School of Doctoral Studies in Biological Sciences
University of South Bohemia in České Budějovice
Faculty of Science

Population genetics of the fish tapeworm *Wenyonia virilis* (Caryophyllidea: Caryophyllaeidae) and its fish host *Synodontis schall* (Siluriformes: Mochokidae)

Ph.D. thesis

RNDr. Dagmar Jirsová

Supervisor: Mgr. Miloslav Jirků, Ph.D.
Institute of Parasitology, Biology Centre CAS, České Budějovice

Consultant: doc. RNDr. Jan Štefka, Ph.D.
Institute of Parasitology, Biology Centre CAS, České Budějovice &
Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

České Budějovice 2017

This thesis should be cited as:

Jirsová, D, 2017: Population genetics of the fish tapeworm *Wenyonia virilis* (Caryophyllidea: Caryophyllaeidae) and its fish host *Synodontis schall* (Siluriformes: Mochokidae). Ph.D. Thesis. University of South Bohemia, Faculty of Science, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic, 94 pp.

Annotation:

The presented thesis consists of three papers/manuscripts (one published, one under review, one manuscript) on population genetic aspects of a host-parasite model, caryophyllidean tapeworm *Wenyonia virilis* and mochokid catfish *Synodontis schall*, in recently separated drainage basins, Lake Turkana and the Nile River. Three main topics are addressed herein: (i) intra- and inter-population genetic variability in and among hosts and parasites, (ii) comprehensive assessment of host model taxonomic status using multiple approaches, (iii) comparison of parasite intraspecific phenotypic with population genetic pattern. Two different genetic markers were applied to address these topics – mtDNA (*coxI*) and whole genome scanning method (AFLP).

Declaration:

I hereby declare that this thesis is based on my own work and all other sources of information have been acknowledged. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my PhD thesis, in full to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

Date... Signature...

This thesis originated from a partnership of Faculty of Science, University of South Bohemia, and Institute of Parasitology, Biology Centre of the ASCR, supporting doctoral studies in the Parasitology study program.

Přírodovědecká
fakulta
Faculty
of Science

Financial:

Individual parts of this thesis were financially supported by one, or combination of several, of the following grants: Czech Science Foundation (P505/12/G112), Grant Agency of Academy of Cciences of the Czech Republic (KJB600960813), the Institute of Parasitology (RVO: 60077344) and the Grant Agency of University of South Bohemia (037/2011/P, 04-135/2010/P).

Acknowledgment:

I would like to express my gratitude to all the people who participated in this research and made it possible to finish this thesis.

Special thanks goes to my both supervisors, Dr. Miloslav Jirků and Dr. Jan Štefka, for their kind and patient attitude, handling my impatience and me as a human being. I am very thankful for the given opportunity to be part of fieldwork search team in Kenya and for all the amazing experience and unforgettable memories; it was truly life changing experience.

I would like to further thank to Dr. Lukas Rüber and Dr. Kai Winkelmann for their hospitality and help during my stay in National History Museum in London.

I am very thankful to University of South Bohemia and the Biology Centre (Institute of Parasitology) for providing excellent theoretical knowledge and practical training.

I am also very thankful for the support of my family and friends, especially, Mirka Soldánová, Corinna Benz and Martin Kašný for the listening of my endless whining.

List of papers and author’s contribution

The thesis is based on the following papers and manuscripts (listed chronologically):

Paper 1

Jirsová D, Štefka J, Jirků M. 2017. Discordant population histories of host and its parasite : A role for ecological permeability of extreme environment ? *PLoS One*:1–25. *IF = 2.806*

Dagmar Jirsová particpated in field sampling, conducted sequencing, phylogenetic and population genetic analyses, and drafted the manuscript.

Paper 2

Jirsová D, Štefka J, Blažek R, Malala JO, Lotuliakou DE, Mahmoud ZN, Jirků M. Hidden diversity and overlooked morphological variability of Synodontis schall: from taxonomic deflation to recognition of cryptic species? (submitted and under review *Molecular Phylogenenetics and Evolution*).

Dagmar Jirsová particpated in field sampling, conducted sequencing, phylogenetic analyses, and drafted the manuscript.

Paper 3

Jirků M, Štefka J, **Jirsová D**. Morphological variability mirrors population genetic structure in fish tapeworm: a role for salinity in an isolated limnic system? (manuscript in advanced preparation).

Dagmar Jirsová was responsible for phylogentic and AMOVA analyses and participated in manuscript writing.

Co-authors agreement

The senior and corresponding authors of the manuscripts included in this thesis, hereby confirm that DJ contributed significantly to these publication, according to the statement above:

Miloslav Jirků   Jan Štefka

Table of Contents

1. General introduction 1

2. Objectives of the study 4

3. Geographic and organismal models 5

3.1. Geographic model – general considerations 5

3.1.1. The Nile 5

3.1.2. Lake Turkana 5

3.2. Model organisms 8

3.2.1. Host model - *Synodontis schall* (Bloch et Schneider, 1801) 8

3.2.2. Parasite model - *Wenyonia virilis* Woodland,1923 9

4. Summary of results and discussion with respect to the partial aims 11

4.1. Opposing population histories: a fast parasite in a stationary host 11

4.2. No cospeciation: a diversifying parasite in a stationary host 12

4.3. Mitochondrial markers: more informative than whole genome markers (AFLP) 13

4.4. Intricate host identity: cryptic sibling and unexpected phenotypic variability 13

4.5. Matching phenotypic and population genetic patterns: a role for contrasting lacustrine environments? 14

5. Conclusions and future perspectives 16

5.1. Conclusions 16

5.2. Future perspectives 16

6. Thesis References 18

7. List of papers 26

Paper 1 26

Paper 2 53

Paper 3 73

Appendix I - Amplified Fragment Length Polymorphism (AFLP) 88

I.I. AFLP utilization 88

I.II. Principles of AFLP 89

Appendix II – Curriculum vitae 91

1. General introduction

Most of the current studies on tapeworms are focused on morphology, phylogeny and phylogeography. The main topics of published studies are descriptions of newly discovered species and species redescriptions, generally using combination of molecular markers and morphological data, or whole genome and mitochondrial genome descriptions. Despite the ubiquitous nature, diversity and medical significance of cestodes, the population genetic studies on tapeworms are still not as common as the above mentioned types of studies, although the importance of population genetics and host-parasite comparative studies is being mentioned by some authors (Huyse *et al.* 2005; Poulin *et al.* 2011; Møller *et al.* 2013). To date, studies on genetic characterization of tapeworm populations were largely limited to species of medical and/or economic interest (Poulin & Morand 2000; Sharma *et al.* 2013; Addy *et al.* 2017; Sanpool *et al.* 2017). Although a fair number of genes has been identified as suitable targets for studies on phylogeny of parasites (Olson & Tkach 2005; Chambrier *et al.* 2015), far less has been done in search of reliable markers to study their population genetics. Therefore, there is still a large area of parasitology to explore, especially in taxa in which biogeography and population structure were heavily affected by the human activity (Zarlenga *et al.* 2014), as well as those in which biogeographically significant phenomena might be expected (see below).

Genes most widely used for phylogenetic studies are small (SSU) and large (LSU) subunit ribosomal DNA (Olson *et al.* 2001; Olson & Tkach 2005), whereas mitochondrial genes are the most suitable markers for studies on population inter- and intra-relationships (Freeland 2005; Zhao *et al.* 2009; Sanpool *et al.* 2017). Although for the specific states of populations, e.g. ancestral polymorphisms and recently divided taxa, the nuclear genes still provide more specific information about lineage splitting/fusing (Carling & Brumfield 2008; Garrick *et al.* 2014; Good *et al.* 2015), only in the combination with mitochondrial genes the whole historical and population story is complete (Day *et al.* 2013; Pinton *et al.* 2013; Garrick *et al.* 2014). Mitochondrial DNA (mtDNA) provides valuable markers for studies on flatworms (Le *et al.* 2000; Kmentová *et al.* 2016). Earlier studies using partial cytochrome oxidase subunit I (*coxI*) gene (Bowles *et al.* 1995; McManus & Bowles 1996) and NADH dehydrogenase 1 (*nad1*) gene (Sørensen *et al.* 1998; Attwood *et al.* 2002; Troell *et al.* 2006; Lier *et al.* 2008) helped to understand the population genetic relationships among species, and their popularity persisted till now (Jabbar *et al.* 2011; Yanagida *et al.* 2012; **Paper 1**). For exemple Zhao *et al.* (2009) in their study on *Schistosoma japonicum* detected differences between populations in China based on three mitochondrial DNA regions, cytochrome c oxidase subunit 3 (*cox3*), NADH dehydrogenase subunits 4 and 5 (*nad4* and *nad5*). In conclusion, analyses of mt *cox3*, *nad4* and *nad5* sequences revealed genetic variability among *S. japonicum* isolates from different endemic regions in China. The combined sequences of *cox3*, *nad4* and *nad5* were useful for re-construction of phylogenetic relationships among *Schistosoma* spp., as well as among *S. japonicum* isolates from different endemic provinces in China. These results demonstrated that mt markers are useful for studying the inter- and intra-population genetic variations in *S. japonicum*, and have implications for studies of population biology, molecular epidemiology, and genetic structure of *S. japonicum* (Zhao *et al.* 2009).

However, despite the undeniable quality of provided information from mitochondrial genes, the popularity of single gene studies is retreating and with the appearance of new multiple genes screening technologies (e.g. RAD-Seq, Illumina platform), more complex studies are becoming much more frequent. Multiple genes screening methods are getting more popular for diagnostics in clinical parasitology, gradually replacing microscopy-based approaches

(Verweij & Rune Stensvold 2014; van Lieshout & Roestenberg 2015). Also, in phylogeny and population genetics, these approaches are increasing and deepening the knowledge about historical processes, lineage splitting/fusing etc. (e. g. Brabec *et al.* 2015; Doležalová *et al.* 2015; Gupta *et al.* 2016).

Molecular markers are great tools for studying phylogenetic relationships between species and evolutionary processes such as speciation, providing great opportunities for comparative studies and tests of host-parasite specificity and/or host-parasite evolution in general. Nowadays, the molecular markers are mainly used in taxonomy and phylogeny for genetic descriptions of taxa and assessment of evolutionary relationships between them In order to reveal and explain evolutionary processes in populations.

To avoid errors and misinterpretations, it is crucial to determine taxonomically both parasites and hosts involved in a given study. The species/taxon determination is sometimes difficult due to the limited availability of morphological characters and/or samples. In case of parasites, adults (of particular sex in many cases) which are often necessary for morphology-based taxonomic assignment, are frequently unavailable, and field samples sometimes consist only of eggs, larval stages and/or fragmentary parasite tissues. Therefore, genetic identification of the samples is by far the most applied molecular approach in parasitology, not only for species determination but also for the application of the appropriate treatment and/or preventive measures (McManus & Bowles 1996; Tham *et al.* 1999; Gupta *et al.* 2016).

Parasites and hosts interact along an evolutionary history on many different levels (Brooks & McLennan 1993; Huyse *et al.* 2005). Therefore, part of the parasite population dynamics is highly associated with that of its host (Nieberding & Olivieri 2007; Toit *et al.* 2013). Although it is of enormous importance to properly identify hosts of studied parasites (**Paper 2**), molecular approach to this matter is still surprisingly scarce among parasitologists. As a result, morphological identification of hosts by parasitologists remains a common practice, leaving behind unknown amount of misleading literature information (**Paper 3**).

Reconstruction of the historical biogeography and identification of major genetic subdivisions within species are major objectives of phylogeography in general (Avise 2000). Furthermore, comparative phylogeography can identify historically and evolutionarily significant geographical regions (Bermingham & Moritz 1998; Riddle *et al.* 2000; Arbogast & Kenagy 2001; Nieberding & Morand 2006; Neiswenter & Riddle 2011). Comparative phylogeographic studies involving both parasite and its host provide better understanding of the ecology and evolution of single species by recognizing the impact of host population genetic structure on that of the parasite (Bouzid *et al.* 2008; Brouat *et al.* 2011; Kvičerová & Hypša 2013; Levin & Parker 2013).To find traditional phylogeographical studies on parasites is a difficult task. The few exceptions are represented especially by medically and economically important endoparasites, such as *Fasciolides magna* and *Schistosoma japonicum* (Králová-Hromadová *et al.* 2011, Zhao *et al.* 2009) and some ectoparasites, e.g. *Gyrodactylus gongae* (Huyse *et al.* 2017), parasitic insects, especially lice and flies (e.g. Štefka *et al.* 2011; Levin & Parker 2013). Phylogenetic/phylogeographic comparative methods provide tools for studying trait evolution history and trait covariance while accounting for non-independence of data collected across species. Hence, this approach is ideal to assess the depth of host-parasite relationship. Parasites are closely tied up to their hosts and therefore to their evolutionary past and population dynamics. Repeated speciation of host taxon leads to congruent phylogenetic relationships among its parasites (cophylogeny) (Page 2003; Clayton *et al.* 2004) and it might be expected that parasites and hosts would share similar patterns. However, parasites often have shorter generation times and, in some cases, faster mutation rates than their hosts, which can lead to greater population differentiation in the parasite

relative to the host (Page *et al.* 1998). Morevover, it might be also expected, that cophylogeny should have been much stronger in case of endoparasites than in ectoparasites due to their more intimate physiological bond with the host. However, the vast majority of existing studies on host-parasite comparative phylogeography is focused on insect ectoparasites (e.g. Štefka *et al.* 2011; Toit *et al.* 2013; Cruaud & Rasplus 2016). Although, ectoparasites might not seem ideal model organisms, their relashionships with hosts are much more complex than might be expected due to various evolutionary events, e.g. host switching and adaptative radiation (Bandilla *et al.* 2008; Hahn *et al.* 2015; Vanhove *et al.* 2015). In addition to that, given the presumably faster evolution of parasites compared to their hosts, phylogeographical and population patterns should be much more pronounced in parasites even if their generation time is similar to that of the host (**Paper 1**). Interestingly, although commonly refferred to in the literature, the evidence of (presumably) fast evolving parasites vs. slow evolving hosts is being limited exclusively to ectoparasites and Apicomplexa (Hafner *et al.* 1994; Cohuet *et al.* 2010; Štefka *et al.* 2011; Khan *et al.* 2013). As a result, no evidence of this phenomenon was available for metazoan endoparasites until recently (**Paper 1**).

Phylogenetic studies that are focused on cospeciation between parasites and hosts have improved our knowledge and understanding of parasite evolution (Page 2003; Engelstädter & Bonhoefferkesh 2009). Likewise, much can be gained by testing (dis)similarities of phylogeographical structure of single host-parasite counterparts, a topic which became the most studied phenomenon among host-parasite interactions studies (Page 2003; Clayton *et al.* 2004; Criscione & Blouin 2007; Attwood 2010; **Paper 1**). However, it is obvious that additional comparative host-parasite phylogeographical studies are needed to allow full understanding of the microevolutionary processes in populations that result in interesting macroevolutionary patterns such as cospeciation e.g. sharing of similar phylogeographical patterns between hosts and parasites (Criscione & Blouin 2005, 2007; Criscione *et al.* 2005; Morand *et al.* 2006).

Interactions among species are fundamental drivers of macroevolutionary patterns of diversification. One of the most relevant examples of interspecific interactions is the co-evolutionary pattern apparent between hosts and parasites. In host–parasite interactions, co-speciation, i.e. the joint speciation between host and its parasite, can arise if the two organisms share a common evolutionary history, so that the parasite follows the speciation events of its host (Brooks 1979; Page 2003; Huyse *et al.* 2005, 2017; Hahn *et al.* 2015). Common history is favored by traits that strengthen the host–parasite relationship, such as a high level of host specificity and vertical transmission of parasites among hosts (Page 2003; Nieberding & Olivieri 2007; Brunner *et al.* 2017).

Yet, co-speciation is an ongoing process; thus co-differentiation between lineages of a host and its parasite species should be already detectable at the intraspecific level. Since speciation often takes place between allopatric lineages (Schluter 2001), co-differentiating lineages of a host species and its parasite are expected to be allopatric and the corresponding host and parasite lineages are expected to cover similar geographic areas (Nieberding *et al.* 2004). In other words, at the intraspecific level, it can be supposed that a common history is best observed between a host and its parasite whose genetic lineages cover similar allopatric areas (Nieberding *et al.* 2004; Nieberding & Olivieri 2007).

2. Objectives of the study

During the Holocene near past, Lake Turkana, a young inland sea in geological terms, was repeatedly part of the Nile drainage. As a result, Turkana's limnic biota is essentially of Nilotic origin. Within the scope of ichthyoparasitological research of the Lake Turkana conducted since 2006 by the supervisor of the present study, numerous parasites, both new and known to science, were recorded. There seemed to be growing evidence that Turkana is not just a piece of the Nile temporarily cut off its maternal basin. Host-parasite communities of Turkana and the Nile proved different and similar at the same time. It became obvious, that the two allied basis represent interesting biogeographic model. As the surprisingly diverse fish parasitofauna of the lake has been revealed, a tempting question worth addressing arose: How different from each other are the conspecific populations of parasites, and possibly their hosts, in the two sibling basins? Were the five millennia of their separation, and/or the specific Turkana environment, enough to induce any kind of discernible signal of their differentiation?

Partial objectives of the research were specified as follows:

1) Evaluation of intra- and inter-population genetic variability of parasite and its host within and between the Nile and Turkana basins, and assessment of factors potentially responsible for the observed patterns, e.g. selection pressures, bottlenecks, salinity

2) Detection of population genetic patterns in parasite and/or its host indicating on-going evolutionary processes acting on millennial time scale and potentially leading to speciation

3) Comparison of utility of mitochondrial versus whole genome markers

4) Reliable host taxon determination using combination of molecular and morphological data

5) Assessment of presence/absence of correlation between morphological and genetic variability of parasite

3. Geographic and organismal models

3.1. Geographic model – general considerations

The geographic model consisting of two (bio)geographic sub-systems was proposed as follows: the Nile basin representing a non-isolated and fully freshwater basin *vs.* Lake Turkana representing an isolated endorheic basin with marked salinity gradient.

Additional criteria were specified for the two geographic sub-systems as follows: i) absence of extant aquatic link between the two basins; ii) model organisms common in both subsystems; iii) presence of contrasting freshwater and highly saline (brackish) environments, i.e. marked salinity gradient, present in the isolated sub-system; iv) isolated sub-system of sufficient size to omit size-dependent bias and stochastic processes; v) recent disconnection of the two geographic sub-systems.

In the Late Pleistocene 35-15 thousand years ago (kya), Lake Turkana almost completely dried out. During the Holocene peak of the African humid period 11.5-10.5 kya and for a brief periods 9.0, 6.6 and 5.2 kya, the lake spilled over the Lotikipi Plain to the north-west, overflowed into White Nile (Bloszies *et al.* 2015) and became a part of Nile-Chad-Niger limnic system (Bloszies *et al.* 2015, see **Paper 1, Fig. 1**). Re-establishment of limnic biota most probably dates to this early to mid Holocene period (compare Johnson & Malala 2009 and Bloszies *et al.* 2015). During the drier, lower level periods, similar to the present, the Nile-Turkana link disappeared and the Turkana basin became endorheic (Harvey & Grove 1982; Nyamweru 1989; Ricketts & Anderson 1998). The near desiccation of Lake Turkana during the Late Pleistocene, followed by temporary connection with the Nile is the reason for the recent origin of its limnic biota, the faunistic similarity to the Nile and relatively low levels of endemism (~60 fish spp. including 10 endemics, Hopson 1982). Currently, the arid Lotikipi plain separating the Nile and Turkana basins lacks permanent aquatic environments.

3.1.1. The Nile

The Nile River is the longest river in the world (for comprehensive review, see Dumont 2009). There are no natural barriers restricting fish dispersal along a remarkably long section of the Nile River between the Murchinson Falls in Uganda and more than 3.000 km (air distance) distant delta of the Nile in Egypt. As a result of permeable nature of the river system, much of the fish diversity is well represented in all its parts, so that differences in composition of fish communities at different (even the most distant) sites are far more influenced by local habitat conditions (e.g. presence of large swamps) than by distance, biogeographic patterns or climatic differences. The fish fauna of the Nile River belongs to the Nilo-Sudan ichthyological province, comprising 128 species in 27 families (excluding connected lakes: Albert, Edward/George, Tana and Victoria/Kyoga); of the fish families present, the most diverse are Cyprinidae (25 spp.), Mochokidae and Mormyridae (both 15 spp.) and Cichlidae (10 spp.) representing slightly over 50% of the overall fish diversity (Witte *et al.* 2009).

The Nile sampling localities: Atbarah River, Khashm al Girba, 14.9194 N, 35.9019 E; Blue Nile, Sennar, 13.5436 N, 33.6367 E; White Nile, Kostí, 13.1722 N, 32.6722 E

3.1.2. Lake Turkana

Lake Turkana is the largest desert lake on Earth measuring ~240 by 13-50 km, with surface area 6,750 km^2, volume 203 km^3 and average/maximum depth 31/114 m. Due to alkaline Quaternary and Tertiary lavas of the rift floor, Turkana is the most alkaline lake of the eastern

branch of EARS (Fuchs 1939). It constitutes a centre of endorheic drainage, the Turkana basin (130,000 km^2), with mean temperatures 31-33°C and annual rainfall < 200 mm. Approximately 98% of Turkana's water comes from Ethiopian highlands via the Omo River. There is no outlet and with reduced inflows and high evaporation the chloro-carbonate alkaline water is subject to seasonal water level fluctuations (1-3 m annually, > 20 m during 20. century). The lake is becoming increasingly saline along the Omo River delta (0.07) in the north, to highly saline (brackish) towards the south (Kalokol 1.78, El-Molo Bay 1.89). The presented salinity values are original data obtained just after the peak of rainy season during September 2-26th 2008, i.e. during the season when lowest possible salinity values are to be expected due to increased inflow from Ethiopian highlands. Based on the salinity values present, the major part of Lake Turkana outside the Omo delta area is classified as brackish. Hopson (1982) lists some 60 fish species from Turkana - of these, 50 species are shared with the Nile, remaining 10 species are endemics.

The highly saline majority of the lake harbours depleted lacustrine communities mostly devoid of aquatic and littoral vegetation (Fig. 1). Fish community of the open water areas of Turkana differ markedly from those of other great African lakes, being dominated by generalist species such as mochokid catfish *Synodontis schall* (Bloch et Schneider, 1801), possibly as a consequence of harsh environmental conditions, especially rapid lake level fluctuations (Kolding 1993; Muška *et al.* 2012). The Omo River delta (~30 km wide) including adjacent parts of the lake represents the only spatially significant freshwater habitat in the whole Turkana limnic system. Small pockets of freshwater habitats are present in estuaries of seasonal affluents, the Kerio River and Turkwell River, at the central part of Turkana's west coast. The main habitats presents in the Omo delta are *Typha* sp. dominated wetlands and narrow lakeshore floodplains harbouring relatively rich Nilotic biota (Fig. 1). The fish communities in Omo delta and the Kerio delta, include multiple fish taxa sensitive to increased salinity and thus absent from the open lake. Fish mostly restricted to these freshwater habitats include for example Arapaimidae (*Heterotis niloticus*), Gymnarchidae (*Gymnarchus niloticus*), Mormyridae (*Mormyrus kannume*, *Hyperopisus bebe*), Polypteridae (*Polypterus bichir*, *Polypterus senegalus*) and Protopteridae (*Protopterus aethiopicus*).

Lake Turkana sampling localities: El-Molo Bay (saline - brackish), 2.8322 N, 36.6958 E; Kalokol (saline - brackish), 3.558611 N, 35.915833 E; Omo River delta, Todonyang (freshwater), 4.4517 N, 35.9439 E.

Additional sampling localities of minor importance: Central Island (saline - brackish), 3.495833 N, 36.040278 E; Kerio River delta (freshwater), 2.974813 N, 36.173469 E.

Gibe III-V – farewell to Turkana as we know it?

In this study, an evidence of evolutionary significance of Turkana's isolated limnic system is shown. A few other existing studies revealed presence of endemic fishes and remarkable parasite taxa (Hopson 1982; Moravec *et al.* 2009a; b). Turkanaland and lower Omo pose one of the very last pieces of the continent where indigenous cultures exist in the context of African wilderness, still at places inhabited by the continent's megafauna. A controversial Gibe III Dam is being filled up on upper Omo, another two - Gibe IV and V, are under construction or pending confirmation. The purpose: hydropower, large-scale irrigation schemes, flood prevention. In fact, the floods to be avoided are actually essential for the region's indigenous communities relying on flood-depending agricultural systems, Turkana's limnic system needs the floods too, and the large-scale irrigation of the arid area is likely to consume much of limited water resources before they reach the lake. Inevitably, the Gibe cascade will pose a plethora of adverse effects on the tiny Omo River and the lake it keeps

alive. Desiccation of the lake, together with reduced water supply and absence of floods are likely to affect the uniquely preserved indigenous cultures of Turkanaland. So far self-sustaining socio-economies of culturally diverse tribes relying on floods and fish resources might be ruined in a couple of decades. Despite the very explicit near-past experience of what happens with heavily exploited hydrological systems in arid endorheic basins, the issue remains largely neglected, and virtually unknown by international community. Turkana might well be on its way to its Late-Pleistocene ground zero state and became the Aral of Africa. These circumstances add further significance to this and other studies on Turkana's biota as these might soon become important sources of baseline data on pre-damming conditions, allowing assessment of the impact of the Gibe cascade on Turkana unique and fragile ecosystem.

Figure 1. Principal lacustrine habitats of the Lake Turkana, Kenya. Above: Saline southern part of the lake in the lava desert 7 km south-west from Loiyangalani; South Isalnd and western shore of the lake are visible at the background. Below: Freshwater habitat in the Omo River delta 5 km north of Todonyang; view over the shoreline floodplain towards *Typha* sp. dominated wetland at the background. Compare the densely vegetated freshwater littoral habitat of the delta with the bare, vegetation free brackish locality above. The remarkably contrasting environments are inhabited by a single unstructured population of catfish host *Synodontis schall*, yet multiple markedly divergent subpopulations of its tapeworm parasite *Wenyonia virilis*. Photo: Miloslav Jirků.

3.2. Model organisms

Following requirements were proposed for both host and parasite model: wide distribution, high abundance and reliable availability in both Nile and Turkana basin and at all sampling localities.

For the parasite model, additional criteria were set as follows: narrow host specificity (to avoid influence of different host), complex life cycle (more likely to be influenced by environment including salinity), well resolved taxonomy.

3.2.1. Host model - *Synodontis schall* (Bloch et Schneider, 1801)

The Mochokidae (Actinopterygii, Siluriformes) is an endemic African family of freshwater catfish containing 211 species in nine genera (Eschmeyer *et al.* 2017). With 131 nominal species currently recognized, *Synodontis* Cuvier, 1817 is the most speciose mochokid genus and is widely distributed in all African ichthyological provinces except for Maghreb and South Africa (Poll 1971; Paugy & Roberts 2003; Froese & Pauly 2017). Available knowledge on most *Synodontis* spp. is limited to morphology and distribution (e.g. Poll 1971; Sands 1983; Musschoot & Lalèyè 2008) and first molecular phylogenetic studies were published during the last decade (Day & Wilkinson 2006; Koblmüller *et al.* 2006; Day *et al.* 2013; Pinton *et al.* 2013). Despite the taxonomic complexity and limited knowledge, no published literature combining morphological and molecular analyses apparently exists on *Synodontis* to date. Morphological determination of multiple *Synodontis* spp. is difficult without extensive taxonomical experience and there are only a few reliably assigned *Synodontis* GenBank sequences available for genetic determination. As a result, much information available in form of various literature species accounts, museum catalogues, GeneBank entries and field collection lists, remain blurred by uncertainties about taxonomic identity of some *Synodontis* spp. involved.

As it is "traditionally" recognized, *Synodontis schall* (Fig. 2), the model host species of the present study, is considered the most widely distributed mochokid, supposedly being widespread, common and abundant throughout the whole Nilo-Sudanian ichthyological province. It is reported from practically all West African basins (except the coastal basins of Upper Guinea province), and further to the east throughout the Niger, Chad, Nile and Turkana basins (Poll 1971; Hopson 1982; Hopkins *et al.* 2007; Musschoot & Lalèyè 2008; Pinton *et al.* 2013). Among *Synodontis* spp. of eastern Nilo-Sudan, *S. schall* might be confused with the morphologically very similar *Synodontis frontosus* Vaillant, 1895. According to the literature, the two species might co-occur in the Nile, Chad and Turkana basins (Hopson 1982; Paugy & Roberts 2003). The diagnostic features of the two species are similar and partly overlapping and distinguishing between them in the field is a notoriously confusing task. Due to the reasons mentioned above, combined approach using morphological and molecular tools was used to determine species identity of *Synodontis* samples included in the present study (see **Paper 2** for details).

Figure 2. Model host of the present study, *Synodontis schall* is medium sized (up to 35 cm standard length in L. Turkana), generalist, mostly carnivorous (in Turkana), bentho-pelagic fish inhabiting all habitats and parts of Turkana, including the open waters of the deepest parts of the lake. Depending on site and season, the species is most abundant in the depths between 10-25 m; exceptionally high aggregations of *S. schall* were occassionally found in Omo delta area in the lake water/river water interface; all information in this legend obtained from Hopson (1982). Photo: David Modrý, Lake Turkana, Kenya.

3.2.2. Parasite model - *Wenyonia virilis* Woodland,1923

The order Caryophyllidea van Beneden in Carus, 1863 (41 genera, ~150 species), to which the model parasite belongs, includes intestinal parasites of freshwater teleost fishes, mainly cyprinids, catfishes and catostomids. The only exception from the general pattern is genus *Archigetes*, members of which parasitize oligochaetes. All members are characterized by monopleuroid (no proglottisation) and monozoic (single set of reproductive organs) body plan (Mackiewicz 1994). The monopleuroid body is relatively small (up to 95 mm in *Khawia sinensis* Hsü, 1935), narrow, and without any internal or external segmentation. Although scolex morphology varies greatly between families, its morphology is always simple, devoid of sclerotized structures (hooks, hooklets, spines etc.), but possessing specialized attachment structures (e.g. loculi, bothria, folds, frills, acetabula) (Mackiewicz 1972; Oros *et al.* 2010; Schaeffner *et al.* 2011). All representatives of the order are hermaphrodites with a male and a female set of reproductive organs, eggs are operculate and mostly unembryonated. The life cycle involves one oligochaete intermediate host and fish definitive host. Intermediate host is infected by ingestion of eggs or oncosphere that develops into plerocercoid in the intermediate host coelom. In case of caryophyllid genus *Caryophyllaeus* including common parasites of carp and other Eurasian cyprinids, the intermediate hosts are tubificids and other oligochaetes that are castrated when infected (Goater et al. 2014). The definitive hosts are mostly benthic-feeding siluriform or cypriniform freshwater fishes, in whose gut the plerocercoid finishes the development into adult stage (Mackiewicz 1972). Caryophyllidea have been reported from Africa, Asia, Australia, Europe, and North America (Mackiewicz 1994). Although the Caryophyllidea as a group is widely distributed on all continents with highest diversity in the Nearctic region, they are absent in the Neotropics. Most caryophyllidean taxa are highly endemic, mostly infecting single host genus within a single biogeographic region, a pattern indicating long history of their association with, and adaptation to, particular host taxa (Mackiewicz 1972).

The genus *Wenyonia* Woodland,1923 is the only caryophyllidean taxon in which genital pore is located at the anterior part of body. As shown by taxonomic (re)analyses of available voucher and type material from various geographic regions, all known *Wenyonia* spp. are exclusively parasites of freshwater catfish of the genus *Synodontis* Cuvier, 1817 endemic to Africa. *Wenyonia* spp. is widespread throughout Sub-Saharan Africa, being absent or remaining unknown in four ichthyological provinces: Maghreb, South Africa, East coast and Zambezi. Absence of *Wenyonia* in Maghreb and South African ichthyological provinces obviously reflects absence of the host genus, whereas the lack of records in the East coast and Zambezi provinces possibly reflects undersampling. Life cycle of *Wenyonia* spp. remains unknown, but probably similar to other fish caryophyllideans as described above. Unlike all other caryophillideans, four body regions are recognized in *Wenyonia* spp.: scolex, neck, body and postovarian region, the latter being absent in all other representatives of Caryophyllidea (Schaeffner *et al.* 2011). The scolex is equipped with number of crevices and creases, sometimes tremed longitudinal furrows, which are considered the most determining feature for mode of attachment to host mucosa (Ibraheem & Mackiewicz 2006). The main diagnostic features used for *Wenyonia* spp. indentification include: i) scolex morphology, ii) disposition of testes, iii) arrangement of vitelline follicles throughout the body, iv) length ratios of the body parts (Schaeffner *et al.* 2011). Currently, six *Wenyonia* spp. are considered valid: *Wenyonia acuminata* Woodland,1923 (Nilo-Sudan and Congo; the geographic terms refer to ichthyological provinces from which the species has been reported); *Wenyonia longicauda* Woodland,1923 (Upper Guinea); *Wenyonia minuta* Woodland, 1923 (eastern Nilo-Sudan); *Wenyonia synodontis* Ukoli, 1965 (Nilo-Sudan); *Wenyonia virilis* (Nilo-Sudan); *Wenyonia youdeoweii* Ukoli, 1965 (Nilo-Sudan). The most recent comprehensive revision of the genus *Wenyonia* has been published by Schaeffner *et al.* (2011).

Wenyonia virilis, type species of the genus, and the parasite model of the present sturdy, is the most widespread and perhaps most common of all *Wenyonia* spp. It is reportedly distributed across numerous basins throughout the Nilo-Sudan ichthyological province - from Gambia in West Africa, across the Niger basin, to the Nile and Turkana basins in east Africa, and up to lower Nile in Egypt in the north-east of Africa (Schaeffner *et al.* 2011). In eastern part of the Nilo-Sudan, the dominant host species of *W. virilis* is its type host *Synodontis schall* (Bloch et Schneider, 1801); another eleven *Synodontis* spp. are reported as hosts of *W. virilis* by Schaeffner *et al.* (2011). Apart from the wide geographic range, other typical life-history traits of *W. virilis* include relatively high prevalence in most localities (frequently over 20% or even 50%), and sometimes high abundances, i.e. tens of specimens per individual host (Schaeffner *et al.* 2011, original unpublished data). Unlike other congeners, *W. virilis* shows remarkable morphological variability both within and between different localities and populations. Schaeffner *et al.* (2011) highlighted remarkable morphological variability in *W. virilis*: "...especially in the shape of the body, proportions and size of individual body regions, shape of the scolex and distribution of the vitelline follicles in the postovarian region". However, a few years earlier, the same author recognized five *W. virilis* morphotypes showing gradual differentiation rather than clearly distinct morphological-units, and considered them conspecific based on general scolex morphology and sequence analysis of 22 specimens (Schäffner 2009). These five morphotypes described by Schäffner (2009) based on material from the Nile River and Lake Turkana were used in the present study to characterize phenotype variability of the newly obtained material from the same localities.

4. Summary of results and discussion with respect to the partial aims

4.1. Opposing population histories: a fast parasite in a stationary host

Mitochondrial (*coxI*) and whole genome (AFLP see Appendix I) markers were used to assess intra- and inter-population genetic variability in parasite and its host from different sampling localities (**Paper 1, Fig. 1 and Tab. 1**) distributed in Turkana and Nile basins. As revealed by analysis of *coxI* data, parasites showed significantly more variability (**Paper 1, Fig. 4a**) compared to the hosts (**Paper 1, Fig. 4b**). Majority of host data clustered into a single haplotype represented in all sampled localities, while parasite data clustered into groups corresponding mostly to the geographic origin. This result is congruent with the long discussed theory of faster mutation rates in the parasite, which supposedly leads to greater population differences compared to its host (Page *et al.* 1998).

The contrasting status of individual parasite (sub)populations was revealed. *Wenyonia virilis* samples from the freshwater and saline Turkana localities exhibited a radial branching pattern with the main haplotype surrounded by satellite low frequency haplotypes - a trait typical for young expanding populations, while populations from the Nile contained (with one exception) only haplotypes organized into networks separated by multiple mutations without star-like patterns - a trait typical for older populations in equilibrium (**Paper 1, Fig. 4a**). Importantly, Turkana haplotypes originate almost exclusively either from saline or freshwater samples and in all hosts examined, only tapeworms belonging either to freshwater or saline haplotype were present. Such absence of mixed infections probably indicates presence of reproductively isolated tapeworm (sub)populations in the lake. The sporadic occurrences of freshwater haplotypes in saline localities, and vice versa, were always detected in sole (individual) hosts, and might be well explained by unrestricted migration of hosts throughout the lake (Hopson 1982). However, the general population genetic pattern indicates absence of gene flow between Turkana parasite subpopulations in freshwater and saline localities, implying that contrasting environmental conditions are more challenging for the parasite than for the host. This is in contrast with previously performed studies and well established concepts where both host and parasite are equally affected by environmental changes (Mostowy & Engelstädter 2011; Brunner *et al.* 2017).

