distinction between historical and ancient DNA, in both cases, the extraction procedures are designed to maximize the capture of small DNA fragments, which are usually lost when applying standard protocols.

This study aimed to analyse only short fragments of hDNA, derived from mitochondrial genes, but, as the field of *museomics* keeps moving forward, advances in DNA sequencing technology begin to allow the sequencing of entire genomes of specimens archived in collections (e.g., Hung et al., 2014; van der Valk et al., 2019). Following the assessment of the species' status, future studies should therefore apply genomic approaches to museum *Paratilapia* specimens also to assess the potential loss of genetic diversity and the degree of inbreeding due to population declines (Jensen et al., 2022), as already done for the Eastern Gorilla (*Gorilla beringei graueri*) in a recent study (van der Valk et al., 2019). Additionally, the genetic information should be considered in a definition of management and conservation units to prevent potential conservation plans from failing (e.g., Mortiz, 1999; Fraser & Bernatchez, 2001; Funk et al., 2012). Accounting for genetic differences (nuclear and mitochondrial) among populations is crucial to define conservation units within a species, which need to be treated independently in conservation plans. For this reason, even if a single *Paratilapia* species will be identified in future studies, the species' population structure potentially recovered with molecular data could warrant the distinction among populations with specific evolutionary features.

5. CONCLUSION

This preliminary study is the first in which a molecular phylogeny of the genus *Paratilapia* includes one of the syntype specimens. The lack of available genetic data of the syntypes specimens of *Paratilapia* is owed, not only to the challenging application of the necessary *museomics* procedures, but also to difficulties in obtaining tissue samples from the museums where these specimens are stored. For this study, the Naturalis Museum of Leiden provided samples of the syntypes of *P. polleni* Bleeker, 1868 and *P. typus* (Bleeker, 1878*)*, while the syntypes of *P. bleekeri* Sauvage, 1882, requested from the Musèum National d'Historie Naturelle of Paris, remained unavailable at the time when the molecular analyses were conducted in Bergen, but are included in a follow-up study.

However, even though only one syntype specimen of *P. polleni* could be included in this study, the genetic analyses exhibit a coherent phylogenetic placement of individuals collected on the island of Nosy Be. The phylogenetic position of these specimens clustering together with the syntype of *P. polleni*, leads to confirm their species status, a fundamental step to assess the specific diversity within the genus

Paratilapia. Improvements of the phylogenetic results achieved in this study will be available soon, since three syntypes specimens of *P. bleekeri* were recently analysed in Bergen, thanks to a second collaboration with Prof. Straube and his research team.

A full resolution of the taxonomy of these threatened Malagasy cichlids, even though challenging due to the necessary use of museum specimens, is extremely relevant, both for increasing our current knowledge about the phylogenetic history of the species within the genus *Paratilapia*, and for enabling the development of conservation strategies.

6. ACKNOWLEDGEMENTS

First, I acknowledge the Erasmus+ programme for providing the funding for my internship in Olso, where I have been working on my master's project at the Natural History Museum (University of Oslo).

I would like to sincerely thank Prof. Chiara Papetti, internal supervisor from the University of Padova, Prof. Michael Matschiner, supervisor from the University of Oslo, and the brilliant PhD student Thore Koppetsch. The contribution of these people was remarkably important for this thesis, since they have been monitoring the development of my project during this experience abroad.

I am also thankful to all the people being part of the Sex and Evolution Research Group (SERG) of the University of Olso, which had integrated me as a member of their team for the six months of my traineeship.

I also thank the DNA lab team of the University of Olso: the lab manager Jarl Andreas Anmarkrud and the head engineers Audun Schrøder-Nielsen and Lisbeth Thorbek.

Further, I would like to thank Prof. Nicolas Straube, the fish collection curator of the University of Bergen, Louise Lindblom and Kenneth Meland, who shared their knowledge about *museomics* protocols.

I am also thankful to Esther Dondorp from the Naturalis Museum in Leiden, who provided the museum specimens analysed in this study, and to Agnes Dettai, fish collection's curator of the Muséum National d'Histoire Naturelle of Paris.

