Jihočeská univerzita v Českých Budějovicích
Přírodovědecká fakulta

Ph.D. Thesis

Host specificity, genetic variability and genealogy in populations of model parasite species

Jan Štefka

Supervisor: Prof. RNDr. Václav Hypša, CSc.
Co-supervisor: Prof. RNDr. Tomáš Scholz, CSc.

Laboratory of Molecular Phylogeny and Evolution of Parasites,
Institute of Parasitology, Biology Centre, ASCR

České Budějovice 2008

Ph.D. Thesis:

Štefka, J. 2008: Host specificity, genetic variability and genealogy in populations of model parasite species. Ph.D. Thesis, in English - 98 p., Faculty of Science, The University of South Bohemia, České Budějovice, Czech Republic.

Anotace:

The component of host specificity in parasites is considered one of the key features in respect to both evolutionary processes and epidemiology. Most parasitic species are restricted to narrow host spectrum. Consequently, the broad-host-range parasites often constitute unrecognized assemblages of morphologically identical but genetically isolated lineages with strict host specificity. The study analyzes the intraspecific structure, genetic variability and genealogy in the populations of two widespread parasites (tapeworm *Ligula intestinalis* and louse *Polyplax serrata*) and investigates the effect of host specificity and/or geographic isolation on genetic differences.

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb., v platném znění souhlasím se zveřejněním své disertační práce, a to v nezkrácené podobě – v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou – elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách.

České Budějovice, 15. 4. 2008

Jan Štefka

Acknowledgement/poděkování:

Rád bych poděkoval svým školitelům Václavu Hypšovi a Tomáši Scholzovi za školitelské vedení a rady, za poskytnuté zázemí a tvůrčí volnost při práci. Dík patří všem spoluobyvatelům Laboratoře molekulární fylogeneze a evoluce parazitů za přátelské prostředí. Dík patří také mé rodině, především Daně, za trpělivost a podporu při mém „nikdy nekončícím" studiu.

I would like to thank all my colleagues and friends abroad, who participated in the research carried out in the frame of the thesis. Namely Wafa Bouzid, Géraldine Loot and Sovan Lek from Toulouse, France and John Gilleard and his lab from Glasgow, Scotland for their help with *Ligula intestinalis* projects. Thanks belong also to Mišo Stanko and his colleagues from Košice, Slovakia, for help with obtaining *Polyplax serrata* louse material.

This work was supported by grants LC06073 and MSM600766605801 (Ministry of education, Czech Republic), projects 524/03/H133 and 524/04/0342 (Grant Agency of the Czech Republic), Institute of Parasitology (research projects Z60220518 and LC 522), and grants 12/2004, 37/2006 (University of S. Bohemia). Research stays in the laboratory of Dr. J. Gilleard (Glasgow) and in the Laboratoire Evolution Diversité Biologique (Toulouse) were funded by a Socrates Erasmus stipendium and by a Barrande project (MSMT), respectively.

Author agreement:

Part of this thesis is a manuscript accepted for publication:

Bouzid, W. Štefka, J*., Hypša V., Lek, S., Scholz, T., Legal, L., Ben Hassine, O.K., Loot G. 2008. Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). *Int. J. Parasitol.*

*Jan Štefka's contribution to this manuscript was equal to the contribution of Wafa Bouzid.

"I agree to the above statements"

Toulouse, France 10.4. 2008

Wafa Bouzid, MSc.

EVOLUTION & DIVERSITE BIOLOGIQUE-EDB
UMR 5174 - Université Paul Sabatier/CNRS
118 route de Narbonne - Bt 4R3 b2
31062 Toulouse cedex 9 - France
Tél. 05 61 55 73 28 - Fax 05 61 55 73 27

Stanovisko spoluautorů – prohlášení k publikační aktivitě Jana Štefky.

Prohlašuji, že Jan Štefka se převažující měrou podílel na vzniku níže uvedených publikací.

Štefka, J., Gilleard, J.S., Grillo, V., Hypša, V. (2007) Isolation and characterization of microsatellite loci in the tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea). Mol. Ecol. Notes, 7: 794-796.

Štefka, J., and V. Hypša. 2008. Host specificity and genealogy of the louse Polyplax serrata on field mice, Apodemus species: A case of parasite duplication or colonisation? Int. J. Parasitol., 38: 731-741.

České Budějovice, 15.4. 2008

Prof. RNDr. Václav Hypša, CSc.

Contents

1. Introduction

- **1.1** Evolution of host specificity 1
- **1.2** Demographic events in the course of evolution 5
- **1.3** Model organisms 7
- **1.4** Objectives 9

2. Summary of the results

- **2.1** The *Ligula* model 9
- **2.2** The *Polyplax* model 12

3. Conclusions and future prospects 13

4. References 14

Attached publications

- **Manuscript No. 1:** Bouzid, W., Štefka, J., Hypša V., Lek, S., Scholz, T., Legal, L., Ben Hassine, O.K., Loot G. 2008. Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). Int. J. Parasitol., in press 23
- **Manuscript No. 2:** Štefka, J., Gilleard, J.S., Grillo, V., Hypša, V. 2007. Isolation and characterization of microsatellite loci in the tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea). Mol. Ecol. Notes, 7: 794-796 53
- **Manuscript No. 3:** Microsatellite analysis of genetic structure of the endoparasite *Ligula intestinalis* in relation to biogeography and host specificity (to be submitted to Molecular Ecology) 57
- **Manuscript No. 4:** Štefka, J., Hypša, V. 2008. Host specificity and genealogy of the louse *Polyplax serrata* on field mice, *Apodemus* species: A case of parasite duplication or colonisation? Int. J. Parasitol., 38: 731-741 85

Curriculum vitae 97

1. Introduction

1.1 Evolution of host specificity

Parasitism represents probably the most widespread life-strategy. Parasites constitute an important component of all major ecosystems found on Earth, their presence affects the course of inter- and intra-specific competition between their hosts, and they thus take a significant part in balancing the ecological equilibrium and maintaining biodiversity of various environments (e.g. Freeland and Boulton, 1992; Yan, 1996; Hudson and Greenman, 1998; Poulin, 1999, Hernandes and Sukhdeo, 2008). The constant struggle between the host and parasite and their interaction with the surrounding environment are supposed to form two different strategies of parasitic life. Differing in the width of the host spectrum, parasites maximise their fitness either as generalists or specialists. Both strategies possess specific advantages and drawbacks depending on the conditions. Generalists achieve reproduction success through the capability to infect wide spectrum of host species, whereas specialists make better use of individual infections (Poulin, 1998; Poulin and Mouillot, 2004). Several cases of shift in the width of the host range during the course of evolution were identified by comparing host-parasitic associations on the macro-evolutionary scale. It was shown that both directions are possible; examples of widening (e.g. Stireman, 2005) as well as narrowing (e.g. Nosil, 2002; Poulin et al, 2006) the host spectra were provided.