Although the parasite and host population structures reconstructed using *coxI* are extremely divergent, evidence of geographically determined events shared between the two organisms was detected by analyses of their mt datasets. Population genetic statistics (haplotype and nucleotide diversities, Tajima's, Fu and Li's tests) showed that Turkana populations are genetically smaller and not in mutation-drift equilibrium, which is typical for populations that experienced a recent bottleneck (**Paper 1, Tab. 4**). Moreover, testing of migration theories showed that the White Nile is an ancestral population for both organisms and the Blue Nile and Turkana were colonized separately (**Paper 1, Fig. 2 and Tab. 5**). According to these results, Lake Turkana might have been colonized multiple times and different parts of the lake might be populated by tapeworms originating from different colonization waves, most

probably during different Nile-Turkana reconnections during the Holocene. This interpretation is further supported by revealed presence of genetically clearly distinct *W. virilis* subpopulations in freshwater and saline (brackish) areas of the lake (**Paper 1, Tab. 5**). The multiple origin scenario was also supported by the Principle Coordinates Analysis (PCoA) of parasite AFLP data. Separation of samples from freshwater and saline samples can be seen in the upper left part and lower left part of the PCoA diagram in **Paper 1** (**Fig. 5a,b**). Similarities between host and parasite data were also observed for the results of population genetic structure based on AFLP data (**Paper 1, Fig. 6a,b**). Clusters, which are less frequent in the Nile, show increased proportional presence in Turkana. This is possibly concordant with the effect of drift (increasing frequency of some clusters after a bottleneck), but does not give support for the origin, because clusters with increased frequency in the Nile should be statistically more likely to colonize Turkana than clusters with low frequency. The formation of Lake Turkana locality-specific (freshwater vs. saline) clusters might support a gradient of salinity as environmentaly-driven selection force, making isolation by adaptation more plausible explanation of the detected segregation of tapewom populations than isolation by distance (Orsini *et al.* 2013; Sexton *et al.* 2014; Spurgin *et al.* 2014). The salinity might affect the Turkana tapeworms indirectly by means of restricted distribution of alternative intermediate hosts (if different subpopulations use different intermediate host) in the freshwater or saline parts of the lake, in a way similar to other platyhelminths with complex life cycles, e.g. *Schistosoma mansoni* and its intermediate snail hosts *Biophalaria* spp. in South America (Paraense 1986; Morgan *et al.* 2005; Webster *et al.* 2013). The observed increased frequency in Turkana of parasite clusters less frequent in the Nile might be associated with multiple colonization events and different frequencies of clusters at the time of each colonization wave, as well as with temporal changes of population sizes (Papkou *et al.* 2016). Although not fully compatible, the outcomes of parasite *coxI* and AFLP data analyses suggest similar evolutionary patterns. When compared to the host data, parasite data show much more structure and better resolution of evolutionary patterns (multiple colonizations, bottleneck effect and ancestral population). Finally, obtained results represent probably the very first evidence supporting the long discussed theory of supposedly rapid parasite evolution opposed by slower evelution of its host, a phenomenon documented thus far only in ectoparasitic arthropods and apicomplexan protists (Cohuet *et al.* 2010; Khan *et al.* 2013).

4.2. No cospeciation: a diversifying parasite in a stationary host

As shown above, the host population structure for both markers shows markedly low variability and extent of evolutionary processes compared to the parasite, thus, only parasite data are further discussed. Compared with *coxI* data, the AFLP data provided better resolution of deeper population structure. Even though AFLP data analyses were not entirely satisfying and collection of more data would be required for further studies (see also point 3 below), the results imply presence of locality-specific genetic patterns (**Paper 1, Fig. 5a,b**), as well as formation of a specific clusters exclusive either for the freshwater or saline parts of Lake Turkana (**Paper 1, Fig. 6a – magenta**). In the environmentally contrasting freshwater

and saline parts of the Lake Turkana, the tapeworm subpopulations might be subject to different environmental pressures, which might, especially given the apparent lack of gene flow between the subpopulations, induce allopatric differentiation in a sympatric system, i.e. divergence of parasite subpopulations in freshwater and saline environments of a single lake in which no barriers exist to dispersal of the host (and the parasite alike). This interpretation seems to be supported by the robust evidence of spatial overlap, and thus physical contact, between the tapeworm subpopulations as reflected by sporadic, but repeated and reliably confirmed records of hosts infected with tapeworms from the saline subpopulation in freshwater part of lake and vice versa. Under such circumstances, the clear divergence of the tapeworm subpopulations and the lack of gene flow between them are of special interest, implying some form of environmental non-permeability for the parasite (but not its host), possibly associated with its complex life cycle.

4.3. Mitochondrial markers: more informative than whole genome markers (AFLP)

AFLP did not provide an optimal level of resolution (**Paper 1**). The AFLP data alone were not fully convincing with respect to clear distinction between the effect of IBA (isolation by adaptation) and IBD (isolation by distance). Moreover, compared to the mt marker, the AFLP did not separate clearly distinct, only distantly related species neither in the *Wenyonia* spp. nor in the *Synodontis* spp. dataset (**Paper 1, Fig. 5a,b,c,d, and Fig. 6a,b**). However, the combination of both types of markers helped us to reveal and confirm multiple colonization of Lake Turkana from the Nile. Moreover, AFLP data also showed on-going separation of freshwater and saline parasite subpopulations in Turkana and formation of Turkana specific cluster (**Paper 1, Fig. 5a,b and Fig. 6a**). Despite these findings, the overall image of evolution in populations was much less clear when using only AFLP data without analysis of *coxI* sequences. AFLP mostly supported the *coxI* outcomes and provided additional information about deeper population structure. Moreover, the signal of the ancestral gene flow contained in the mtDNA data was similar for both organisms and provided valuable information about historical changes in species distributions. Therefore, the suitability of AFLP data to elucidate one of the main questions of this study, i.e. significance of the salinity gradient itself, remains ambiguous. Thus, using other molecular markers (e.g. GWAS) for further studies of this host-parasite system would be required in order to reveal detailed population patterns and to make inference about the role of the environment in population differentiation.

4.4. Intricate host identity: cryptic sibling and unexpected phenotypic variability

Apart from several well recognizable *Synodontis* spp. in the Nile, two very similar species, *S. schall* and *Synodontis frontosus*, i.e. easily mistakable potential hosts for *Wenyonia* tapeworms, were reported from both the Nile and Lake Turkana. As the study of the present host-parasite system required reliable host determination, the whole *Synodontis* dataset from

both basins was subjected to thorough morphological and genetic analysis (diagnostic morphological traits vs. AFLP + the original *coxI* dataset enlarged and supplemented with *cytb* and *RAG2* genes). Morphological analysis of the *Synodontis* sample (n=138) in the field (specimens examined, photographed and tissue samples taken, no vouchers collected) revealed not only typical *S. schall*-like and *S. frontosus*-like morphotypes, but also high incidence of intermediate forms. Haplotype network revealed the presence of only one species in Turkana (*S. schall*) and two species in the Nile sample (*S. schall* + morphologically well distinguished control *Synodontis nigrita*) (**Paper 2, Fig. 4b**). All analyses of the enlarged *coxI* (n=131), *cytb* (n=96) and *RAG2* (n=23) datasets, supplemented with all available GenBank data for respective genes, revealed equivocally a presence of only a single *Synodontis* sp. in Lake Turkana. All *S. schall*-like + *S. frontosus*-like specimens, including intermediate forms, from the Nile clustered together with those from Turkana in all analyses performed for all genes (**Paper 2, Figs. 3, 4, 6, 7, 8**). The cluster containing clearly conspecific specimens from Turkana and the Nile was assigned to *S. schall* sensu stricto as it is defined by the type locality in the Nile. Based on the results, morphological characteristics of *S. frontosus* were considered intraspecific variability of *S. schall*, revealing its yet overlooked phenotypic variability, and synonymization of the two nominal taxa was suggested (though not formally performed due to absence of vouchrs), indicating priority of the specific name *S. schall* over *S. frontosus*. In addition, mitochondrial data (*coxI*, *cytb*) revealed a strong biogeographical signal in distribution *S. schall*-like + *S. frontosus*-like catfishes (both original and GenBank data) throughout the Nilo-Sudan ichthyolocial province. The two distinct, yet unrecognized clades, *S. schall* sensu stricto and *S. schall* sensu lato, most likely restricted to the eastern (e.g. Nile, Turkana, Chad) and western (e.g. Senegambia, Niger, Chad) parts of the Nilo-Sudan ichthyological province, respectively. The western sensu lato cluster most probably represents a cryptic species, yet unrecognised due to the absence of discernible morphological signal. Contact zone of both entities was identified in the Chad basin. These findings highlight the intricate nature of comparative host-parasite datasets in which special emphasis should be put on host determination. The fact that authors of redescription of *S. schall* (Musschoot & Lalèyè 2008) did not detect any morpphological signal among 105 specimens from throughout Nilo-Sudan by statistic analysis implies that thorough revision of the Nilo-Sudan *S. schall*-like catfishes based on both morphological and genetic approaches is needed.

4.5. Matching phenotypic and population genetic patterns: a role for contrasting lacustrine environments?

Individual tapeworms from the present Nile-Turkana *W. virilis* sample were assigned to five morphotypes described previously by Schaeffner (2009) based on material from the same localities. To assess potential correlations between phenotype variability and population genetic structure, morphotype data were compared with corresponding *coxI* and AFLP datasets and the observed patterns cross verified by AMOVA analyses of both types of data. AMOVA analysis of *coxI* data confirmed absence of statistical differences within individual

sequence datasets comprising specimens assigned to different morphotypes. No difference has been found both between sequence datasets corresponding with two very similar morphotypes, and their datasets were pooled (D+G). Significant difference between D+G sequence dataset and a sequence dataset corresponding with morphotype E was found. Datasets of two other morphotypes (F, H) were excluded from further analyses due to small sample sizes. The two major well sampled morphotypes (D+G and E) were mapped over population genetic data for both markers (*coxI* and AFLP), revealing strong correlation between morphotype distribution and population genetic structure. The outcome showed clear affinity of each of the two morphotypes either to the freshwater or saline tapeworm subpopulation, i.e. freshwater or saline environments of Turkana (**Paper 3, Figs. 1 and 2, Tab. 3**). Observed patterns were cross verified by conducting AMOVA analyses of both types of data with identical outcomes (**Paper 3, Tabs. 3**). Due to the origin of Turkana *W. virilis* population in the Nile (**Paper 1**), presence of only a single host in our Nile-Turkana sample (**Paper 2**) and occurrence of all (but one, not analyzed) morphotypes at all (but two, extremely low sample sizes) major sampling localities in both the Nile and Turkana (**Paper 3, Tab. 1, Fig. S1**), the effect of different hosts as well as geographic origin of samples on observed patterns could be dismissed. The phenotypic variability and occurrence of different morpho-varieties is common in caryophyllidean cestodes (e.g. Barčák *et al.* 2014; Bazsalovicsová *et al.* 2014; Hanzelová *et al.* 2015). However, the majority of studies compared cestodes from different hosts or vastly distant localities (e.g. France, Portugal, Ukraine, Slovakia, Finland, Barčák *et al.* 2014). However, the presented results show that geographic origin, i.e. long distances between sampling localities, and different host species are might not necessarily be responsible for phenotypic patterns in caryophyllideans. The coinciding absence of gene flow between *W. virilis* Turkana subpopulations (**Paper 1**) and the apparent match between phenotype variability and population structure further supports hypothesis proposed in **Paper 1**, evolutionary processes involving allopatric differentiation (freshwater vs. saline environments) in a sympatric system (same host species with unlimited dispersal, same limnic system) are ongoing in the Lake Turkana. The documented level of divergence within Turkana tapeworm population is striking, especially taking into account that the young Turkana limnic system is of early-mid Holocene origin, i.e. ~11.5-5.2 millennia of age only, and the remarkable patterns observed reflect slightly over five millennia of isolation (for Turkana's history, see Bloszies *et al.* 2015).

5. Conclusions and future perspectives

5.1. Conclusions

Obtained results for host-parasite system showed that both model organisms originate from the White Nile and both populations experienced bottleneck during the Lake Turkana colonization. The colonization events are better reflected in the parasite dataset. The population structure of the parasite implies segregation of its conspecific subpopulations in different, environmentally contrasting (freshwater and saline) areas of the limnic system, as well as apparent lack of gene flow between the subpopulations inhabiting freshwater and saline parts of the lake. This stands opposed to the patterns obtained from host data, which assume (probably repeated) fusions of colonization waves during the Holocene Nile-Turkana reconnection events. By comparing organisms with same or similar generation time, the performed analyses of host-parasite system confirmed the long dicussed teory that parasites evolve much more rapidly and have faster mutation rates than their hosts (Page *et al.* 1998). In the present study, first evidence of this phenomenon was provided for endoparasitic metazoans. Although AFLP and mtDNA showed similar patterns in both host and its parasite, mtDNA provided a remarkably higher resolution of population structure, historical gene flow, evolutionary patterns, as well as inter- and intra-specific genetic differentiation.

Morphological and molecular examination of *Synodontis* samples from Turkana and the Nile aided with the all available sequence data from other *Synodontis* spp., showed that only one *Synodontis* species is present in Lake Turkana. Our results not congruent with literature reports on a presence of two morphologically similar congeners *S. schall* and *S. frontosus* in Lake Turkana and the Nile (Hopson 1982; Day *et al.* 2009). All individuals morphologically identified either as *S. schall* or *S. frontosus* in our sample from both basins clustered together in performed all analyses for all sequenced genes. By analysing and comparing mt and nuclear genes we proved that the apparent occurrence of only one species is not a consequence of introgression and/or lineage sorting. In addition, our analyses revealed highly probable presence of a cryptic species hidden under the nominal species *S. schall* with a strong geographical affinity to western part of the Nilo-Sudan ichtyological province and apparently overlapping with *S. schall* sensu stricto (as defined by type locality) in the Chad basin. Therefore we hint for thorough revision and detailed examination of morphological features and population genetic patterns of *S. schall*-like mochokids in Nilo-Sudan.

Our study proved that *W. virilis* is a well-defined monophyletic species with a high level of phenotypic variability. All five morphotypes were recorded in both basins at all major sampling localities, therefore the geographic origin or environmental factors are not a priori decisive with respect to their presence/absence in particular part of the lake. Most prevalent and best sampled morphotypes present in Turkana (D+G and E) showed high affinity either to freshwater or saline parts of the Lake Turkana, and correlated with population structure. Compared to findings of population genetic structure, where AFLP data did not provide optimal level of resolution when compared with mt data, in this case, outcomes from both markers are almost identical and showed that different morphotypes are significantly correlated with population genetic structure of Turkana parasite populations.

5.2. Future perspectives

Overall, results obtained in **Paper 1** show that subsequent population genetics studies should include different molecular markers (e.g. microsatellites, RAD-seq) to cross check our outcomes and to clearly answer the question of salinity gradient impact as the selection pressure on parasite population structure (isolation by distance vs isolation by adaptation).

Furthermore, different parasite taxa from the same host might be included in the study to provide insight into how distantly related members of the same taxonomic class respond to the same or similar environmental conditions and historical events. Among potential candidates *Proteocephalus synodontis* (Cestoda, Proteocephalidae), parasitizing the same host in the same basins, appears suitable. However, inclusion of parasite taxa other than Cestoda could reveal completely different scenarios of population genetic response to the peculiar environment and short-term isolation. Other parasites of *S. schall* present in both Turkana and the Nile include especially Trematoda (*Sandonia sudanensis*, *Glossidium pedatum*). Among other suitable host-parasite models present in the model geographic area and fulfilling the criteria specified in this study, especially *Labeo horiae-Nematobothrium labeonis*, i.e. common and abundant cyprinid and highly host-specific didymozoid trematode, might pose a good comparative model.

Furthermore, *W. virilis* morphotypes and its population structure in the Nile should be investigated in more detail to asses geographic structuring of the populations and morphotype frequencies in maternal population.

The importance of correct host species determination was shown in **Paper 2**, highlighting one of the underestimated aspects of parasitology. During fieldwork, fish are most likely to be categorized by parasitologists not by ichthyologists. This fact might not only lead to incorrect species identification, but moreover, also contribute to an artificial host species list in various parasites (as shown in practice in **Paper 2**). This type of error might affect future studies on parasite diversity, leading to both over- and underestimations of host spectra or diversity of parasites. Simply, proper host determination should become essential part of studies on parasite populations and, molecular identification of hosts should become the standard procedure to avoid further errors.

Owing to the fact that population genetic studies are conducted with multiple samples of one species from the same locality, it is not an exception that during that intensive sampling cryptic species or species complexes are discovered (Brown *et al.* 2007; Elmer *et al.* 2007; Paterson *et al.* 2016). Furthermore, population genetic studies combine the knowledge about biota and the geological and climatic history with adequate sample size and taxonomic description; therefore the discovery of cryptic species is much more accurate (Beheregaray & Caccone 2007; Detwiler *et al.* 2010; Nadler & De Leon 2011). The prevalence of cryptic species, even in well studied animals like the giraffe (Brown *et al.* 2007) or tortoise (Russello *et al.* 2005), highlights the importance of combining multiple approaches in order to describe species complexity.

6. Thesis References

Addy F, Wassermann M, Banda F *et al.* (2017) Genetic polymorphism and population structure of *Echinococcus ortleppi*. *Parasitology*, **144**, 450–458.

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819–825.

Attwood SW (2010) Studies on the Parasitology, Phylogeography and the Evolution of Host-Parasite Interactions for the Snail Intermediate Hosts of Medically Important Trematode Genera in Southeast Asia. *Advances in Parasitology*.

Attwood SW, Upatham ES, Meng XH, Qiu DC, Southgate VR (2002) The phylogeography of Asian *Schistosoma* (Trematoda: Schistosomatidae). *Parasitology*, **125**, 99–112.

Avise JC (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.

Bandilla M, Hakalahti-Sirén T, Valtonen ET (2008) Patterns of host switching in the fish ectoparasite *Argulus coregoni*. *Behavioral Ecology and Sociobiology*, **62**, 975–982.

Barčák D, Oros M, Hanzelová V, Scholz T (2014) Phenotypic plasticity in *Caryophyllaeus brachycollis* Janiszewska, 1953 (Cestoda: Caryophyllidea): Does fish host play a role? *Systematic Parasitology*, **88**, 153–166.

Bazsalovicsová E, Kralová-Hromadová I, Brabec J *et al.* (2014) Conflict between morphology and molecular data: A case of the genus *Caryophyllaeus* (Cestoda: Caryophyllidea), monozoic tapeworms of cyprinid fishes. *Folia Parasitologica*, **61**, 347–354.

Beheregaray LB, Caccone A (2007) Cryptic biodiversity in a changing world. *Journal of biology*, **6**, 9.

Bermingham E, Moritz C (1998) Comparative phylogeography: concepts and applications. *Molecular Ecology*, **7**, 367–369.

Bloszies C, Forman SL, Wright DK (2015) Water level history for Lake Turkana, Kenya in the past 15,000years and a variable transition from the African Humid Period to Holocene aridity. *Global and Planetary Change*, **132**, 64–76.

Bouzid W, Stefka J, Hypsa V *et al.* (2008) Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). *International Journal for Parasitology*, **38**, 1465–79.

Bowles J, Blair D, McManus DP (1995) A molecular phylogeny of the human schistosomes. *Molecular phylogenetics and evolution*, **4**, 103–109.

Brabec J, Waeschenbach A, Scholz T, Littlewood DTJ, Kuchta R (2015) Molecular phylogeny of the Bothriocephalidea (Cestoda): Molecular data challenge morphological classification. *International Journal for Parasitology*, **45**, 761–771.

Brooks DR (1979) Testing the Context and Extent of Host-Parasite Coevolution. *Systematic Biology*, **28**, 299–307.

Brooks DR, McLennan DA (1993) *Parasites and the language of evolution Smithsonian Institution Press, Washington, District of Columbia Burnham*. Parascript.

Brouat C, Tatard C, Machin A *et al.* (2011) Comparative population genetics of a parasitic nematode and its host community: the trichostrongylid *Neoheligmonella granjoni* and *Mastomys rodents* in southeastern Senegal. *International Journal for Parasitology*, **41**,

1301–9.

Brown DM, Brenneman RA, Koepfli K-P *et al.* (2007) Extensive population genetic structure in the giraffe. *BMC biology*, **5**, 57.

Brunner FS, Anaya-Rojas JM, Matthews B, Eizaguirre C (2017) Experimental evidence that parasites drive eco-evolutionary feedbacks. *Proceedings of the National Academy of Sciences*, **114**, 3678–3683.

Carling MD, Brumfield RT (2008) Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in Passerina buntings. *Genetics*, **178**, 363–377.

Chambrier A de, Waeschenbach A, Fisseha A, Scholz T, Mariaux J (2015) A large 28S rDNA-based phylogeny confirms the limitations of established morphological characters for classification of proteocephalidean tapeworms (Platyhelminthes, Cestoda). *ZooKeys*, **500**, 25–59.

Clayton DH, Bush S, Jahnson K (2004) Ecology of Congruence: Past Meets Present. *Systematic Biology*, **53**, 165–173.

Cohuet A, Harris C, Robert V, Fontenille D (2010) Evolutionary forces on *Anopheles*: what makes a malaria vector? *Trends in Parasitology*, **26**, 130–136.

Criscione CD, Blouin MS (2005) Effective sizes of macroparasite populations: a conceptual model. *Trends in Parasitology*, **21**, 212–7.

Criscione CD, Blouin MS (2007) Parasite phylogeographical congruence with salmon host evolutionarily significant units: Implications for salmon conservation. *Molecular Ecology*, **16**, 993–1005.

Criscione CD, Poulin R, Blouin MS (2005) Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology*, **14**, 2247–57.

Cruaud A, Rasplus JY (2016) Testing cospeciation through large-scale cophylogenetic studies. *Current Opinion in Insect Science*, **18**, 53–59.

Day JJ, Bills R, Friel JP (2009) Lacustrine radiations in African *Synodontis* catfish. *Journal of Evolutionary Biology*, **22**, 805–17.

Day JJ, Peart CR, Brown KJ *et al.* (2013) Continental diversification of an African catfish radiation (Mochokidae: *Synodontis*). *Systematic Biology*, **62**, 351–365.

Day JJ, Wilkinson M (2006) On the origin of the *Synodontis* catfish species flock from Lake Tanganyika. *Biology Letters*, **2**, 548–552.

Detwiler JT, Bos DH, Minchella DJ (2010) Revealing the secret lives of cryptic species: Examining the phylogenetic relationships of echinostome parasites in North America. *Molecular Phylogenetics and Evolution*, **55**, 611–620.

Doležalová J, Vallo P, Petrželková KJ *et al.* (2015) Molecular phylogeny of anoplocephalid tapeworms (Cestoda: Anoplocephalidae) infecting humans and non-human primates. *Parasitology*, **142**, 1278–1289.

Dumont HJ (2009) A Description of the Nile Basin, and a Synopsis of Its History, Ecology, Biogeography, Hydrology, and Natural Resources. *The Nile*, 1–21.

Elmer KR, Dávila J a, Lougheed SC (2007) Cryptic diversity and deep divergence in an upper Amazonian leaflitter frog, *Eleutherodactylus ockendeni*. *BMC evolutionary biology*, **7**,

247.

Engelstädter J, Bonhoefferkesh S (2009) Red Queen Dynamics with Non-Standard Fitness Interactions. *PLoS Computational Biology*, **5**, e1000469.

Eschmeyer WN, Fricke R, van der Laan R (2017) Catalog of fishes: genera, species, references. (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 22/07/2017 (WN Eschmeyer, R Fricke, R van der Laan, Eds,).

Freeland JR (2005) *Molecular Ecology*. John Wiley & Sons Ltd., West Sussex.

Froese R, Pauly D (2017) Fishbase. *World Wide Web electronic publication. www.fishbase.org, version (02/2017).*

Garrick RC, Benavides E, Russello M a *et al.* (2014) Lineage fusion in Galápagos giant tortoises. *Molecular Ecology*, **23**, 5276–5290.

Good JM, Vanderpool D, Keeble S, Bi K (2015) Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution*, **69**, 1961–1972.

Gupta H, Srivastava S, Chaudhari S *et al.* (2016) New molecular detection methods of malaria parasites with multiple genes from genomes. *Acta Tropica*, **160**, 15–22.

Hafner M, Sudman P, Villablanca F *et al.* (1994) Disparate rates of molecular evolution in cospeciating hosts and parasites. *Science*, **265**, 1087–1090.

Hahn C, Weiss SJ, Stojanovski S, Bachmann L (2015) Co-speciation of the ectoparasite *Gyrodactylus teuchis* (Monogenea, Platyhelminthes) and its salmonid hosts. *PLoS ONE*, **10**, 1–20.

Hanzelová V, Oros M, Barčák D *et al.* (2015) Morphological polymorphism in tapeworms: redescription of *Caryophyllaeus laticeps* (Pallas, 1781) (Cestoda: Caryophyllidea) and characterisation of its morphotypes from different fish hosts. *Systematic Parasitology*, **90**, 177–190.

Harvey CPD, Grove AT (1982) A Prehistoric Source of the Nile. *The Geographical Journal*, **148**, 327–336.

Hopkins CD, Lavoué S, Sullivan JP (2007) The Fresh and Brackish Water Fishes of Lower Guinea, West-Central Africa. In: *The Fresh and Brackish Water Fishes of Lower Guinea, West-Central Africa* (eds Stiassny ML, Teugels GG, Hopkins CD), pp. 729–751.

Hopson A (1982) *Lake Turkana; A Report on the Finding of the Lake Turkana Project 1972–1975, Vols. 1-6*. Overseas Development Administration, London.

Huyse T, Oeyen M, Larmuseau MHD, Volckaert FAM (2017) Co-phylogeographic study of the flatworm *Gyrodactylus gondae* and its goby host Pomatoschistus minutus. *Parasitology International*, **66**, 119–125.

Huyse T, Poulin R, Théron A (2005) Speciation in parasites: a population genetics approach. *Trends in Parasitology*, **21**, 469–75.

Ibraheem MH, Mackiewicz JS (2006) Scolex development, morphology and mode of attachment of *Wenyonia virilis* Woodland, 1923 (Cestoidea, Caryophyllidea). *Acta Parasitologica*, **51**, 51–58.

Jabbar A, Narankhajid M, Nolan MJ *et al.* (2011) A first insight into the genotypes of *Echinococcus granulosus* from humans in Mongolia. *Molecular and Cellular Probes*, **25**, 49–54.

Jirsová D, Štefka J, Jirků M (2017) Discordant population histories of host and its parasite : A role for ecological permeability of extreme environment ? *PLoS ONE*, **12**, 1–25.

Johnson TC, Malala JO (2009) The Nile: Origin, Environments, Lymnology and Human Use. In: *Lake Turkana and its link to the Nile*, pp. 287–302. Springer Science+Business Media BV.

Khan SM, Reece SE, Waters AP, Janse CJ, Kaczanowski S (2013) Why are male malaria parasites in such a rush?: Sex-specific evolution and host-parasite interactions. *Evolution, medicine, and public health*, **2013**, 3–13.

Kmentová N, Gelnar M, Mendlová M *et al.* (2016) Reduced host-specificity in a parasite infecting non-littoral Lake Tanganyika cichlids evidenced by intraspecific morphological and genetic diversity. *Nature Publishing Group*, 1–15.

Koblmüller S, Sturmbauer C, Verheyen E, Meyer A, Salzburger W (2006) Mitochondrial phylogeny and phylogeography of East African squeaker catfishes (Siluriformes: *Synodontis*). *BMC evolutionary biology*, **6**, 49.

Kolding J (1993) Trophic interactions and community structure at two different periods of Lake Turkana, Kenya: a comparison using the ECOPATH II box model. In: *Trophic Models of Aquatic Ecosystems: ICLARM* (eds Christensen V, Pauly D), pp. 116–123.

Králová-Hromadová I, Bazsalovicsová E, Štefka J *et al.* (2011) Multiple origins of European populations of the giant liver fluke *Fascioloides magna* (Trematoda: Fasciolidae), a liver parasite of ruminants. *International Journal for Parasitology*, **41**, 373–383.

Kvičerová J, Hypša V (2013) Host-Parasite incongruences in iodent *Eimeria* suggest significant role of adaptation rather than cophylogeny in maintenance of host specificity. *PLoS ONE*, **8**.

Le TH, Blair D, McManus DP (2000) Mitochondrial genomes of human helminths and their use as markers in population genetics and phylogeny. *Acta Tropica*, **77**, 243–256.

Levin II, Parker PG (2013) Comparative host-parasite population genetic structures: obligate fly ectoparasites on Galapagos seabirds. *Parasitology*, 1–9.

Lier T, Johansen M V., Hjelmevoll SO, Vennervald BJ, Simonsen GS (2008) Real-time PCR for detection of low intensity *Schistosoma japonicum* infections in a pig model. *Acta Tropica*, **105**, 74–80.

van Lieshout L, Roestenberg M (2015) Clinical consequences of new diagnostic tools for intestinal parasites. *Clinical Microbiology and Infection*, **21**, 520–528.

Mackiewicz JS (1972) Caryophyllidea (Cestoidea): a review. *Experimental Parasitology*, **31**, 417–512.

Mackiewicz JS (1994) Order Caryophyllidea van Beneden in Carus, 1863. In: *Keys to cestodes parasites of vertebrates* (eds Khalil LF, Jones A, Bray RA), pp. 21–43. CAB International, Wallingford, U.K.

McManus DP, Bowles J (1996) Molecular genetic approaches to parasite identification: Their value in diagnostic parasitology and systematics. *International Journal for Parasitology*, **26**, 687–704.

Møller AP, Merino S, Soler JJ *et al.* (2013) Assessing the effects of climate on host-parasite interactions: A comparative study of european birds and their parasites. *PLoS ONE*, **8**.

Morand S, Krasnov BR, Poulin R (2006) *Micromammals and macroparasites: From evolutionary ecology to management* (S Morand, BR Krasnov, R Poulin, Eds,). Springer-Verlag Tokyo.

Moravec F, Charo-Karisa H, Jirků M (2009a) *Philometrids* (Nematoda: Philometridae) from fishes of Lake Turkana, Kenya, including two new species of *Philometra* and erection of *Afrophilometra* gen. *Folia Parasitologica*, **56**, 41–54.

Moravec F, Jirků M, Charo-Karisa H, Mašová Š (2009b) *Mexiconema africanum* sp. n. (Nematoda: Daniconematidae) from the catfish *Auchenoglanis occidentalis* from Lake Turkana, Kenya. *Parasitology Research*.

Morgan JAT, Dejong RJ, Adeoye GO *et al.* (2005) Origin and diversification of the human parasite *Schistosoma mansoni*. *Molecular Ecology*, **14**, 3889–3902.

Mostowy R, Engelstädter J (2011) The impact of environmental change on host-parasite coevolutionary dynamics. *Proceedings. Biological sciences / The Royal Society*.

Muška M, Vašek M, Modrý D *et al.* (2012) The last snapshot of natural pelagic fish assemblage in Lake Turkana, Kenya: A hydroacoustic study. *Journal of Great Lakes Research*, **38**, 98–106.

Musschoot T, Lalèyè P (2008) Designation of a neotype for *Synodontis schall* (Bloch and Schneider, 1801) and description of two new species of *Synodontis* (Siluriformes: Mochokidae). *Journal of Natural History*, **42**, 1303–1331.

Nadler SA, De Leon GPP (2011) Integrating molecular and morphological approaches for characterizing parasite cryptic species: implications for parasitology. *Parasitology*.

Neiswenter S a., Riddle BR (2011) Landscape and climatic effects on the evolutionary diversification of the *Perognathus fasciatus* species group. *Journal of Mammalogy*, **92**, 982–993.

Nieberding CM, Morand S (2006) Comparative phylogeography: The use of parasites for insights into host history. In: *Micromammals and Macroparasites: From Evolutionary Ecology to Management* (eds Morand S, Krasnov BR, Poulin R), pp. 277–293. Springer-Verlag Tokyo.

Nieberding C, Morand S, Libois R, Michaux JR (2004) A parasite reveals cryptic phylogeographic history of its host. *Proceedings of the Royal Society B: Biological Sciences*.

Nieberding CM, Olivieri I (2007) Parasites: proxies for host genealogy and ecology? *Trends in Ecology and Evolution*, **22**, 156–165.

Nyamweru C (1989) New evidence for the former extent of the Nile drainage system. *The Geographical Journal*, **155**, 179–188.

Olson PD, Littlewood DTJ, Bray RA, Mariaux J (2001) Interrelationships and Evolution of the Tapeworms (Platyhelminthes: Cestoda). *Molecular Phylogenetics and Evolution*, **19**, 443–467.

Olson PD, Tkach V V. (2005) Advances and trends in the molecular systematics of the parasitic platyhelminthes. *Advances in Parasitology*, **60**, 165–243.

Oros M, Scholz T, Hanzelová V, Mackiewicz JS (2010) Scolex morphology of monozoic

cestodes (Caryophyllidea) from the palaearctic region: A useful tool for species identification. *Folia Parasitologica*, **57**, 37–46.

Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L (2013) Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983–5999.

Page RDM (2003) *Tangled trees: Phylogeny, Cospeciation, and Coevolution*. University of Chicago press, Chicago.

Page RDM, Lee PL, Becher SA, Griffiths R, Clayton DH (1998) A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. *Molecular phylogenetics and evolution*, **9**, 276–93.

Papkou A, Gokhale CS, Traulsen A, Schulenburg H (2016) Host–parasite coevolution: why changing population size matters. *Zoology*, **119**, 330–338.

Paraense WL (1986) Distribuicao dos caramujos no Brasil. In: *Medernos conhecimentos sobre a Esquistossomose mansonica* (eds Reis FA, Faria I, Katz N), pp. 117–127. Suplemento dos Anais de 1983/84 da Academia Mineira de Medicina, Belo Horizonte 1986.

Paterson ID, Mangan R, Downie DA *et al.* (2016) Two in one: cryptic species discovered in biological control agent populations using molecular data and crossbreeding experiments. *Ecology and Evolution*, **6**, 6139–6150.

Paugy D, Roberts TR (2003) Mochokidae. p. In: *C. Lévêque, D. Paugy and G.G. Teugels (eds.) Faune des poissons d'eaux douce et saumâtres de l'Afrique de l'Ouest, Tome 2. Coll.*, pp. 195–268. Faune et Flore tropicales 40. Musée Royal de l'Afrique Centrale, Tervuren, Belgique, Museum National d'Histoire Naturalle, Paris, France and Institut de Recherche pour le Développement, Paris, France. 815 p.

Pinton A, Agnèse JF, Paugy D, Otero O (2013) A large-scale phylogeny of *Synodontis* (Mochokidae, Siluriformes) reveals the influence of geological events on continental diversity during the Cenozoic. *Molecular phylogenetics and evolution*, **66**, 1027–40.

Poll M (1971) Révision des Synodontis africains (famille Mochokidae). *Annales – Musée royal de l'Afrique centrale. Zoology*, **191**, 1–497.

Poulin R, Krasnov BR, Mouillot D (2011) Host specificity in phylogenetic and geographic space. *Trends in parasitology*, **27**, 355–61.

Poulin R, Morand S (2000) The diversity of parasites. *The Quarterly review of biology*, **75**, 277–293.

Ricketts RD, Anderson RF (1998) A Direct comparison between the historical record of lake level and the δ18O signal in carbonate sediments from Lake Turkana, Kenya. *Limnology and Oceanography*, **43**, 811–822.

Riddle BR, Hafner DJ, Alexander LF (2000) Comparative phylogeography of Baileys' pocket mouse (*Chaetodipus baileyi*) and the *Peromyscus eremicus* species group: Historical Vicariance of the Baja California Peninsular Desert. *Molecular Phylogenetics and Evolution*, **17**, 161–172.

Russello MM, Glaberman S, Gibbs JP *et al.* (2005) A cryptic taxon of Galápagos tortoise in conservation peril. *Biology letters*, **1**, 287–90.

Sands DD (1983) *Catfishes of the World. Vol. 2: Mochokidae*. Dunure Enterprises, Dunure,

Scotland.

Sanpool O, Rodpai R, Intapan PM *et al.* (2017) Genetic diversity of *Taenia saginata* (Cestoda: Cyclophyllidea) from Lao People's Democratic Republic and northeastern Thailand based on mitochondrial DNA. *Parasites & Vectors*, **10**, 141.

Schaeffner BC, Jirků M, Mahmoud ZN, Scholz T (2011) Revision of *Wenyonia* Woodland, 1923 (Cestoda: Caryophyllidea) from catfishes (Siluriformes) in Africa. *Systematic parasitology*, **79**, 83–107.

Schäffner B (2009) Monozoic tapeworms (Cestoda: Caryophyllidea) of the genus *Wenyonia*, parasites of *Synodontis* catfishes in Africa. Universität Karlsruhe (TH).

Schluter D (2001) Ecology and the origin of species. TL - 16. *Trends in ecology & evolution*.

Sexton JP, Hangartner SB, Hoffmann AA (2014) Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, **68**, 1–15.

Sharma M, Fomda BA, Mazta S *et al.* (2013) Genetic diversity and population genetic structure analysis of *Echinococcus granulosus* sensu stricto complex based on mitochondrial DNA signature. *PLoS ONE*, **8**, 1–8.

Sørensen E, Drew A., Brindley PJ *et al.* (1998) Variation in the sequence of a mitochondrial nadh dehydrogenase i gene fragment among six natural populations of *Schistosoma japonicum* from Chinafn1. *International Journal for Parasitology*, **28**, 1931–1934.