I am grateful to Adrian Indermaur, Head of the Animal Facilities from the University of Basel, who clarified the origin of two *Paratilapia* specimens included in this study (PP and TZTIS35).

Finally, I would like to thank Prof. Miguel Vences for resolving the initial issues linked to the use of an unspecific reverse primer for the PCR-amplification of the 16S marker.

7. REFERENCES

Albering, J. (2006). *Paratilapia* sp. „Fianarantsoa ". *DCG-Informationen*, 37(11), 255-261.

Ali, J. R., & Aitchison, J. C. (2008). Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through latest Eocene (166–35 Ma). *Earth-Science Reviews*, 88(3-4), 145-166. doi: 10.1016/j.earscirev.2008.01.007.

Ali, J. R., & Hedges, S. B. (2023). The colonisation of Madagascar by land-bound vertebrates. *Biological Reviews*. doi: 10.1111/brv.12966.

Ali, J. R., & Huber, M. (2010). Mammalian biodiversity on Madagascar controlled by ocean currents. *Nature*, *463*(7281), 653-656. doi: 10.1038/nature08706.

Allgayer, R. (1998). Descriptions de *Lamena nourissati* sp. n. genre et espèce nouveaux, endémiques de Madagascar (Teleostei: Etroplinae). *Revue Française des Cichlidophiles*, *179*, 7-17.

Allgayer, R. (1996). Description d'une espèce nouvelle du genre *Paretroplus* Bleeker (Teleostei: Cichlidae) de Madagascar. *Revue française des Cichlidophiles* No. 159: 6-20.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410. doi: 10.1016/S0022-2836(05)80360-2.

Arnoult, J. (1960). Sur une nouvelle espèce de poisson Malgache (Cichlidae), *Paretroplus kieneri* n. sp. *Bulletin du Muséum National d'Histoire Naturelle (2)*, *32*(4), 305-307.

Azuma, Y., Kumazawa, Y., Miya, M., Mabuchi, K., & Nishida, M. (2008). Mitogenomic evaluation of the historical biogeography of cichlids toward reliable dating of teleostean divergences. *BMC Evolutionary Biology*, 8, 1-13. doi: 10.1186/1471-2148-8-215.

Belluardo, F., Scherz, M. D., Santos, B., Andreone, F., Antonelli, A., Glaw, F., ... & Crottini, A. (2022). Molecular taxonomic identification and species-level phylogeny of the narrow-mouthed frogs of the genus *Rhombophryne* (Anura: Microhylidae: Cophylinae) from Madagascar. *Systematics and Biodiversity*, 20(1), 1-13. doi: 10.1080/14772000.2022.2039320.

Benstead, J. P., De Rham, P. H., Gattolliat, J. L., Gibon, F. M., Loiselle, P. V., Sartori, M., ... & Stiassny, M. L. (2003). Conserving Madagascar's freshwater biodiversity. *BioScience*, 53(11), 1101-1111. doi: 10.1641/0006-3568(2003)053[1101:CMFB]2.0.CO;2.

Bleeker, P. (1868). Description de trois espèces inédites de chromidoïdes de Madagascar. Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen Afdeeling Natuurkunde (Ser. 2) v. 2: 307–314.

Bleeker, P. (1878). Sur deux espèces inédites de Cichloides de Madagascar. Verslagen en Mededeelingen der Koninklijke Akademie van Wetenschappen. Afdeeling Natuurkunde (Ser. 2) v. 12: 192-198, Pl. 3.

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., ... & Di Palma, F. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, 513(7518), 375-381. doi: 10.1038/nature13726.

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature methods*, 12(1), 59-60. doi: 10.1038/nmeth.3176.

Burbrink, F. T., Ruane, S., Kuhn, A., Rabibisoa, N., Randriamahatantsoa, B., Raselimanana, A. P., ... & Raxworthy, C. J. (2019). The origins and diversification of the exceptionally rich gemsnakes (Colubroidea: Lamprophiidae: Pseudoxyrhophiinae) in Madagascar. *Systematic Biology*, 68(6), 918-936. doi: 10.1093/sysbio/syz026.