Regardless the strategy no parasitic species is capable of infecting all available hosts and is always restricted to more or less confined range of hosts. Such restriction, called host specificity, is implicitly considered an inherent feature of parasitic organisms, as witnessed by the frequent inclusion of this parameter into the descriptions of parasitic species. The concept of host specificity is connected to the purely theoretical programs, such as investigation of speciation modes in free-living and parasitic organisms, and the origin of biodiversity, as well as to many practical aspects, for instance it has important epidemiological implications (Paterson and Viney, 2000; Galvani, 2003; Wilson et al., 2005). Despite that, the origin, degree and evolution of parasite specificity, remains largely unresolved and is one of the most debated topics in parasitology (Brooks and McLennan, 1993; Thompson, 1994; Clayton and Moore, 1997; Poulin, 2007). The issue includes two main components. At the long-time evolutionary scale, it is represented by co-evolutionary studies based on phylogenetic methods (Brooks and McLennan, 1993; Page and Holmes, 1998, Brooks and Ferrao 2005). From the point of view of population genetics, the gene flow and the genetic structure are of main interest, and the classical methods of population genetics are employed (e.g. Nadler, 1990; Nadler et al., 1995; Abderrazak et al., 1999; Leo et al., 2005; Wang et al., 2006).

1.1.1 Macro-evolutionary concept of co-speciation

Phylogenetic research in the area of host-parasite associations and host-specificity lead to the concept of co-speciation between the hosts and parasites. Traditionally, two scenarios of such evolutionary histories are distinguished (Kellogg, 1913 and 1913b), and both of them were enhanced by application of new molecular tools (for review see Klassen, 1992; Brooks and McLennan, 1993; Clayton and Moore, 1997). The first scenario characterizes the evolutionary pattern as a strict congruence between the phylogenies of the two counterparts due to a series of co-speciations. The second approach, which is characterized by phylogeny-independent distribution, was formerly adopted by researchers studying relationships of plants and phytophagous insects. Insects typically show narrow host specificity however without any clear co-speciation signal, instead, the distribution of insects on the host plants is formed by a process termed ecological fitting (Janzen, 1985). This scenario assumes the present pattern of

host-insect distribution as a result of the evolution of host resources and the evolution of insect abilities to exploit these resources. Such associations, originated by ex post adaptation, must inevitably lead to largely decoupled and hence incongruent phylogenies of the two counterparts. Similar phylogenetic incongruences were also recognized in parasites and they are usually attributed to host switches and other evolutionary events, particularly species duplications and sorting events (Brooks and McLennan, 1993; Page and Holmes, 1998). An array of methods has been developed to compare the host and parasite phylogenies and to establish the underlying evolutionary scenario (Brooks and McLennan, 1993; Page 1994; Charleston, 1998; Johnson et al., 2001a; Legendre et al., 2002, Brooks and Ferrao, 2005). Although the co-speciation model is applied to already differentiated and well-defined species of hosts and parasites, it has been recognized that the overall co-evolutionary pattern is often influenced by the genetic structure of both host and parasite populations. Moreover, the typical co-speciation methods can only be reliably applied when the parasite species/strains/host-races are well-defined monophyletic lineages. When applied at subspecific level these methods make no assumptions about how the tree topology is affected by population processes of genetic drift, migration, recombination, or fluctuations in population size.

1.1.2 Intraspecific genetic variability: determinants of the gene flow

A more complex question arises if the co-evolutionary thinking is applied to the population level, where intraspecific genetic structure is taken into account. The issue of intraspecific variability of parasites is closely connected to the debate of species concept in parasitology. Parasitic organisms are often extensively modified by their mutualistic way of life and many of the important adaptations and fine scale variability are hidden to the prism of a regular observer's eye Therefore, the delimitation of parasitic species has for a long time been recognized as methodologically and conceptually challenging problem (e.g., Thompson and Lymbery, 1990; Lymbery and Thompson, 1996; De Meeus et al., 1998). Several concepts have been proposed; neither of them seems to be universally applicable to parasites (Perkins, 2000; Kunz, 2002). In context of these theoretical inconveniences, the concept of host specificity is helpful, because it provides an important character of a species trait; width of the host spectrum is narrower among populations of a single species than among different species of parasites. In general, it is supposed that among the many generalist parasites reported from a broad range of hosts, at least some may constitute assemblages of morphologically undistinguishable but genetically distinct host-specific species or strains. Although the phenomenon of "cryptic species" or "species complexes" is not new to biology, only the application of molecular techniques proved the existence of such assemblages of broad-host-range parasites (Jousson et al., 2000; Demanche et al., 2001, Smith et al., 2006). Within young and/or morphologically undifferentiated forms, such strains may persist through time and become a potential source of new parasitic species (Bush, 1969). Such host-mediated speciation has been suggested as an evolutionary process alternative to a geographical isolation and typical for parasitic taxa (Brooks and McLennan, 1993). Of course, examples of inverse situation to these cryptic lineages exist, and parasites originally supposed to represent several different species with narrow host specificities are sometimes found to come from populations of a single species with intensive genetic exchange (e.g. Blair et al, 1997; Grillo et al., 2007).