Spurgin LG, Illera JC, Jorgensen TH, Dawson DA, Richardson DS (2014) Genetic and phenotypic divergence in an island bird: Isolation by distance, by colonization or by adaptation? *Molecular Ecology*, **23**, 1028–1039.

Štefka J, Hoeck PEA, Keller LF, Smith VS (2011) A hitchhikers guide to the Galápagos: co-phylogeography of Galápagos mockingbirds and their parasites. *BMC Evolutionary Biology*, **11**, 284.

Tham JM, Lee SH, Tan TMC, Ting RCY, Kara UAK (1999) Detection and species determination of malaria parasites by PCR: Comparison with microscopy and with ParaSight-F and ICT Malaria Pf tests in a clinical environment. *Journal of Clinical Microbiology*, **37**, 1269–1273.

Toit N, Van Vuuren BJ, Matthee S, Matthee CA (2013) Biogeography and host-related factors trump parasite life history: Limited congruence among the genetic structures of specific ectoparasitic lice and their rodent hosts. *Molecular Ecology*, **22**, 5185–5204.

Troell K, Engström A, Morrison D a, Mattsson JG, Höglund J (2006) Global patterns reveal strong population structure in *Haemonchus contortus*, a nematode parasite of domesticated ruminants. *International journal for parasitology*, **36**, 1305–16.

Vanhove MPM, Pariselle A, Van Steenberge M *et al.* (2015) Hidden biodiversity in an ancient lake: phylogenetic congruence between Lake Tanganyika tropheine cichlids and their monogenean flatworm parasites. *Scientific Reports*, **5**, 13669.

Verweij JJ, Rune Stensvold C (2014) Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. *Clinical Microbiology Reviews*, **27**, 371–418.

Webster BL, Webster JP, Gouvras AN *et al.* (2013) DNA 'barcoding' of *Schistosoma mansoni* across sub-Saharan Africa supports substantial within locality diversity and geographical separation of genotypes. *Acta Tropica*, **128**, 250–260.

Witte F, van Oijen MJP, Sibbing FA (2009) Fish fauna of the Nile. In: *The Nile: Origin, Environments, Limnology and Human Use.* (ed Dumont H), pp. 647–675. Springer.

Yanagida T, Mohammadzadeh T, Kamhawi S *et al.* (2012) Genetic polymorphisms of *Echinococcus granulosus* sensu stricto in the Middle East. *Parasitology International*, **61**, 599–603.

Zarlenga DS, Hoberg E, Rosenthal B, Mattiucci S, Nascetti G (2014) Anthropogenics: Human influence on global and genetic homogenization of parasite populations. *Journal of Parasitology*, **100**, 756–772.

Zhao GH, Mo XH, Zou FC *et al.* (2009) Genetic variability among *Schistosoma japonicum* isolates from different endemic regions in China revealed by sequences of three mitochondrial DNA genes. *Veterinary Parasitology*.

7. List of papers

Paper 1

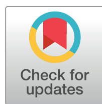
RESEARCH ARTICLE

Discordant population histories of host and its parasite: A role for ecological permeability of extreme environment?

Dagmar Jirsová[1,2]*, Jan Štefka[1,2], Miloslav Jirků[2]

1 Faculty of Science, University of South Bohemia, Branišovská, České Budějovice, Czech Republic,
2 Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská, České Budějovice, Czech Republic

* dasa.jirsova@gmail.com



Check for updates

Abstract

Biogeographical and ecological barriers strongly affect the course of micro-evolutionary processes in free living organisms. Here we assess the impact of a recently emerged barrier on populations of limnic fauna. Genetic diversity and population structure in a host-parasite system (*Wenyonia virilis* tapeworm, *Synodontis schall* catfish) are analyzed in the recently divided Turkana and Nile basins. The two basins, were repeatedly connected during the Holocene wet/dry climatic oscillations, following late Pleistocene dessication of the Turkana basin. Mitochondrial DNA sequences for *cytochrome oxidase I* gene (*cox I*) and a whole genome scanning method—amplified fragment length polymorphism (AFLP) were employed. A total of 347 *cox I* sequences (representing 209 haplotypes) and 716 AFLP fragments, as well as 120 *cox I* sequences (20 haplotypes) and 532 AFLP fragments were obtained from parasites and hosts, respectively. Although results indicate that host and parasite populations share some formative traits (bottlenecks, Nilotic origin), their population histories/patterns differ markedly. Mitochondrial analysis revealed that parasite populations evolve significantly faster and show remarkably higher genetic variability. Analyses of both markers confirmed that the parasites undergo lineage fission, forming new clusters specific for either freshwater or saline parts of Lake Turkana. In congruence with the geological history, these clusters apparently indicate multiple colonisations of Lake Turkana from the Nile. In contrast, the host population pattern indicates fusion of different colonisation waves. Although fish host populations remain connected, saline habitats in Lake Turkana (absent in the Nile), apparently pose a barrier to the gene flow in the parasite, possibly due to its multi-host lifecycle, which involves freshwater annelids. Despite partially corroborating mitochondrial results, AFLP data was not sufficiently informative for analyzing populations with recently mixed biogeographic histories.

OPEN ACCESS

Citation: Jirsová D, Štefka J, Jirků M (2017) Discordant population histories of host and its parasite: A role for ecological permeability of extreme environment? PLoS ONE 12(4): e0175286. https://doi.org/10.1371/journal.pone.0175286

Editor: Heike Lutermann, University of Pretoria, SOUTH AFRICA

Received: July 12, 2016

Accepted: March 23, 2017

Published: April 10, 2017

Copyright: © 2017 Jirsová et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data are available from the Dryad database doi:10.5061/dryad.jq205 for researchers who meet the criteria for access to confidential data.

Funding: The authors acknowledge the Czech Science Foundation (P505/12/G112), the Institute of Parasitology (RVO: 60077344) and the Grant Agency of University of South Bohemia (037/2011/P, 04-135/2010/P). Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, under the

programme "Projects of Large Research, Development, and Innovations Infrastructures". The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The cyclic Pleistocene climatic oscillations severely affected all biomes. Alternating hot-humid and cool-dry periods lasting for tens of millennia caused repeated contraction/expansion and/or fragmentation of vegetational belts and hydrological systems [1]. Freshwater habitats and taxa were among the most severely affected [2–4]. Contractions/fragmentations of hydrological systems during the cold-dry periods resulted in patchy distribution of limnic biota, which remained restricted to areas with continuous freshwater habitats. In the tropics such humid refugia allowed localized persistence and accumulation of both aquatic and terrestrial taxa during the critical cold-dry periods and these are recognized as biodiversity and/or endemism centres analogous to the Pleistocene refugia of the Holarctic [4,5]. During the cold-dry periods, desiccation, shrinkage and fragmentation affected all major drainages and lakes worldwide, making some of them isolated and endorheic and consequently, in arid areas, subject to salinization. Yet, the significance of African Pleistocene refugia for speciation processes remains unclear, since many of the radiations associated with them are of Miocene or older ages (African vertebrates: phylogenetic evidence e.g. [6–8]; indirect paleontological evidence e.g. [9,10]) and because there is only limited information available on millennial-scale genetic processes in metazoans [11,12].

Lundberg and colleagues [13], further supported by [14], proposed a 'hydrogeological hypothesis' based on the finding that Neotropical fish diversity might be a result of geologically and climatically controlled palaeohydrological changes. Recently, similar relationships between phylogenetic history and hydrogeological changes have been revealed in two Afrotropical fish genera, an alestid *Hydrocynus* and a mochokid *Synodontis* [8,15,16]. The latter two studies highlighted the significance of two geological structures, the Central African shear zone and the East African rift system, as the drivers of diversification and subsequent allopatric speciation.

Mechanisms of allopatric speciation by geology- and climate-driven isolation/fragmentation of drainages are well explained. On the other hand, the effects of isolation-related hydrological changes on limnic biotas could be more complex. Despite the obvious complementarity of lake isolation and salinization observed in numerous contemporary lakes in arid regions (in Africa for example lakes Chad, Natron and Turkana), the significance of salinity itself for the intraspecific diversification and/or speciation remains virtually unknown. As a result, allopatric speciation mostly serves as a universal explanation of diversity and biogeography of African freshwater organisms. One of the few exceptions is the well-known phenomenon of sympatric speciation in African cichlids [12,17].

The present study provides an analysis of evolutionary processes on a millennial scale, 5 thousand years ago (kya) following the Holocene split of initially continual populations of fish and their parasites in geo-historically and biogeographically allied Nile and Turkana basins, following the Late Pleistocene (near) dessication of Lake Turkana and its reestablishment and expansion during the African humid period [18,19]. The main aim of the study is to show that metazoan evolutionary processes can be detected within short time periods (millennia) on a population genetic level [20,21]. Based on this evidence we predict to see an emerging population structure between the two recently divided units (Nile and Turkana). In addition, we test the hypothesis, that apart from the well-understood effects of isolation (by fragmentation, distance etc.), salinization of limnic systems also poses a significant evolutionary force that can shape population structure on short spatial scales and might facilitate diversification in otherwise strictly freshwater organisms [22]. These hypotheses are assessed by testing two assumptions: a) the Nile represents a source/ancestral biogeographic unit from which Lake Turkana is derived, and b) there is a population genetic signal mirroring a recent (African humid period

~11–5 kya) establishment and subsequent five millennia of isolation of extant biota of Lake Turkana, and/or its marked environmental gradient of salinity. In order to deepen the knowledge on the degree of genetic changes taking place on millennial time scales, this study addresses the following partial aims: i) evaluation of intra- and inter-population genetic variability of hosts and parasites within and between the Nile and Turkana basins and assessment of factors potentially responsible for the observed patterns, e.g. selection pressures, bottlenecks, salinity; (ii) evaluation, whether or not the parasite and/or host population genetic patterns indicate on-going evolutionary processes potentially leading to speciation; (iii) comparison of the rate of evolution between host and parasite in connection with differences in the depth of population structure.

Methods

Geographic model

We proposed the following geographic model consisting of two (bio)geographic sub-systems: a non-isolated and fully freshwater basin, and an isolated endorheic basin with marked salinity gradient. Our criteria were specified for the two geographic sub-systems: i) absence of aquatic link; ii) model organisms common in both subsystems; iii) clinal salinity gradient present in the isolated sub-system; iv) isolated sub-system of sufficient size to omit size-dependent bias; v) recent disconnection of the two sub-systems.

Approximately 12.5 kya, the Victoria Nile, an endorheic lake until then, and the White Nile system merged and became part of the main Nile system to form its contemporary extent [23]. Lake Turkana is the largest desert lake on Earth measuring ~240 by 13–50 km, with surface area 6,750 km^2, volume 203 km^3 and average/max. depth 31/114 m. In the Late Pleistocene 35–15 kya, Lake Turkana almost completely dried out. During the Holocene, the peak of the African humid period was 11.5–10.5 kya, followed by brief periods (9.0, 6.6 and 5.2 kya), when the lake spilled over the Lotikipi Plain to the north-west, overflowed into White Nile [19] and became a part of Nile-Chad-Niger limnic system (Fig 1). During the drier, lower level periods, similar to the present, the Nile-Turkana link disappeared and the Turkana basin became endorheic [24–26]. The near desiccation of Lake Turkana during the Late Pleistocene, followed by temporary connection with the Nile is the reason for the recent origin of its limnic biota, the faunistic similarity to the Nile and relatively low levels of endemism (~60 fish spp. including 10 endemics, [27]). Currently, the arid area separating the Nile and Turkana basins lacks permanent aquatic environments.

There is no outlet and with reduced inflows and high evaporation the chloro-carbonate alkaline water is subject to significant water level fluctuations (1–3 m annually, $>$ 20 m during 20th century). The lake is becoming increasingly saline along the Omo River delta (0.07) in the north, to highly saline (brackish) towards the south (Kalokol 1.78, El-Molo 1.89) (original unpublished data, values obtained after the peak of rainy season 2–26 September 2008). The Omo delta harbours rich Nilotic biota, whereas the major saline part of Turkana possesses depleted communities [27].

Model organisms

Model organisms were selected according to their wide distribution, high abundance and reliable availability in both basins and at all localities. Additional criteria for the parasite model included narrow host specificity, complex life cycles (more likely to be influenced by salinity) and well resolved taxonomy.

The endemic African catfish genus *Synodontis* Cuvier (Siluriformes: Mochokidae) comprises ~120 nominal species [30]. Most species are freshwater, benthopelagic, potamodromous

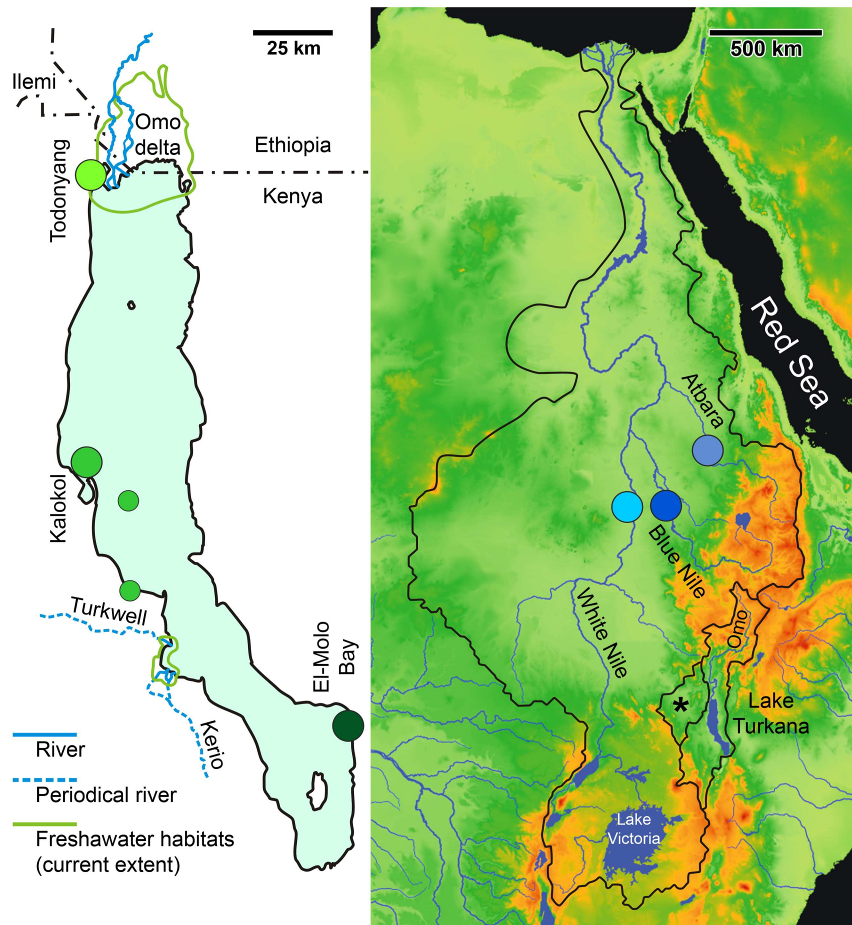


Fig 1. Maps of Lake Turkana (left) and the Nile basin (right) showing the locations of our sampling localities and the outlines of the separate basins (based on [28,29] and original observations). The asterisk indicates the current endorheic Lotikipi Plain, which is being temporarily flooded during exceptionally rainy years. During the wet Pleistocene climatic phases and the African humid period of the Holocene, the Lotikipi Plain repeatedly posed an aquatic link between the Nile and Turkana basins.

https://doi.org/10.1371/journal.pone.0175286.g001

and omnivorous, feeding on detritus, algae, plants, invertebrates, small fish and carrion. The life history is not well known, albeit *Synodontis* spp. can mature sexually before one year of age (M. Jirků & R. Blazek, pers. obs.). *Synodontis schall*, the only representative of the genus in Lake Turkana, was pre-selected as a model host due to its high abundance in both basins and association with several host-specific parasite taxa. *Synodontis nigrita* was selected as an outgroup due to its clear distinctiveness from its congeners and its availability in the study area.

The endemic African tapeworm genus *Wenyonia* Woodland, 1923 (Caryophyllidea: Caryophyllaeidae) comprises six species exclusively parasitizing *Synodontis* hosts. Our model, *Wenyonia virilis* Woodland, 1923, is the most widespread and abundant representative and unlike its congeners, present in both basins [31]. The life cycle of *Wenyonia* is unknown but probably similar to other fish caryophyllideans, i.e. larval development in freshwater oligochaetes, no second intermediate host, maturation in the fish definitive host. In the related genus *Khawia*, maturation in the fish takes 1.5–6 months [32,33]. Adult longevity is also unknown but probably less than a year.

Table 1. Sampling localities, coordinates and sample sizes.

Locality	Coordinates	Wenyonia	Synodontis
Lake Turkana			
El-Molo Bay—saline	2.832222 N, 36.695833 E	110	30
Kalokol—medium salinity	3.558611 N, 35.915833 E	1	4
Kerio River delta—medium salinity	2.974813 N, 36.173469 E	0	6
Central Island—medium salinity	3.495833 N, 36.040278 E	0	1
Todonyang, Omo delta—feshwater	4.451667 N, 35.94388 E	170	50
Nile River			
White Nile in Kostí	13.172222 N, 32.672222 E	75	26
Blue Nile in Sennar	13.543611 N, 33.636667 E	7	21
Atbarah River in Khashm el-Girba	14.919444 N, 35.901944 E	5	0
Totals (n)		368	138

https://doi.org/10.1371/journal.pone.0175286.t001

Specimen collection and DNA extraction

Parasites and fish hosts were collected from Lake Turkana and the Nile at eight localities (Table 1). The Turkana localities were chosen to cover the salinity gradient from freshwater to saline environments. Fish examined were obtained in fish markets and/or from fishermen under supervision of Kenya Marine and Fisheries Research Institute authorities. Since fish were destined for human consumption, no permission was needed to collect these fish. Fish that were still alive when examined were killed by dorsal pithing (spinal cord and blood vessels cut immediately behind the head), a method compliant with Kenyan, Czech and European legislation. All fish were euthanized and processed for dissection shortly after purchase, no further manipulation was involved.

A small piece of tissue from the worm and host liver was cut off, washed extensively in physiological saline solution and stored in 96% ethanol. The remaining part of each cestode was fixed with hot 4% buffered formalin and stained with Schuberg's hydrochloric carmine solution [34]. Vouchers were deposited in the helminthological collection of the Institute of Parasitology BC-CAS, České Budějovice, Czech Republic (*W*. *virilis*, IPCAS C-503; *W*. *minuta*, IPCAS C-571; *W*. *youdeoweii*, IPCAS C-573). Genomic DNA was extracted using a standard phenol chloroform extraction method [35].

PCR

Specific primers, Wen-cox F3 (`5′ AGAGAGCGGTTACTGCTAATAA 3′`) and Wen-cox R3 (`5′ ATAATGAAAGTGCGCTACTACAAATCA 3′`), were designed using sequences of related taxa from GenBank for amplifying and sequencing a partial region (~ 1100 bp) of cytochrome oxidase I gene (*cox I*) in *Wenyonia*. A PPP Master Mix (Top Bio) was used to prepare PCR reactions containing 12.5 μl of the PPP Master mix, 10 μl of ultra pure water, 5 pM of each PCR primer and 1.5 μl of isolated DNA (50 to 150ng). Cycling conditions were as follows: denaturation for 5min at 94˚C, followed by 30 cycles of 30s at 94˚C, 35s at 50˚C, 1min 45s at 72˚C, a final extension at 72˚C for 10min. All products were verified on a 1% agarose gel and purified using exonuclease I (Exo I) and shrimp alkaline phosphatase (SAP) [36]. Purified DNA was sequenced by Macrogen Inc. (Amsterdam, Netherlands) on a 3730 DNA Analyzer.

In fish, the same mtDNA locus (*cox I*) was selected for a comparison of genetic diversity with the parasites. Primers for amplification of partial *cox I* sequence (~800 bp), FishF2 and VF2_t1, were adopted from [37]. The PCR mixture contained 12.5 μl of the PPP Master mix, 10.5 μl of ultra pure water, 5 pM of each PCR primer and 0.5 μl of isolated DNA (100 to

170ng). Cycling conditions were as follows: denaturation for 2 min at 94˚C, followed by 35 cycles of 30s 94˚C, 30s at 54˚C, 1min at 72˚C, and completed by 10min at 72˚C (see 37). All products were verified on an agarose gel, purified and sequenced using PCR primers as in *Wenyonia*.

Amplified fragment length polymorphism

Restriction and ligation of genomic DNA was performed using AFLP Core Reagent Kit (Invitrogen) and following the manufacturer's manual. To reduce the number of fragments in the mixture, ligated DNA fragments were amplified by PCR (Pre-selective PCR) using primers complementary to the adapter and restriction site sequence extended with an additional selective nucleotide at their 3'—ends (EcoRI-A and MseI-C) [38]. The PCR mixture consisted of 2 μl of restriction-ligation product, 2.5 μl 5x MangoTaq™ Coloured Reaction Buffer, 0.75 μl $MgCl_2$ (50mM), 0.625 μl of EcoRI-A and MseI-C primers (10pmol), 0.1 μl BSA (1mg/μl), 1.25 μl dNTPs (8mM), 4.6 μl ultra pure water and 0.05 μl MangoTaq™ DNA Polymerase (Bioline) (5U/μl). Cycling conditions for preselective-PCR were as follows: denaturation for 1min at 95˚C, followed by 45 cycles of 15s at 95˚C, 15s at 55˚C, 30s at 72˚C; and completed by 7min at 72˚C. Fragments were verified on a 1% High Resolution Agarose gel (Invitrogen). This step was followed by selective amplification with primers containing 3 selective nucleotides at their 3'—end. Forward primers were labeled with different fluorescent colours on their 5'—end (EcoRI-ANN* and MseI-CNN) [38] to allow multiplexing in fragment analysis. Selective PCR consisted of 2μl of Pre-selective PCR product diluted by ultra pure water 1:10, 2.5 μl of 5x MangoTaq Buffer—coloured, 0.5 μl $MgCl_2$ (50mM), 0.05 μl of EcoRI-ANN* fluorescent labeled primer (10pmol), 0.07 μl of MseI-CNN (10pmol), 0.1 μl of BSA (1mg/μl), 0.65 μl of dNTPs (8mM), 6.52 μl of MQ water and 0.1 μl of MangoTaq DNA Polymerase (Bioline) (5U/μl). Cycling conditions for selective PCR were as follows: denaturation for 2min 94˚C, followed by 10 cycles of 20s at 94˚C, 30s at 66˚C (decreasing by 1˚C every cycle for 10 cycles), 2min at 72˚C; then 20 cycles of 20s 94˚C, 30s at 55˚C, 2min at 72˚C and completed by 15min at 60˚C. Amplified fragments were checked on 1% High Resolution Agarose gel (Invitrogen) and analyzed on ABI3730XL sequencer by Macrogen Inc. (Seoul, South Korea). Primer combinations (Table 2) were chosen according to available literature and published results for both *Wenyonia* [39–42] and for *Synodontis* [12,43,44].

Phylogenetic analysis

Sequences were assembled, inspected for errors and aligned using GENEIOUS Pro software package version 6.1, software MAFFT [45,46]. Due to low variability of fish sequences, genetic distances in the host dataset varied between 0 and 1.5%. Hence, phylogenetic trees were reconstructed only for the parasite dataset. Suitable models of molecular evolution were selected

Table 2. Primer combinations used for AFLP.

***Synodontis* spp.**		***Wenyonia* spp.**	
EcoRI-ACA*	MseI-CTC	EcoRI-ACA*	MseI-CTC
EcoRI-ACT*	MseI-CTA	EcoRI-ACT*	MseI-CTA
EcoRI-ACG*	MseI-CAA	EcoRI-ACG*	MseI-CAA
-	-	-	MseI-CAG
-	-	-	MseI-CAC

Asterisk indicates fluorescent-labelled primers.

https://doi.org/10.1371/journal.pone.0175286.t002

using jMODELTEST 0.1.1. [47]. The model with the best likelihood was chosen using AIC criteria and phylogenetic trees were reconstructed using Bayesian inference (BI) in the program MrBayes v. 3.2.2. [48,49]. MrBAYES analyses were run for 15 million MCMC generations, with four chains and three independent runs and were performed under the HKY+G model of molecular evolution. Convergence was checked in AWTY [50]. Maximum-likelihood (ML) phylogeny was generated using the PhyML 3.0 software [51] under the HKY+G search parameters [52]. Initial trees were generated by an improved version of the neighbor-joining algorithm (BIONJ) [53]. Reliability of branching patterns within trees was tested by the bootstrap method with 1000 resamplings

As suggested for the coding genes (e.g. [54]), we assessed best-fit partitioning schemes for cox I using PartitionFinder [55]. Optimal partitioning schemes were set for MrBayes to perform BI runs and for all available models to run ML topologies. The best models were selected according to the AIC criteria as specified in Table 3.

BI analyses were performed by software MrBAYES, run for 20 million MCMC generations, with four chains, three independent runs and were performed under the selected partition models; the coherence of each run was checked using software TRACER v1.5 [56]. ML analyses were carried out in the program GARLI 2.0 [57] under the selected substitution models and with automatic termination conditions.

All phylogenetic topologies (BI and ML) of *Wenyonia* dataset were rooted using sequences of *W*. *minuta*, representing the closest available sequence from Caryophyllidae.

Haplotype networks

To examine the evolutionary relationships among haplotypes in populations of each parasite and host species statistical parsimony networks based on pairwise differences were constructed using PopART v1.7 [58].

Population genetics statistics - mtDNA

To characterize the diversity of populations and their demographic history we performed several population genetic statistics for parasite and host mitochondrial datasets. Statistics of nucleotide genetic diversity (*Pi*), Haplotype diversity (*Hd*), Tajima's test (*D*) and Fu & Li's test (*D*) were calculated using DNASP v5 [59]. The statistical significance of *D* values were tested by 10,000 coalescent simulations [59].

Gene flow and migration

The direction of historical gene flow was explored using coalescent based software MIGRATE v3.2.16 [60]. Different pathways of migration and colonization were tested to identify the ancestral population and to assess the direction of gene flow. Twelve and ten alternative migration hypotheses were tested for the host and parasite dataset, respectively (Fig 2a and 2b). Although representing similar scenarios, the hypotheses for hosts and parasites are not entirely identical. Only two parasite individuals were obtained from the medium salinity area in Lake

Table 3. Substitution models for nucleotide data partitions of the *Wenyonia* spp. coxI dataset selected using the AIC in PartitionFinder for BI and ML runs.

Codon position	MrBAYES models	All models
1st	F81+I+G	F81+I+G
2nd	GTR+G	TrN+G
3rd	GTR+I+G	TrN+I+G

https://doi.org/10.1371/journal.pone.0175286.t003

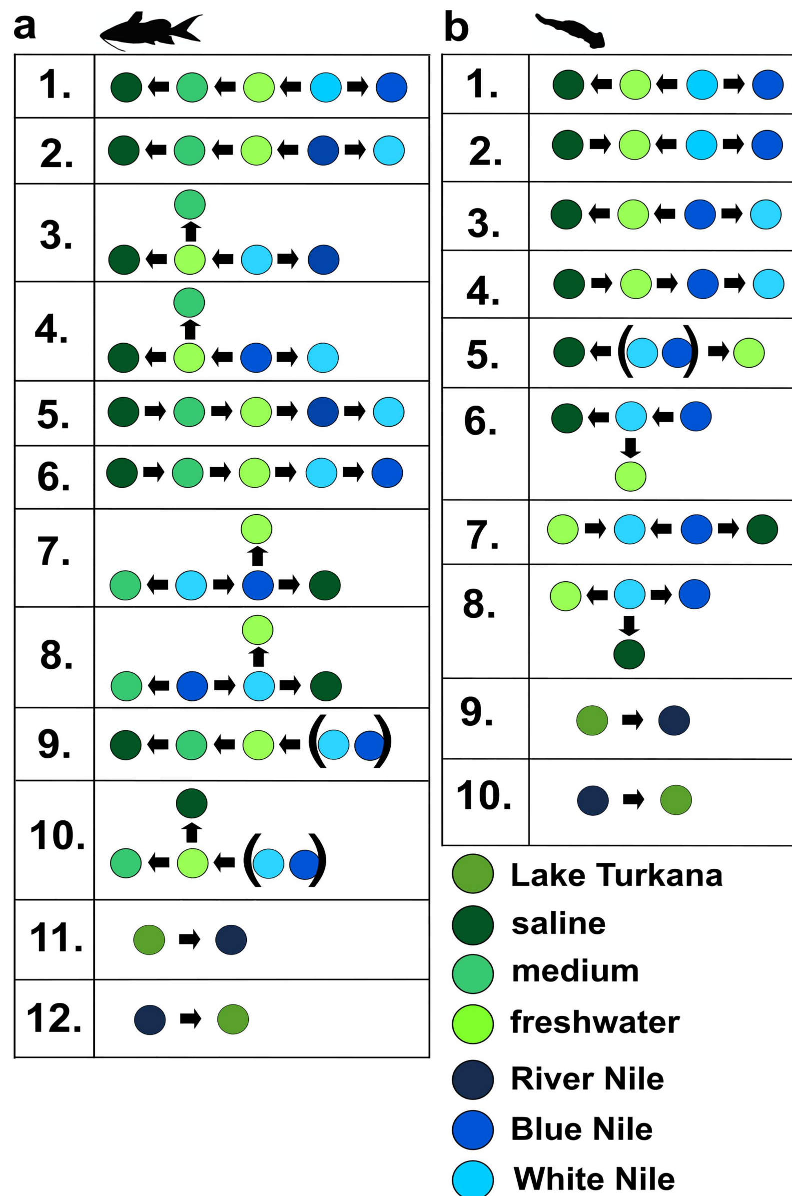


Fig 2. Migration hypotheses tested by software MIGRATE, datasets represented here are a) host; b) parasite. Arrows depict direction of historical gene flow.

https://doi.org/10.1371/journal.pone.0175286.g002

Turkana (Kalokol) and thus this locality was excluded from the analysis of the parasite dataset. In the host dataset, hypotheses 1–10 represent scenarios of gene flow between individual localities, whereas hypotheses 11 and 12 represent general patterns of gene flow between Lake Turkana and the Nile. The same approach was adopted for the *W*. *virilis* dataset (see Fig 2b). Mitochondrial datasets were analyzed under HKY+G Model in the Bayesian mode of the program. Migrate runs were initiated using the following values of θ (theta) (minimum 0.0, maximum 1.0 and δ = 0.1) and M (migration) (minimum 0.0, maximum 1000.0 and δ = 100). Heating was set for five Multiple Markov Chains (1.0, 1.5, 3.0, 10000.0, 100000.0) and a swapping interval between chains of 1. The performance of the MCMC processes was checked for coherence and sufficient effective sample sizes (ESS) in Tracer v1.5 [56]. According to the ESS

values obtained from preliminary runs, 90 million and 200 million generations were selected for *Wenyonia* and *Synodontis* datasets final analyses. Marginal likelihoods of different migration hypotheses were compared, and the best performing hypothesis was selected using Bayes Factors [61,62].

AFLP data analysis

Analysis of the AFLP fragment data was performed using the GeneMapper Software v3.7 (Applied Biosystems). The thresholds were set for each primer combination (S1 Table). Each threshold was set as a 5% value of the highest peak fluorescence intensity. Selected peaks were manually checked for presence of any 'noise'. All peaks above the threshold limit were exported, obtaining 6038 and 2169 loci for the parasite and host datasets, respectively. The output was transformed to a binary format using MICROSOFT EXCEL 2011 and then converted to GenePop file format using GenAlEx v6.5 [63]. Selectively neutral alleles suitable for analysis of population structure were filtered from loci under putative selection using MCHEZA [64]. The initial analysis was performed with 100 000 simulations and with default settings for confidence intervals, Θ, β-a, β-b and critical frequency. Candidates included for balancing and positive selection were removed from the dataset and the same simulation was run again. This procedure resulted in 716 loci for the *Wenyonia* dataset and 532 loci for *Synodontis* dataset. Principal Coordinate Analysis (PCoA), which was based on Nei's genetic distances of markers, was computed in GenAlEx to visualize any emerging patterns of genetic structure in populations. To quantify the level of genetic variation between populations explained by geography were carried out in GenAlEx using AMOVA. Calculations were performed for host and parasite data from the Turkana freshwater and saline localities. The significance of the estimations was obtained with 999 permutations of the data. To estimate the number of population clusters or groups (K) in the dataset a Bayesian analysis was carried out using the software STRUCTURE v2.3.3 [65]. STRUCTURE was run using 15 independent calculations for each K, with values of K from one to ten for *Wenyonia* and from one to nine for *Synodontis*. The number of MCMC generations and burn in for both datasets was 3 million and 300,000, respectively. Runs were performed for both models of population history (admixture and noadmixture). Mean value of posterior probabilities, L (K), was determined for each K value. Since L(K) did not reveal a sharp peak, the rate of change between posterior probabilities of successive runs (delta K statistics, [66]) was calculated to allow estimation of the optimal value of K. Multiple runs generated by STRUCTURE were analyzed using the cluster matching and permutation program CLUMPP v1.1.2 [67] and graphically displayed as bar plots in Distruct [68].

Results

A total of 347 sequences (990 bp) representing 209 haplotypes were obtained from *Wenyonia* spp. as follows: *W. virlis*—148 samples from the Turkana-freshwater site, 102 from the Turkana-saline site, 50 from the White Nile, 7 samples from the Blue Nile, 1 sample from the Atbarah River; 4 samples of *W. minuta*; 13 samples of *W. youdeoweii*. For *Synodontis* spp.,120 sequences (604 bp) representing 20 haplotypes were obtained: 32 samples from the Turkana-saline site (*W. virilis* prevalence 43.75%), 39 from the Turkana-freshwater site (prevalence 17.94%), 10 from the Turkana-medium site (prevalence 20%), 14 samples from the White Nile (prevalence 61.11%), 21 samples from the Blue Nile (prevalence 17.24%); 6 samples of *S. nigrita* (*Wenyonia* absent).

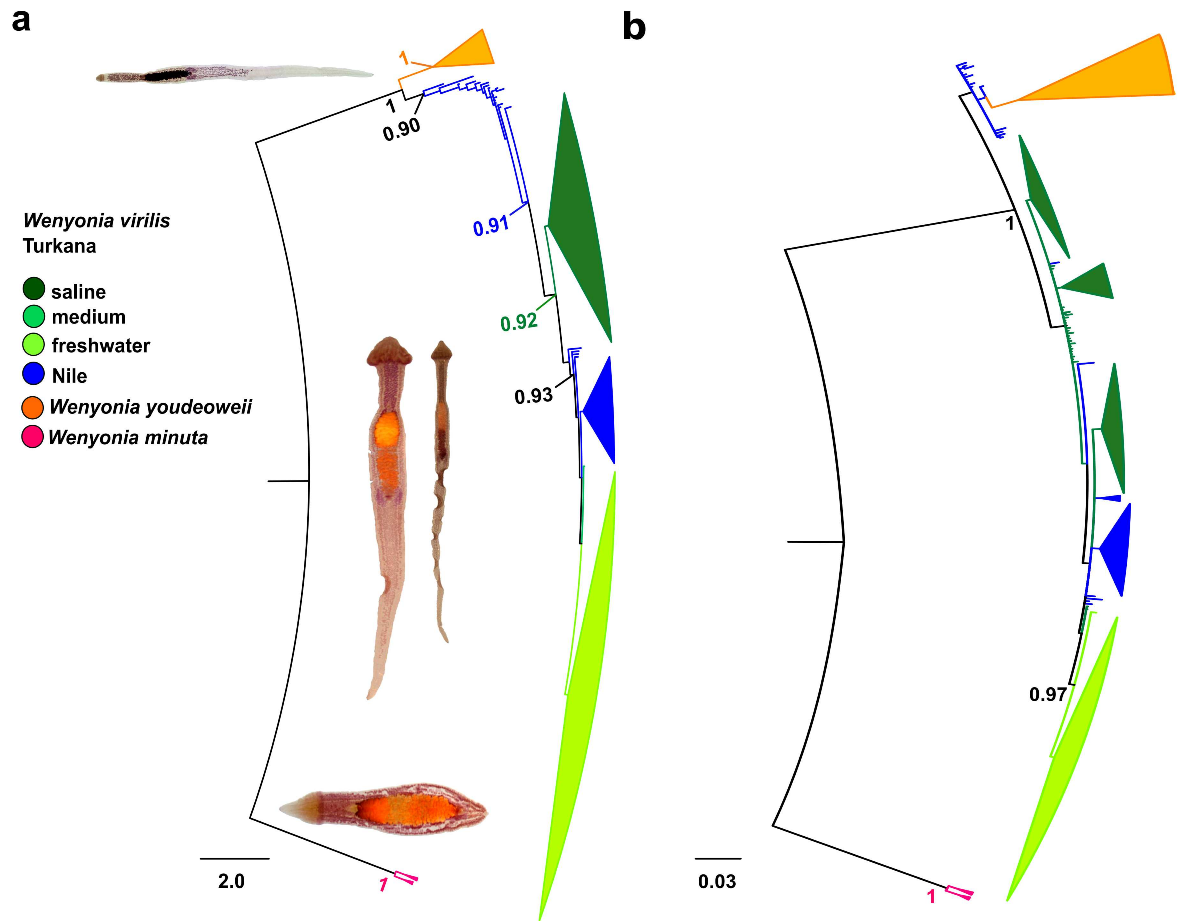


Fig 3. Phylogenetic tree of the *cox I Wenyonia* spp. computed by MrBayes for a) the model of molecular evolution; b) the best-fit partitioning schemes. Branches are collapsed into groups, which correspond to sampling locations. Statistical support for each group was obtained from the Bayesian posterior probability.

https://doi.org/10.1371/journal.pone.0175286.g003

Phylogenetic analysis

The results for analyses performed under the model of molecular substitution (MMS) and partition schemes (PS) differed in the position of *W*. *youdeoweii* in obtained phylogenies (Fig 3). In MMS, *W*. *youdeoweii* formed a clade separate from *W*. *virils* and *W*. *minuta*, whereas under PS settings, *W*. *youdeoweii* clustered within *W*. *virilis*.