Burney, D. A., Burney, L. P., Godfrey, L. R., Jungers, W. L., Goodman, S. M., Wright, H. T., & Jull, A. T. (2004). A chronology for late prehistoric Madagascar. *Journal of Human Evolution*, 47(1-2), 25-63. doi: 10.1016/j.jhevol.2004.05.005.

Catala, R. (1977). Poissons d'eau douce de Madagascar par René Catala (suite). *Rev. fr. Aquariol*, 1, 27-32.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), i884-i890. doi: 10.1093/bioinformatics/bty560.

Cichocki, F. P. (1976). Cladistic history of cichlid fishes and reproductive strategies of the American genera *Acarichthys*, *Biotodoma*, and *Geophagus*, Vol. 1. Unpublished PhD Thesis. University of Michigan, Ann Arbor.

Crottini, A., Harris, D. J., Miralles, A., Glaw, F., Jenkins, R. K., Randrianantoandro, J. C., ... & Vences, M. (2015). Morphology and molecules reveal two new species of the poorly studied gecko genus *Paragehyra* (Squamata: Gekkonidae) from Madagascar. *Organisms Diversity & Evolution*, 15, 175-198. doi: 10.1007/s13127-014-0191-5.

Crottini, A., Madsen, O., Poux, C., Strauß, A., Vieites, D. R., & Vences, M. (2012). Vertebrate time-tree elucidates the biogeographic pattern of a major biotic change around the K–T boundary in Madagascar. *Proceedings of the National Academy of Sciences*, 109(14), 5358-5363. doi: 10.1073/pnas.1112487109.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M. T., Weihmann, A., Nickel, B., ... & Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences*, 110(39), 15758-15763. doi: 10.1073/pnas.1314445110.

Emerson, B. C. (2002). Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular ecology*, 11(6), 951-966. doi: 10.1046/j.1365-294X.2002.01507.x.

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), 3047-3048. doi: 10.1093/bioinformatics/btw354.

Fraser, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular ecology*, 10(12), 2741-2752. doi: 10.1046/j.0962-1083.2001.01411.x.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in ecology & evolution*, 27(9), 489-496. doi: 10.1016/j.tree.2012.05.012.

Gansauge, M. T., & Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nature protocols*, 8(4), 737-748. doi: 10.1038/nprot.2013.038.

Gansauge, M. T., Gerber, T., Glocke, I., Korlević, P., Lippik, L., Nagel, S., ... & Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic acids research*, 45(10), e79-e79. doi: 10.1093/nar/gkx033.

Ganzhorn, J., Lowry, P., Schatz, G., & Sommer, S. (2008). The biodiversity of Madagascar: One of the world's hottest hotspots on its way out. *Oryx*, 35(4), 346-348. doi:10.1046/j.1365-3008.2001.00201.x.

Gilbert, M. T. P., Haselkorn, T., Bunce, M., Sanchez, J. J., Lucas, S. B., Jewell, L. D., ... & Worobey, M. (2007). The isolation of nucleic acids from fixed, paraffin-embedded tissues–which methods are useful when?. *PloS one*, 2(6), e537. doi: 10.1371/journal.pone.0000537.

Goodman, SM & Benstead, JP (eds). (2003). The natural history of Madagascar. The University of Chicago Press, Chicago & London. *Revue d'Écologie (La Terre et La Vie)*, 60(3), 292-292.

Gottfried, M. D., Randriamiarimanana, L. L., Rabarison, J. A., & Krause, D. W. (1998). Late Cretaceous fish from Madagascar: implications for Gondwanan biogeography. *Journal of African Earth Sciences*, 27(1), 91-91.

Groombridge, B. & Jenkins, B. (1998). *Freshwater Biodiversity: A Preliminary Global Assessment*. Cambridge (United Kingdom): World Conservation Monitoring Centre (WCMC), World Conservation Press. WCMC Biodiversity Series no. 8.

Hoffman, E. A., Frey, B. L., Smith, L. M., & Auble, D. T. (2015). Formaldehyde crosslinking: a tool for the study of chromatin complexes. *Journal of Biological Chemistry*, 290(44), 26404-26411. doi: 10.1074/jbc.R115.651679.