The importance of hosts influence on the creation of genetically distinct lineages has been assumed for a long time, however search for the ecological determinants involved in the process was realised rather as "a priori" predictions instead of conclusions based on empirical studies. For instance, parasite populations were traditionally assumed to possess low genetic diversity and reduced heterozygosity as a result of inbreeding (Price, 1980). This view reflects the ecology of parasites' distribution in space. The host represents an unstable habitat that is

distributed unevenly in the environment, and the colonization of a new host individual is thus to a certain degree a matter of chance. As a result, parasite species were supposed to be fractionated into intrahost communities with minimal genetic exchange between the populations. The prediction of low genetic variability, however, proved false. Most parasites exhibit levels of population variability comparable to the free living organisms and genetic variability of parasites with high infections intensity often exceeds the variability of their hosts (Whiteman et al., 2007; Hunt et al, 2008; Redman et al., 2008). The example of phytophagous insects, which served well as a clue when applying the ecological fitting theory onto parasites (see above), was misleading when used to predict the genetic character of parasite populations. In contrast to insects, which frequently inhabit a single host plant for many generations, parasites usually release their offspring into the exterior and the gene pool exchange in parasites is therefore more extensive. An analogy to the plant-insect system among parasites is provided by lice, which complete several recurrent generations parasitizing single host. However, it was shown that even genetic diversity within louse species may significantly exceed the diversity of their hosts. For instance, no population genetic structure was seen in a Galapago hawk (*Buteo galapagoensis*), probably as result of a recent bottleneck. On the contrary, its ectoparasites (two species of louse and a hippoboscid fly) were found to be a good estimator of population structure of the host. The pattern of genetic variability in the parasites followed the colonization history of the hawk among particular islands of the archipelago (Whiteman and Parker, 2005; Whiteman et al., 2007). A factor responsible for such difference in the level of genetic variability may be the short generation time of parasites compared to their hosts, which results in higher mutation rate (Hafner et al, 1994; Johnson et al., 2003, McGauhran et al., 2008). Hence, following a bottleneck the genetic variability in neutral markers may be sooner restored in parasites than in their hosts.

The view of parasite populations as genetically homogeneous assemblages also stems from another preconception. A parameter expected to reduce the genetic variability within some groups of parasites is asexual reproduction involved in their life cycles. Among metazoan parasites such reproduction is usually realised in form of clonal propagation of larval stages (Trematoda, some Cestoda) or as parthenogenesis in adult females (Nematoda). Similar genetic outcome is supposed to be produced by self-fertilization in hermaphroditic adults (Trematoda, Cestoda). In particular, it was assumed that clonal reproduction of trematodes inside the intermediate host decreases the genetic variability of adults in the definite host, which is infected by batches of genetically identical clones (Criscione et al., 2005). However, as was shown in species studied using molecular markers, this rule is not always applicable. The level of intra-host genetic variability did not differ from the level seen among different host individuals (Criscione and Blouin, 2006). Furthermore, at least some of these helminths are capable of mate choice (Trouvé et al., 1999), which provides opportunity for evolution of specific mating behaviour to avoid inbreeding. Conversely, geographical isolation was the main factor conditioning the population structure in some of these parasites (Criscione and Blouin, 2007). As was shown in free-living Bdelloid Rotifers, strict clonality tends to increase the heterozygosity within populations due to accumulation of mutations in alleles, which do not undergo recombination, a process termed Meselson effect (Judson and Normark, 1996; Welch and Meselson, 2000). Thus, under certain ratio of mixed clonal and sexual reproduction the level of heterozygosity in the population is indistinguishable from outcome of strict sexual reproduction (Balloux et al., 2003).

Different situation arises in cestodes which posse's hermaphroditic adults but (with the exception of some Cyclophillid tapeworms) they are not capable of clonal propagation in larvae. Self-fertilization is an extreme form of inbreeding, hence its impact on the genetic character of populations depends on the actual ratio between selfing and outcrossing in the

population of adults. With the exception of economically important taxa, the data on predominating mode of reproduction in cestodes are limited. Populations of *Echinococcus granulosus*, were found to be extremely homozygotic; no signs of outcrossing between the adults were found using allozymic and SSCP data (Lymbery et al., 1997; Haag et al., 1999). Larval stages of this species are known to possess narrow host specificity and clonal reproduction; genetic strains associated with various species of domesticated and wildlife animals were found. It is therefore possible that such extreme reproduction mode is an adaptation to preventing segregation of specific combinations of alleles. On the contrary, high levels of selfing are usually connected with loss of fitness due to inbreeding depression. Decrease of hatching capacity, ability to infect intermediate hosts and to compete outbred progeny was demonstrated in inbred progeny of other cestodes (Milinski, 2006).

On the other hand, experimental evidence supports the view that despite the strong disadvantages selfing may prevail under certain conditions. For instance, selfing was preferred when two *Schistocephalus solidus* adults of differing body mass were coupled in the experiments of Luscher and Milinski (2003). Furthermore, when given a choice the adults preferred to mate with their sibblings over genetically less related individuals (Schjorrig and Jäger, 2007). Explanation of such behaviour follows a simple logic. Through inbred crossing the parasite maximises the number of its own gene copies which is passed to the next generation. In the case of two unequally sized worms, outcrossing may be disadvantageous for the larger individual, since it is allowed to fertilize smaller amount of eggs than its less developed counterpart. According to Schjorrig and Jäger (2007) the increased amount of parasite's own genes in the inbred progeny may surpass the disatvantage of lower viability of the progeny. Such phenomenons, if proven as common in nature, would strongly affect the genetic character of parasite populations and their evolution. For instance, limited outcrossing would enhance the effects of genetic drift and accelerate fixation of population specific alleles. Unfortunatelly, any data based on the examination of natural populations of *S. solidus* or related cestodes lack to support or oppose the experimental observations and to evaluate the predominating form of reproduction.

Such examples show that various complex models of host-parasite association have to be explored before the observed genetic patterns can be reliably explained. The models should include information on evolutionary history of both counterparts, their geographical variability and important bionomical features (including the mode of reproduction in clonal or hermaphroditic organisms). Unfortunately, so far, only few studies have focused on interpreting intraspecific genetic structure in relation to all of these factors. In respect to host-specificity, the models examined indicate that host spectrum may significantly influence the intraspecific genetic structure and creation of barriers preventing the gene flow (Nadler et al., 2000; McCoy et al., 2001; Johnson et al., 2002, McCoy et al, 2005). They also show that the patterns of parasite distribution and specificity may vary among related parasite taxa due to the difference in various biological characteristics of the host and the parasites (host feeding preference, parasite mobility, etc.). One of the best-documented examples involves chewing lice of the genera *Columbicola* and *Physconelloides* parasitizing several species of doves. Difference between the levels of host-specificity and gene flow detected in both genera has been attributed to the difference of their dispersal biology (Johnson et al., 2002). Among bionomical features effectively restricting the gene flow, an important role is played by the site specificity of the parasite. For example, two cryptic species of *Plasmodium* infecting Caribbean lizards were found to differ in the preference between haemocytes and leucocytes (Perkins, 2000; Perkins, 2001). Similar effect of site preference on the distribution of parasite species was seen also in one genus of Monogenea (Šimková et al., 2004). On the contrary, other examples show that the contribution of the host or parasite bionomy on the population