Regardless of the differences in position of *W*. *youdeoweii*, the information obtained for *W*. *virilis* was similar in both analyses. Samples of *W*. *virilis* formed four groups. The resulting groups almost exclusively contained samples from either the Nile or the Lake Turkana. Neither of the basins formed mutually monophyletic clades, the Turkana groups were embedded in the Nilotic clades. The most basal *W*. *virilis* clades comprised exclusively samples from the Nile. The group derived from them was formed by individuals from the saline part of Lake Turkana. A second Nile cluster was located in the centre of the tree crown. The phylogenetically most derived group consists of samples from the freshwater part of Lake Turkana. This result supports the hypothesis that the Nile represents the ancestral population and the Turkana basin was subjected either to repeated colonization, or to bottlenecks reducing ancestral diversity of the lineages.

The unstable position of *W*. *youdeoweii* was also revealed by ML analyses (S1 Fig). In the ML analysis run under MMS, *W*. *youdeoweii* clustered inside *W*. *virilis* clade, while under PS, *W*. *youdeoweii* formed a separate clade sister to the *W*. *virilis* cluster. Regardless of the conflicting results in both types of analyses, the informative value of the ML results was less reliable due to very low bootstrap support.

Haplotype networks

Specimens of *W*. *minuta* and *W*. *youdeoweii* were clearly separated from those of *W*. *virilis* (Fig 4a). Apart from the remarkably high number of different haplotypes, i.e. the high ratio between the number of haplotypes and the number of individuals (194:347), the haplotype network revealed a considerable complexity of the haplotype grouping (Fig 4a). Individuals of *W*. *virilis* fell into four main groups. Samples from the freshwater Turkana locality exhibited a radial branching pattern with the main haplotype surrounded by satellite low frequency haplotypes—a trait typical for young expanding populations. In contrast, the populations from the Nile contained (with one exception) only haplotypes organized into networks separated by multiple mutations without a star-like patterns—a pattern typical for older populations in equilibrium (Fig 4a groups 2 and 3). Interestingly, groups 1 and 4 from Lake Turkana showed strong affinity either to the freshwater, or the saline part of the lake. Moreover, one Lake Turkana freshwater haplotype also occurred in the saline part of the lake, but not vice versa. Samples from the Nile fell into two groups (2 and 3), but four individuals were also placed in group 4 otherwise containing only samples from the saline part of Lake Turkana.

A total of 120 sequences representing 20 haplotypes of *Synodontis* spp. were used to create the haplotype network (Fig 4b). Six individuals of *S*. *nigrita* produced a separate cluster with four haplotypes, whilst the rest of the 114 sequences of *S*. *schall* from the Nile and Lake Turkana clustered together. The two networks were separated by 55 mutations. In contrast to the genetically diverse parasite, the host network was distinctly less variable. The central *Synodontis* haplotype in which the majority of specimens clustered was represented at all localities in both basins, only two haplotypes that were separate from the central haplotype by more than two mutations were found (Fig 4b).

Population genetic analysis of mtDNA sequences

The differences between the host and parasite populations seen in their haplotype networks were also reflected in their population genetic statistics (Table 4). Results of the DNA polymorphism statistics for *W*. *virilis* showed high values of haplotype and nucleotide diversity and small differences between the Turkana and Nile basin. In *Synodontis*, values of all statistics were considerably lower for Turkana populations and both Tajima's and Fu and Li's tests showed significant negative values. The results indicate that the Turkana populations are genetically smaller and expanding, which is typical for populations subjected to a recent bottleneck.

Gene flow and migration

The likelihood scores and probabilities of the hypotheses tested using MIGRATE software are shown in Table 5. Despite the fact that one of the analyses of the *Synodontis* dataset favoured the hypotheses assuming a gene flow from Lake Turkana (hypothesis 11), an analysis of with the dataset split into individual localities revealed that the Lake Turkana population originated in the White Nile (hypothesis 3, Table 5). In contrast, for *Wenyonia* spp. the direction of gene flow was the same for both analyses. The results supported the hypothesis of gene flow from the Nile to Lake Turkana (hypothesis 10) and independent colonization of the Blue Nile as

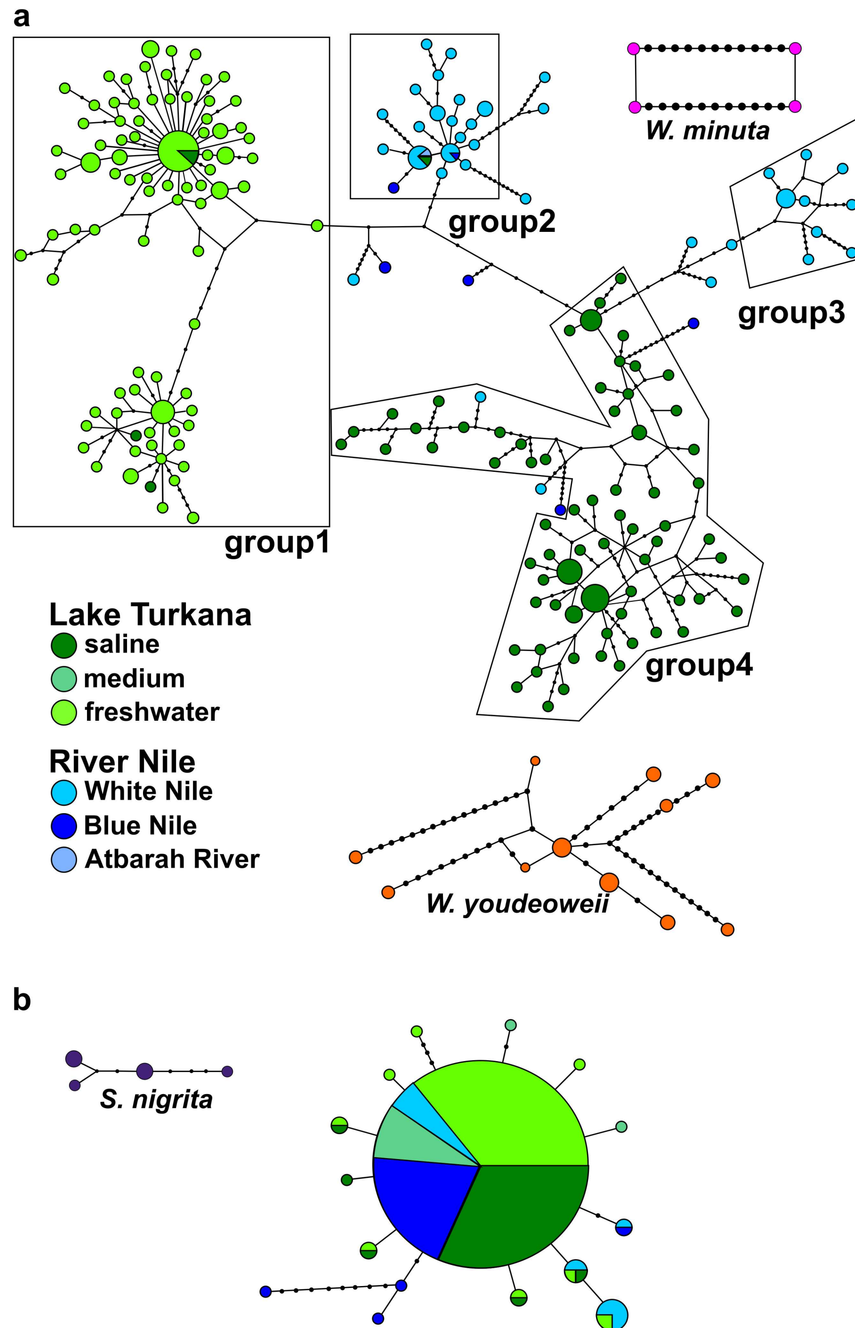


Fig 4. Haplotype network for the *cox I Wenyonia* spp. dataset (a) and for the *cox I Synodontis* spp. dataset (b) computed by PopART v1.7. Most groups were separated from each other by a maximum of six mutations. The exception was group 4, which was separated by ten mutations from group2 and eleven mutations from group3. Samples of *S. nigrita* clustered together and created a separate network. The sizes of haplotypic nodes are relative to the sample size.

https://doi.org/10.1371/journal.pone.0175286.g004

well as of the freshwater and saline areas of Lake Turkana from the White Nile (hypothesis 8). Similarly, according to the gene flow and migration analyses of *Synodontis*, the ancestral population was located in White Nile from where the colonization wave spread to the Blue Nile and Lake Turkana (hypothesis 3). Within Lake Turkana, the freshwater area was colonized first,

Table 4. Results of population genetic statistics of mtDNA for both parasites and hosts.

	Wenyonia virilis		*Synodontis* spp.	
	Nile	Turkana	Nile	Turkana
Haplotype diversity	Hd = 0.98	Hd = 0.986	Hd = 0.631	Hd = 0.389
Nucleotide diversity	π = 0.016	π = 0.015	π = 0.004	π = 0.001
Tajima's D	D = -1.211*	D = -1.271*	D = -1.37	D = -2.40*
Fu and Li's D	D = -3.123	D = -4.56	D = -1.73	D = -4.00*

Asterisk indicates values significant at P<0.05.

https://doi.org/10.1371/journal.pone.0175286.t004

while saline and medium salinity areas were populated separately. However, this colonization sequence is not completely testable for parasites due to their low prevalence in the medium salinity part of Lake Turkana (host n = 23, only 2 infected).

Population genetic structure—AFLP

AFLP data were obtained for 368 parasites using 15 primer combinations and for 138 hosts using 9 primer combinations (Table 2). The final dataset contained 716 and 532 fragments for parasite and host, respectively. The analysis performed for parasites (Fig 5a and 5b) produced a pattern indicating possible formation of multiple clusters. Surprisingly, all three tapeworm species clustered together, with an apparent tendency towards segregation of most of the individuals from the saline part of Lake Turkana and also for separation of Lake Turkana from the Nile. A separate, smaller cluster containing exclusively Turkana samples was clearly subdivided into two groups, according to their origin either in the freshwater, or saline sites. Note that the larger cluster comprises samples from all localities, but also contains *W*. *youdeoweii* and *W*. *minuta* samples, which did not create separate units (compare with mtDNA analysis Figs 3 and 4). The smaller cluster contained exclusively samples from Lake Turkana (bottom left). The cluster is clearly subdivided into two groups, according to their origin either from the

Table 5. Results of the *Synodontis* spp. and *Wenyonia* spp. migration hypotheses tested with the coalescent based software MIGRATE, showing comparisons of marginal likelihoods and model probabilities for each dataset.

	Synodontis spp.		*Wenyonia* spp.	
	Marginal LH	Model Probability	Marginal LH	Model Probability
H1	-2098.6	$1.55\ e^{-222}$	-26193.9	$7.36\ e^{-277}$
H2	-2551.8	0	-27100.8	0
H3	**-1843.2**	**1**	-32718.3	0
H4	-2455.6	0	-26171.8	$1.26\ e^{-257}$
H5	-1878.0	$5.61\ e^{-31}$	-27224.7	0
H6	-2178.9	$2.71\ e^{-292}$	-26527.2	0
H7	-1879.7	$1.94\ e^{-32}$	-26254.0	0
H8	-2003.1	$1.30\ e^{-139}$	**-25876.0**	**1**
H9	-2201.9	0	-28132.6	0
H10	-2180.9	$4.33\ e^{-294}$	**-26607.9**	**1**
H11	**-1536.4**	**1**	-	-
H12	-1990.3	0	-	-

The best-supported hypotheses are in bold; H1-Hn refer to corresponding hypothesis {Marginal Likelihoods were used to calculate the Bayes factors (model probability) and for comparing alternative models of gene flow}.

https://doi.org/10.1371/journal.pone.0175286.t005

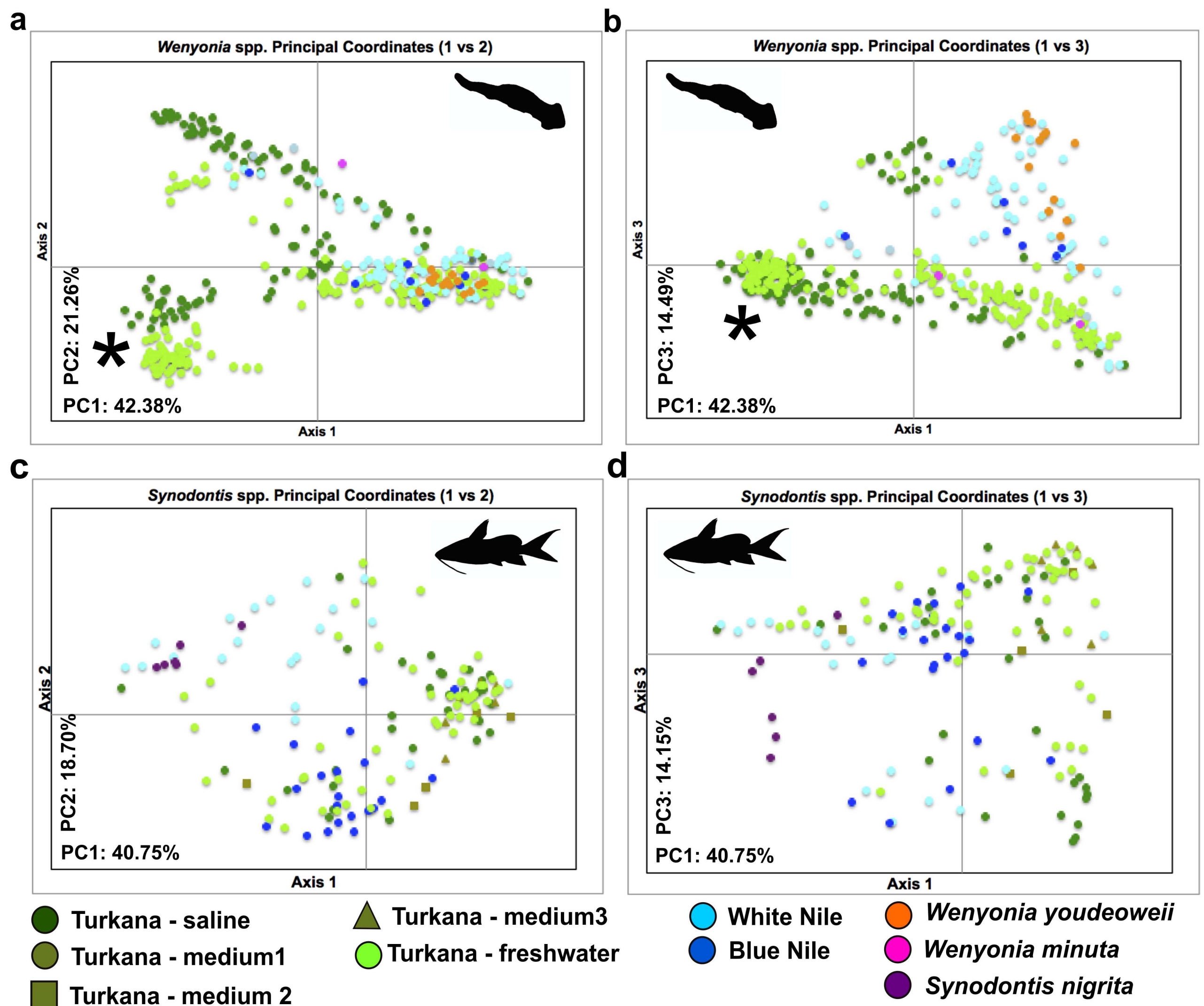


Fig 5. Results of PCoA for (a, b) *Wenyonia* spp. and (c, d) *Synodontis* spp. based on the genetic distances generated by GenAlEx software. Results are presented for the first three axes, because the second and the third axis explained similar levels of variability. In PCoA, all individuals were labelled with different colours based on their geographical or species origin. Asterisks (a, b) indicate clearly separated clusters comprising solely *Wenyonia* Turkana samples. *Synodontis* samples (c, d) from all localities clustered together without forming distinguishable groups. *S. nigrita* individuals clustered together and even showed a tendency to create a separate group (d). Turkana—medium 1–3 refers to three different localities: 1, Central Island; 2, Kerio River delta; 3, Kalokol.

https://doi.org/10.1371/journal.pone.0175286.g005

freshwater, or saline part of the lake. When axis 1 and axis 3 were compared, the majority of the samples from Lake Turkana formed a separated group (Fig 5b—bottom half).

Results of the AFLPs PCoA analysis are different from the picture of the population structure obtained from mtDNA. PCoA clustering of individuals did not corroborate the separation of *W. virilis* specimens into mtDNA haplogroups, but showed some differentiation between the Turkana and Nile populations (Fig 5a and 5b). Moreover, AFLP markers did not separate *W. youdeoweii* and *W. minuta* from *W. virilis* in any of the performed analyses (Figs 5 and 6a). The AMOVA analyses based on F_{ST} calculation (Infinite Alleles Model) revealed that the majority of variation occurred within the sample sites for all datasets (90% for parasite, 93% for host in the whole dataset; 89% for the parasite, 98% for the host in Lake Turkana dataset).

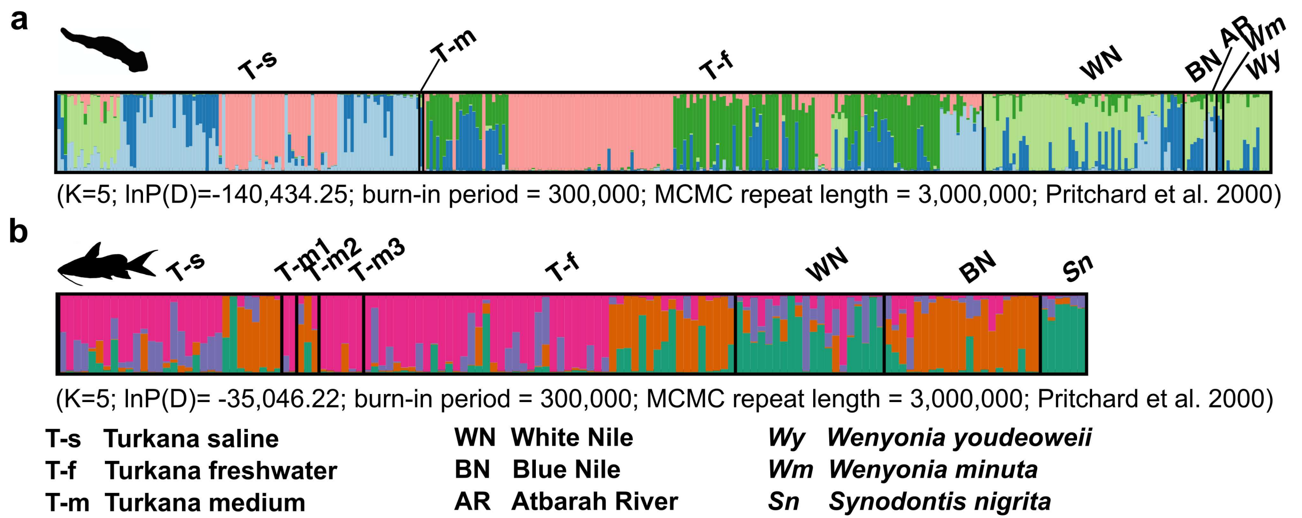


Fig 6. Individual-based cluster representation of all sampled *Wenyonia* spp. (a) and *Synodontis* spp. (b) as revealed by the Bayesian inference of population structure. Each colour represents one assumed population cluster K. Multiple coloured bars display an individual's estimated membership proportion in more than one population (q), i.e. admixture. Turkana—medium 1–3 refers to three different localities: 1, Central Island; 2, Kerio River delta; 3, Kalokol.

https://doi.org/10.1371/journal.pone.0175286.g006

Among sites the population structure (Turkana vs. Nile) explained only 10% of variance for *Wenyonia virilis* and 7% of variation for *Synodontis* spp. Among sites the population structure between the Turkana freshwater and saline part explained 11% variation for *Wenyonia* virilis and 2% for *Synodontis* spp. (Table 6).

Table 6. Results of the AMOVA based on F_{ST} for *Wenyonia virilis* and *Synodontis* spp. AFLP data from Lake Turkana and the Nile; Turkana freshwater and saline part. Estimations are based on 999 permutations and the significant at P<0.001.

Source of varition	*df*	Sum of squares	Mean squares	Estimated variance	Explained variance (%)
			Turkana vs Nile		
			Wenyonia virilis		
Among sample sites	1	1559.575	1559.575	12.753	10%
Within sample sites	349	40106.447	114.918	114.918	90%
Total	350	41666.023		127.672	100%
			Synodontis spp.		
Among sample sites	1	374.602	374.602	5.301	7%
Within sample sites	130	9738.262	74.910	74.910	93%
Total	131	10112.864		80.211	100%
			Turkana freshwater part vs saline part		
			Wenyonia virilis		
Among sample sites	1	1902.670	1902.670	13.434	11%
Within sample sites	278	30110.370	108.311	108.311	89%
Total	279	32013.039		121.744	100%
			Synodontis spp.		
Among sample sites	1	132.709	132.709	1.551	2%
Within sample sites	78	5813.353	74.530	74.530	98%
Total	79	5946.063		76.082	100%

https://doi.org/10.1371/journal.pone.0175286.t006

The PCoA for the host dataset (Fig 5c and 5d) revealed a similar pattern to the haplotype network of *cox I* sequences. The lack of population structure observed in AFLP corresponds with mtDNA data.

Population structure was further analyzed using the STRUCTURE software. The optimal number of genetic clusters (K) was estimated by using the delta K statistics (DK). The number of clusters in the parasite dataset was K = 5 for the admixture model of population history (Fig 6a) and K = 4 for the non-admixture model of population history (see S2 and S3 Figs). As in the PCoA analysis, *W*. *youdeoweii* and *W*. *minuta* were not separated from *W*. *virilis* and shared clusters with samples from the Nile (Fig 6a). The Nile population was characterized by the predominant occurrence of one cluster (lime green), supplemented with two other clusters (light and dark blue). Samples from Lake Turkana shared clusters with the Nile (light blue, dark blue, less frequently lime green), but a unique Turkana cluster was present with high frequency (magenta). Samples from the freshwater part also frequently contained another dark green cluster, which was almost absent in Nile samples. The population from the saline area also contains a unique cluster (magenta). Clusters shared between the saline population and the Nile are different from the freshwater part of Lake Turkana (light blue, lime green, dark blue).

Clusters, which are less frequent in the Nile (light blue, dark green), showed an increased proportional presence in Lake Turkana, concordant with the hypothesis of the Nilotic origin of the Turkana populations. The presence of different Nile clusters in the freshwater and saline parts of Lake Turkana might reflect multiple colonization events. Similar to the PCoA, AFLP clusters obtained in STRUCTURE did not match with specific mtDNA haplogroups. *Wenyonia youdeoweii* and *W*. *minuta* did not separate from *W*. *virilis*.

In the host dataset, results distinguishing the four clusters (admixture population model; K = 5 for no-admixture historical model S3 Fig) received the highest peak in the DK statistics (S2 Fig). All four clusters were present at all localities from the Nile and Lake Turkana albeit with different frequencies (Fig 6b). The cluster with the highest frequency in Lake Turkana (pink) has the lowest frequency in the Nile, whereas the main clusters in the Nile (green and purple) were less frequently observed in Lake Turkana, a pattern similar to that revealed in the parasite. Interestingly, the two Nile branches, Blue and White Nile, differed in the composition of their clusters too. Contrary to the PCoA analysis, *S*. *nigrita* did not show a tendency to cluster separately and shared clusters with *Synodontis* samples from the White Nile.

Discussion

Two different markers, AFLP and mtDNA, were used to assess the levels of genetic diversity, population structure and shared evolutionary patterns in a host-parasite system in two recently divided basins. Evidence of geographically determined events shared between the two organisms was revealed in the form of a historical bottleneck and Nilotic origin of the Lake Turkana populations. Additionally, an indirect impact of salinity gradient on the structure of Lake Turkana parasite populations was indicated.

Fast evolving parasites vs. stationary hosts

The markedly faster rate of molecular evolution reflected by an increased sequence diversity of *W*. *virilis*, when compared to its host possessing similar generation time (one year), is striking. We were not able to provide an exact calculation of the difference in the mutation rates due to the character of our data. Methods used for such estimations rely on clade to clade comparisons (e.g. *BEAST [69]), whereas our fish samples lacked any such structure. Despite that, it is clear that the mutation rates must differ dramatically (sequence variability for *Wenyonia virilis*

dataset is between 100–95.5% vs. 100–97.3% for *Synodontis*). An increased mutation rate in parasites is a common feature [70,71] and a shorter generation time of the parasite is the most common explanation for the discrepancy [72]. However, in our case, organisms with similar generation times were studied. Instead, several other evolutionary factors might be involved. For example, it was suggested that higher migration rates of the host could increase mutation rate in the maladapted populations of the parasite [73,74]. However, multiple factors, such as differences in effective population sizes, species-specific historical events, or a higher life expectancy of the host, may also play a role. Although frequently referred to in the literature [75–77], the phenomenon of the slowly evolving hosts opposed by rapid evolution in their parasites remained surprisingly poorly documented until recently. To our knowledge, the *Wenyonia-Synodontis* model represents the only example documented in metazoan endoparasites, as most other examples include exclusively ectoparasites and Apicomplexa [70,72,78,79]. Hence, when supplemented with more genes and species, our system might provide an alternative model for the studies of host-parasite coevolution.

Opposing population histories of host and its parasite

Both mtDNA analyses showed that tapeworms from the Nile formed two separate groups (Figs 3 and 4a). Each of them was phylogenetically more closely related to one of the two Lake Turkana (sister) groups than to each other. Interestingly, each of the two Lake Turkana groups was almost exclusively restricted either to the freshwater or the saline part of the lake (Fig 4a). This pattern might indicate multiple colonization waves of Lake Turkana from the Nile and the individual haplogroups might be considered to be result of different colonization events. Moreover, the apparent absence of gene flow between the saline and freshwater Turkana haplogroups implies lineage fission. Alternatively, the haplogroups could be interpreted as remnants of the initial ancestral polymorphism introduced during colonization, which was later reduced through genetic drift or selection.

A similar pattern to AFLP data was seen in the PCoA results. Despite poor population structure resolution the clear separation of a cluster comprising solely Lake Turkana samples (Fig 5a and 5b asterisk) from a cluster containing samples from both basins provides further support for multiple migration events. Thus, both types of data—mitochondrial and total genomic—provide an indication of multiple migration events. This correlates with the well-documented Holocene re-connections of the Nile and Turkana basins, and provides a plausible explanation for the occurrence of multiple, genetically differentiated haplogroups in each basin (Fig 4a).

The AFLP data were structured into 5 clusters; only one of them (magenta) was unique and occurred in both the saline and the freshwater part of Lake Turkana. The abundance of the unique Turkana cluster (Fig 6a) was not as high in the saline part as it is in the freshwater area. This could indicate that saline and freshwater environments might have different effects on populations, e.g. by means of different osmotic conditions or differences in the spectrum of available intermediate annelid hosts.

In contrast to the parasite, the multiple colonization events in Lake Turkana coupled with the marked similarity of the Nile and Turkana *Synodontis* populations revealed by both mtDNA and AFLP data might be well explained by genomic admixture among multiple founding populations, i.e. lineage fusion. A phenomenon characteristic for some taxa in re-colonized insular biogeographic systems, such as in giant tortoises *Chelonoidis becki* from Galápagos Islands [80,81]. Thus, in our model, the lineage fusion pattern in the host is virtually opposite to the lineage fission, i.e. ongoing subpopulation segregation in the parasite.

Signs of parasite's allopatric differentiation in a sympatric system?

Following the end of the Pleistocene megadrought hypersaline conditions, the Turkana basin underwent severe environmental oscillations. During the Holocene freshwater periods, *Wenyonia virilis* populations were most probably widespread, whereas under current conditions they are restricted to areas where suitable, presumably freshwater, intermediate hosts are available. This assumption is strongly supported by the near absence of *Wenyonia* spp. in the medium salinity part of Lake Turkana that lacks permanent freshwater habitats. The presence of freshwater springs in the saline part of Lake Turkana [18], together with the common occurrence of *Wenyonia* tapeworms, confirm that at least locally a suitable (presumably freshwater) intermediate host is also present in the saline part of the lake. As there are no physical obstacles to host migration in Lake Turkana, isolation by adaptation can better explain the extent of the tapeworm diversification rather than isolation by distance [82–84]. The restricted distribution and the lack of substantial gene flow between the 'freshwater' and 'saline' *Wenyonia* (haplo) groups might be caused by their different environmental requirements, possibly including the role of the intermediate host(s).

The revealed restriction of the 'freshwater' subpopulation to a freshwater refugium of the Omo delta (and perhaps lower Omo) could be explained by highly saline conditions throughout most of the open lake. However, restriction of the 'saline' *Wenyonia* population to a particular part of the lake, makes adaptation of the two *W*. *virilis* subpopulations to the saline or freshwater conditions *per se* unlikely.

The documented, but quantitatively slight, geographical overlap between freshwater and saline haplogroups, i.e. sporadic occurrence of tapeworms of the freshwater haplogroup in the saline part of the lake, might be explained by migration of hosts within the lake [27]. This interpretation is supported by two different phenomena observed in our model: i) the apparent lack of actual gene flow between freshwater and saline haplogroups, and ii) the complete absence of mixed infections based on a robust dataset comprising 347 tapeworms and 120 hosts analyzed where all individual hosts were infected by tapeworms belonging either to the freshwater or saline haplogroup only. Importantly, common co-infections with up to three *Wenyonia* spp. in an individual *Synodontis* hosts in both basins [31] make competitive exclusion unlikely to be responsible for this pattern.

In fact, the pattern observed in the Lake Turkana parasite population might reflect an uneven or non-overlapping distribution of different intermediate host species, to which individual tapeworm subpopulations might have adapted.

Hence, our data suggest that the salinity causes, though possibly indirectly, the apparent ecological and near-complete geographical segregation of parasite subpopulations. This might indicate that within a single (isolated) limnic system, salinity might facilitate sympatric speciation by means of ecological segregation of subpopulations triggered by the presence of physiochemically contrasting freshwater and saline environments. Analogical processes leading to the emergence of a salinity gradient affected many African basins, and thus could have a fundamental, yet overlooked, influence on the evolution of limnic organisms throughout the Sub-Saharan Africa and elsewhere.

The indication of multiple, at least two, colonization events has been shown herein by both markers. Two migration events are reflected in our dataset by the occurrence of multiple mt haplotypes in individual localities. Importantly, AFLP data are showing on-going segregation in freshwater and saline parasite (sub)populations that were probably established prior to the last colonization. However, this separation is still incomplete and population patterns shared between the Nile and Lake Turkana indicate that the last ~11 ky were not long enough to completely segregate the tapeworm populations.

Our finding of almost exclusively 'freshwater' and 'saline' tapeworm subpopulations with no apparent gene flow between them represents the first described case of allopatric differentiation (lineage fission) in a clearly sympatric system in the case of endoparasites (documented in ectoparasites only [85,86]).

mtDNA versus AFLP

Compared to previous studies, which successfully used AFLP for a variety of purposes including evolutionary studies of populations, drug sensitivity or detection of inbred individuals [39,87–90], AFLP did not provide an optimal level of resolution in our study. Using solely AFLP data, we were not fully convinced that we could clearly distinguish whether the effect of IBA is overpowering IBD (i.e. isolation by adaptation or isolation by distance) in differentiating Lake Turkana populations.

Moreover, individual *Wenyonia* and *Synodontis* species were clearly differentiated by mtDNA data. Contrary to our expectations, the AFLP data reflected the interspecific genetic differentiation obtained from mtDNA only to a very limited degree, distinguishing *S. nigrita* in the host dataset but failing to clearly separate any of the three *Wenyonia* species. Similar to outcomes documented in insects [91] and attributed to deep genetic divisions, in which species/subspecies separated on the basis of mtDNA data are not always reflected in the nuclear genome. To some degree, the lack of differentiation in multilocus samples could be caused by undersampling (e.g. *W. minuta* with only 2 specimens available). However, it is certainly not the case in *W. youdeoweii* for which a representative population sample was obtained. Thus, it remains unknown whether the AFLP differences found in the Lake Turkana (sub)populations are directly associated with particular loci that are important for the development of differentiation/segregation between 'freshwater' and 'saline' populations, or not. Therefore, although we cannot rule out that the observed changes in genetic diversity have little or nothing to do with segregation, it is possible that we in fact demonstrate early evolutionary changes associated with it.

Although the AFLP data showed similar patterns in both organisms, the picture of population structure obtained from these data was much less clear. In particular, an increased frequency of some clusters from the Nile was seen in Lake Turkana and vice versa and unlike mtDNA, markers did not reveal any corresponding geographical pattern among *Wenyonia* samples from Lake Turkana. Moreover, the signal of the ancestral gene flow contained in the mtDNA data was similar for both organisms and provided valuable information about historical changes in species distributions. Therefore, the suitability of AFLP data to elucidate one of the main questions of this study, i.e. significance of the salinity gradient itself, remains unanswered. Compared with other studies this study is based on a sufficiently high number of isolated loci (e.g. 716 fragments for parasite and 532 for host; vs. 229 and 987 AFLP loci for Nematodes [39,92]; 731 and 237 loci for plants [93,94]; 672 loci for *Plasmodium* spp. [95]). Despite the large number of fragments analysed in both hosts and parasites, co-dominant markers like microsatellites or SNP could potentially provide more informative data. The low informativeness might be caused by suboptimal performance of the primers used for selective PCR in our assay. For example, PCR performance of the selective primers with 3 additional bases could have been too strict and therefore many fragments, which are shared within a population or a lineage, may have been missed. As a result, datasets might proportionately contain a low number of shared and a high number of rare loci/fragments, making detection of genetic patterns identifying different populations difficult.

Conclusions

Despite the common origin of the Lake Turkana populations in the Nile, the population genetic pattern observed in hosts implies the fusion of multiple colonisation waves, contrasting

with the lineage fission (segregation) in its parasite. This pattern might be explained by ecological permeability of the specific environment in which the saline part of Turkana poses a barrier for a parasite dependent on freshwater intermediate host(s), while allowing free dispersal of the definitive fish host. The parasite populations exhibited markedly higher molecular evolutionary rates compared to the hosts, providing a rare example of the long-discussed theory of rapid evolution in parasites compared to their hosts. Although AFLP and mtDNA showed similar patterns in both host and its parasite, mtDNA provided a remarkably higher resolution of population structure, historical gene flow, evolutionary patterns, and inter- and intra-specific genetic differentiation.

Supporting information

S1 Table. Discordant population histories—Fish and tapeworms. Thresholds were individually set for each AFLP primer combinations.
(XLSX)

S1 Fig. Discordant population histories—Fish and tapeworms. Phylogenetic tree of *cox I Wenyonia* spp. for a) the model of molecular evolution computed by PHYML; b) the best-fit partitioning schemes carried out in Garli. Branches are collapsed into groups, which correspond to sampling locations. Statistical support for each group was generated from Likelihood bootstrap proportions.
(TIFF)

S2 Fig. Discordant population histories—Fish and tapeworms. ΔK results for historical models admixture of *Wenyonia* spp. (a) and *Synodontis* spp. (b) and noadmixture of parasite (c) and host (d).
(TIFF)

S3 Fig. Discordant population histories—Fish and tapeworms. Individual-based cluster representation of all sampled *Wenyonia* spp. (a) and *Synodontis* spp. (b) as revealed by Bayesian inference of population structure. Each colour represents one assumed population cluster K. Multiple coloured bars display an individual's estimated membership proportion in more than one population (q), i.e. noadmixture. Turkana—medium 1–3 refers to three different localities: 1, Central Island; 2, Kerio River delta; 3, Kalokol.
(TIFF)

Acknowledgments

We are obliged to Tomáš Scholz, Julius Lukeš, Miroslav Oborník and Milan Jirků (Institute of Parasitology, BC-CAS, České Budějovice), D. Modrý (UVPS Brno) and M. Gelnar (Masaryk University, Brno) for support. Sampling in Kenya (M.J. and D.J., 2006–2010) would not have been possible without H. Charo-Karisa, D. Lotuliakou and J. Malala from the Kenya Marine and Fisheries Research Institute, which issued corresponding permits. Our thanks are also due to Father S.J. Ochieng and Father Fernando Aguire (Our Lady Queen of Peace Todonyang Catholic Mission, Kenya), field assistants Molo, Severino and Geoffrey from Loiyangalani and Turkana fishermen. We thank Bjoern C. Schaeffner (University of Melbourne) for photographs. Sampling in the Nile (M.J., 2008) was possible thanks to our Sudanese colleagues Zuheir N. Mahmoud (University of Khartoum) and A. Osman (White Nile Fisheries Research Station, Kostí). Our thanks are also due to reviewers and editor, who significantly aided the final version of manuscript. The authors acknowledge the Czech Science Foundation (P505/12/G112), the Institute of Parasitology (RVO: 60077344) and the Grant Agency of University

of South Bohemia (037/2011/P, 04-135/2010/P). Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, under the programme "Projects of Large Research, Development, and Innovations Infrastructures". The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: MJ JS.