Hung, C. M., Shaner, P. J. L., Zink, R. M., Liu, W. C., Chu, T. C., Huang, W. S., & Li, S. H. (2014). Drastic population fluctuations explain the rapid extinction of the passenger pigeon. *Proceedings of the National Academy of Sciences*, 111(29), 10636-10641. doi: 10.1073/pnas.1401526111.

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., ... & Tappu, R. (2016). MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS computational biology*, 12(6), e1004957. doi: 10.1371/journal.pcbi.1004957.

IUCN. (2020). The IUCN Red List of Threatened Species. Version 2020-1. https://www.iucnredlist.org.

Jensen, E. L., Díez-del-Molino, D., Gilbert, M. T. P., Bertola, L. D., Borges, F., Cubric-Curik, V., ... & Vernesi, C. (2022). Ancient and historical DNA in conservation policy. *Trends in Ecology & Evolution*. doi: 10.1016/j.tree.2021.12.010.

Jones, J. P., Rasamy, J. R., Harvey, A., Toon, A., Oidtmann, B., Randrianarison, M. H., ... & Ravoahangimalala, O. R. (2009). The perfect invader: a parthenogenic crayfish poses a new threat to Madagascar's freshwater biodiversity. *Biological Invasions*, *11*, 1475-1482. doi: 10.1007/s10530-008-9334-y.

Katoh, K., & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Briefings in bioinformatics*, 9(4), 286-298. doi: 10.1093/bib/bbn013.

Kiener, A. (1963). Poissons, pêche et pisciculture à Madagascar. Nogent-sur-Marne (France) : *Centre Technique Forestier Tropical*. (24): 1-244.

Kiener, A., & Mauge, M. (1966). Contributions à l'étude systématique et écologique des poissons Cichlidae endémiques de Madagascar. *Mémoires du Muséum National d'Histoire Naturelle, Séries A, Zoologie*.40 : 51-99.

Kocher, T. D., Conroy, J. A., McKaye, K. R., Stauffer, J. R., & Lockwood, S. F. (1995). Evolution of NADH dehydrogenase subunit 2 in East African cichlid fish. *Molecular phylogenetics and evolution*, 4(4), 420-432. doi: 10.1006/mpev.1995.1039.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences*, 86(16), 6196-6200. doi: 10.1073/pnas.86.16.6196.

Krause, D. W., Hoffmann, S., Hu, Y., Wible, J. R., Rougier, G. W., Kirk, E. C., ... & Rahantarisoa, L. J. (2020). Skeleton of a Cretaceous mammal from Madagascar reflects long-term insularity. *Nature*, 581(7809), 421-427. doi: 10.1038/s41586-020-2234-8.

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26(5), 589-595. doi: 10.1093/bioinformatics/btp698.

Loiselle, P. V. (2011). How many Marakely update. *Buntbarsche Bulletin −The journal of the American Cichlid Association*. (267):18-24.

Loiselle, P. V., & Stiassny, L. J. (1993). How many Marakely? *Buntbarsche Bulletin −The journal of the American Cichlid Association.* (157): 2-8.

Longrich, N. R., Tokaryk, T., & Field, D. J. (2011). Mass extinction of birds at the Cretaceous–Paleogene (K–Pg) boundary. *Proceedings of the National Academy of Sciences*, 108(37), 15253-15257. doi: 10.1073/pnas.1110395108.

Losos, J. B., & Ricklefs, R. E. (2009). Adaptation and diversification on islands. *Nature*, *457*(7231), 830-836. doi: 10.1038/nature07893.

Maréchal, C. & Poll, M. (1991). *Paratilapia*. p. 358. In J. Daget, J.-P. Gosse, G.G. Teugels and D.F.E. Thys van den Audenaerde (eds.) Check-list of the freshwater fishes of Africa (CLOFFA). ISNB, Brussels; MRAC, Tervuren; and ORSTOM, Paris. Vol. 4.

Masters, J. C., De Wit, M. J., & Asher, R. J. (2006). Reconciling the origins of Africa, India and Madagascar with vertebrate dispersal scenarios. *Folia Primatologica*, *77*(6), 399-418. doi: 10.1159/000095388.