structure is not always straightforward. For instance, a microsatellite study of the human lice of genus *Pediculus* did not reveal any gene flow between head-specific and body-specific populations. The isolation was maintained even between populations sampled from single host individual (Leo et al., 2005). Such site-mediated division supports the existence of two traditionally recognized (sub)species (*P. humanus corporis* and *P. humanus capitis*). However, the situation became far more complicated when lice sampled at multiple localities were studied using several mtDNA and nuclear loci (Reed et al. 2004; Leo and Barker, 2005). The lice turned up to form two genetically distant lineages, but the division did not reflect the main bionomical features (head vs. body) in any trivial manner. While one lineage represents a monophyletic "capitis" form distributed mainly in the New World and associated with native inhabitants of the area, the other lineage is an admixture of both "capitis" and "corporis" forms with a worldwide distribution. This example stresses the need for maximal complexity in the models of host-parasite coevolution at the population level.

1.2 Demographic events in the course of evolution

Due to their life style, parasite populations experience in course of their evolution various demographic changes that are reflected in present constitution of their gene pools. Typically, the parasite populations are likely to suffer from range fragmentation and/or fluctuations of population sizes. For example, host taxa that occur in sympatry today and share the same parasitic species may have been isolated in the past. Thus, genetic variability of the parasites and occurrence of possible host-specific lineages may be an imprint of past allopatric speciation rather than reflection of any recent host-mediated process. The ability to detect such past demographic processes is an important prerequisite for rigorous evaluation of the genetic variability distribution in populations of parasites.

1.2.1 History of populations: genealogy and coalescence theory

Population genetics provides analytic tools to describe the current shape of the population structure. Deliminating the barrier to gene flow, and estimating the rate of migration and gene exchange are the "alfa and omega" of traditional studies of population genetics. Wright's F statistics, a widely used estimator in such studies, provides a good approximation of the depth of genetic structuring between populations, but it suffers several drawbacks. First, it is incapable to distinguish between the ongoing gene flow and outcomes of various historical demographic processes. Further, F statistics is not suitable tool for analysing DNA-sequence data because it requires data on frequencies at several independent loci to produce reliable estimates (Slatkin and Maddison, 1989). To overcome these problems, a new concept, the coalescence theory (Kingman, 1982 a, b; Tavaré, 1984), combining the approaches of both phylogenetics and population genetics, was introduced into the studies of microevolutionary events during the last decades. In contrast to classical population genetic models which were built to infer the future properties of populations coalescent is retrospective. It traces all alleles of a given gene back to the ancestral copy shared by all members of the population. Such point is recognized as the most recent common ancestor (MRCA). The time axis was reversed in the coalescent, but the rationale behind it is similar to that of the classical population genetics and it stems from the theory of genetic drift. Under the process of genetic drift the rate of fixation of an allele in a population is a function of introduction of new alleles (i.e. the rate of mutations) and of the population size. Usually there are no reasons to expect changes of mutation rate in time and then the dynamics of the process is dependent only on the number of alleles and population size at the point of start (the process is faster in small populations). Likewise, in the coalescent, which moves back in time, the expected number of generations to the MRCA corresponds to the starting number of alleles in the population (and again the time

to the MRCA is shorter in populations of small size). The usefulness of the coalescent to reveal various population events rests on the fact that any violation of the stochasticity of the process (e.g. selection, recombination, migration, size fluctuations) changes the course of the time axis from present to the MRCA.

The coalescent based approach and the methods reconstructing intraspecific genealogies provide powerful analytical device to investigate recent history of population structure and to determine such events as the population range expansion or fragmentation (Hudson, 1990; Tempelton et al., 1995; Crandall and Templeton, 1996; Nee et al., 1996; Emerson et al., 2001). An array of statistical methods was developed to facilitate processing of DNA data. When analysing the population structure a statistical parsimony (Templeton et al., 1992; Crandall et al., 1994, Clement et al., 2000) is first performed to build networks of haplotypes instead of traditional phylogenetic trees. Then, using the nested clade analysis (Tempelton et al., 1987; Templeton and Sing, 1993), the haplotypes are organized into nested clades. Such nested design allows testing hypotheses on both the spatial and historical character of populations using several statistical procedures: the categorical test (Templeton and Sing, 1993) and the clade distance analysis (Templeton et al., 1995; Tempelton, 1998; Posada et al. 2000, Panchal, 2007). Finally, the demographic character of populations can be further tested using a variety of coalescent-based softwares (e.g., Kuhner et al., 1998, Beerli and Felsenstein, 2001, Slatkin, 2001; Mailund et al., 2005). Recently, the coalescent approach was also applied to extract the historical information from microsatellite data using the Bayesian paradigm (Wilson et al., 2003; Kuhner, 2006, Kuhner and Smith, 2007). Even though one of the methods, the clade distance analysis, recently attracted much criticism due to its vulnerability to the Type I Error (Knowles, 2004; Petit, 2008), it still remains one of the most preferred methods in historical demography. Regarding this flaw its results should always be interpreted in context with other coalescent based or population genetic methods.

1.2.2 Methodological pitfalls

Compared to the traditional phylogenetics, the investigations of intraspecific genealogies present several difficulties. First, such studies are demanding in terms of sufficient quantity of the studied material. Populations are not monophyletic evolutionary units, thus intraspecific variability produces variance in genetic data, which may obscure the inferred ecological or evolutionary patterns. To evaluate the variance in the data numerous individuals have to be sampled. Extensive sampling of parasites is usually difficult, because many of them occur with low prevalence. Hence, it is not surprising that relatively few genealogical data are available for parasites (for reviews see Criscione et al., 2005; Nieberding and Olivieri, 2007).