Data curation: DJ.

Formal analysis: DJ.

Funding acquisition: MJ DJ JS.

Investigation: DJ MJ.

Methodology: MJ JS DJ.

Project administration: MJ.

Resources: MJ.

Supervision: MJ JS.

Validation: JS.

Visualization: DJ MJ.

Writing – original draft: DJ JS MJ.

Writing – review & editing: DJ JS MJ.

References

1. Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. Philos Trans R Soc Lond B Biol Sci. 2004; 359: 183–195; discussion 195. https://doi.org/10.1098/rstb.2003.1388 PMID: 15101575
2. Behrensmeyer AK, Todd NE, Potts R, Mcbrinn GE. Late Pliocene Faunal Turnover in the Turkana Basin, Kenya and Ethiopia. Science. 1997; 278: 1589–1594. PMID: 9374451
3. McKee JK. Faunal Turnover Rates and Mammalian Biodiversity of the Late Pliocene and Pléistocene of Eastern Africa. Paleobiology. 2001; 27: 500–511.
4. Weber W, White L, Vedder A, Naughton-Treves L. The African rain forest: climate and vegetation. In: Weber W, Veder A, Simons Morland H, White LJT, Hart TB, editors. African Rain Forest Ecology and Conservation An Interdisciplinary Perspective. Yale University Press; 2001. pp. 3–29.
5. Hewitt GM. Some genetic consequences of ice ages, and their role, in divergence and speciation. Biol J Linn Soc. 1996; 58: 247–276.
6. Matthee CA, Tilbury CR, Townsend T. A phylogenetic review of the African leaf chameleons: genus *Rhampholeon* (Chamaeleonidae): the role of vicariance and climate change in speciation. Proc Biol Sci / R Soc. 2004; 271: 1967–1975.
7. Mariaux J, Lutzmann N, Stipala J. The two-horned chamaeleons of East Africa. Zool J Linn Soc. 2008; 152: 367–391.
8. Day JJ, Peart CR, Brown KJ, Friel JP, Bills R, Moritz T. Continental diversification of an African catfish radiation (Mochokidae: Synodontis). Syst Biol. 2013; 62: 351–365. https://doi.org/10.1093/sysbio/syt001 PMID: 23302956
9. Otero O, Gayet M. Palaeoichthyofaunas from the Lower Oligocene and Miocene of the Arabian Plate: Palaeoecological and palaeobiogeographical implications. Palaeogeogr Palaeoclimatol Palaeoecol. 2001; 165: 141–169.
10. Stewart KM. The freshwater fish of Neogene Africa (Miocene-Pleistocene): Systematics and biogeography. Fish Fish. 2001; 2: 177–230.

11. Elmer KR, Reggio C, Wirth T, Verheyen E, Salzburger W, Meyer A. Pleistocene desiccation in East Africa bottlenecked but did not extirpate the adaptive radiation of Lake Victoria haplochromine cichlid fishes. Proc Natl Acad Sci U S A. 2009; 106: 13404–13409. https://doi.org/10.1073/pnas.0902299106 PMID: 19651614
12. Kautt AF, Elmer KR, Meyer A. Genomic signatures of divergent selection and speciation patterns in a "natural experiment", the young parallel radiations of Nicaraguan crater lake cichlid fishes. Mol Ecol. 2012; 21: 4770–86. https://doi.org/10.1111/j.1365-294X.2012.05738.x PMID: 22934802
13. Lundberg JG, Marshall LG, Guerrero J, Horton B, Malabarba MCSL, Wesselingh F. The stage for neotropical fish diversification a history of tropical south american rivers. Phylogeny and Classification of Neotropical Fishes. 1998. p. 603.
14. Montoya-Burgos JI. Historical biogeography of the catfish genus *Hypostomus* (Siluriformes: Loricariidae), with implications on the diversification of Neotropical ichthyofauna. Mol Ecol. 2003; 12: 1855–1867. PMID: 12803637
15. Goodier SAM, Cotterill FPD, O'Ryan C, Skelton PH, de Wit MJ. Cryptic Diversity of African Tigerfish (Genus *Hydrocynus*) Reveals Palaeogeographic Signatures of Linked Neogene Geotectonic Events. PLoS One. 2011; 6.
16. Pinton A, Agnèse J-F, Paugy D, Otero O. A large-scale phylogeny of *Synodontis* (Mochokidae, Siluriformes) reveals the influence of geological events on continental diversity during the Cenozoic. Mol Phylogenet Evol. 2013; 66: 1027–40. https://doi.org/10.1016/j.ympev.2012.12.009 PMID: 23274216
17. Wagner CE, Harmon LJ, Seehausen O. Cichlid species-area relationships are shaped by adaptive radiations that scale with area. Ecol Lett. 2014; 17: 583–92. https://doi.org/10.1111/ele.12260 PMID: 24602171
18. Johnson TC, Malala JO. The Nile: Origin, Environments, Lymnology and Human Use. Lake Turkana and its link to the Nile. Springer Science+Business Media BV.; 2009. pp. 287–302.
19. Bloszies C, Forman SL, Wright DK. Water level history for Lake Turkana, Kenya in the past 15,000years and a variable transition from the African Humid Period to Holocene aridity. Glob Planet Change. Elsevier B.V.; 2015; 132: 64–76.
20. Friesen VL, Smith AL, Gómez-Díaz E, Bolton M, Furness RW, González-Solís J, et al. Sympatric speciation by allochrony in a seabird. Proc Natl Acad Sci U S A. 2007; 104: 18589–18594. https://doi.org/10.1073/pnas.0700446104 PMID: 18006662
21. Rolshausen G, Segelbacher G, Hobson KA, Schaefer HM. Contemporary Evolution of Reproductive Isolation and Phenotypic Divergence in Sympatry along a Migratory Divide. Curr Biol. Elsevier Ltd; 2009; 19: 2097–2101.
22. Berg PR, Jentoft S, Star B, Ring KH, Knutsen H, Lien S, et al. Adaptation to low salinity promotes genomic divergence in Atlantic Cod (*Gadus morhua* L.). Genome Biol Evol. 2015; 7: 1644–1663. https://doi.org/10.1093/gbe/evv093 PMID: 25994933
23. Dumont HJ. A Description of the Nile Basin, and a Synopsis of Its History, Ecology, Biogeography, Hydrology, and Natural Resources. Nile. 2009; 1–21.
24. Harvey CPD, Grove AT. A Prehistoric Source of the Nile. Geogr J. 1982; 148: 327–336. Available: http://www.jstor.org/stable/633150
25. Nyamweru C. New Evidence for the Former Extent of the Nile Drainage System. Geogr J. Blackwell Publishing on behalf of The Royal Geographical Society (with the Institute of British Geographers); 1989; 155: 179–188. Available: http://www.jstor.org/stable/635059
26. Ricketts RD, Anderson RF. A Direct Comparison Between the Historical Record of Lake Level and the δ18O Signal in Carbonate Sediments from Lake Turkana, Kenya. Limnol Oceanogr. American Society of Limnology and Oceanography; 1998; 43: 811–822. Available: http://www.jstor.org/stable/2839176
27. Hopson A. Lake Turkana; A Report on the Finding of the Lake Turkana Project 1972–1975, Vols. 1–6. Overseas Development Administration, London; 1982.
28. Camberlin P. Nile Basin Climates. The Nile: Origin, Environments, Limnology and Human Use. 2009. pp. 307–333.
29. Barrows TT, Williams MAJ, Mills SC, Duller GAT, Fifield LK, Haberlah D, et al. A White Nile megalake during the last interglacial period. Geology. 2014; 42: 163–166.
30. Froese R, Pauly D. FishBase [Internet]. Version 2015.8. 2014. p. www.fishbase.org. www.fishbase.org
31. Schaeffner BC, Jirků M, Mahmoud ZN, Scholz T. Revision of *Wenyonia* Woodland, 1923 (Cestoda: Caryophyllidea) from catfishes (Siluriformes) in Africa. Syst Parasitol. 2011; 79: 83–107. https://doi.org/10.1007/s11230-011-9290-2 PMID: 21544709
32. Mackiewicz JS. Caryophyllidea (Cestoidea): a review. Exp Parasitol. 1972; 31: 417–512. PMID: 4558418

33. Scholz T, Brabec J, Kráĺová-Hromadová I, Oros M, Bazsalovicsová E, Ermolenko A, et al. Revision of *Khawia* spp. (Cestoda: Caryophyllidea), parasites of cyprinid fish, including a key to their identification and molecular phylogeny. Folia Parasitol (Praha). 2011; 58: 197–223.

34. Scholz T, Hanzelová V. Tapeworms of the genus *Proteocephalus* Weinland, 1858 (Cestoda: Proteocephalidae), parasites of fishes in Europe. Praha Acad Stud AV ČR. 1998; 119 pp.

35. Sambrook J, Russell DW. Molecular Cloning—Sambrook & Russel. Human Mutation. 2001.

36. Werle E, Schneider C, Renner M, Völker M, Fiehn W. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res. 1994; 22: 4354–4355. PMID: 7937169

37. Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN. Universal primer cocktails for fish DNA barcoding. Mol Ecol Notes. 2007; 7: 544–548.

38. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, et al. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 1995; 23: 4407–14. Available: http://www.ncbi.nlm.nih.gov/pubmed/10206688 PMID: 7501463

39. Tandon R, Lyons ET, Tolliver SC, Kaplan RM. Effect of moxidectin selection on the genetic variation within *Cylicocyclus nassatus* based on amplified fragment length polymorphism (AFLP). Int J Parasitol. 2005; 35: 813–9. https://doi.org/10.1016/j.ijpara.2005.02.004 PMID: 15925599

40. de Gruijter JM, Polderman AM, Dijkshoorn L, Roberts H, Ziem J, Kunwar CB, et al. AFLP fingerprinting for the analysis of genetic diversity within *Necator americanus*. Mol Cell Probes. 2006; 20: 317–21. https://doi.org/10.1016/j.mcp.2006.03.004 PMID: 16716566

41. Nejsum P, Grøndahl C, Murrell KD. Molecular evidence for the infection of zoo chimpanzees by pig *Ascaris*. Vet Parasitol. 2006; 139: 203–10. https://doi.org/10.1016/j.vetpar.2006.02.025 PMID: 16567044

42. Patra KP, Ramu T, Hoti SL, Pragasam GS, Das PK. Identification of a molecular marker for genotyping human lymphatic filarial nematode parasite *Wuchereria bancrofti*. Exp Parasitol. 2007; 116: 59–65. https://doi.org/10.1016/j.exppara.2006.11.011 PMID: 17250828

43. Mattersdorfer K, Koblmüller S, Sefc KM. AFLP genome scans suggest divergent selection on colour patterning in allopatric colour morphs of a cichlid fish. Mol Ecol. 2012; 21: 3531–44. https://doi.org/10.1111/j.1365-294X.2012.05634.x PMID: 22625655

44. Xu D, Lou B, Xu H, Li S, Geng Z. Isolation and Characterization of Male-Specific DNA Markers in the Rock Bream *Oplegnathus fasciatus*. Mar Biotechnol (NY). 2013; 15: 221–9.

45. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012; 28: 1647–1649. https://doi.org/10.1093/bioinformatics/bts199 PMID: 22543367

46. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30: 3059–3066. PMID: 12136088

47. Posada D. jModelTest: Phylogenetic model averaging. Mol Biol Evol. 2008; 25: 1253–1256. https://doi.org/10.1093/molbev/msn083 PMID: 18397919

48. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. Science. 2001; 294: 2310–2314. https://doi.org/10.1126/science.1065889 PMID: 11743192

49. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 2003; 19: 1572–1574. PMID: 12912839

50. Nylander JAA, Wilgenbusch JC, Warren DL, Swofford DL. AWTY (are we there yet?): A system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics. 2008; 24: 581–583. https://doi.org/10.1093/bioinformatics/btm388 PMID: 17766271

51. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol. 2003; 52: 696–704. PMID: 14530136

52. Hasegawa M, Kishino H, Yano TA. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol. 1985; 22: 160–174. PMID: 3934395

53. Gascuel O. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol. 1997; 14: 685–695. PMID: 9254330

54. Brandley MC, Schmitz A, Reeder TW. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst Biol. 2005; 54: 373–90. https://doi.org/10.1080/10635150590946808 PMID: 16012105

55. Lanfear R, Calcott B, Ho SYW, Guindon S. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol. 2012; 29: 1695–1701. https://doi.org/10.1093/molbev/mss020 PMID: 22319168

56. Rambaut A, Drummond AJ. Tracer V1.5. http://beast.bio.ed.ac.uk/Tracer. 2009;

57. Zwickl DJ. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Philosophy. 2006; 115. (http://www.zo.utexas.edu/faculty/antisense/garli/Garli.html)
58. Leigh JW, Bryant D. Popart: Full-Feature Software for Haplotype Network Construction. Methods Ecol Evol. 2015; 6: 1110–1116.
59. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25: 1451–2. https://doi.org/10.1093/bioinformatics/btp187 PMID: 19346325
60. Beerli P. MIGRATE: documentation and program, part of LAMARC. http://evolution.genetics.washington.edu/lamarc.html. 1997;
61. Kass RE, Raftery AE. Bayes Factors. J Am Stat Assoc. American Statistical Association; 1995; 90: 773–795.
62. Beerli P, Palczewski M. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. Genetics. 2010; 185: 313–26. https://doi.org/10.1534/genetics.109.112532 PMID: 20176979
63. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics. 2012; 28: 2537–2539. https://doi.org/10.1093/bioinformatics/bts460 PMID: 22820204
64. Antao T, Beaumont MA. Mcheza: a workbench to detect selection using dominant markers. Bioinformatics. 2011; 27: 1717–8. https://doi.org/10.1093/bioinformatics/btr253 PMID: 21505032
65. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 2009; 9: 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x PMID: 21564903
66. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005; 14: 2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x PMID: 15969739
67. Jakobsson M, Rosenberg NA. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007; 23: 1801–1806. https://doi.org/10.1093/bioinformatics/btm233 PMID: 17485429
68. Rosenberg NA. DISTRUCT: A program for the graphical display of population structure. Mol Ecol Notes. 2004; 4: 137–138.
69. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007; 7: 214. https://doi.org/10.1186/1471-2148-7-214 PMID: 17996036
70. Štefka J, Hoeck PEA, Keller LF, Smith VS. A hitchhikers guide to the Galápagos: co-phylogeography of Galápagos mockingbirds and their parasites. BMC Evol Biol. 2011; 11: 284. https://doi.org/10.1186/1471-2148-11-284 PMID: 21966954
71. Johnson KP, Allen JM, Olds BP, Mugisha L, Reed DL, Paige KN, et al. Rates of genomic divergence in humans, chimpanzees and their lice. Proc Biol Sci. 2014; 281: 20132174. https://doi.org/10.1098/rspb.2013.2174 PMID: 24403325
72. Hafner M, Sudman P, Villablanca F, Spradling T, Demastes J, Nadler S. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science. 1994; 265: 1087–1090. PMID: 8066445
73. Kaltz O, Shykoff JA. Local adaptation in host—parasite systems. Heredity (Edinb). 1998; 81: 361–370.
74. Gandon S, Michalakis Y. Local adaptation, evolutionary potential and host-parasite coevolution: interactions between migration, mutation, population size and generation time. J Evol Biol. 2002; 15: 451–462.
75. Frank SA. Evolution of host-parasite diversity. Evolution (N Y). JSTOR; 1993; 47: 1721–1732.
76. Day T, Graham AL, Read AF. Evolution of parasite virulence when host responses cause disease. Proc Biol Sci. 2007; 274: 2685–2692. https://doi.org/10.1098/rspb.2007.0809 PMID: 17711836
77. Best A, White A, Boots M. The Implications of Coevolutionary Dynamics to Host-Parasite Interactions. Source Am Nat. 2009; 173: 779–791. Available: http://www.jstor.org/stable/10.1086/598494%5Cnhttp://www.jstor.org/page/info/about/policies/terms.jsp%5Cnhttp://www.jstor.org
78. Cohuet A, Harris C, Robert V, Fontenille D. Evolutionary forces on *Anopheles*: what makes a malaria vector? Trends Parasitol. Elsevier Ltd; 2010; 26: 130–136.
79. Khan SM, Reece SE, Waters AP, Janse CJ, Kaczanowski S. Why are male malaria parasites in such a rush?: Sex-specific evolution and host-parasite interactions. Evol Med public Heal. 2013; 2013: 3–13.
80. Emerson BC, Faria CMA. Fission and fusion in island taxa—serendipity, or something to be expected? Mol Ecol. 2014; 23: 5132–5134. https://doi.org/10.1111/mec.12951 PMID: 25330853
81. Garrick RC, Benavides E, Russello MA, Hyseni C, Edwards DL, Gibbs JP, et al. Lineage fusion in Galápagos giant tortoises. Mol Ecol. 2014; 23: 5276–5290. https://doi.org/10.1111/mec.12919 PMID: 25223395

82. Sexton JP, Hangartner SB, Hoffmann AA. Genetic isolation by environment or distance: Which pattern of gene flow is most common? Evolution (N Y). 2014; 68: 1–15.

83. Orsini L, Vanoverbeke J, Swillen I, Mergeay J, De Meester L. Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. Molecular Ecology. 2013. pp. 5983–5999. https://doi.org/10.1111/mec.12561 PMID: 24128305

84. Spurgin LG, Illera JC, Jorgensen TH, Dawson DA, Richardson DS. Genetic and phenotypic divergence in an island bird: Isolation by distance, by colonization or by adaptation? Mol Ecol. 2014; 23: 1028–1039. https://doi.org/10.1111/mec.12672 PMID: 24447303

85. Hughes J, Kennedy M, Johnson KP, Palma RL, Page RDM. Multiple cophylogenetic analyses reveal frequent cospeciation between pelecaniform birds and *Pectinopygus* lice. Syst Biol. 2007; 56: 232–251. https://doi.org/10.1080/10635150701311370 PMID: 17464880

86. Koop JAH, DeMatteo KE, Parker PG, Whiteman NK. Birds are islands for parasites. Biol Lett. 2014; 10: 6–10.

87. Otsen M, Hoekstra R, Plas ME, Buntjer JB, Lenstra JA, Roos MH. Amplied fragment length polymorphism analysis of genetic diversity of *Haemonchus contortus* during selection for drug resistance. Int J Parasitol. 2001; 31: 1138–1143. PMID: 11429180

88. Van Der Veer M, De Vries E. Genetic intrapopulation variation, revealed by amplified fragment length polymorphism, within a population of the trichostrongylid nematode *Cooperia oncophora*. Exp Parasitol. 2003; 104: 70–73. PMID: 12932763

89. Roos MH, Otsen M, Hoekstra R, Veenstra JG, Lenstra JA. Genetic analysis of inbreeding of two strains of the parasitic nematode *Haemonchus contortus*. Int J Parasitol. 2004; 34: 109–115. PMID: 14711596

90. Simo G, Cuny G, Demonchy R, Herder S. *Trypanosoma brucei gambiense*: study of population genetic structure of Central African stocks using amplified fragment length polymorphism (AFLP). Exp Parasitol. 2008; 118: 172–80. https://doi.org/10.1016/j.exppara.2007.07.010 PMID: 17850792

91. Dasmahapatra KK, Elias M, Hill RI, Hoffman JI, Mallet J. Mitochondrial DNA barcoding detects some species that are real, and some that are not. Mol Ecol Resour. 2010; 10: 264–273. https://doi.org/10.1111/j.1755-0998.2009.02763.x PMID: 21565021

92. Folkertsma RT, Rouppe van der Voort JN, de Groot KE, van Zandvoort PM, Schots A, Gommers FJ, et al. Gene pool similarities of potato cyst nematode populations assessed by AFLP analysis. Mol Plant Microbe Interact. 1996; 9: 47–54. PMID: 8589422

93. Gaudet M, Jorge V, Paolucci I, Beritognolo I, Mugnozza GS, Sabatti M. Genetic linkage maps of *Populus nigra* L. including AFLPs, SSRs, SNPs, and sex trait. Tree Genet Genomes. 2007; 4: 25–36.

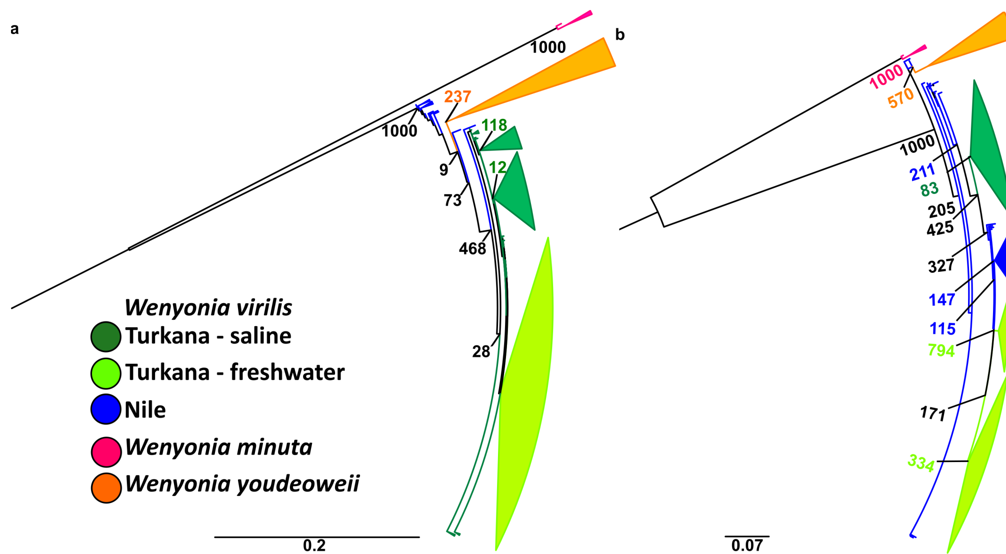
94. Frascaroli E, Schrag TA, Melchinger AE. Genetic diversity analysis of elite European maize (*Zea mays* L.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. Theor Appl Genet. 2013; 126: 133–41. https://doi.org/10.1007/s00122-012-1968-6 PMID: 22945268

95. Martinelli A, Hunt P, Cheesman SJ, Carter R. Amplified fragment length polymorphism measures proportions of malaria parasites carrying specific alleles in complex genetic mixtures. Mol Biochem Parasitol. 2004; 136: 117–122. PMID: 15478791

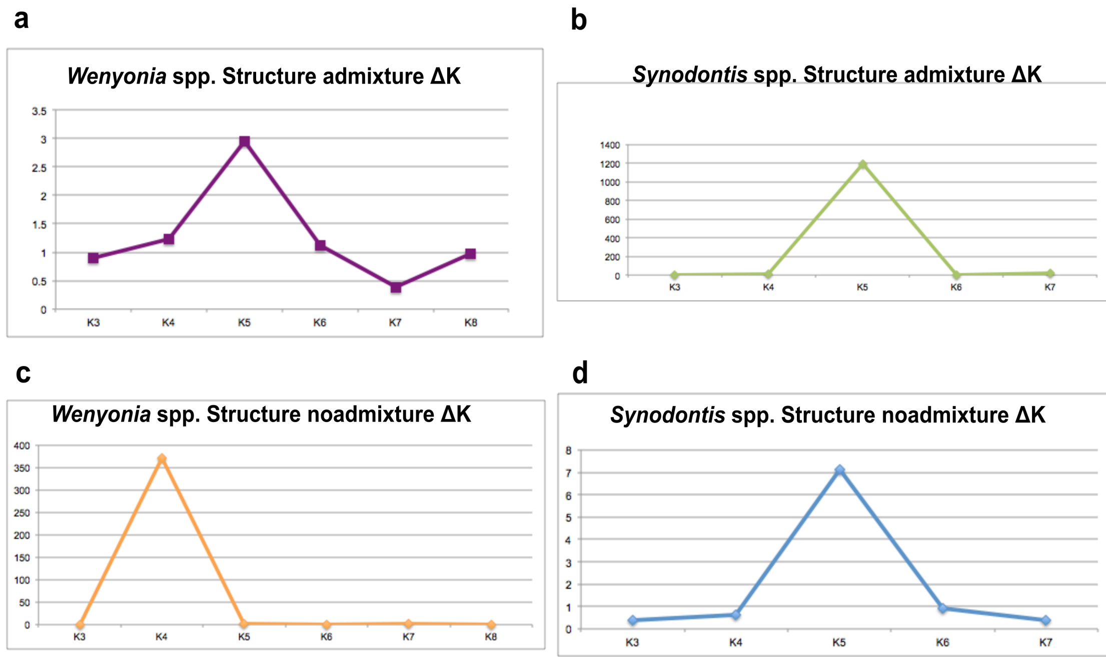
Supplementary data

S1 Table. Thresholds were individually set for each AFLP primer combinations.

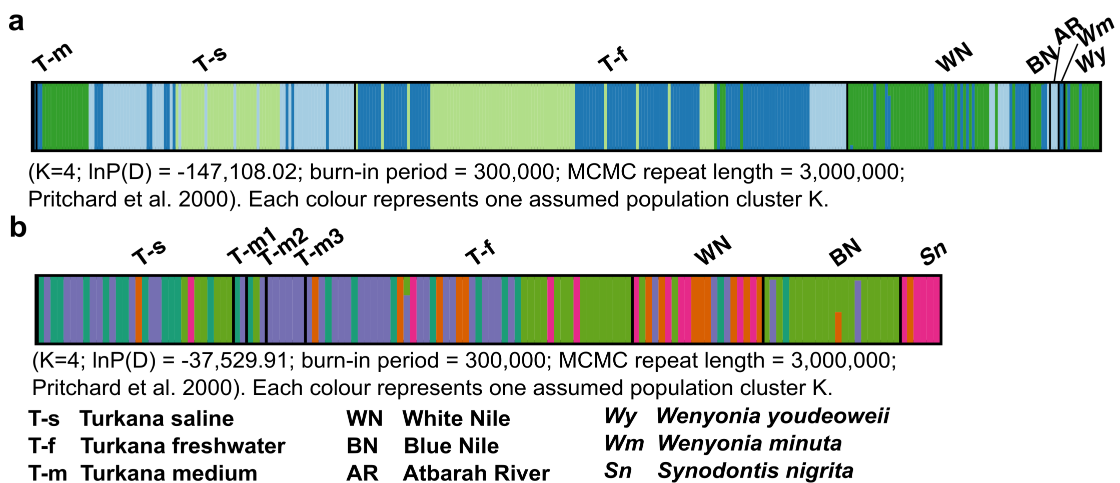
	Wenyonia spp.					*Synodontis* spp.		
	MseI-CTC	MseI-CTA	MseI-CAA	MseI-CAG	MseI-CAC	MseI-CTC	MseI-CTA	MseI-CAA
EcoRI-ACA*	150	150	150	100	150	150	200	200
EcoRI-ACT*	150	150	200	150	200	150	150	150
EcoRI-ACG*	150	150	150	100	200	200	200	150



S1 Figure. Phylogenetic tree of *cox I Wenyonia* spp. for a) model of molecular evolution computed by PHYML; b) best-fit partitioning schemes carried out in Garli. Branches are collapsed into groups, which correspond to sampling locations. Statistical supports for each group were generated from Likelihood bootstrap proportion.



S2 Figure. ΔK results for historical models admixture of *Wenyonia* spp. (a) and *Synodontis* spp. (b) and noadmixture of parasite (c) and host (d).



S3 Figure. Individual-based cluster representation of all sampled *Wenyonia* (a) and *Synodontis* (b) as revealed by Bayesian inference of population structure. Each colour represents one assumed population cluster K. Multiple coloured bars display an individual's estimated membership proportion in more than one population (q), i.e. noadmixture. Turkana – medium 1-3 refers to three different localities: 1, Central Island; 2, Kerio River delta; 3, Kalokol.

Paper 2

Hidden diversity and overlooked morphological variability of *Synodontis schall*: from taxonomic deflation to recognition of cryptic species?

Dagmar Jirsová[1,2], Jan Štefka[1,2], Radim Blažek[3,4], John O. Malala[5], David E. Lotuliakou[5], Zuheir N. Mahmoud[6], Miloslav Jirků[2]

[1]Faculty of Science, University of South Bohemia, Branišovská 1760, 370 05 České Budějovice, Czech Republic

[2]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská 31, 370 05 České Budějovice, Czech Republic

[3]Institute of Vertebrate Biology, Czech Academy of Sciences, Květná 8, 603 65 Brno, Czech Republic

[4]Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

[5]Kenya Marine and Fisheries Research Institute, Lake Turkana Station, P.O. Box 205, 305 00 Lodvar, Kenya

[6]Department of Zoology, Faculty of Science, University of Khartoum, 111 15 Khartoum, Sudan

Key words: Africa, *coxI, cytb*, frontosus, Nile, phylogeny, synonymy, taxonomy, Turkana

Correspondence: Dagmar Jirsová; E-mail: dasa.jirsova@gmail.com

Abstract

Examination of 138 specimens of *Synodontis schall* (Mochokidae, Siluriformes) from the Nile River and Lake Turkana revealed presence of both *S. schall*-like and *S. frontosus*-like morphotypes, with clinal variation between them. Phylogenetic and haplotype analyses based on 131 *coxI* (565bp), 96 *cytb* (973bp) and 23 *RAG2* (896bp) sequences from the Nile-Turkana population, plus additional GenBank sequences of *Synodontis* spp., showed that both morphotypes and intermediate forms are conspecific. The results imply probable synonymy of *S. frontosus* with *S. schall*. Strong biogeographical signal has been revealed among supposedly widely distributed, homogenous and conspecific *S. schall*-like catfish of the Nilo-Sudanian ichthyological province using the mitochondrial dataset. *Synodontis schall* sensu stricto, as defined by type locality in the Nile, is apparently restricted to eastern part of the Nilo-Sudanian ichthyological province (e.g. Nile, Turkana, Chad), while *S. schall* sensu lato from western parts (e.g. Senegambia, Niger, Chad) most probably represents a cryptic taxon unrecognized thus far due to the absence of tenable morphological differences. Apparently there is an overlap of the eastern "sensu stricto" and western "sensu lato" *S. schall* in the Chad basin.

1. Introduction

The Mochokidae (Actinopterygii, Siluriformes) is an endemic African family of freshwater catfish containing 211 species in nine genera (Eschmeyer et al. 2017). With 131 nominal species currently recognized, *Synodontis* Cuvier, 1817 is the most speciose mochokid genus and is widely distributed in all African ichthyological provinces except for Maghreb and South Africa (Poll 1971; Paugy and Roberts 2003; Froese and Pauly 2017). Available knowledge on most *Synodontis* spp. is limited to morphology and distribution (e.g. Poll 1971; Sands 1983; Musschoot & Lalèyè 2008), while the first molecular phylogenetic studies were published during the last decade (Day and Wilkinson 2006; Koblmüller et al. 2006; Day et al. 2013; Pinton et al. 2013).

Within the scope of research on fish parasites in the Nile River and Lake Turkana, we examined a total of 138 *Synodontis* specimens. Two of three *Synodontis* spp. in our sample, *Synodontis nigrita* Valenciennes, 1840 and *Synodontis serratus* Rüppell, 1829 (both absent in Lake Turkana) could be readily morphologically determined. The third species *Synodontis schall* (Bloch et Schneider, 1801) showed deviation from both data in the literature, as well as from differential diagnosis and the neotype proposed recently by Musschoot & Lalèyè (2008). Among *Synodontis* spp. occurring in the Nile and Turkana basins, *S. schall* might be confused with the morphologically very similar *Synodontis frontosus* Vaillant, 1895. According to the literature, *S. schall* is common and abundant in both the Nile River and Lake Turkana (e.g. Poll 1971; Lévêque *et al.* 1991; Bailey 1994; Paugy & Roberts 2003). Although the two species might co-occur in both basins, *S. frontosus* is considered absent in Lake Turkana itself, reportedly being restricted to its main source, the Omo River (Hopson 1982). Diagnostic features of both species are similar and partly overlapping, therefore, distinguishing between the two in the field is a difficult task leaving room for errors.

Due to the above mentioned uncertainties and because *S. schall* is our model in parasitological studies requiring proper host determination (Jirsová et al. 2017), we decided to base the taxonomical assignment of our Nile-Turkana *S. schall* specimens on a phylogenetic approach aided with an assessment of selected morphological traits. Herein, we assess the identity of *S. schall* from the Nile and Turkana basins by means of phylogenetic and haplotype analyses of a representative set of *Synodontis* spp. *coxI*, *cytb* and *RAG2* sequences, revealing yet unrecognized biogeographical pattern among the Nilo-Sudanian *S. schall*-like catfish.

2. Material and Methods

2.1 Sampling and fish determination

Sampled fish were obtained from local fishermen in seven different localities in the Nile and Turkana basins (*S. schall* n=120, *S. nigrita* n=8, *S. serratus* n=10, Total: 138; see Table 1 for details). The two Nile localities cover both the White (Kostí) and Blue Nile (Sennar). The five Turkana localities cover the salinity gradient present in the lake, i.e. the freshwater Omo River delta (Todonyang) and the saline (brackish) main part of the lake (central Turkana saline: Kalokol, Central Island; central Turkana freshwater refugium: Kerio River delta; south Turkana: El-Molo Bay). A small piece of liver tissue from each individual was cut off and preserved in 96% ethanol for molecular analyses. Fish were routinely measured, photo-documented and their mouthparts including tooth plates and barbels were cut off and preserved in 4% formaldehyde for further morphological examination. Fish were destined for human consumption, therefore, no permission was needed to collect these fish. Specimens

that were still alive when examined were killed by dorsal pithing (spinal cord and blood vessels cut immediately behind the head), a method congruent with Kenyan, Sudanian, Czech and European legislation.

Upon field examination, all fish were determined based on following distinguishing combinations of features: *S. nigrita*: body, head and fins with black spots, inner mandibular barbles bearing thick and tuberculate ramifications, maxillary barbels with a well visible broad black membrane, short and small adipose fin; *S. serratus*: maxillary barbels with a very large light-colored membrane, humeral process pointed, non-keeled, distinctive serrations on both edges of pectoral spine, prolonged rostrum (Bailey 1994; Paugy and Roberts 2003); for *S. schall*, see Table 2 and Results and Discussion sections.

2.2 DNA extraction, PCR & sequencing

DNA was isolated using the phenol chloroform method described in Sambrook & Russell (2001). Two mitochondrial genes and one nuclear gene were used for phylogenetic analyses, cytochrome oxidase I (*coxI*), cytochrome b (*cytb*) and recombination activating gene 2 (*RAG2*), respectively (see Table S1 in supplementary data for a complete list of sequences and other details). Amplification protocol and PCR primer FishF2 and VF2_t1 for partial sequence of *coxI* gene (~800 bp) were adopted from Ivanova *et al.* (2007). PCR reaction contained 12.5 µl of PPP Master mix, 10.5 µl of ultra pure water, 5 pM of each PCR primer and 0.5 µl of isolated DNA (100 to 170ng). Primer sequences (5'-GAC-TTGAAGAACCACCGTTG-3' forward and 5'-TTTAGAATTCTGG CTTTGGGAG-3' reverse, ~ 800bp) to amplify part of *cytb* and PCR conditions were adopted from Pinton *et al.* (2013). Primers for *RAG2* amplification (forward 5'-TGY TAT CTC CCA CCT CTG CGY TAC C-3' and reverse 5'-TCA TCC TCC TCA TCK TCC TCW TTG TA-3', ~ 1100bp) and reaction conditions were adopted from Sullivan *et al.* (2006). Amplification reactions for *cytb* and *RAG2* contained the same PCR reagents and amount of DNA as mentioned above. All products were verified on a 1% agarose gel and purified using enzymes exonuclease I (Exo I) and shrimp alkaline phosphatase (SAP) (Werle et al. 1994). Purified fragments were directly sequenced by Macrogen Inc. (Amsterdam, Netherlands). The same primers used for PCR amplification were used for sequencing.