Matschiner, M. (2019). Gondwanan vicariance or trans-Atlantic dispersal of cichlid fishes: a review of the molecular evidence. *Hydrobiologia*, 832, 9-37. doi: 10.1007/s10750-018-3686-9.

Matschiner, M., Böhne, A., Ronco, F., & Salzburger, W. (2020). The genomic timeline of cichlid fish diversification across continents. *Nature communications*, 11(1), 5895. doi: 10.1038/s41467-020-17827-9.

Matschiner, M., Musilová, Z., Barth, J. M., Starostová, Z., Salzburger, W., Steel, M., & Bouckaert, R. (2017). Bayesian phylogenetic estimation of clade ages supports trans-Atlantic dispersal of cichlid fishes. *Systematic biology*, 66(1), 3-22. doi: 10.1093/sysbio/syw076.

McGuire, J. A., Cotoras, D. D., O'Connell, B., Lawalata, S. Z., Wang-Claypool, C. Y., Stubbs, A., ... & Iskandar, D. T. (2018). Squeezing water from a stone: high-throughput sequencing from a 145-year old holotype resolves (barely) a cryptic species problem in flying lizards. *PeerJ, 6*, e4470. doi: 10.7717/peerj.4470.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*, 37(5), 1530-1534. doi: 10.1093/molbev/msaa015.

Moritz, C. (1999). Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas*, 130(3), 217-228. doi: 10.1111/j.1601-5223.1999.00217.x.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature,* 403(6772), 853-858. doi: 10.1038/35002501.

Palumbi, S. R. (1996). Nucleic acids II: the polymerase chain reaction. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), *Molecular Systematics*, 2nd edn. Sinauer, Sunderland, MA, pp. 205–247.

Pellegrin, J. (1904). Contribution à l'étude anatomique, biologique et taxinomique des poissons de la famille des Cichlidés. *Mémoires de la Société zoologique de France*, 16: 41-399.

Pellegrin, J. (1929). Cichlidés de Madagascar recueillis par M. Georges Petit; description d'une espèce nouvelle. *Bulletin de la Société Zoologique de France*, *54*, 252-255.

Rakotomamonjy, L., & Oliarinony, R. (2018). Phylogenetic, morphological and systematic studies of three genera of Cichlidae (Teleosts fish, Perciformes) endemic of Madagascar. *Biodiversity Int J*, 2(6), 538-543. doi: 10.15406/bij.2018.02.00110.

Raxworthy, C. J., & Smith, B. T. (2021). Mining museums for historical DNA: advances and challenges in museomics. *Trends in Ecology & Evolution,* 36(11), 1049-1060. doi: 10.1016/j.tree.2021.07.009.

Razafitsiferana, T. (2015). Etude comparative entre les proprietes physico-chimiques des eaux souterraines, des eaux de surface et des eaux marines de la region diana, dans le district de Nosy Be. Pg. 5. PhD Thesis, University of Antananarivo, Department of Mineral Chemistry and Physical Chemistry.

Reddy, S., Driskell, A., Rabosky, D. L., Hackett, S. J., & Schulenberg, T. S. (2012). Diversification and the adaptive radiation of the vangas of Madagascar. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1735), 2062-2071. doi: 10.1098/rspb.2011.2380.

Reinhardt, A. L., Kasper, T., Lochner, M., Bliedtner, M., Krahn, K. J., Haberzettl, T., ... & Montade, V. (2022). Rain forest fragmentation and environmental dynamics on Nosy Be Island (NW Madagascar) at 1300 cal BP is attributable to intensified human impact. *Frontiers in Ecology and Evolution*, 9*,* 783770. doi: 10.3389/fevo.2021.783770.

Reinthal, P.N., & Stiassny, M.L.J. (1997). Revision of the Madagascan genus Ptychochromoides (Teleostei: Cichlidae), with description of a new species. *Ichthyological Exploration of Freshwaters.* 7(3-4):353-368.

Rham, P. D., & Nourissat, J. C. (2004). The endemic cichlids of Madagascar (English translation). French Cichlid Association, St. Julia, France. 191 pp.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature biotechnology*, 29(1), 24-26. doi: 10.1038/nbt.1754.