Second, due to the short evolutionary time of the studied processes, most of the common phylogenetic markers, such as ribosomal RNAs, do not provide sufficient information. Therefore, for these studies of recent-events, the fast evolving genes have to be determined and employed. Among the most useful types of sequences are the internal transcribed spacers (ITS) and microsatellites, the short series of repetitive sequences (e.g., Depaquit et al., 2002; Galazzo et al., 2002; Worheide et al., 2002; Rosenberg et al., 2002, Mallon et al., 2003). Also mtDNA sequencing and restriction mapping appeared to be a useful marker exhibiting high level of variability on population level in various organismal groups (e.g., Villá et al. 1999; Tarjuelo et al., 2001; Johnson et al., 2003; Stevanovitch et al., 2004, Fleischer et al., 2008).

Third, it was shown that, in closely related lineages, the gene phylogeny may not correspond to the organism phylogeny (Page and Charleston, 1997). Several genetic processes may cause this phenomenon, ancestral interspecific polymorphism and differential extinction of alleles are

probably the most common. A convincing example has been provided in the study of host-specific lineages of human and pig infecting parasites of the genus *Ascaris* (Anderson and Jaenike, 1997; Anderson, 2001). Another problem stems from the presence of multiple different copies of some genetic markers distributed throughout the genome. An existence of such paralogue genes is a potential problem for all phylogenetic studies, but it is particularly dangerous at the intraspecific level. For instance, the ITS sequences, which are among the most frequently used markers in low-level phylogenies, were found to possess considerable level of intragenomic variability in many species of plants and metazoans (e.g. Wendel et al, 1995; Buckler et al., 1997; O'Donnell and Cigelnik, 1997; Gasser et al., 1998). Since the level of intraindividual variability within ITS usually does not exceed the intraspecific variability, these paralogues usually do not affect the interspecific phylogeny. However, when unrecognized, their presence may distort intraspecific genealogies (Leo and Barker, 2002). Despite this flaw, no attempt has been made to consider this problem when interpreting the results in many studies, where ITS sequences were employed (e.g. Almeida-Artigas et al., 2000; Mas-Coma et al., 2001; Depaquit et al., 2002; Bargues et al., 2006). Thus, to resolve a reliable phylogeny at the subspecific level, several unlinked genetic markers have to be analyzed and their phylogenies compared.

1.3 Model organisms

The current state of knowledge provides only limited information on intraspecific structure and the significance of both host-preference and geography. As discussed above, it now becomes obvious that more data on various types of host-parasite associations is needed to establish a reliable theory concerning the host-specificity, parasite distribution and the evolutionary or epidemiological consequences. In this thesis, I explore two different types of parasites, endoparasite and ectoparasite, with intention to investigate and compare intraspecific genetic variability and structure of their populations, and to evaluate the significance of the two evolutionary factors, the host-specificity and geographic isolation. As model species, I selected the tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea) and the sucking lice *Polyplax serrata* (Anoplura: Polyplacidae).

1.3.1 Ligula intestinalis

L. intestinalis is a diphyllobothridean (Kuchta et al, 2008) tapeworm species with wide Holarctic distribution and a three-host life-cycle. The first host is presented by a crustacean copepode. The coracidium larva swallowed by the copepod evolves into procercoid larva within two to three weeks (Dubinina, 1980). Secondary larval stage, the plerocercoid, infects several freshwater planctonophagous fish species. The records of secondary host species span enormous range of fish taxa, representing mainly the families Cyprinidae, Catostomidae, Percidae, Siluridae, Esocidae and Osmeridae (Szalai et al., 1989; Loot et al., 2001a; Shields et al., 2002; Morgan, 2003; Innal et al. 2007). Plerocercoid stage inside the fish host represents the longest phase in the parasite life cycle (10 months or more, Dubinina, 1980), while inside its definitive host, the fish-eating bird, the worm quickly reaches maturation and starts breeding. The whole infection of definitive hosts lasts no longer than 5-6 days (Dubinina, 1980).

Duration of the plerocercoid stage inside the fish host imposes strong pressure on the evolution of specific adaptations. Parasitic castration (Arme 1975), suppress of the host growth and increased mortality (Loot et al., 2001b and 2002b) as well as behavioural modifications (Barber et al. 2000; Brown et al. 2001; Loot et al., 2002a) were described as results of

L. intestinalis infection. These deep effects render plerocercoid the most important stage of the *Ligula* lifecycle in respect to evolution and maintenance of host-specificity.

The existence of distinct strains or even species within *L. intestinalis* has been suggested by several authors (Harris and Wheeler, 1974; Arme, 1975, 1997; Kennedy and Burrough, 1981). On the contrary, the allozyme analysis performed by McManus (1985) failed to find any host-related polymorphism among populations. The idea of strain variation was later supported by Arme (1997) who found differences in pathology between two fish hosts (*Gobio gobio* vs. *Rutilus rutilus*). Most recently, cryptic genetic variability among *Ligula* specimens from these two hosts was revealed by Olson et al. (2002), who studied sequences of the ITS-1 and ITS-2 rDNA regions. A relatively high level of genetic polymorphism in *Ligula* was observed also by Luo et al. (2003), who compared European and Chinese specimens of *Ligula* with related genus *Digramma*. When addressing intraspecific variability of *L. intestinalis*, another ligulid species has to be taken in account. Although *Ligula colymbii* seems to be ecologically isolated from *L. intestinalis* by different host spectrum (fish host – loach, *Cobitis taenia*, definite host – grebes, Podicipediformes, Dubinina, 1980), its taxonomic validity was recently questioned due to its paraphyletic nature in respect to *L. intestinalis* (Logan et al., 2004). All these data indicate that *L. intestinalis* and other species of the genus are likely to form an assemblage of lineages/species with complex genetic structure affected by various ecological and historical parameters.

1.3.2 Polyplax serrata

Lice of the genus *Polyplax* are ectoparasites found predominantly on various rodents. The species *P. serrata*, although reported from a relatively broad spectrum of hosts (two families of rodents and one family of insectivores) is typically and most frequently associated with mice of the genus *Apodemus*. From the geographic point of view, *P. serrata* is a widely distributed parasite, known from numerous localities throughout the Palaearctic (e.g. Durden and Musser 1994, Krištofík 1999, Krištofík and Dudich 2000), where it colonizes various species of locally distributed *Apodemus* mice.