2.3 Phylogenetic and haplotype analyses

Obtained sequences were assembled and inspected for errors using GENEIOUS Pro software package version 6.1 (Kearse *et al.* 2012). In addition to our original sequences (Tab. S1), we used GenBank sequences of *Synodontis* spp. to interpret our results in a broader phylogenetic context (see Table S2 in supplementary data for complete list of sequences and additional information). Alignments were constructed in the MAFFT v. 6 (Katoh et al. 2002) and corrected manually using GENEIOUS Pro software package version 6.1. Best-fit models of molecular evolution were selected using AIC criteria in jMODELTEST v2.1.4 (Posada 2008). Phylogenetic analyses were performed using Bayesian inference (BI) in MrBayes v3.2.2. (Huelsenbeck et al. 2001; Ronquist and Huelsenbeck 2003) and maximum likelihood (ML) in PHYML v3.0 (Guindon and Gascuel 2003). BI analyses were performed under the same conditions for all datasets - using 5 million MCMC replications, with four chains and three independent runs. The following nucleotide substitutional models were used: GTR+I+G for *coxI* data, GTR+G for *cytb* and for HKY+I+G for *RAG2* data. ML analyses were run using bootstrap analyses with 1000 replications and models mentioned above. All phylogenetic

trees were rooted using GenBank sequences of *Microsynodontis* sp. (HF565849 for *coxI*, DQ886604 for *cytb,* HF565741 for *RAG2*).

To examine the population and dataset variability, statistical parsimony networks based on pairwise differences were constructed using PopART v1.7 (Leigh and Bryant 2015), using settings for TCS networks (Clement et al. 2000). Separate datasets containing only sequences of *S*. *schall* sensu stricto and *S*. *schall* sensu lato clades as revealed by phylogenetic analyses were analysed.

3. Results

3.1 Morphological determination

Standard length of examined *S. schall* specimens ranged from 142 to 300 mm (mean ± SD; 209.8 ± 33.2 mm). Weight ranged from 75 to 783 g (mean ± SD, 227.1 ± 118.9 g). *Synodontis schall* from the Nile and Turkana showed remarkable variation in body shape, proportions, coloration, as well as diagnostic features (see Fig. 1 and Table 2 for details). We observed a phenotypic gradient from typical *S. schall* morphotype (Fig. 1a,d,i,k), which dominated the sample, to typical *S. frontosus* morphotype (Fig. 1c,h,j,m), which was less frequent. These observations, together with the common occurrence of intermediate forms hampered reliable morphology-based determination of numerous specimens. In the majority of Nile-Turkana specimens, there was an obvious, blunt, low, relatively wide, longitudinally striated keel on the ventral side of the humeral process (Fig. 1d,e,h) fitting *S. schall*. Only in some specimens the keel was absent (Fig. 1f,g), resembling that of *S. frontosus*. The humeral process shape showed clinal variation from *S. schall*-like, i.e. margins straight and/or dorsally concave and/or ventrally convex (Fig. 1d,e) to *S. frontosus*-like, i.e. straight both dorsally and ventrally (Fig. 1h). Clinal variation was observed also in anterior pectoral spine serrations, from *S. schall*-like, i.e. fine, sharp, distinct (Fig. 1i) to *S. frontosus*-like, i.e. barely discernible (sometimes due to pigmentation in dark-colored specimens) (Fig. 1j). Remarkable clinal variation was present also in maxillary barbels from *S. schall*-like, i.e. devoid of a basal membrane or with a hardly visible slat-like rudiment (Fig. 1k,l) to *S. frontosus*-like, i.e. distinct, well developed dark basal membrane (Fig. 1m). Mandibular teeth numbered from 26 to 46 in total covering the range for both *S. schall* (18-39) and *S. frontosus* (33-48) (Table 2). When only a single main row of mandibular teeth was considered the count was 21 to 40. Mandibular teeth seem to grow up from the outer margin of tooth plate towards the mouth interior, where the old ones lose their sharp tips and eventually disappear, while new ones start to grow from the outer side (Fig. 2). The number of new mandibular teeth emerging before the main teeth row varied from 1 to 12.

In three main Turkana sampling localities, the ratio of *S. schall*/*S. frontosus*/intermediate morphotypes was as follows: El-Molo Bay 35/16/7, Kalokol 19/11/0, Todonyang 32/31/14, implying somewhat higher relative incidence of *S. frontosus*-like morphotype in the freshwater part of Turkana (Chi-square test = 10.3; df =4; $P < 0.05$) (Tab. S3).

In addition, several other features corresponding either to *S. schall* or *S. frontosus* have been observed, but not systematically recorded. For example, pointed snout combined with narrower interorbital region and eyes well within lateral margin of the head in dorsal view (as in *S. schall*) vs. snout more broadly rounded with wider interorbital region and orbits impinging on the lateral outline of the head in dorsal view (as in *S. frontosus*) (sensu Hopson 1982). The two states are well demonstrated by the two specimens showing *S. schall* and *S. frontosus* morphotype in Fig. 1a and Fig. 1c, respectively.

3.2 Phylogenetic analyses

In total, 131 sequences of *coxI* (565bp), 96 *cytb* (973bp) and 23 *RAG2* (896bp) were obtained from the Nile River and Lake Turkana and used for phylogenetic analyses together with *Synodontis* spp. sequences retrieved from GenBank. The sequence data were deposited in GenBank under accession numbers *coxI*: KY483652-KY483782, *cytb*: KY483783-KY483878, *RAG2*: MF136700-MF136718, MF150161- MF150164. As BI and ML phylogenetic analyses provided similar outcome, only BI final trees are presented with posterior probabilities as branch supports.

In *coxI* analysis, sequences of 81 *Synodontis* spp. were analyzed (Fig. 3). As expected, our samples of *S. nigrita* and *S. serratus* from the Nile clustered with conspecific sequences from GenBank (see Fig. S1 for composition of the collapsed clades). All Nile-Turkana *S. schall* samples clustered together and created a well-supported group. According to the predominant morphological patterns observed in the Nile-Turkana sample and location of the type locality of *S. schall* in the Nile, we named this clade "*S. schall* sensu stricto". Importantly, all GenBank sequences labelled *S. frontosus* (HF565883–89) and one sequence of *S. caudovittatus* (HF565865), all from the Nile, also clustered in this clade in all analyses. On the contrary, GenBank sequences labelled *S. schall* (HF565950, HF565953-56), *Synodontis ouemeensis* HF565953, all from western part of Nilo-Sudanian and Eburneo-Ghanian provinces, together with sequences labelled *Synodontis* aff. *bastiani*, *Synodontis* aff. *haugi* HF565950, HF565954-56 (pet trade), created a separate group with high nodal support. Due to the clear separation from the "*S. schall* sensu stricto" and taxonomic assignment of most its sequences, we named the clade "*S. schall* sensu lato".

Dataset for *cytb* included sequences representing 52 *Synodontis* spp. (Fig. 4). Our sequences clustered with all GenBank *S. frontosus* sequences (FM878850–54, HF566018–23), as well as *S. schall* sequences (EU781905–07, EU781915) and EU781915 from the Nile and/or Chad basin. This clade clearly corresponds with the *S. schall* sensu stricto clade, hence its name in the tree. The rest of *S. schall* sequences (EU781902–04, EU781908–14, EU781916) from Senegambia, Niger and Chad basin created a group corresponding to the *S. schall* sensu lato clade, closely related to, but clearly distinct from the *S. schall* sensu stricto clade. In conclusion, analyses of both mitochondrial markers revealed very similar patterns with strong biogeographic signals, indicating phylogenetic distinctiveness and geographic affinity of two clades, i.e. eastern *S. schall* sensu stricto and western *S. schall* sensu lato (Figs. 3, 4, 5).

Analysis of the nuclear gene was less informative. A total of 19 *RAG2* sequences of *Synodontis* spp. was analysed (Fig. 6) together with 102 sequences from Genbank. Our *S. schall* sensu stricto samples from the Nile and Turkana comprised an internally unresolved cluster together with the samples of *S. frontosa* (HF565870-74), *S. caudovittatus* (HF565755), *S. violacea* (HF565838) and S. aff. bastiani (HF565751) from Eastern Nilo-Sudan province, as well as *S. schall* (HF565817-19) from Western Nilo-Sudan province. Although sequences of *S. serratus* from our sample and from GenBank clustered together with high support, the *S. serratus* branch was placed within the *S. schall* clade, probably as a result of low sequence variability in the whole dataset (100-94.5%).

3.3 Haplotype networks

Haplotype networks were constructed for the mitochondrial genes, but not for the nuclear gene (*RAG2*), because sufficient numbers of sequences from Western Nilo-Sudan were only available for mitochondrial genes.

Samples created two well-separated clusters corresponding to the *S. schall* sensu stricto and *S. schall* sensu lato clades previously revealed in phylogenetic analyses and showing virtually identical patterns (Fig. 7, 8). Both mitochondrial genes proved good determining markers with high accuracy for the sensu stricto and sensu lato clades, although somewhat higher levels of variability in the dataset/population and more haplotypes were observed in *cytb*.

While data for *coxI* created one big central haplotype with minimal presence of other more derived haplotypes, *cytb* showed a more complicated network structure with some haplotypes divided by more than ten mutations. Although the two networks might appear very different, similar phenomena can be seen in both datasets, most remarkably sharing of the same haplotypes between the Nile and Turkana basins.

4. Discussion

4.1 Identity of Nile-Turkana S. schall*-like catfish*

In the Nile and Turkana basins, both typical *S. schall* morphotype and typical *S. frontosus* morphotype, as well as a phenotypic gradient between the two, have been identified. Regardless of morphological patterns, phylogenetic and haplotype analyses based on robust datasets rendered all the Nile-Turkana samples clearly conspecific, i.e. both *S. schall*-like and *S. frontosus*-like. It is to be noted, that the taxonomical incongruence between taxonomic assignment of GenBank sequences and our phylogenetic analyses clearly indicated both *S. schall* and *S. frontosus* morpho-species among samples revealed herein as conspecific as identified by other authors (see original sequence assignments retained in Fig. 4 and Table S2). This result renders features presumably differentiating *S. schall* and *S. frontosus* untenable. Especially the relatively frequent presence of distinct dark membranes on maxillary barbels is worth attention, because its “absence, or a presence of only hardly visible rudiment” (sensu Paugy & Roberts 2003 and Musschoot & Lalèyè 2008) was considered a basic feature distinguishing *S. schall* from *S. frontosus*. Given the genetic uniformity of the Nile-Turkana sample on one hand, and its remarkable morphological variability, together with obvious overlap of presumably differential features on the other, we assume that diagnostic features of *S. frontosus* merely describe part of the intraspecific variability of *S. schall*. In other words, explicit evidence of remarkable intraspecific variability provided herein renders the two nominal species *S. schall* and *S. frontosus* probably synonymous, hence requiring taxonomical reassessment.

Based on following criteria, we assign the Nile-Turkana population as represented in our sample to *S. schall* sensu stricto: i) the most prevalent morphotype matching *S. schall*, ii) phylogenetic affinity (see below), iii) geographic origin matching with the type locality in the Nile and iv) with respect to possible synonymy, the priority of the specific name *S. schall* over *S. frontosus*. The interpretation is indirectly further supported by the supposed absence of *S. frontosus* in Lake Turkana itself, being presumably limited to Omo River (as highlighted by Hopson 1982 and other authors). Although in our Turkana sample *S. frontosus* morphotype was not restricted to freshwater Omo delta, it was clearly most prevalent here.

Concurrent morphological and molecular sampling throughout the Nile basin, including terra typica of both *S. schall* and *S. frontosus*, i.e. Nile River in Assouan and in Khartoum, respectively, is necessary for sound taxonomical measures (we presume synonymy of *S.*

frontosus and *S. schall*). Unfortunately, studies of other authors do not aid further resolution, because even recent taxonomic studies on Nilo-Sudanian *Synodontis* spp. lack molecular phylogenetic analysis, and vice versa, all available phylogenetic studies lack clues for determination of sampled specimens, and/or even the exact origin of the samples (e.g. Musschoot & Lalèyè 2008; Day *et al.* 2013).

4.2 Hidden diversity of S. schall*-like catfish of the Nilo-Sudanian ichtyological province*

Unlike the morphological determination, the results of phylogenetic and haplotype analyses were strikingly straightforward. Phylogenetic analyses of both mitochondrial markers equivocally revealed *S. schall* sensu stricto and *S. schall* sensu lato clades, each comprising a mixture of samples assigned by different authors to *S. schall*, *S. frontosus* or other taxa. Regardless of GenBank taxonomic assignment, the two well supported clades, *S. schall* sensu stricto and sensu lato, exclusively comprise either samples from eastern part of the Nilo-Sudanian ichthyological province (Nile, Turkana, Chad), or those from its western part (e.g. Senegambia, Niger, Chad), respectively. The taxonomically incongruent, but geographically non-random phylogenetic pattern reveals a strong biogeographical signal among the supposedly homogenous, conspecific and widely distributed *S. schall*-like catfish. By chance, the combined Nile-Turkana and GenBank dataset comprises all major limnic systems of the Nilo-Sudanian province, i.e. rivers Gambia, Senegal, Volta, Niger, Nile (including Lake Albert), endorheic basins of Chad and Turkana, as well as Eburneo-Ghanian (sub) province. As such, the dataset is robust enough to allow reliable conclusions.

In contrast, phylogenetic analyses of the nuclear gene (*RAG2*) did not corroborate the results obtained from mitochondrial genes. Unfortunately, the lack of representative data from the Western Nilo-Sudanian province, together with the generally low variability of the gene did not allow clear distinction of mitochondrial groups. As a result *S. schall* clustered with several *Synodontis* spp. (e.g. *S. caudovittatus*, *S. serratus*, *S. violacea*) into a single unresolved (polytomic) clade. However, although not being informative, the nuclear pattern is not in conflict with mitochondrial data either. The analysis did not exhibit interspecific mitochondrial introgression, which would provide an alternative explanation to the apparent conspecificity of *S. schall* and *S. frontosa*. Complete fixation of mitochondrial DNA introgressed from one species to another is rare, but is sometimes found, for example recently in chipmunks (Good et al. 2015). Analyses of more variable nuclear loci would be required to completely exclude such a possibility in *S. schall* and *S. frontosa*. However, given the extent of the sampled populations in Nile and Turkana, the high sequence variability found in *cytb* and the uncertainty of morphological traits distinguishing the two species, we do not consider the "introgression alternative" probable.

Although long considered morphologically uniform and widely distributed throughout the Nilo-Sudanian ichtyological province from the Atlantic to Indian Ocean (e.g. Paugy & Roberts 2003; Musschoot & Lalèyè 2008), our results imply that *S. schall* sensu stricto, as defined by location of type locality in the Nile, is restricted to the province's eastern part, i.e. Nile, Turkana and Chad basins, and possibly elsewhere in the north-east Africa (Fig. 6). On the other hand, morphologically identical *S. schall* sensu lato from the western part of the Nilo-Sudanian province probably represents cryptic species apparently more closely related to congeners from the same region than to its eastern sibling. Interestingly, there seems to be an overlap of distribution ranges of the eastern "sensu stricto" and the western "sensu lato" forms in the Chad basin, probably as a result of independent and temporally distinct colonization events.

4.3 Conclusions

In conclusion, while *S. frontosus* might be merely a junior synonym of *S. schall*, populations of *S. schall* from the western part of Nilo-Sudanian province probably represent distinct species, which remained unrecognized due to a lack of morphological signal. A candidate name *Synodontis gambiensis* Günther, 1864 (Terra typica: Gambia), currently considered a junior synonym of *S. schall*, is available for eventual taxonomical amendment of the western “sensu lato” clade.

An interesting, rather unusual taxonomical case has emerged from this study, since concurrent taxonomical deflation (nominal species synonymization) and taxonomical split (cryptic species formal recognition) might be inevitable in the future for the long recognized congeners under scrutiny. First, the two nominal species *S. schall* and *S. frontosus* apparently represent a single polymorphic species (the species epithet *schall* having taxonomical priority over *frontosus*). Second, not only is *S. frontosus* conspecific with *S. schall* sensu stricto (as defined by the type locality), but also *S. schall*, as it is currently defined, apparently includes two non-sister phylogenetic units, i.e. cryptic species, hidden under this name.

Acknowledgements

Sampling in Kenya would not be possible without support from Kenya Marine and Fisheries Research Institute. Our thanks are due to Turkana fishermen, field assistants and chiefs, as well as Father S.J. Ochieng and Father Fernando Aguire, Our Lady Queen of Peace Todonyang Catholic Mission. Sampling in the Nile was possible thanks to White Nile Fisheries Research Station, Kostí. We thank Thomas Vigliotta, Cornell University, for comments on the manuscript. Authors greatly acknowledge Czech Science Foundation (P505/12/G112), Institute of Parasitology (RVO: 60077344) and Grant Agency of University of South Bohemia (037/2011/P, 04-135/2010/P). Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures".

5. References

Bailey RG. 1994. Guide to the fishes of the River Nile in the Republic of the Sudan. J. Nat. Hist.

Clement M, Posada D, Crandall K a. 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. [Internet] 9:1657–1659. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11050560

Day JJ, Peart CR, Brown KJ, Friel JP, Bills R, Moritz T. 2013. Continental diversification of an African catfish radiation (Mochokidae: *Synodontis*). Syst. Biol. 62:351–365.

Day JJ, Wilkinson M. 2006. On the origin of the *Synodontis* catfish species flock from Lake Tanganyika. Biol. Lett. 2:548–552.

Eschmeyer WN, Fricke R, van der Laan R. 2017. Catalog of fishes: genera, species, references. (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp). Electronic version accessed 22/07/2017.Eschmeyer WN, Fricke R, van der Laan R, editors.

Froese R, Pauly D. 2017. Fishbase. World Wide Web Electron. Publ. www.fishbase.org, version (02/2017). [Internet]. Available from: www.fishbase.org, version (02/2017)

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Hopson A. 1982. Lake Turkana; A Report on the Finding of the Lake Turkana Project 1972–1975, Vols. 1-6. Overseas Development Administration, London

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP. 2001. Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. Science (80-.). 294:2310–2314.

Ivanova N V., Zemlak TS, Hanner RH, Hebert PDN. 2007. Universal primer cocktails for fish DNA barcoding. Mol. Ecol. Notes 7:544–548.

Jirsová D, Štefka J, Jirků M. 2017. Discordant population histories of host and its parasite : A role for ecological permeability of extreme environment ? PLoS One:1–25.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30:3059–3066.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Koblmüller S, Sturmbauer C, Verheyen E, Meyer A, Salzburger W. 2006. Mitochondrial phylogeny and phylogeography of East African squeaker catfishes (Siluriformes: *Synodontis*). BMC Evol. Biol. 6:49.

Leigh JW, Bryant D. 2015. Popart: Full-Feature Software for Haplotype Network Construction. Methods Ecol. Evol. [Internet] 6:1110–1116. Available from: http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12410/abstract%5Cnhttp://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12410/full%5Cnhttp://onlinelibrary.wiley.com/store/10.1111/2041-210X.12410/asset/mee312410.pdf?v=1&t=ietwemkc&s=9b33a185ce363ca7148777c

Lévêque C, Paugy D, Teugels GG. 1991. Annotated check-list of the freshwater fisltes of the Nilo-sudan river basins , in Africa. Changes 24:131–154.

Musschoot T, Lalèyè P. 2008. Designation of a neotype for *Synodontis schall* (Bloch and Schneider, 1801) and description of two new species of Synodontis (Siluriformes: Mochokidae). J. Nat. Hist. 42:1303–1331.

Paugy D, Roberts TR. 2003. Mochokidae. p. In: C. Lévêque, D. Paugy and G.G. Teugels (eds.) Faune des poissons d'eaux douce et saumâtres de l'Afrique de l'Ouest, Tome 2. Coll. Faune et Flore tropicales 40. Musée Royal de l'Afrique Centrale, Tervuren, Belgique, Museum National d'Histoire Naturalle, Paris, France and Institut de Recherche pour le Développement, Paris, France. 815 p. p. 195–268.

Pinton A, Agnèse JF, Paugy D, Otero O. 2013. A large-scale phylogeny of *Synodontis* (Mochokidae, Siluriformes) reveals the influence of geological events on continental diversity during the Cenozoic. Mol. Phylogenet. Evol. [Internet] 66:1027–1040. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23274216

Poll M. 1971. Révision des Synodontis africains (famille Mochokidae). Ann. – Musée R. l'Afrique Cent. Zool. 191:1–497.

Posada D. 2008. jModelTest: Phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Sambrook J, Russell DW. 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harb. Lab. Press. Cold Spring Harb. NY [Internet]:999. Available from: http://books.google.com/books?id=YTxKwWUiBeUC&printsec=frontcover%5Cnpapers2://publication/uuid/BBBF5563-6091-40C6-8B14-06ACC3392EBB

Sands DD. 1983. Catfishes of the World. Vol. 2: Mochokidae. Dunure Enterprises, Dunure, Scotland

Sullivan JP, Lundberg JG, Hardman M. 2006. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. Mol. Phylogenet. Evol. [Internet] 41:636–662. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16876440

Werle E, Schneider C, Renner M, Völker M, Fiehn W. 1994. Convenient single-step, one tube purification of PCR products for direct sequencing. Nucleic Acids Res. 22:4354–4355.

Tables and figures

Table 1. Sampling localities and *Synodontis* spp. sample sizes.

Locality	Coordinates	n		
Nile River		*S. schall*	*S. nigrita*	*S. serratus*
White Nile in Kostí	13.172222 N, 32.672222 E	17	6	7
Blue Nile in Sennar	13.543611 N, 33.636667 E	21	2	3
Lake Turkana				
El-Molo Bay - saline	2.832222 N, 36.695833 E	45	absent	absent
Kalokol - medium salinity	3.558611 N, 35.915833 E	2	absent	absent
Kerio River delta – freshwater	2.974813 N, 36.173469 E	8	absent	absent
Central Island - medium salinity	3.495833 N, 36.040278 E	2	absent	absent
Todonyang, Omo delta - freshwater	4.451667 N, 35.94388 E	25	absent	absent
Totals (n)		120	8	10

Table 2. Distinguishing features of *Synodontis schall* and *Synodontis frontosus*; data obtained within the scope of this study compared with the following literature: [1]Hopson (1982), [2]Bailey (1994), [3]Paugy & Roberts. (2003), [4]Musschoot & Lalèyè (2008). Note the intermediate states in the present Nile-Turkana sample and its overlap with literature data for both *S. schall* and *S. frontosus*.

Assessed diagnostic features	This study - Nile-Turkana (figure references given)	Literature	
	S. schall	*S. schall*	*S. frontosus*
Humeral process ventral keel	usually present, more or less pronounced (1d-f,h), sometimes absent (1f,g)	non-keeled[1] vs. slightly-keeled[3] vs. usually non-keeled[4]	non-keeled[1,3]
Humeral process surface sculpturing	granulose with striated ventral keel	striated[1] vs. granulose[3]	granulose[1,3]
Humeral process shape	upper margin straight (1d), concave (1e) or wavy (1g), lower margin straight (1d,g,h) or convex (1e,f), posterior tip sometimes curved (1f)	pointed[3], upper margin straight or a little concave, lower margin straight or a little convexe[4]	deep, pointed[3]
Mandibular teeth	26-46 (21-40 in single row)	23-33[1], 24-32[2], 24-39[3], 18-32[4]	36-48[1,3], 33-48[2]
Pectoral spine anterior serrations	distinct, fine, sharp (1i), or barely discernible (1j)	Fine, sharp, obvious[1,3]	barely discernible[1,3]
Maxillary barbel membrane	none, rudimentary, or distinct dark-colored (1k-m)	none[1,3] vs. none or hardly visible[2,4]	distinct dark-colored[1,2,3]

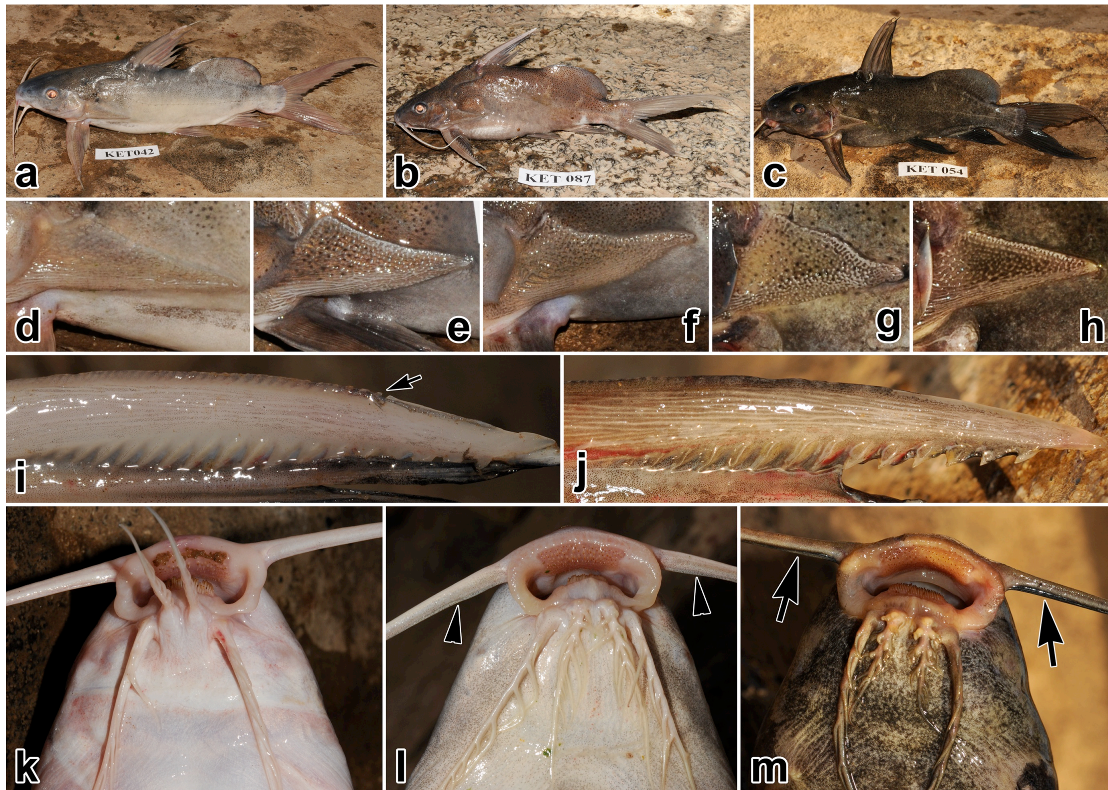


Figure 1. Live specimens of *Synodontis schall* from Lake Turkana, Kenya. Corresponding humeral processes, pectoral spines and mouthparts are shown for two specimens representing typical *S. schall* morphotype (Fig. 1a,d,i,k) and typical *S. frontosus* morphotype (1c,h,j,m); remaining pictures originating from different individuals show additional variation. **a.** Todonyang (Omo River delta); **b.** El-Molo Bay (southern Lake Turkana); **c.** Todonyang; **d-h.** humeral process variants (see Tab. 2 for details); **i.** pectoral spine with well visible fine sharp anterior serrations (starting from small arrow towards left side of picture); **j.** pectoral spine with anterior serrations barely discernible (partly due to pigmentation); **k.** maxillary barbel without membrane; **l.** rudimentary slat-like form (arrowheads); **m.** distinct dark membrane (arrows). Note: Depicted specimens show numerous dark specks in various body parts, but specimens almost lacking specks might occur too.

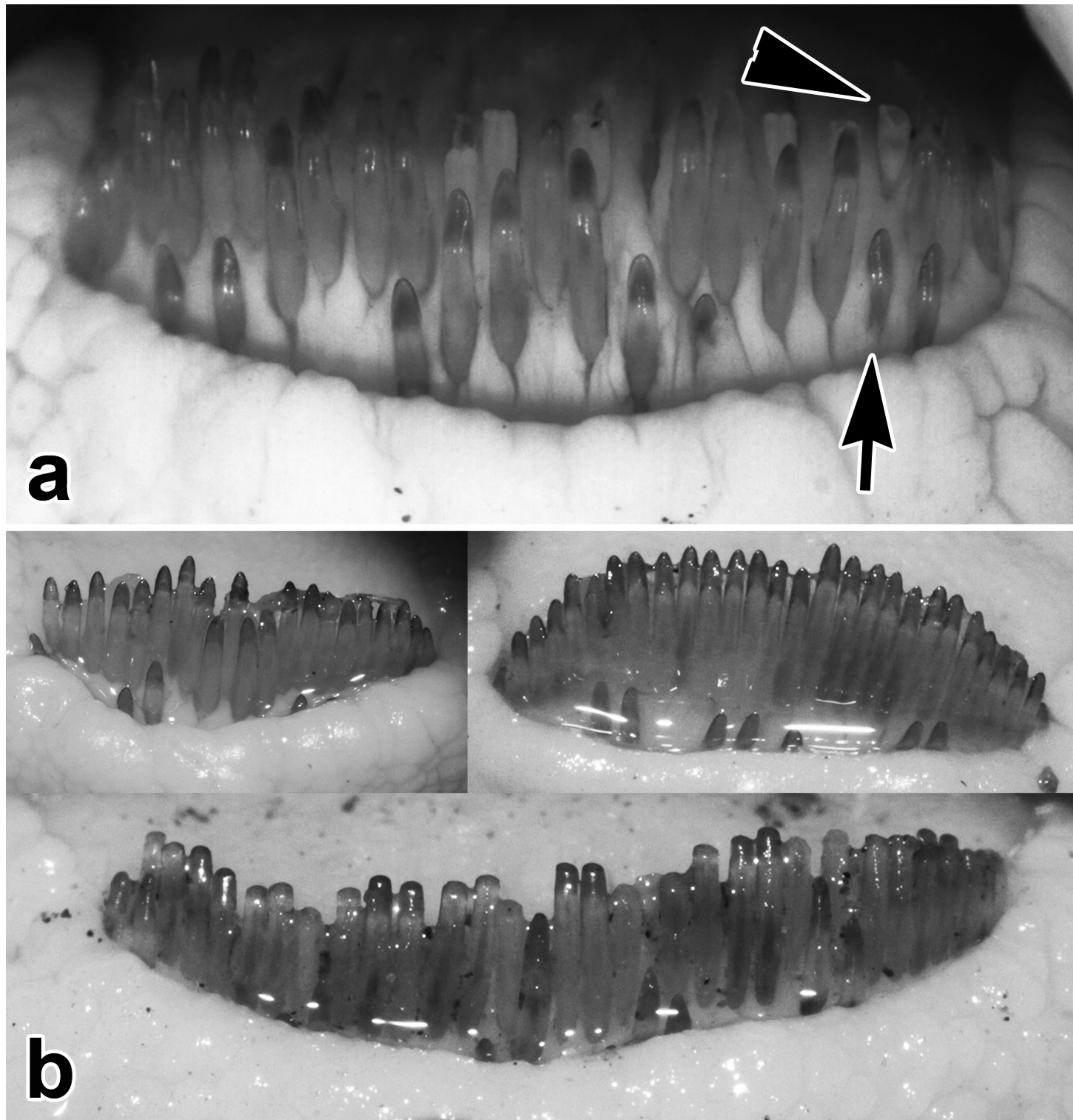


Figure 2. *Synodontis schall* mandibular dentition, Lake Turkana, Kenya. **a.** high magnification image showing new teeth emerging at the outer margin of the tooth line (arrowhead, only one marked), main row of fully grown teeth with sharp tips (unmarked), and old teeth that already lost their sharp tips positioned at the inner margin of the tooth row (arrow, only one marked); **b.** composite image showing variability of teeth arrangement and number: 28 (upper left), 37 (upper right), 39 (bottom).

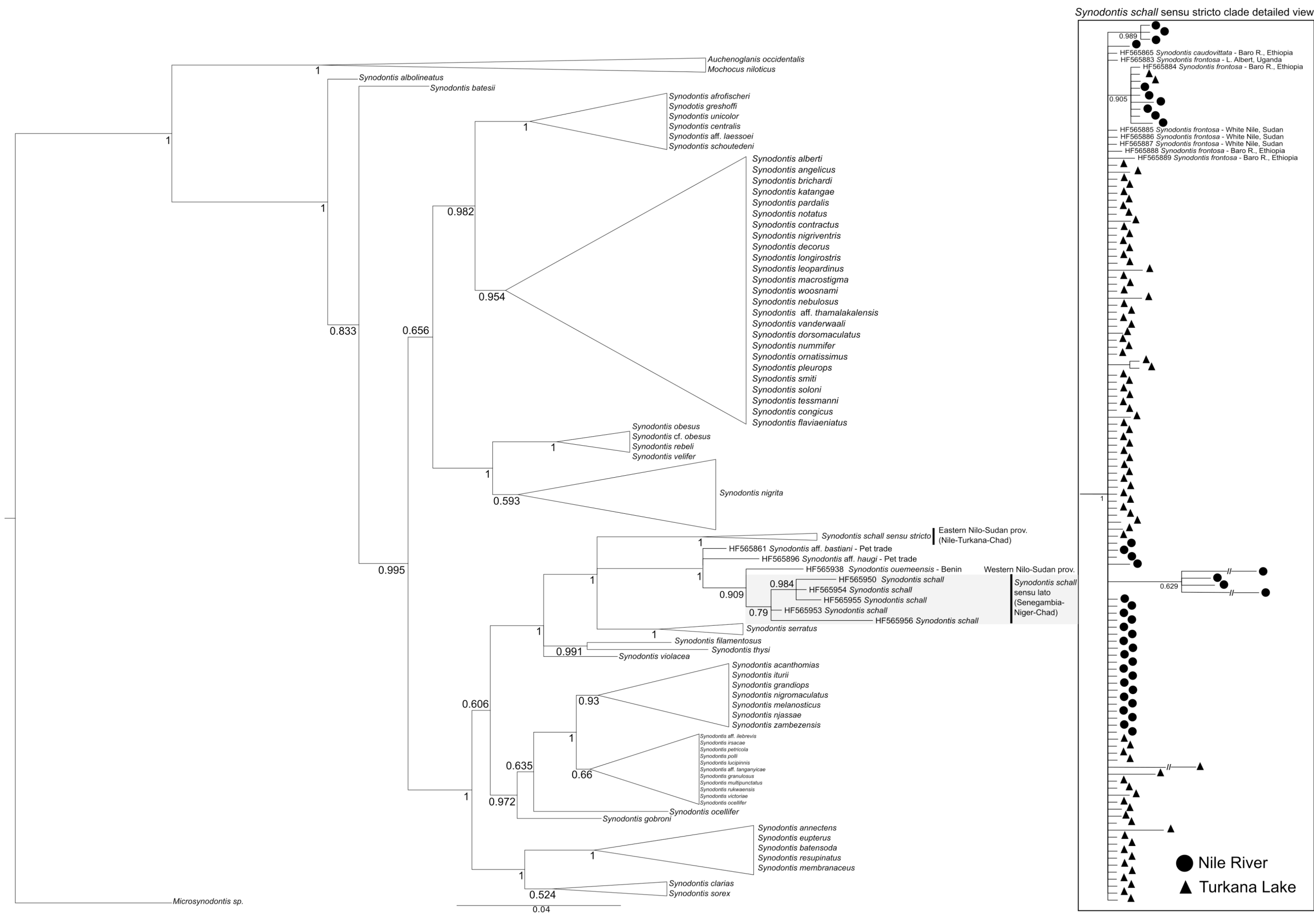


Figure 3. Bayesian phylogeny for *coxI.* Posterior probabilities are shown as branch supports Tree is rooted with *Microsynodontis* sp. sequence. Samples of *S. schall* sensu stricto are collapsed; detailed view of sensu stricto clade structure is presented with sampling localities. Valid names of *Synodontis* spp. following (Froese & Pauly, 2017)are used for convenience. Original assignments retrieved from GenBank, together with sequence accession numbers are provided in Figure S2 and Tables S1-S2.