Rohland, N., Siedel, H., & Hofreiter, M. (2004). Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *Biotechniques*, 36(5), 814-821. doi: 10.2144/04365ST05.

Ruane, S., & Austin, C. C. (2017). Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Molecular ecology resources*, 17(5), 1003-1008. doi: 10.1111/1755-0998.12655.

Salzburger, W. (2018). Understanding explosive diversification through cichlid fish genomics. *Nature Reviews Genetics*, 19(11), 705-717. doi: 10.1038/s41576-018-0043-9.

Samonds, K. E., Godfrey, L. R., Ali, J. R., Goodman, S. M., Vences, M., Sutherland, M. R., ... & Krause, D. W. (2012). Spatial and temporal arrival patterns of Madagascar's vertebrate fauna explained by distance, ocean currents, and ancestor type. *Proceedings of the National Academy of Sciences*,109(14), 5352-5357. doi: 10.1073/pnas.1113993109.

Samonds, K. E., Godfrey, L. R., Ali, J. R., Goodman, S. M., Vences, M., Sutherland, M. R., ... & Krause, D. W. (2013). Imperfect isolation: factors and filters shaping Madagascar's extant vertebrate fauna. *PLoS One*, 8(4), e62086. doi: 10.1371/journal.pone.0062086.

Sauvage, H. (1891). Histoire physique, naturelle et politique de Madagascar. *Histoire Naturelle de Poissons*. Imprimerie Nationale, Paris. 16, 1-543.

Šimková, A., Řehulková, E., Rasoloariniaina, J. R., Jorissen, M. W., Scholz, T., Faltýnková, A., ... & Vanhove, M. P. (2019). Transmission of parasites from introduced tilapias: a new threat to endemic Malagasy ichthyofauna. *Biological Invasions*, 21, 803-819. doi: 10.1007/s10530-018-1859-0.

Sparks, J. S. & Reinthal, P. N. (1999). *Paretroplus maromandia*, a new cichlid fish from the northwest of Madagascar. Occasional Papers of the Museum of Zoology University of Michigan No. 727: 1-18.

Sparks, J. S. (2002). *Paretroplus dambabe*, a new cichlid fish (Teleostei: Cichlidae) from northwestern Madagascar, with a discussion on the status of *P. petiti*. *Proceedings of the Biological Society of Washington*, *115*(3), 546-563.

Sparks, J. S. (2004). Molecular phylogeny and biogeography of the Malagasy and South Asian cichlids (Teleostei: Perciformes: Cichlidae). *Molecular Phylogenetics and Evolution*, 30(3), 599-614. doi: 10.1016/S1055-7903(03)00225-2.

Sparks, J. S. (2008). Phylogeny of the cichlid subfamily Etroplinae and taxonomic revision of the Malagasy cichlid genus *Paretroplus* (Teleostei: Cichlidae). *Bulletin of the American Museum of Natural History*, 2008(314), 1-151. doi: 10.1206/314.1.

Sparks, J. S. (2001). Phylogeny and biogeography of the Malagasy and South Asian cichlid fishes (Teleostei: Perciformes: Cichlidae), including a survey of the freshwater fishes of Madagascar. Unpublished PhD Thesis, University of Michigan, Ann Arbor.

Sparks, J. S. (2023). Pg. 1283 in: Goodman, S.M. (ed.): The New Natural History of Madagascar. – Princeton University Press (Princeton / NJ).

Sparks, J. S., & Stiassny, M. L. J. (2003). Introduction to the freshwater fishes. *The Natural History of Madagascar. SM Goodman and JP Benstead (eds.)*, 849-863. University of Chicago Press, Chicago.

Stiassny, M. L. (1996). An overview of freshwater biodiversity: with some lessons from African fishes. *Fisheries, 21*(9), 7-13. doi: 10.1577/1548-8446(1996)021<0007:AOOFB>2.0.CO;2.

Stiassny, M. L. J., & Raminosoa, N. (1994). The fishes of the inland waters of Madagascar. *Annales-Musee Royal de l'Afrique Centrale. Sciences Zoologiques (Belgium) ISSN 0770-4666*.