Suitability of *Polyplax – Apodemus* association as a model for population genetics study is supported by the fact that the population structure and the colonization history of the Europe has been well established for *Apodemus* species (Filippucci et al. 2002; Michaux et al. 2004; Michaux et al. 2005a, b; Catteau, 2006). In particular, it was demonstrated that *A. sylvaticus* and *A. flavicollis*, which are the most abundant hosts in Europe, differ in their glacial histories. Populations of both species were fragmented into several refugia, some shared by both species, but in most cases geographically separated. Populations of *A. sylvaticus* in the western refugia experienced strong bottleneck, while eastern populations of *A. sylvaticus* and populations of *A. flavicollis* remained relatively intact (Michaux et al., 2005b). After the glaciation receded the populations melted during the Europe recolonisation, with several mtDNA lineages recognizable in both species of *Apodemus*. Such complex demographical history is likely to affect at various levels population structuring of both the mouse hosts and their parasites.

1.4 Objectives

The main aim of the thesis is to explore genetic structure, intraspecific variability and genealogical relationships within populations of the selected models. The results will be used as basis for evaluation of the influence of host-specificity and geographic isolation on gene flow and structure in populations of the parasites. Particular objectives of the thesis are as follows:

1. To map genetic variability across and within the strains of *Ligula intestinalis* and *Polyplax serrata*.

2. To evaluate gene flow within and among the host-associated and geographically isolated sets of parasites

3. To reconstruct demographic and evolutionary events within the populations

4. To interpret observed patterns with respect to biology and evolutionary history of the hosts

2. Summary of the results

2.1 The *Ligula* model

Two approaches were selected to reconstruct the phylogeny and population structure of the first model. Sequences of three genetic markers (two mtDNA genes and a ribosomal DNA spacer) were employed in **Manuscript no. 1**, and microsatellite markers were chosen for the further analysis. Isolation of microsatellite primers and basic characterization of the new markers are described in **Manuscript no. 2.** Results of the microsatellite analysis interpreted in the context of the patterns obtained from sequence-based data are provided in **Manuscript no. 3**. The following section summarizes the results and conclusions of these three studies.

Over 100 *Ligula* specimens of worldwide sampling were sequenced for selected genes. Sequential data of another 30 specimens were retrieved from GeneBank and incorporated into the analyses (MS no 1, Tables 1 and 3). The pattern of population structure based on the sequences of cytochrome oxidase subunit 1 (COI) revealed deep level of divergence between lineages from separate geographic areas (MS no. 1; Fig. 1). Each of the haplotypes was associated only with a single geographic area (Canada, Mexico, Ethiopia, Australia, China and Euro-Mediterranean); hence no gene flow has taken place among the main lineages recently. The same pattern of genetically distant lineages was observed in the analysis of cytochrome B sequences (COB) (data not shown); however, the phylogeny was less resolved, probably due to higher level of saturation. Such pattern proved that geography is the main feature structuring populations of the parasite on a global scale.

Two sympatric lineages were recovered within the Euro-Mediterranean area. Plerocercoids representing the first and more frequently encountered lineage (named clade A) were retrieved from phylogenetically derived species of cyprinid fish (Leuciscinae, Alburninae). Geographically the populations of clade A stretched across the whole area from the European part of Russia to Tunisia in northern Africa. Representatives of the second lineage formed two monophyletic strains within the lineage designated as clade B (Fig. 1 in MS no. 1). The host spectrum of the second lineage was restricted to phylogenetically basal cyprinids (*Gobio*,

Barbus, *Rhodeus*). The latter lineage was less represented among the studied material, however it probably occupies similarly extent area as the clade A lineage, considering the distances between the sampled localities (Algeria, Ireland, Czech Rep., Poland). The fish hosts seem to play important role in maintaining genetic divergence between the two lineages: the specificity to different spectra hosts was almost absolute, only one specimen of the clade A lineage was retrieved from basal cyprinid (Carassius, Ukraine). Although rare events of crossings between the two lineages cannot be excluded, as suggested by the pattern of ITS2 variability (see below), the genetic differences between the lineages are deep and they do not exhibit any substantial mixing despite sharing the same definitive host during reproduction. Sympatric occurrence of adults from both lineages in grebes (*Podiceps cristatus*), was found at one locality.

The pattern revealed by analysing the internal transcribed spacer 2 (ITS2) was well compatible with the mtDNA results. However, occasional incongruences were identified and attributed to occurrence of intragenomic variability in ITS2 (MS no. 1; Fig. 3). Clusters of rDNA genes are organized into arrays of hundreds to thousands copies, which are sometimes dispersed on several chromosomes (Buckler et al., 1997; Ganley and Kobayashi, 2007). Under such conditions it is practically impossible to search for all paralogue copies of ITS2 within the genome. However, in order to make an approximate estimate of the extent of intragenomic variability within individual genomes, PCR products of several samples were cloned into plasmids and sequenced. The level of variability within a single individual usually did not exceed the variability of a given mtDNA lineage, although several cases of divergent ITS2 paralogues from single individual were found (MS no.1; Table 4). Three possible explanations of this phenomenon were suggested; introgression through rare events of hybridization between the lineages was found the most plausible after mapping the ITS2 paralogues on the COI tree (MS no. 1; Fig. 2).

Different genetic character of the two European lineages was further supported by demographic patterns revealed using several statistics (MS no. 1; Fig.3, Tables 5 and 6). Whereas clade A lineage exhibited relatively steep population growth, the results for clade B populations showed a constant population size. Differences were related to the host spectra; larger population size in the clade A was attributed to the fact that it parasitizes hosts with large population densities, compared to the clade B. The population growth can also reflect recent uptake of a new host lineage (derived cyprinids) or post-glacial propagation of the populations, as observed in other species of European fauna (e.g. Michaux et al., 2005b; Pinho et al., 2007). In contrast to the deep divergence between individual lineages, the level of intra-lineage variability was relatively shallow. In particular, the populations of the most extensively sampled clade A were further characterized using concatenated matrix of COI and COB sequences. The historic and geographic pattern was retrieved using the method of Nested Clade Analysis (Templeton et al., 1995). Several events of isolation by distance or continuous range expansion were identified (MS no. 1; Fig.4, Table 7), but the potential of mtDNA markers to disclose geography or host dependent population structure inside the lineages of *Ligula* was largely limited. In parasites with complex life-cycles, migration with host possessing the highest capabilities of mobility is assumed to reduce genetic variation among populations (Poulin, 2007). Fish-eating birds, the definitive hosts of *Ligula*, were suggested to homogenize the *Ligula* populations as they cross relatively large distances during their seasonal migrations (Curry-Lindahl, 1980). Similar example was provided by Figuerola et al. (2005) for freshwater invertebrates in Northern America with their eggs distributed by waterfowl migrations.