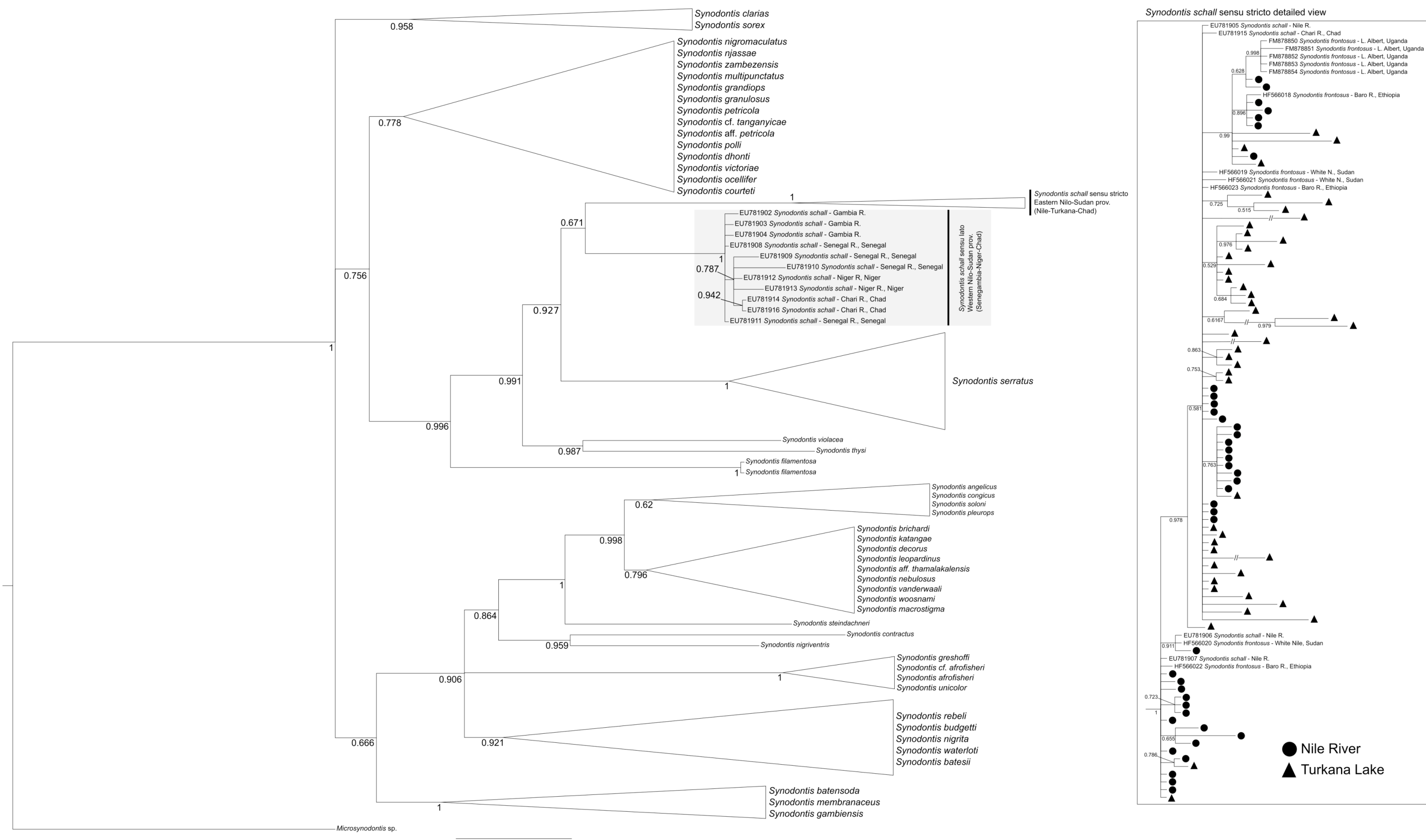


Figure 4. Bayesian phylogeny for *cytb.* Posterior probabilities are shown as branch supports Tree is rooted to *Microsynodontis* sp. sequence. Samples of *S. schall* sensu stricto are collapsed; detailed view of sensu stricto clade structure is presented with sampling localities. Valid names of *Synodontis* spp. following Froese & Pauly (2017) are used for convenience. Original assignments retrieved from GenBank, together with sequence accession numbers are provided in Figure S2 and Tables S1-S2.

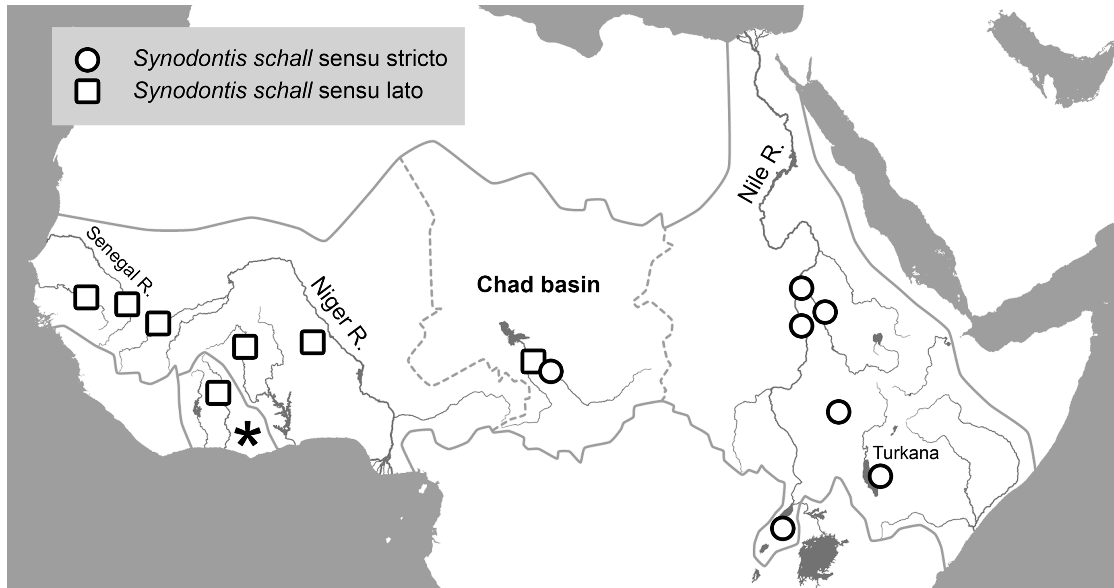


Figure 5. Map of the Nilo-Sudanian ichthyological province (solid grey line) showing geographical origin of *Synodontis schall* sensu stricto and sensu lato samples as revealed by phylogenetic analyses. Dashed line indicates western and eastern limits of Chad basin; asterisk indicates Eburneo-Ghanian province; only major lakes are shown outside the area of interest.

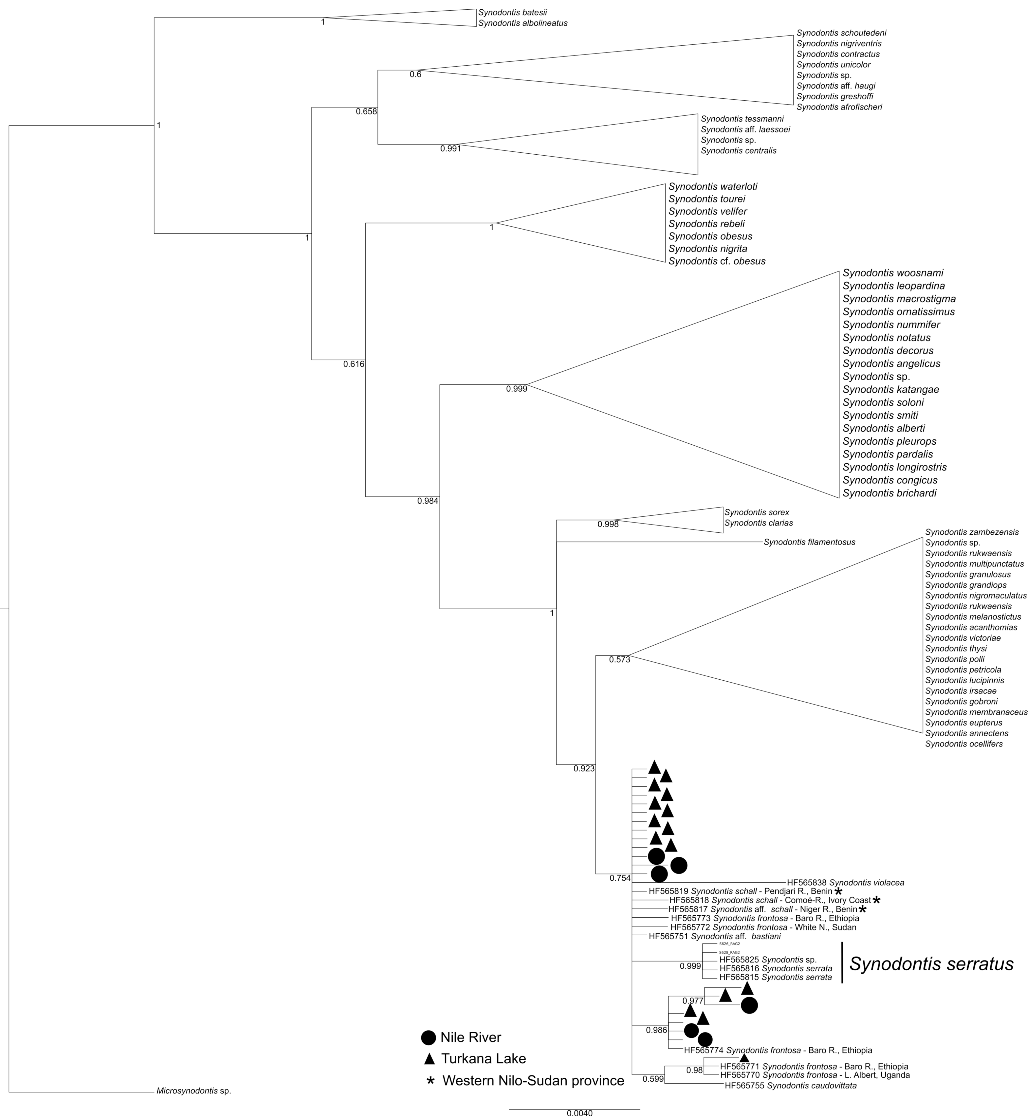


Figure 6. Bayesian phylogeny for *RAG2.* Posterior probabilities are shown as branch supports Tree is rooted with *Microsynodontis* sp. *Synodontis schall* sensu stricto clade is displayed in and branches for other *Synodontis* species are collapsed for better visualisation. Valid names of *Synodontis* spp. following Froese & Pauly (2017) are used for convenience. Original assignments retrieved from GenBank, together with sequence accession numbers are provided in Figure S3 and Tables S1-S2.

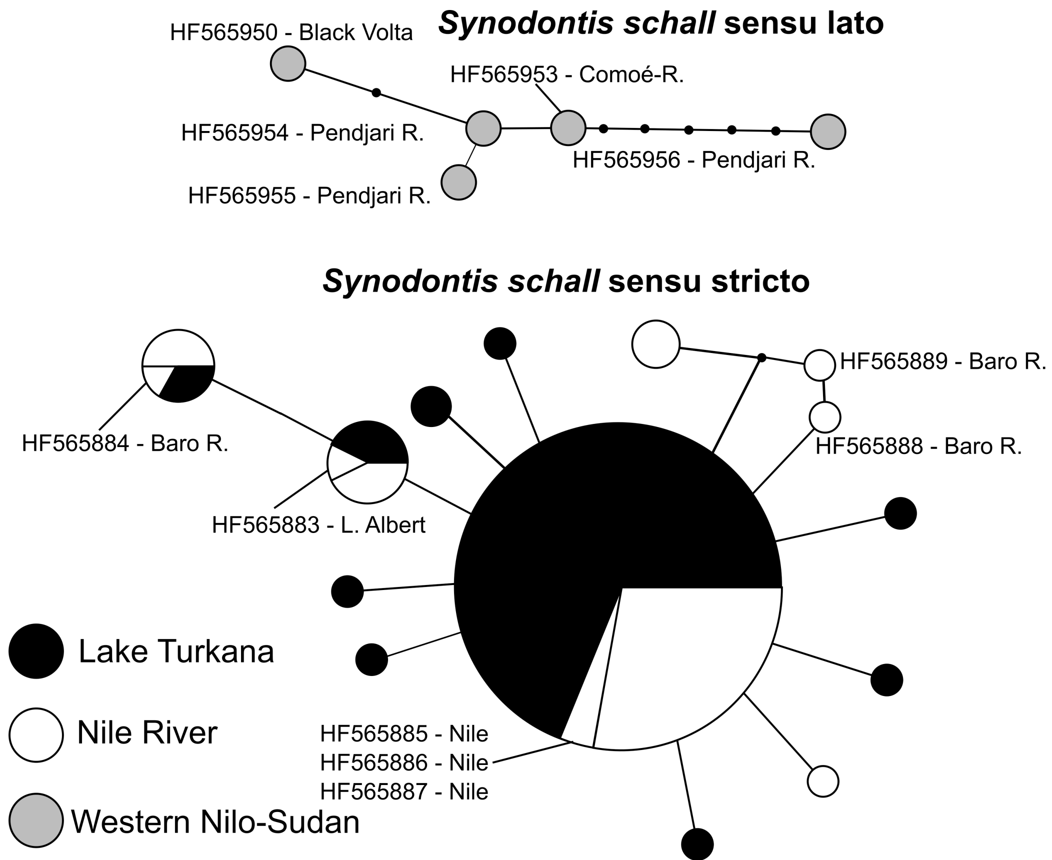


Figure 7. Haplotype networks for coxI constructed in software PopART v1.7. Networks were constructed only for samples clustering in S. schall sensu stricto and sensu lato clades. Sizes of haplo-nodes are relative to the sample size. The two networks were separated by 28 steps.

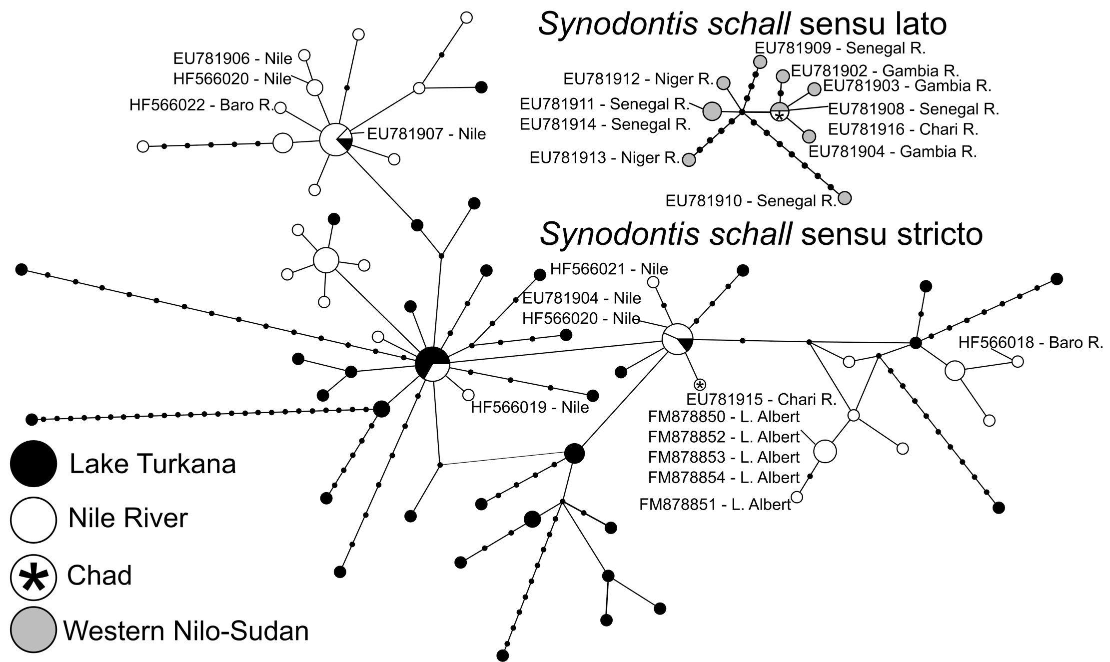


Figure 8. Haplotype networks for cytb constructed in software PopART v1.7. Networks were constructed only for samples clustering in S. schall sensu stricto and sensu lato clades. Sizes of haplo-nodes are relative to the sample size. The two networks were separated by 51 steps.

Supplementary data

Table S1. Sequences of *Synodontis* spp. from the Nile and Turkana deposited in GenBank, including detailed information on localities. Data provided in separate file.

Table S2. List of GenBank *coxI*, *cytb* and *RAG2 Synodontis* spp. sequences used in phylogenetic and haplotype analyses Data provided in separate file.

Table S3. Results of chi-square test for spatial distribution of *S. schall*, *S. frontosus* and intermediate morphotypes.

obs.		El Molo	Kalokol	Todonyang	totals(marginals)
	1	35	19	32	86
	2	16	11	31	58
	3	7	0	14	21
		58	30	77	165

obs. frekv.		El Molo	Kalokol	Todonyang	null (expected)
	1	0,603	0,63	0,415	0,52
	2	0,276	0,36	0,402	0,35
	3	0,121	0	0,181	0,13

expected counts		El Molo	Kalokol	Todonyang	
	1	30,23	15,63	40,13	
	2	20,38	10,54	27,06	
	3	7,38	3,81	9,8	

		El Molo	Kalokol	Todonyang	
χ2		0,753	0,723	1,648	
		0,944	0,019	0,571	
		0,019	3,818	1,8	
		1,716	4,561	4,019	**10,297**
P(=0,05; df=4)		**9,488**			

Figure S1. Phylogenetic trees for *coxI* a), *cytb* b) and *RAG2* c) as revealed by BI analyses, showing detailed branching with posterior probabilities as branch supports. Final trees were created by software MrBayes. All trees are rooted to *Microsynodontis* sp. sequence. Sequence data of *S. schall* sensu stricto obtained within the scope of this study are presented with GenBank numbers. Data provided in separate file.

Figure S2. Phylogenetic trees for *coxI* a), *cytb* b) and *RAG2* c) as revealed by ML analyses, showing detailed branching with bootstraps as branch supports. Final trees were created by software MrBayes. All trees are rooted with *Microsynodontis* sp. sequence. Sequence data of *S. schall* sensu stricto obtained within the scope of this study are presented with GenBank numbers. Data provided in separate file.

Paper 3

Morphological variability mirrors population genetic structure in fish tapeworm: a role for salinity in an isolated limnic system?

Miloslav Jirků[1], Jan Štefka[1,2], Dagmar Jirsová[1,2]

[1]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská 31, 370 05 České Budějovice, Czech Republic;

[2]Faculty of Science, University of South Bohemia, Branišovská 1760, 370 05 České Budějovice, Czech Republic

Key words: AFLP, Africa, Cestoda, *coxI*, Nile, phylogeny, Turkana, *Wenyonia virilis*

Correspondence: Miloslav Jirků; E-mail: miloslav.jirku@seznam.cz

Running title: Tapeworm morphology vs. population genetic pattern

Abstract

Recent research revealed markedly complex population structure in *Wenyonia virilis* (Cestoda, Caryophyllidae) parasitizing *Synodontis schall* (Siluriformes, Mochokidae) in Lake Turkana. To assess eventual matches between its phenotypic and population genetic variability, a total of 298 individuals was assigned to one of five predefined morphotypes; 167 coxI sequences (~988bp) and AFLP data from 233 individuals (716 fragments) were analyzed phylogeographically and by Principle Coordinate Analysis, respectively; results from both markers were cross-verified by AMOVA. Two undersampled morphotypes were excluded from analyses, while datasets of two other morphotypes showed no statistical differences, their datasets were pooled and treated as identical. The two remaining morphotypes mapped over population genetic data for both molecular markers showed strong correlation between morphotype distribution and population genetic structure. The analysis revealed clear affinity of both morphotypes to particular tapeworm subpopulations restricted either to the freshwater or saline habitats in Turkana. The effect of different hosts and geographic origin could be dismissed, therefore, the match between phenotype and population structure might indicate allopatric differentiation in a sympatric system in Lake Turkana.

Introduction

Morphological features defining individual species show some degree of variability, often reflecting certain life conditions. In parasites, the same genotype may produce different phenotypes under differing conditions, such as different host, host immune system, geographic origin, infection intensity (crowding effect), which might be reflected by phenotypic plasticity and/or fecundity (Stunkard 1957, Poulin 2007).

High level of phenotypic plasticity and/or intraspecific variability was repeatedly reported in numerous groups of parasites, including caryophyllidean cestodes (Hanzelová *et al.* 2015). These presumably the earliest evolved 'true' tapeworms (Eucestoda) are parasites of freshwater fish worldwide (except Neotropics) and are unique among eucestodes by the monozoic body plan (Mackiewicz 1994). In these studies, the morphological variability was usually found to be associated either with different geographical origin or specific host affinity (Barčák *et al.* 2014, Bazsalovicsová *et al.* 2014). In contrast, taxonomic significance of phenotypic variability was overestimated by multiple authors as reflected by extensive synonymy of some caryophyllidean taxa, especially in the Indomalayan Region (Ash *et al.* 2011, 2012).

The endemic African caryophyllidean genus *Wenyonia* Woodland, 1923 comprises six currently recognized species occurring exclusively in catfishes of the genus *Synodontis* Cuvier (Siluriformes: Mochokidae) throughout almost all Africa.

The type species of the genus, *Wenyonia virilis* Woodland, 1923, is the most widespread, being reportedly distributed across numerous basins from Gambia in West Africa, to the Nile and Turkana basin in east Africa and up to lower Nile in Egypt, north-east Africa (Schaeffner *et al.* 2011). In the eastern part of the Nilo-Sudanian ichthyological province, the dominant (and type) host species of *W. virilis* is *Synodontis schall* (Bloch et Schneider, 1801). Apart from the wide geographic range, other typical life-history traits of *W. virilis* include relatively high prevalence in most localities, and sometimes high abundances, i.e. tens of specimens per individual host (Schaeffner *et al.* 2011; unpublished data).

Unlike other congeners, *W. virilis* shows remarkable morphological variability both within and between different localities/populations. In *W. virilis* sample from the Nile and Turkana basins, Schäffner (2009) recognized five morphotypes (gradually differentiated rather than clearly distinct morpho-units well distinguished from each other). Although Schäffner (2009) considered the morphotypes conspecific and taxonomically non-significant, he documented the enormous variability *per se*. Subsequently, based on another extensive *W. virilis* sample from the same localities, Jirsová *et al.* (2017) revealed remarkably structured genetic composition of *W. virilis* populations by identifying distinct subpopulations with no gene flow between them, which were restricted either to the freshwater or saline parts of Lake Turkana. This observation arose the question whether or not the marked phenotypic and genotypic patterns correlate.

Herein, we present an analysis of intraspecific variability in *W. virilis* by contrasting morphology of extensive tapeworm sample with genetic information from two molecular markers, mitochondrial gene for *cytochrome oxidase I* (*coxI*) and whole genome information using amplified fragment length polymorphism (AFLP). Another congeneric and sympatrically occurring species, *Wenyonia minuta* Woodland, 1923 and *Wenyonia youdeoweii* Ukoli, 1965 were included in the phylogenetic analysis to confirm their distinctiveness. By comparison of molecular and morphological data we assess matches between morphological and population genetic variability in an attempt to better understand the population-level patterns in a recently isolated tapeworm population.

Material and methods

Sampling and morphotype assignment

All tapeworm samples in this study originate from a single host species *Synodontis schall* (Bloch et Schneider, 1801) (Siluriformes: Mochokidae), from Lake Turkana in Kenya and the Nile River in the Sudan; individual hosts were identified molecularly and morphologically (see Jirsová *et al.*, **Paper 2**). Upon dissection, a small piece of tissue from the tapeworm caudal (postovarian) region was cut off and stored in 96% ethanol for DNA isolation (see Table 1 for sampling details). The remaining part of each specimen was fixed with hot 4% buffered formalin; subsequently, all specimens were stained with Mayer's hydrochloric carmine solution and mounted as permanent slides for morphological identification (Scholz & Hanzelová 1998). All individual *W. virilis* specimens were assigned to five morphotypes defined originally by Schäffner (2009). The morphological traits used for morphotype assignments included body shape, scolex morphology and arrangement of testes (see Table 2 for details). Several traits used by Schäffner (2009) for morphotype characterization were not considered due to their tendency to reflect either conditions of fixation procedure (widest body part), or degree of ontogenetic development (character of testicular and uterine regions, nature of ovarian arms); distribution of postovarian vitelline follicles was not considered because parts of the postovarian regions of all analyzed specimens were collected for DNA extraction. Voucher specimens were deposited in the Helminthological Collection, Institute of Parasitology, Biology Centre of the CAS, České Budějovice, Czech Republic, Cat. Nos.: *W. minuta* IPCAS C-571, *W. virilis* IPCAS C-503, *W. youdeoweii* IPCAS C-573.

Model limnic system, Lake Turkana in arid northern Kenya, is an endorheic lake in eastern arm of the East African rift system. The limnic system of Turkana is composed of two contrasting environments: a freshwater part along the Omo River delta in the very north of the lake, and the saline (brackish, alkaline) central and southern parts of the lake (see **Paper 1** for details).

Molecular markers

Genomic DNA was extracted using a standard phenol chloroform extraction method (Sambrook & Russell 2001). Sequences of *coxI* and AFLP data for this study were adopted from previously published dataset by Jirsová *et al.* (2017). As we were unable to obtain corresponding DNA and morphological data for all specimens, only those samples from the original dataset assigned to morphogtypes with at least one molecular marker (*coxI* and/or AFLP) were included in final analyses (see Table 1 for sample sizes). All sequences used in this study were deposited in GenBank (accession numbers: MF563685 – MF563852 *W. virilis*, MF563853 – MF563868 *W. youdeoweii*, MF563869 – MF563872 *W. minuta*).

Morphotype distribution analysis

Genetic differentiation between and among morphotypes was assessed by analysis of molecular variance analysis (AMOVA) using the software package Arlequin v3.5 for *coxI* data (Excoffier *et al.* 2005) and AMOVA (Infinite Alleles Model) in software GenAlex v6.5 for AFLP data (Peakall & Smouse 2012). Calculations were performed only for samples with determined morphotypes; therefore, juvenile samples were omitted (Table 1). Both datasets were split into five groups according to their morphotype determination and F_{ST} pairwise analyses were conducted, with 10,000 permutations and significant value P= 0.05 for *coxI* data and estimations for AFLP data were based on 999 permutations and significant at P <

0.001. If the groups showed zero or negligible differences between sequence data, groups were combined into one bigger unit and additional analyses were performed. AMOVA was conducted only for Turkana samples, whereas samples from the Nile were not included due to small sample size.

Phylogenetic analysis

Sequential data were assembled and inspected for errors using GENEIOUS Pro software package version 6.1 (Kearse *et al.* 2012). Alignment was created using software MAFFT (Katoh *et al.* 2005). The best model of molecular evolution was selected using AIC criteria in software jMODELTEST 0.1.1. (Posada 2008). Phylogenetic trees were constructed using Bayesian inference (BI) in the program MrBayes v. 3.2.2. (Huelsenbeck *et al.* 2001; Ronquist & Huelsenbeck 2003). In MrBayes four independent runs were performed under the GTR + I + G model of molecular evolution, with following settings: 5 million MCMC generations and 4 chains. First 500 thousands replications were discarded as a burnin. Convergence of runs was checked in TRACER (Rambaut & Drummond 2009). Final tree was visualised and graphically processed in FigTree v1.4.2 (Rambaut 2009). The presented tree was constructed in software MrBayes. Sequences of the cyclophyllidean cestodes *Taenia hydatigena* (FJ518620), *Bertiella* sp. (JQ771105) and *Rodentocestus* sp. (KJ778921), and the dactylogyrid monogenean *Euryhaliotrematoides grandis* (EU008788) were used as an outgroup to root the tree; available data on *W. virilis* from GenBank (JQ034111-4; Brabec *et al.* 2012) were used in this study as an comparison to our data.

AFLP data analysis

Analysis of the AFLP fragment data was performed using the GeneMapper Software v3.7 (Applied Biosystems). Thresholds were set individually for each primer combination and all peaks were exported, obtaining 6,038 loci for the *W. virilis* dataset. Output was transformed to a binary format using Microsoft Excel 2011 and then converted to GenePop file format using GenAlEx v6.5 (Peakall & Smouse 2012). Selectively neutral alleles suitable for analysis of population structure were filtered from loci under putative selection using MCHEZA (Antao & Beaumont 2011). The initial analysis was performed with 100,000 simulations and with default settings for confidence intervals, Θ, β-a, β-b and critical frequency. Candidates included for balancing and positive selection were removed from the dataset and the same simulation was run again. This procedure resulted in 716 loci. Principal Coordinate Analysis (PCoA), which was based on Nei's genetic distances, was computed in GenAlEx to visualize any emerging patterns of genetic structure in populations.

Results

Morphotype distribution analysis

Results of *coxI* AMOVA analyses showed that there is no statistically significant difference between sequence data from tapeworms of morphotypes D, G and H and between E and F morphotypes (Table 3). Based on this result, the sufficiently sampled morphotypes D and G were considered intra-morphotype variability, their datasets were pooled and involved in following analyses. In contrary, undersampled morphotypes F (n=4) and H (n=6) were eliminated from the general dataset to avoid errors associated with small sample sizes. Resulting analyses of the sequence data on D+G and E morphotypes showed significant

difference between the two datasets (Table 3a). Resulting distribution of the morphotypes D + G and E mostly correlated with the population structure obtained by phylogenetic analyses, implicating their non-random geographic affinity either to the saline or freshwater part of Lake Turkana.

The AMOVA analyses for AFLP data based on F_{ST} calculation (Infinite Alleles Model) revealed that the majority of variation occurred within the sample sites for all datasets (94.3% Tab. 3c). This result provides further evidence of non-random, geographically and population-genetically correlated distribution of major morphotypes in Turkana.

Phylogenetic analysis

We used 167 *W. virilis* sequences, four sequences of *W. minuta* and 16 sequences of *W. youdeoweii*. Samples of *W. youdeoweii* clustered together in a branch with low support within *W. virilis* crown. This pattern could be attributed to the nature of mtDNA, which is more prone to introgression and lineage sorting than nuclear data. Importantly, the position of *W. youdeoweii* did not affect information on relationships between *W. virilis* samples. Mitochondrial data confirmed the presence of three different *Wenyonia* spp. in our dataset (Fig. 1), namely *W. virilis*, *W. youdeoweii* (as a long branch within *W. virilis*) and *W.minuta*. Samples of *W. virilis* created four clusters corresponding with the sampling localities. Sequences of *W. virilis* from GenBank originating from the Nile clustered with our Nile samples. The morphotype D + G was most prevalent in Turkana freshwater locality, the prevailing morphotype for Turkana saline locality was morphotype E (Fig. 1; see Fig. S1 for tree with the five original morphotypes mapped).

AFLP analysis

Morphological and AFLP data are presented for 233 *W. virilis* individuals. Sampling localities and morphotypes were mapped on PCoA results (Fig. 2a, b; see Fig. S2 for the five original morphotypes mapped). Highly prevalent morphotypes at Turkana saline locality was morphotype D+G, whereas E was the most abundant morphotype in Turkana freshwater locality.

Both types of molecular markers (*coxI* and AFLP data) corroborate high prevalence of main morphotypes D + G and E in the freshwater and saline parts of Turkana, respectively. No clear conclusion about morphotypes from the Nile could be made due to the relatively lower number of obtained *W. virilis* samples. Moreover, the final *coxI* results are comparatively less convincing than AFLP data due to the lower number of sequences available.

Discussion

Apart from one morphotype which was not identified in the Nile sample (F), all morphotypes recognized previously by Schäffner (2009) were recorded in both the Nile and Turkana basins. In addition, all morphotypes were identified in every single sampling locality in Turkana, implying that geographic origin or environmental factors are not *a priori* decisive with respect to the presence/absence of individual morphotypes in particular basin and/or in different parts of Lake Turkana. Performed analyses revealed significant correlation between the population genetic structure in Turkana and distribution of two major morphotypes (D+G and E). Although not exclusively, the morphotype D+G characterized by a very elongate and narrow body, with a gradual scolex-body interface (scolex edge), was mostly restricted to the

freshwater part of Lake Turkana. In contrast, the morphotype E with a wide robust body and abruptly delimited scolex-body interface was mostly restricted to the saline (brackish) part of the lake.

Although the scolex edge, i.e. scolex-body interface, might not seem particularly relevant and 'measurable' feature, its two alternative states present in *W. virilis*, i.e. gradual vs. abrupt, are clearly discernible by both light (Fig. 1) and scanning electron microscopy (see Fig. 4a, d in Schaeffner *et al.* 2011, for gradual scolex edge, and their Fig. 4b,c,e for abrupt scolex edge). Scolex is one of the most prominent and functionally significant features of tapeworm external anatomy, facilitating essential attachment to host intestinal mucosa. Scoleces are variable, providing taxonomically important features used commonly for characterization of cestode taxa. As such, scoleces are likely to be prone to environmentally driven functional and perhaps morphological changes. Therefore, it is possible that the scolex variations observed in Turkana *W. virilis* morphotypes might be of evolutionary significance.

In the past, two to three *Synodontis* spp. were morphologically identified in Turkana (e.g. Hopson 1982 and references therein; unpublished data). Just recently, the presence of only a single species of *Synodontis* in Lake Turkana has been confirmed (**Paper 2**). As a result, multiple host misidentifications confusing literature information on host spectra of multiple Turkana fish parasites should be taken into account (de Chambrier *et al.* 2011; Schaeffner *et al.* 2011; Kuchta *et al.* 2012; Mašová 2012; Moravec & Scholz 2017). This observation highlights the importance of proper host determination by combination of morphological and molecular tools whenever uncertainties emerge. *Synodontis schall* is the only member of the genus present in Lake Turkana and thus the only host available for *W. virilis* in that lake (**Paper 2**). This evidence rules out the influence of different host species on the documented population-genetic and/or phenotypic patterns in the Turkana *W. virilis* population.

Being ubiquitous and abundant in all parts of the lake, as well as in terms of biomass and ecological significance, *S. schall* is one of dominant components of Turkana's fish community, making the lake's ichthyofauna unique among those of other great African lakes (Hopson 1982; Muška *et al.* 2012). There are no barriers restricting *S. schall* and its migrations between saline and freshwater habitats of Lake Turkana have been reported (Hopson 1982). This is apparently well reflected by Turkana's genetically non-structured *S. schall* population sharply contrasting with remarkably structured population of its *W. virilis* parasite **(Paper 1, Paper 2**). Therefore, influence of limited host dispersal on population-genetic pattern of its parasite can be ruled out.

Interestingly, all individual *S. schall* examined in Turkana harboured only *W. virilis* belonging either to the freshwater or saline tapeworm subpopulation (**Paper 1**), implying that whenever *S. schall* gets infected in the lake, only one reproducing subpopulation of *W. virilis* (freshwater or saline) is present in a given area. In contrast, sporadic occurrences of freshwater *W. virilis* in the saline part of the lake, and vice versa, indicate bidirectional host mediated dispersal of tapeworms from both subpopulations between the freshwater and saline parts of the lake. Therefore, there is obvious spatial overlap in the occurrence of individuals from freshwater and saline *W. virilis* subpopulations within the lake. Despite this, there is no apparent gene flow between the subpolulations, which are genetically and, as shown here, also morphologically divergent (**Paper 1**; this study). By ruling out the influence of both different host (only one present) and spatial isolation (of the freshwater and saline tapeworm subpopulations), we speculate that the salinity gradient dividing Turkana's limnic system into two ecologically contrasting environments is the single main factor responsible for the pattern.

Our results imply that morphological variability in caryophyllidean tapeworms, which appears taxonomically non-significant, i.e. representing intraspecific variability, might actually reflect population-genetic structure and/or phenotype plasticity. Among caryophyllidean tapeworms, considerable intraspecific morphological variability mirroring phylogenetic pattern has been described in the genus *Caryophyllaeus* Gmelin, 1790 from Palaearctic cyprinids (e.g. Barčák *et al.* 2014; Bazsalovicsová *et al.* 2014; Hanzelová *et al.* 2015). However, unlike in *W. virilis*, variability of species of *Caryophyllaeus* corresponded to different hosts, a pattern markedly different from *W. virilis* populations studied, which are restricted to a single host in a single lake.

In addition to the lack of gene flow between the Turkana freshwater and saline tapeworm subpopulations presented elsewhere (**Paper 1**), the results of this study reveal their morphological divergence. The coincidence further supports the hypothesis proposed recently by Jirsová *et al.* (**Paper 1**) that the population genetic, and as shown herein, also phenotypic, patterns in the Turkana *W. virilis* population reflect evolutionary processes involving allopatric differentiation (freshwater vs. saline environments) in a sympatric system (same host species with unlimited dispersal, same limnic system).

The documented level of divergence within the Turkana population of *W. virilis* is striking. Taking into account the Late Pleistocene (near) desiccation followed by re-establishment of the lake during the Early-Mid Holocene between 11.5-5.2 thousand years ago (Bloszies *et al.* 2015), the remarkable *W. virilis* Turkana population patterns are result of only slightly over five millennia of isolation.

Conclusions

Morphotypes described previously in the fish tapeworm *Wenyonia virilis* were found at all sampled Turkana localities, implying that the geographic origin or environmental factors are not *a priori* decisive with respect to their presence/absence in a particular part of the lake. Distribution of major morphotypes showed a strong affinity to genetically distinct tapeworm subpopulations restricted either to the freshwater or saline parts of Turkana. The influence of different host species as well as restricted tapeworm dispersal can be reliably ruled out, implying a possible major role for salinity in the structuring of population genetic and phenotypic patterns documented in the Lake Turkana tapeworm population.

Acknowledgement

Our thanks are due to Bjoern C. Schaeffner (Laboratório de Helmintologia Evolutiva, Instituto de Biociências, Universidade de São Paulo) for provision of photographs of morphotypes of *W. virilis*. Authors greatly acknowledge the Czech Science Foundation (P505/12/G112), Institute of Parasitology (RVO: 60077344) and Grant Agency of University of South Bohemia (037/2011/P, 04-135/2010/P) for financial support. Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme “Projects of Large Research, Development, and Innovations Infrastructures”.

References

Antao T, Beaumont MA (2011) Mcheza: a workbench to detect selection using dominant markers. *Bioinformatics (Oxford, England)*, **27**, 1717–8.

Ash A, Scholz T, de Chambrier A *et al.* (2012) Revision of *Gangesia* (Cestoda:

Proteocephalidea) in the Indomalayan Region: Morphology, Molecules and Surface Ultrastructure. *PLoS ONE*.

Ash A, Scholz T, Oros M, Levron C, Kar PK (2011) Cestodes (Caryophyllidea) of the Stinging Catfish *Heteropneustes fossilis* (Siluriformes: Heteropneustidae) from Asia. *Journal of Parasitology*.

Barčák D, Oros M, Hanzelová V, Scholz T (2014) Phenotypic plasticity in *Caryophyllaeus brachycollis* Janiszewska, 1953 (Cestoda: Caryophyllidea): Does fish host play a role? *Systematic Parasitology*, **88**, 153–166.