Stiassny, M. L., & Gerstner, C. L. (1992). The parental care behaviour of *Paratilapia polleni* (Perciformes, Labroidei), a phylogenetically primitive cichlid from Madagascar, with a discussion of the evolution of maternal care in the family Cichlidae. *Environmental Biology of Fishes*, 34, 219-233. doi: 10.1007/BF00004768.

Stiassny, M. L., Chakrabarty, P., & Loiselle, P. V. (2001). Relationships of the Madagascan cichlid genus *Paretroplus*, with description of a new species from the Betsiboka River drainage of northwestern Madagascar. *Ichthyological Exploration of Freshwaters*, *12*(1), 29-40.

Straube, N., Lyra, M. L., Paijmans, J. L., Preick, M., Basler, N., Penner, J., ... & Hofreiter, M. (2021). Successful application of ancient DNA extraction and library construction protocols to museum wet collection specimens. *Molecular Ecology Resources*, 21(7), 2299-2315. doi: 10.1111/1755-0998.13433.

Tang, E. P. (2006). Path to effective recovering of DNA from formalinfixed biological samples in natural history collections. In Workshop Summary edn. National Academies Press: Washington, DC.

Thys Van Den Audenaerde, D. F. E. (1968). A preliminary contribution to a systematic revision of the genus *Pelmatochromis* Hubrecht sensu lato (Pisces, Cichlidae). *Revue de zoologie et de botanique africaines*, 77, 349-391.

van der Valk, T., Díez-del-Molino, D., Marques-Bonet, T., Guschanski, K., & Dalén, L. (2019). Historical genomes reveal the genomic consequences of recent population decline in eastern gorillas. *Current Biology*, 29(1), 165-170. doi: 10.1016/j.cub.2018.11.055.

van der Valk, T., Pečnerová, P., Díez-del-Molino, D., Bergström, A., Oppenheimer, J., Hartmann, S., ... & Dalén, L. (2021). Million-year-old DNA sheds light on the genomic history of mammoths. *Nature,* 591(7849), 265-269. doi: 10.1038/s41586-021-03224-9.

Vences M, Freyhof J, Sonnenberg R, Kosuch J, Veith M (2001) Reconciling fossils and molecules: Cenozoic divergence of cichlid fishes and the biogeography of Madagascar. *Journal of Biogeography,* 8: 1091–1099. doi: 10.1046/j.1365-2699.2001.00624.x.

Vences, M., Brown, J. L., Lathrop, A., Rosa, G. M., Cameron, A., Crottini, A., ... & Murphy, R. W. (2017). Tracing a toad invasion: lack of mitochondrial DNA variation, haplotype origins, and potential distribution of introduced *Duttaphrynus melanostictus* in Madagascar. *Amphibia-Reptilia*, *38*(2), 197-207. doi:10.1163/15685381-00003104.

Vences, M., Stützer, D., Rasoamampionona Raminosoa, N., & Ziegler, T. (2022). Towards a DNA barcode library for Madagascar's threatened ichthyofauna. *Plos one,* 17(8), e0271400. doi: 10.1371/journal.pone.0271400.

Vences, M., Wollenberg, K. C., Vieites, D. R., & Lees, D. C. (2009). Madagascar as a model region of species diversification. *Trends in ecology & evolution*, *24*(8), 456-465. https://doi.org/10.1016/j.tree.2009.03.011.

Wit, M. J. D. (2003). Madagascar: heads it'sa continent, tails it's an island. *Annual Review of Earth and Planetary Sciences*, *31*(1), 213-248. doi: 10.1146/annurev.earth.31.100901.141337.

Yoder, A. D., & Nowak, M. D. (2006). Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annu. Rev. Ecol. Evol. Syst.*, *37*, 405-431. doi: 10.1146/annurev.ecolsys.37.091305.110239.

Yoder, A. D., Burns, M. M., Zehr, S., Delefosse, T., Veron, G., Goodman, S. M., & Flynn, J. J. (2003). Single origin of Malagasy Carnivora from an African ancestor. *Nature*, *421*(6924), 734-737. doi: 10.1038/nature01303.