Microsatellite loci were chosen as a multilocus marker for assessing fine scale variability and the predominating mode of reproduction (selfing vs. outcrossing) in the *Ligula* model.

Microsatellite sequences constitute of tandemly repeated di- to oligo-nucleotides which are dispersed randomly throughout the genome. Usefulness of these markers in population studies lies in their high variability. Typically, the number of repeats (i.e. different alleles) varies among individuals within a population, and allelic frequencies differ between the populations. Furthermore, microsatellites are inherited in Mendelian fashion and thus allow for estimation of population structure by employing the hypotheses on Hardy-Weinberg equilibrium and Linkage Disequilibria. On the contrary, due to elevated variability of microsatellite-surrounding sequences, the primers are not usually applicable on a wide spectrum of taxa and have to be isolated separately for each studied species.

Since isolation of novel primers is a time-consuming procedure, we first tested the possibility of cross-amplification using six microsatellite loci isolated for a related tapeworm species *Schistocephalus solidus* (Binz et al., 2000). Unfortunately, none of the loci produced reliable PCR amplicons. Therefore microsatellites containing two repeat motifs, the GA and CA dinucleotides, were searched for using genomic DNA from two individuals of the European strain (**MS no. 2**). To increase efficiency of the search, combination of two methods (Fleischer and Loew, 1995; Koblížková et al., 1998) was employed during the construction of genomic libraries enriched with desired microsatellite motifs. Approximately 80 sequences containing the motifs were retrieved using the specified approach, and PCR primers were constructed in the regions flanking the microsatellites. In 16 of these loci, successful and consistent PCR amplification was obtained in a test population and they were selected for the subsequent population analysis. Most of the loci also showed successful cross-amplification with *Ligula* specimens from the non-European strains (MS no. 3; Table 2).

Microsatellites allow for simultaneous analysis of large sets of samples. Hence, the number of specimens studied in the **MS no. 3** was significantly extended compared to **MS no.1**. Altogether 338 plerocercoids were studied using 15 microsatellite loci (MS no. 3; Table 1). One primer pair, Li.15, was excluded from the study since it yielded large number of null homozygotes.

Global pattern of the variability expressed a structure similar to the mtDNA results, with each geographic area constituting separate lineage (MS no. 3, Figs. 3, 4). However, several differences were observed even at the local scale. Deep level of divergence was found between Irish and Algerian strains of *Ligula* in contrast to partial admixture observed in mtDNA haplotypes (MS no 1; clade B in the Fig. 1). Furthermore, genetic substructuring was recognized also within the lineages largely homogeneous in the mtDNA analyses. Using Bayesian clustering algorithm implemented in SRUCTURE software (Falush et al., 2007), geographic isolation was revealed between European and Tunisian populations of the clade A (MS no. 3; Fig. 6 A), suggesting a role of Mediterranean See as sufficient barrier to gene flow between the populations. Such situation points to the limited potential of mtDNA sequences to reveal recent evolutionary events due to incomplete lineage sorting.

The most surprising finding was division of *Ligula* plerocercoids from Lake Tana (Ethiopia) into two clades differing in their host specificity (MS no. 3; Figs 4 and 6). Distribution of *Ligula* isolates among the two clades reflected phylogenetic distances between their fish hosts. The plerocercoids were retrieved from four species of barbs, which belong to two evolutionary distinct groups (Machordom and Doadrio, 2001). Regarding the split depth between the two barb lineages, their distribution in Africa, and recent formation of the Lake Tana (De Graaf et al., 2007), the two sympatric strains of parasites may represent secondary encounter of lineages that have originally adapted to their host taxa in different areas. However, it should be stressed

that only four barb species, out of hundreds of species living in Africa, were sampled in this study an the situation is thus likely to be much more complex.

Indications of a weak host-related subdivision were also found in two populations of clade A, containing sympatric isolates of *Rutilus* and *Alburnus* hosts (MS no, 4; Appendix 1). Possible role of immune response and selection for specific combination of genes resulting in this pattern was discussed. Despite frequent HW disequilibria, the multilocus character of intrapopulation genetics indicated that outcrossing is a predominating reproduction mode in *Ligula*. Even though low level of selfing or inbred mating could not be excluded in several populations, its impact on gene flow within the populations was negligible. Studying mating behaviour of a related tapeworm, Milinski (2006) showed that preference to selfing is possible under experimental conditions (differences in body mass of the sexual counterparts, mating between genetically related vs. unrelated individuals). While occurrence of such specific situations cannot be excluded in *Ligula*, the observed genetic data suggest that their frequency under natural conditions is probably rare.

2.2 The *Polyplax* model

Altogether 126 *P. serrata* specimens were gathered for the analysis of population structure in the **MS no. 4**. The populations were sampled mainly across the central Europe, and were completed by specimens from several distant localities (MS no. 4, Fig.1) in aim to explore the overall range of genetic diversity. Majority of the analysed material was retrieved from three host species (*Apodemus flavicollis*, *A. sylvaticus* and *A. agrarius*) and few specimens from two other species (*A. uralensis*, *Clethrionomys glareolus*). Samples from two west-European localities were collected from *A. sylvaticus*. Samples from Lake Baikal were retrieved from a locally specific host, *A. peninsulae*. Two genes, cytochrome oxidase subunit I (COI) and elongation factor 1 alpha (EF1α), were selected for the molecular analysis. Both genes have previously been successfully applied to resolve low level phylogenies in several louse taxa (Johnson et al., 2001b, 2002; Whiteman et al. 2007). The sequences of COI were used as the main source of information and a representative proportion of the samples was sequenced in the nuclear marker (EF1α).