Bazsalovicsová E, Kralová-Hromadová I, Brabec J *et al.* (2014) Conflict between morphology and molecular data: A case of the genus *Caryophyllaeus* (Cestoda: Caryophyllidea), monozoic tapeworms of cyprinid fishes. *Folia Parasitologica*, **61**, 347–354.

Bloszies C, Forman SL, Wright DK (2015) Water level history for Lake Turkana, Kenya in the past 15,000years and a variable transition from the African Humid Period to Holocene aridity. *Global and Planetary Change*, **132**, 64–76.

Brabec J, Scholz T, Králová-Hromadová I, Bazsalovicsová E, Olson PD (2012) Substitution saturation and nuclear paralogs of commonly employed phylogenetic markers in the Caryophyllidea, an unusual group of non-segmented tapeworms (Platyhelminthes). *International Journal for Parasitology*, **42**, 259–67.

de Chambrier A, Scholz T, Mahmoud ZN, Mariaux J, Jirku M (2011) Tapeworms (Cestoda: Proteocephalidea) of *Synodontis* spp. (Siluriformes) in Africa: Survey of species and their redescriptions. *Zootaxa*, **14**, 1–14.

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

Hanzelová V, Oros M, Barčák D *et al.* (2015) Morphological polymorphism in tapeworms: redescription of *Caryophyllaeus laticeps* (Pallas, 1781) (Cestoda: Caryophyllidea) and characterisation of its morphotypes from different fish hosts. *Systematic Parasitology*, **90**, 177–190.

Hopson A (1982) *Lake Turkana; A Report on the Finding of the Lake Turkana Project 1972–1975, Vols. 1-6*. Overseas Development Administration, London.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. *Science*, **294**, 2310–2314.

Jirsová D, Štefka J, Jirků M (2017) Discordant population histories of host and its parasite : A role for ecological permeability of extreme environment ? *PLoS ONE*, **12**, 1–25.

Katoh K, Kuma KI, Toh H, Miyata T (2005) MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, **33**, 511–518.

Kearse M, Moir R, Wilson A *et al.* (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

Kuchta R, Burianová A, Ů MJ *et al.* (2012) Bothriocephalidean tapeworms (Cestoda) of freshwater fish in Africa, including erection of Kirstenella n. gen. and description of Tetracampos martinae n. sp. *Zootaxa*, **3309**, 1–35.

Mackiewicz JS (1994) Order Caryophyllidea van Beneden in Carus, 1863. In: *Keys to*

cestodes parasites of vertebrates (eds Khalil LF, Jones A, Bray RA), pp. 21–43. CAB International, Wallingford, U.K.

Mašová Š (2012) Structure of the cephalic end and eggs of female *Cithariniella khalili* Petter , Vassiliadès et Troncy , 1972 (Nematoda : Pharyngodonidae), a parasite of African fishes. , 115–118.

Moravec F, Scholz T (2017) Some nematodes , including two new species , from freshwater fishes in the Sudan and Ethiopia. , 1–19.

Muška M, Vašek M, Modrý D *et al.* (2012) The last snapshot of natural pelagic fish assemblage in Lake Turkana, Kenya: A hydroacoustic study. *Journal of Great Lakes Research*, **38**, 98–106.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*, **28**, 2537–2539.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Poulin R (2007) Evolutionary ecology of parasites: (Second edition). *Evolutionary Ecology of Parasites*.

Rambaut A (2009) FigTree, a graphical viewer of phylogenetic trees. *Institute of Evolutionary Biology University of Edinburgh*.

Rambaut A, Drummond AJ (2009) Tracer V1.5. *Available from http://beast.bio.ed.ac.uk/Tracer*.

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

Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*, 999.

Schaeffner BC, Jirků M, Mahmoud ZN, Scholz T (2011) Revision of *Wenyonia* Woodland, 1923 (Cestoda: Caryophyllidea) from catfishes (Siluriformes) in Africa. *Systematic parasitology*, **79**, 83–107.

Schäffner B (2009) Monozoic tapeworms (Cestoda: Caryophyllidea) of the genus *Wenyonia*, parasites of Synodontis catfishes in Africa. Universität Karlsruhe (TH).

Scholz T, Hanzelová V (1998) Tapeworms of the genus *Proteocephalus* Weinland, 1858 (Cestoda: Proteocephalidae), parasites of fishes in Europe. *Praha: Academia. Studie AV ČR*, 119 pp.

Stunkard HW (1957) Intraspecific Variation in Parasitic Flatworms. *Systematic Zoology*, **6**, 7–18.

Woodland WNF (1923) On some remarkable new forms of Caryophyllaeidae from the Anglo-Egyptian Sudan, and a recision of the families of the Cestodaria. *Quarterly Journal of Microscopical Science*, **67**, 435–472.

Tables and figures

Table 1. Sampling and relative occurrence of *Wenyonia virilis* morphotypes at individual localities.

			Nile						Lake Turkana			
	Totals		Atbarah River 14.9194 N, 35.9019 E		Blue Nile 13.5436 N, 33.6367 E		White Nile 13.1722 N, 32.6722 E		Turkana freshwater part 4.4517 N, 35.9439 E		Turkana saline part 2.8322 N, 36.6958 E	
No. of samples (*coxI*/AFLP)	298 (167/233)	%	2	%	5	%	33	%	123	%	136	%
Morphotype D	56 (30/44)	19	-	-	-	-	10	30	43	35	3	2
Morphotype E	85 (45/58)	29	-	-	2	40	3	9	10	8	70	52
Morphotype F	62 (4/40)	21	-	-	2	40	4	12	14	11	42	31
Morphotype G	55 (50/64)	19	-	-	-	-	10	30	30	24	15	11
Morphotype H	12 (6/11)	4	2	100	1	20	3	9	1	1	5	4
Juvenile	15 (13/15)	5,0	-	-	-	-	3	9	11	9	1	1

Table 2. Key traits used for assignment of *Wenyonia virilis* specimens to morphotypes *sensu* Schäffner (2009).

Morphotype	body shape	Scolex length/width	Scolex edge	Testes	Presence
D	very elongate/ very narrow	≥ 1	gradual	separated	Nile, Turkana
E	elongate/ very wide	< 1	abrupt	merged two layers	Nile, Turkana
F	elongate/ narrow	≤ 1	abrupt	merged	Turkana
G	very elongate/ very narrow	> 1	gradual	separated	Nile, Turkana
H	very short/ very wide	< 1	abrupt	compact multiple layers	Nile, Turkana

Table 3. Results of AMOVA based on FST for *Wenyonia virilis* coxI data a) and b) and for AFLP data c). Morphotypes F and H were excluded pro the following analysis b) due to the low number of samples and morphotypes D and G were combined owing to insignificant difference between the data. AMOVA for AFLP data c) showed significant difference within morphotypes.

a

	Morphotype D	Morphotype E	Morphotype F	Morphotype G	Morphotype H
Morphotype D	–				
Morphotype E	0.339	–			
Morphotype F	0.335	-0.021	–		
Morphotype G	0.018	0.24	0.142	–	
Morphotype H	-0.032	0.211	0.158	-0.034	–

b

	Morphotype D+G	Morphotype E
Morphotype D+G	–	
Morphotype E	0.29882	–

c

Source of varition	df	Sum of squares	Mean squares	Estimated variance	Explained variance (%)
All morphotypes					
Among sample sites	4	1334.812	333.703	6.367	5.627%
Within sample sites	183	19542.464	106.789	106.789	94.373%
Total	187	20877.276		113.156	100%

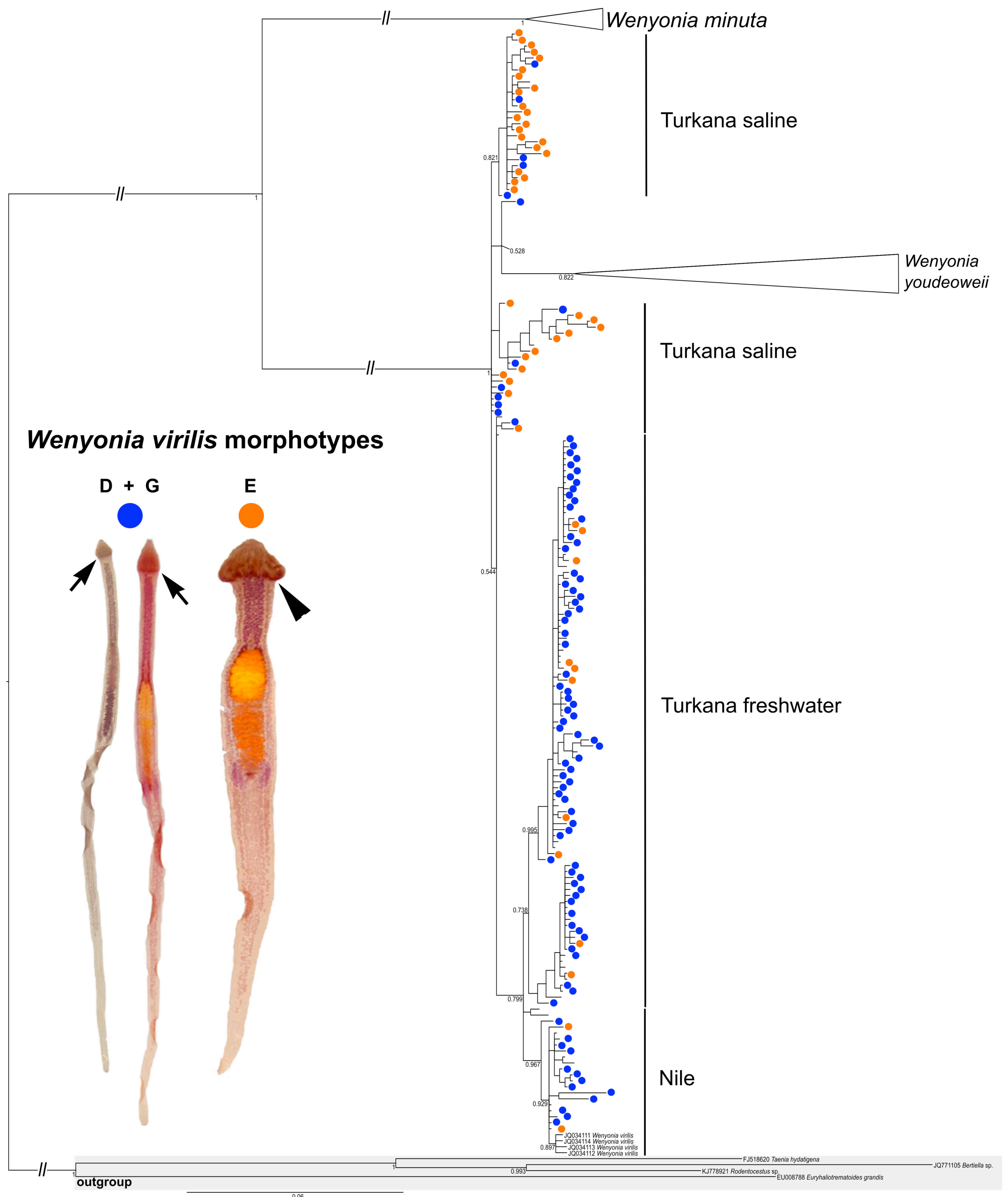


Figure 1. Phylogenetic analysis based on *coxI* gene. Final tree was created by MrBayes; nodal support was estimated as posterior probabilities. Samples of *W. virilis* were morphologically classified into five morphotypes, three of which are included in this figure; different colours correspond to different morphotypes or their groupings (morphotypes D + G and E). Note the markedly divergent body shapes (narrow-elongate vs. robust), and scolex shape (long-narrow vs. short-wide) showed by the typical representatives of particular morphotypes; highlighted is one of traits of apparent significance for morphotype distinction, i.e. gradual (arrows) vs. abrupt (arrowhead) scolex edge.

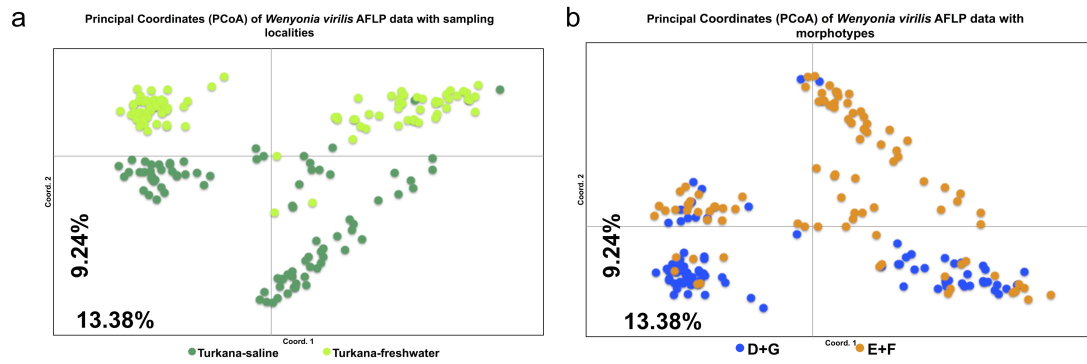


Figure 2. Results of PCoA of *Wenyonia virilis* population based on genetic distances and comparison of two different axes generated by GenAlEx software. Sampling localities (a) and morphotypes (D+G and E+F) (b) were mapped on PCoA results to assess possible correlation between morphology and/or geographical origin of the sampled individuals.

Supplementary material

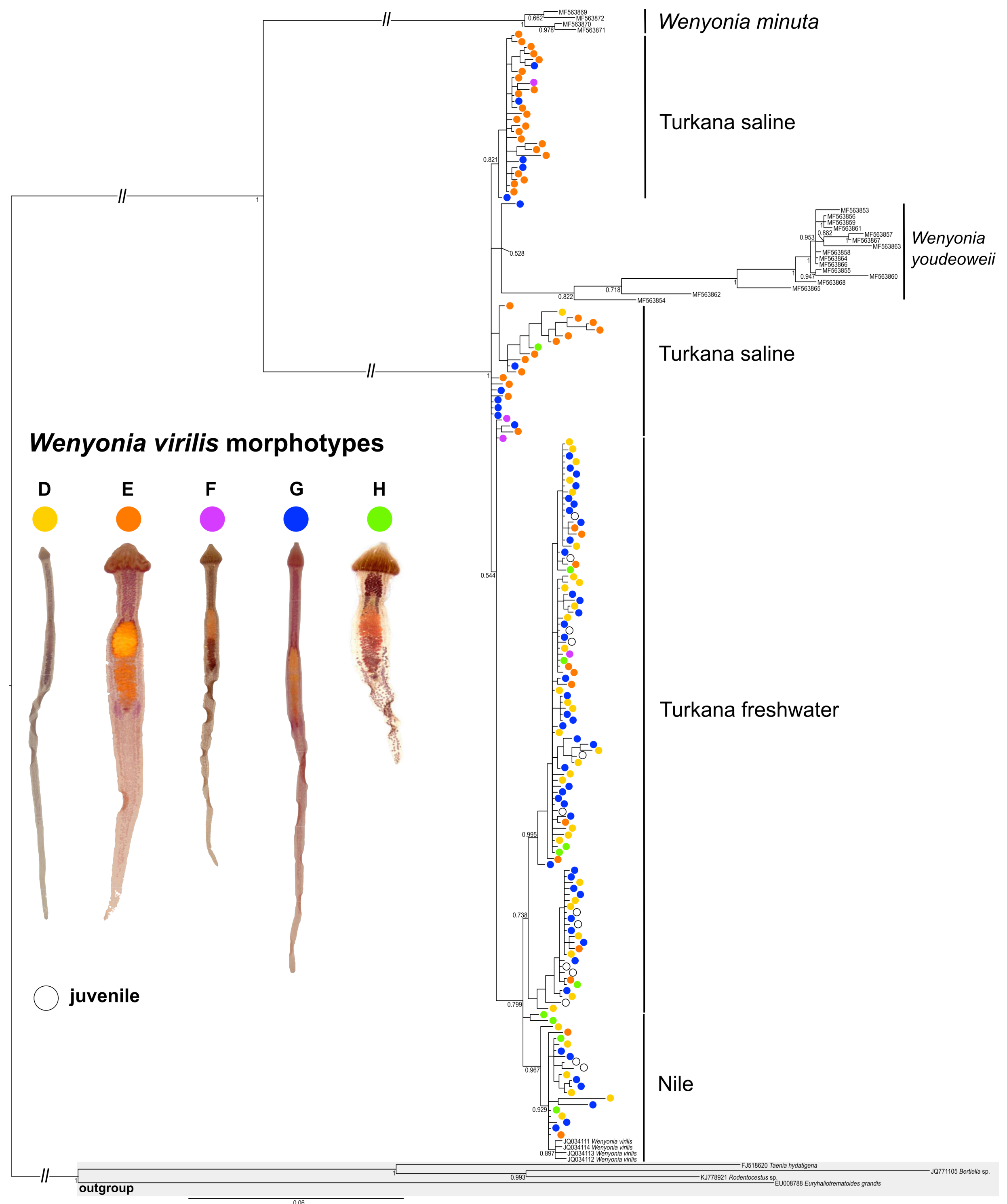


Figure S1. Phylogenetic analysis based on *coxI* gene. Final tree was created by software MrBayes, nodal support was estimated as posterior probabilities. Samples of *W. virilis* were morphologically classified into five morphotypes; different colours correspond to different morphotypes.

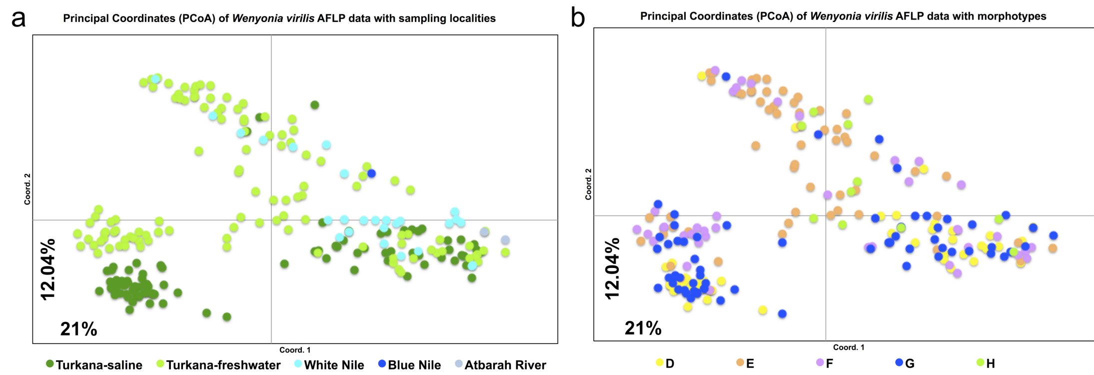


Figure S2. Results of PCoA *W. virilis* based on genetic distances and comparison of two different axes generated by GenAlEx software. Sampling localities (a) and all detected morphotypes (b) were mapped on PCoA results to assess eventual correlation between morphology and/or geographical origin of the sampled individuals.

Table S1. Detailed information about NCBI samples. Data provided in separate file.

Appendix I - Amplified Fragment Length Polymorphism (AFLP)

The molecular markers were used throughout the present study: *coxI* and AFLP. AFLP is not frequently used in parasitological studies; hence the brief summary of AFLP utility and technique principles is provided below.

I.I. AFLP utilization

The AFLP technique is a fingerprinting technique which allows the scanning of multiple loci across the whole genome for the presence of polymorphism(s) produced by point mutations, insertions, deletions, and/or other genetic rearrangements (Vos *et al.* 1995). AFLP has good reliability and reproducibility between laboratories (Gorni *et al.* 2004) and has a wide range of polymorphism detection (Paun & Schönswetter 2012). AFLP was first described by Vos et al. (1995) and since then has been used extensively to study plant genomes. However, since then, AFLP has been also used in genetic studies of several parasitic nematodes (De Gruijter *et al.* 2005; Nejsum *et al.* 2005; Höglund *et al.* 2012; Tydén *et al.* 2013). AFLP has mainly been used by the plant science community, particularly for the generation of polymorphic markers for genetic linkage mapping in a variety of plants (e.g. Blignaut *et al.* 2013; Andersen *et al.* 2016; Tausch *et al.* 2017). However, it has also been used for typing bacteria (e.g. Goudet & Keller 2002; Kütahya *et al.* 2011; van de Vossenberg *et al.* 2013). These markers detect DNA polymorphisms as specific restriction enzymes sites and enable these polymorphisms to be detected at multiple independent restriction sites simultaneously. The genetic similarity of individuals and populations can be inferred from the numbers of AFLP bands that they have in common. Additional information can be obtained by modifying the standard AFLP method to study gene expression and the map of genes. By ligating linkers to digests of cDNA, it is possible to compare the banding patterns of genes that have been expressed, as opposed to the entire genome (Freeland 2005).

Empirical studies in ecology and evolution often depend on accurate assessment of genetic diversity to address questions regarding genetic relatedness among individuals, population structure, phylogenetic relationships and mapping of quantitative trait loci (QTL). Series of techniques and genetic markers have been developed to estimate genetic diversity, but no single technique is universally ideal; each available technique exhibits both strengths and weaknesses. The main disadvantage of AFLP–PCR is the difficulty in identifying homologous markers (alleles), rendering this method less useful for studies that require precise assignment of allelic states, such as heterozygosity analyses. However, because of the rapidity and ease with which reliable, high-resolution markers can be generated, AFLPs are emerging as a powerful addition to the molecular toolkit of ecologists and evolutionary biologists (Mueller and Wolfenbarger 1999). The key feature of AFLP-PCR is its capacity for the simultaneous screening of many different DNA regions distributed randomly throughout the genome. AFLP methods allow the detection of polymorphisms of genomic restriction fragments by PCR amplification (Vos *et al.* 1995).

AFLP markers have been used to uncover cryptic genetic variation of strains, or closely related species, that had been impossible to resolve with morphological or other molecular systematic characters (Dasmahapatra *et al.* 2010; Valdez & Douhan 2012; Withrow & Winker 2014; Tausch *et al.* 2017). AFLP markers have found the widest application in analyses of genetic variation below the species level, particularly in investigations of population structure and differentiation and genetic variation within populations (Tandon *et al.* 2005; Troell *et al.* 2006; Dasmahapatra *et al.* 2008; Masumu *et al.* 2009) Apart from problems of population structure and variation, AFLP markers have been applied to evaluate gene flow and dispersal, outcrossing, introgression and cases of hybridization. The high resolution of AFLP

markers also enables testing for clonal identity between individuals (i.e. absence of recombination), and thus permits inferences about sexual versus asexual modes of reproduction (Mueller & Wolfenbarger 1999; Kollmann *et al.* 2000; Freeland 2005). The capacity of AFLP markers to resolve extremely small genetic differences has been demonstrated in several studies. Because AFLP method can generate many genomewide polymorphic markers without prior sequence knowledge, AFLP is a powerful tool for generating linkage maps. Although AFLP–PCR is not a panacea for all molecular problems in ecology and evolution, it still offers many advantages.

I.II. Principles of AFLP

AFLP is the selective amplification of restriction fragments from a digest of total genomic DNA using the polymerase chain reaction (PCR). The technique was developed by Zabeau & Vos (1993). With AFLP, molecular genetic polymorphisms are identified by the presence or absence of DNA fragments following restriction and amplification of genomic DNA. AFLP technique consists of five steps: DNA digestion, ligation, amplification, selective amplification and fragment analysis (Fig. I). Genomic DNA is first digested by two restriction enzymes. Double-stranded oligonucleotide adapters, homologous to one 5´- or 3´-end of overhangs generated during restriction digestion, are ligated to the DNA fragments. The ligated DNA fragments are amplified by PCR using primers complementary to the adapter and restriction site sequence with additional selective nucleotides at their 3' – end. The use of selective nucleotides reduces the complexity of the mixture to ~ 1/16. Only those fragments with complementary nucleotides extending beyond the restriction site will be amplified by the selective primers under stringent annealing conditions. This step is followed by selective amplification with primers containing 3 selective nucleotides at their 3'—end, as mentioned above primers are complementary to the adapter and restriction site. Forward primers are labeled with different fluorescent colours on their 5'—end to allow poolplexing in the final step, fragment analysis. Alternative subsets of loci can be amplified by using different primer combination with different selective bases. All amplified fragments are analyzed using capillary fragment analysis. The capillary instrument detects presents of fragments in the spectrum of each fluorophore aka fluorescent label, producing an electronic profile of relative fluorescence units versus fragment size, usual fragment size is from 50 bp to 500 bp. Polymorphisms, which are observed as presents or absents of peaks in compared samples, are caused by the gain or loss of a restriction site, a change (e.g. Single Nucleotide Polymorphism) in the selective primer binding site, or a length polymorphism (e.g. indel or variable microsatellite) between the restriction sites. Profiles from multiple individuals are aligned and scored based on the presence (1) or absence (0) of a peak, producing a binary data matrix (Vos *et al.* 1995; Bonin *et al.* 2005; Meudt & Clarke 2007).

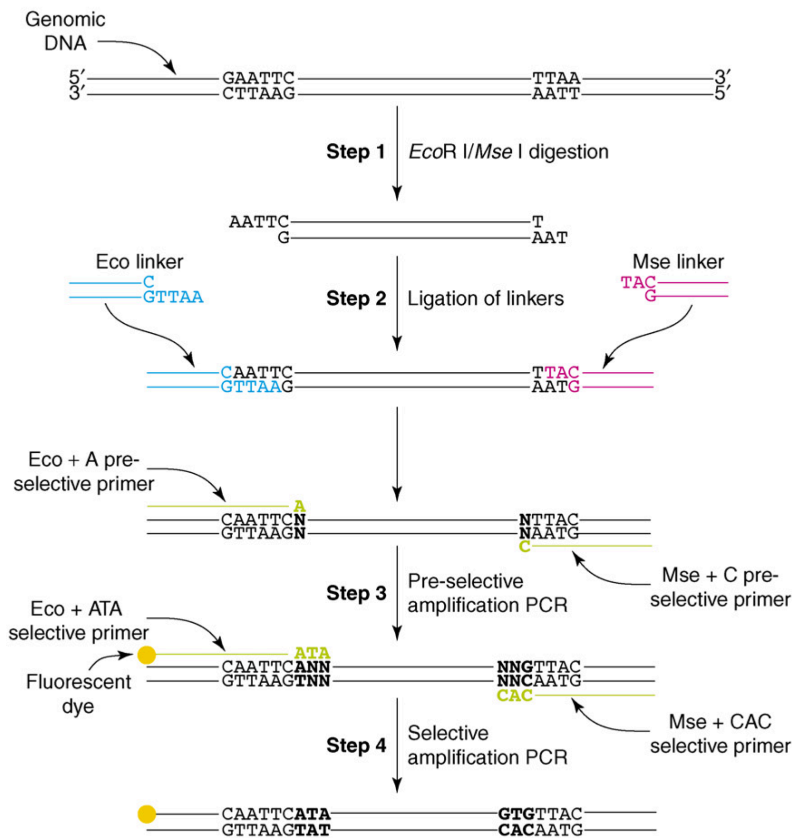


Figure I. Schematic representation of the AFLP technique. Four basic steps of AFLP are described in following order: digestion, ligation, amplification and selective amplification. Genomic DNA is digested by restriction enzymes and adapters are ligated to the restriction fragments. The selected subset of ligated fragments is amplified by PCR, using primers with one selective nucleotide at the 3′—end. This step is followed by selective amplification with primers containing 3 selective nucleotides at their 3'—end and forward primer is labeled with different fluorescent colours on their 5'—end. Adopted from Meudt & Clarke 2007.

Appendix II – Curriculum vitae

Dagmar Jirsová

Date of birth 16th May 1985

Place of birth Praha, Czech Republic

Nationality Czech

Address Department of Forest Botany, Dendrology and Geobiocoenology

Mendel University in Brno
Zemědělská 3, Brno 61300, Czech Republic
Tel: +420 545 134 048, e-mail: dasa.jirsova@gmail.com

Department of Botany and Zoology (Parasitology)

Faculty of Science, Masaryk University

Campus Bohunice, Kamenice 753/5 (A31, 314)

Brno 625 00, Czech Republic

Tel.: +420 549 498 841, e-mail: dasa.jirsova@gmail.com

Education

- **2004-2007:** B.S. in Biology at The University of South Bohemia in Faculty of Science, Department of Parasitology, České Budějovice, Czech Republic. **B.S. Thesis:** Phylogenetic relationships of trematodes of the families Opisthorchiidae and Heterophyidae.

- **2007-2009:** M.S. in Parasitology at The University of South Bohemia in Faculty of Science, Department of Parasitology, České Budějovice, Czech Republic. **MSc. Thesis:** Phylogenetic relationships of small liver flukes of the family Heterophyidae with a view to taxonomy of komplex of the family Ascocotyle.

- **2009-presence: Ph.D.** in Parasitology at The University of South Bohemia in Faculty of Science, Department of Parasitology, České Budějovice, Czech Republic. Ph.D. Thesis: Population genetics of the fish tapeworm *Wenyonia virilis* (Caryophyllidea: Caryophyllaeidae) and its fish host *Synodontis schall* (Siluriformes: Mochokidae).

Research Experience

- **2015-present:** Junior research assistant at Department of Botany and Zoology (Parasitology), Faculty of Science, Masaryk University, Brno, Czech Republic.
- **2015-present:** Research assistant and head of Molecular genetic laboratory at Mendel University in Brno Department of Forest Botany, Dendrology and Geobiocoenology, Brno, Czech Republic.
- **2014-2015:** Junior research assistant at University of Veterinary and Pharmaceutical

Sciences Brno, Department of Pathological Morphology and Parasitology, Czech Republic.

- **2010-2014:** Junior research assistant, Department of Helminthology, Institute of Parasitology, Biology Centre ASCR, České Budějovice, Czech Republic.

Research grants

- *CEITEC Project Open Access* (2016): The consequences of water scarcity and consequent irrigation on stress hormones expression of beech. (D.J. holder of the grant).
- *Grant Agency of South Bohemia University* (Grant n. 037/2011/P): *Real-time Evolution:* Using AFLP to study population genetics and demography of fish and their parasites in recently divided biogeographical units. (D.J. holder of the grant).

Fellowships

- Research stay in The Natural History Museum, London, UK; laboratory of Dr. Lukas Rüber, 1st March-3rd May, 2011.
- Research stay in The Natural History Museum, London, UK; laboratory of Dr. Lukas Rüber, 1st November- 12th December, 2011.

Fieldwork experience

- Sampling of Monogeneans, with a special focus on *Paradiplozoon hemiculteri*, China, 20th October – 11th November 2015.
- Sampling of great apes feces, Dja Reservation, Cameroon, 6th August – 16th September 2014.
- Sampling of fish parasites, Kenya, 28th October – 22nd December 2010.
- Sampling of fish parasites, Kenya, 28th October – 15th November 2009.

Publications

Jirsová D, Štefka J, Jirků M. Discordant population histories of host and its parasite : A role for ecological permeability of extreme environment ? *PLoS One*. 2017; 1–25. doi:10.5061/dryad.jq205

Lawrence AL, Hii S-F, **Jirsová D**, Panáková L, Ionică AM, Gilchrist K, et al. Integrated morphological and molecular identification of cat fleas (Ctenocephalides felis) and dog fleas (Ctenocephalides canis) vectoring Rickettsia felis in central Europe. *Vet Parasitol*. 2015;210. doi:10.1016/j.vetpar.2015.03.029

Hofmannová L, Romeo C, Štohanzlová L, **Jirsová D**, Mazzamuto MV, Wauters LA, et al. Diversity and host specificity of coccidia (Apicomplexa: Eimeriidae) in native and introduced squirrel species. *Eur J Protistol*. 2016;56: 1–14. doi:10.1016/j.ejop.2016.04.008

Gallusová M, **Jirsová D**, Mihalca AD, Gherman CM, D’Amico G, Qablan MA, et al. Cytauxzoon Infections in Wild Felids from Carpathian-Danubian-Pontic Space: Further Evidence for a Different Cytauxzoon Species in European Felids. *J Parasitol*. 2016;102: 1–5. doi:10.1645/15-881

Votýpka J, Rádrová J, Skalický T, Jirků M, **Jirsová D**, Mihalca AD, et al. A tsetse and tabanid fly survey of African great apes habitats reveals the presence of a novel trypanosome lineage but the absence of Trypanosoma brucei. Int J Parasitol. 2015;45. doi:10.1016/j.ijpara.2015.06.005

Kuchta R, Scholz T, Brabec J, **Jirsová D**, Gustinelli A. Bothriocephalidean tapeworms (Cestoda) from the blackfish, *Centrolophus niger* (Perciformes: Centrolophidae). *Folia Parasitol* (Praha). 2008;55:111-121.

Conference papers

- **Jirsová D.**, Šrámová E., Kašný M., Gelnar M. *Eudiplozoon nipponicum*: "Together - chance or destiny?" 16th Evolutionary Biology Meeting at Marseilles, Marseilles, France (2016).
- Jirsová D., Hájíčková M., Oppelt J., Geabuer R. Beech transcriptome data: “I am stressed!”. IUFRO ‘Genomics and Forest Genetics’ Conference, Arcachon, France (2016).
- Vorel J., Roudnický P., Ilgová J., Strnad H., Dvořáková H., Jedličková L., Mikeš L., Brož P., **Jirsová D.**, Strnad H., Leontovyč R., Dzika E., Koubková B., Gelnar M., Kašný M. Exploring the transcriptome of Eudiplozoon nipponicum (Monogenea, Diplozoidae) adult worm. XII. české a slovenské parazitologické dny, Ledeč nad Sázavou, Czech Republic (2016).
- Vorel J., Roudnický P., Ilgová J., Strnad H., Dvořáková H., Jedličková L., Mikeš L., Brož P., **Jirsová D.**, Strnad H., Leontovyč R., Dzika E., Koubková B., Gelnar M., Kašný M. Monogenea: Comparative study of molecules transcribed by mucophagous and hematophagous parasites of fish. In EMOP XII – the 12th European Multicolloquium of Parasitology, Turku, Finland (2016).
- Vorel J., Roudnický P., Ilgová J., Mikeš L., Dvořáková H., Jedličková L., **Jirsová D**, Strnad H., Leontovyč R., Dzika E., Koubková B., Vetešník L., Jurajda P., Gelnar M., Kašný M. Eudiplozoon nipponicum (Monogenea): the molecules transcribed by hematophagous parasite of fish. In 5th ECIP meeting - European Centre of Ichtyoparasitology, Prušánky, Czech Republic (2016).
- Kašný M., Ilgová J., Vorel J., Roudnický P., Mikeš L., Dvořáková H., Jedličková L., **Jirsová D**, Strnad H., Leontovyč R., Dzika E., Koubková B., Vetešník L., Jurajda P., Gelnar M. Monogenea: parasite-host interactions at molecular level. 5th ECIP meeting - European Centre of Ichtyoparasitology, Prušánky, Czech Republic (2016).
- Vorel J., **Jirsová D.**, Ilgová J., Roudnický P., Jedličková L., Dvořáková H., Leontovyč R., Mikeš L.,, Strnad H., Koubková B., Gelnar M., Kašný M. “Omics” and population genetic tools applied on selected species from the class Monogenea. In 4th ECIP meeting - European Centre of Ichtyoparasitology, Vodňany, Czech Republic (2015).
- **Jirsová D.**, Vorel J., Ilgová J., Roudnický P., Strnad H., Jedličková L., Dvořáková H., Skipalová K., Mikeš L., Brož P., Leontovyč R., Dzika E., Koubková B., Gelnar M., Kašný M. Monogenea: From sequences to molecules…, 4th Workshop of European Centre of Ichthyoparasitology, (2015).
- Vorel J., **Jirsová D.**, Ilgová J., Roudnický P., Strnad H., Jedličková L., Dvořáková H., Skipalová K., Mikeš L., Brož P., Leontovyč R., Dzika E., Koubková B., Gelnar M., Kašný M. Terra incognita – sequencing projects of monogeneans parasites. Grid

computing workshop, Brno, Czech Republic (2015).

- **Jirsova D.**, Gelnar M., Kašný M. (2015). Monogenea … what else can be done? 22nd Helminthological Days held at Straz nad Nezarkou, Czech Republic (2015).
- **Jirsová D.**, Ilgová J., Gelnar M., Kašný M. AFLP: handy for Monogenea! The 3rd ECIP meeting - European Centre of Ichtyoparasitology, Skalní mlýn 96, 678 01 Blansko, Czech Republic (2014).
- **Jirsová D.,** Štefka J., Jirků M., Real-time evolution: population genetics of fish and their parasites in recently divided biogeographic units, 20th Helmintological days, Štedronín u Orlíku, Czech Republic (2013).
- **Jirsová D.,** Štefka J., Jirků M., Real-time evolution: population genetics of fish and their parasites in recently divided biogeographic units, 16th Evolutionary Biology Meeting at Marseilles, Marseilles, France (2012).
- **Jirsová D.,** Štefka J., Jirků M., AFLP: Convenient Tool for Revealing Population Patterns and Evolutionary History, Jirsová D., Jirků M., Štefka J., 19th Helmintological days, Czech Republic (2011).
- **Jirsová D.,** Štefka J., Schäffner B., Jirků M., Real-time Evolution: Population genetics and demography of fish and their parasites in recently divided biogeographical units, 18th Helmintological days, Rožnov pod Radhoštěm, Czech Republic (2010).