The Asian samples of *P. serrata* formed separate clade in MP and ML analyses, as was expected regarding the geographic distance separating Baikal Lake region from Europe. On the contrary, the picture of population structure in the European clades was much more complex than expected based on the traditional co-evolutionary presumptions. The samples formed three separate lineages, clades A, B and C, with different levels of host specificity (MS no. 4, Fig 2). Clade A exhibited the widest geographic distribution containing samples from all studied areas and was associated with two hosts *A. sylvaticus* and *A. flavicollis*. In contrast, Clade B possessed narrow host-specificity and contained exclusively the lice collected from *A. flavicollis.* Also its geographic distribution was smaller than in the first clade. Clade C, which occurred most frequently in the eastern part of the studied area, was collected mainly from two hosts *A. agrarius* and *A. uralensis* and with a minimal frequency from *C. glareolus*, *A. flavicollis* and *A. sylvaticus*. Allelism was found in EF1α which hindered usage of this marker in phylogenetic analysis. However, the distribution of polymorphic positions found among sequences of the gene showed that the three clades represent true genetic lineages without recent admixture (MS no 4., Fig 2.).

Demographic analysis of the data was performed using coalescent approach implemented in the Lamarck software (Kuhner, 2006). Estimates of Theta (θ) showed that the detected clades differ with respect to their population sizes (MS no 4, Table 3). Smaller population size

inferred for the clade B corresponds to its narrow host specificity, and vice versa, the clade A lineage inhabiting two hosts possessed larger population size. Surprising were the extremely small population size and steep growth inferred for clade C lineage. This is probably a consequence of a bottleneck connected with a recent colonization of Europe from the east with its principal host *A. agrarius* (Catteau, 2006).

Two different scenarios can explain occurrence of the two genetically isolated strains (clade A and B) sharing one host in sympatry (*A. flavicollis*): separated co-speciation of each lineage with different host, followed by host switch in the clade A, or relatively recent duplication within a single lineage occupying both host species (*A.flavicollis* and *A. sylvaticus*). The duplication of parasite lineages within populations of their host has been postulated an important process in the host-parasite coevolutionary scenarios (Page, 2003). Despite its theoretical importance, such process has only rarely been demonstrated by empirical data. The date of divergence between the European lineages of *Polyplax*, estimated here using a Bayesian approach (Drummond et al., 2003), indicates that the speciation of the hosts (*A. flavicollis* vs. *A. sylvaticus*) preceded speciation of their parasites by approx. 2.5 My. Such divergence suggests the *Polyplax* speciation took place on well formed host species. Hence, duplication event within one fragmented population is a likely explanation of the observed occurrence of two lineages sharing one host in sympatry.

3. Conclusions and future prospects

Molecular analyses of intraspecific genetic structure often show large amount of hidden variability within the studied groups. The origin of population structure is usually attributed to gaps in the geographic distribution or complex evolutionary history of the species. The issue is further extended in parasites, where the possibility of host-dependent structuring adds yet another dimension. Analyses of the two models performed here, revealed high levels of cryptic genetic variability and revealed previously unrecognized lineages. Populations of the model species were structured according to both studied components, the geography and the host origin. However, distribution and the scale of the variability differed among the two models and produced unexpected patterns.

The wide geographic sample of the *Ligula* model was deeply affected by geographic distances on the global scale. On the contrary, the analysis revealed none or very weak structuring on the local scale. Isolation by distance was only detected across the Mediterranean Sea using the microsatellite markers. Homogeneity of the big European metapopulation was attributed mainly to migrations with the bird hosts. Sympatric occurrence of the host-specific lineages was also observed, but with the exception of Ethiopian populations, these lineages represented strains differing in the host spectra of infected fish rather than distinct lineages with narrow host specificity. The effect of both components was far more pronounced in louse populations. Despite the smaller area covered by the sampling, geography was largely imprinted in the population structure of *Polyplax* even on a local scale. Such differences are likely to stem from the lower mobility of rodents compared to birds.

Regardless obvious impact of the geographic isolation, host-specificity was the main force structuring populations in *Polyplax*. Even though multilocus markers were not available (their preparation is planned in the future studies of the author), the distribution of allelic variability of the nuclear marker used (EF1α) suggests that the host-specific mtDNA lineages of *P. serrata* represent distinct populations/strains with no mutual mixing. The life-cycle of

Polyplax, which is closely tight to its mice host, facilitates such strict differentiation. The lice may reproduce for several generations on single host individual and only direct contact between host individuals allows for their transmission. On the contrary, the complex life-cycle of *Ligula*, where free-living and copepode associated stages exist, provides more opportunities to switch between different fish hosts in the course of evolution. Consequently, lineages with relatively wide spectrum of hosts were found.

The obtained patterns suggest the direction of further studies of the two models. The prospects of studying the African populations of *Ligula* have already been mentioned. Identification of the biological traits of the host, responsible for formation of gene flow barriers inside some lineages and lack of the structure in others, is one of the most interesting challenges arising from this study. Further, knowledge on *Polyplax* genetics in glacial refugia of the hosts, and its comparison with the results obtained in this study, would significantly improve the insight into how the shared history affected evolution of the identified host-specific lineages. Moreover, obtaining multilocus markers will allow for testing on finer scale diversification in the polyxeneous lineages of *Polyplax*. Also, a locally focused study of the two lineages with overlapping host spectrum (clades A and B) would provide data interesting from the epidemiological and ecological point of view. For example, it is traditionally postulated that specialist species of parasites are favoured over generalists in infecting individual hosts (Poulin, 1998). Only one louse specimen per infected host individual was studied here, hence evaluating the distribution and infestation rates of the two lineages on particular hosts, which they share, will allow testing such ecologically based evolutionary hypotheses.

4. References

Abderrazak, S.B., Oury, B., Lal, A.A., Bosseno, M.F., Force-Barge, P., Dujardin, J.P., Fandeur T., Molez, J.F., Kjellberg, F., Ayala, F.J. and Tibayrenc M. 1999. *Plasmodium falciparum*: population genetic analysis by multilocus enzyme electrophoresis and other molecular markers. Exp. Parasitol., 92: 232-238

Almeyda-Artigas, R.J., Bargues, M.D., Mas-Coma, S., 2000. ITS-2 rDNA sequencing of *Gnathostoma* species (Nematoda) and elucidation of the species causing human gnathostomiasis in the Americas. J. Parasitol., 86: 537-544

Anderson, T.J.C. 2001. The dangers of using single locus markers in parasite epidemiology: *Ascaris* as case study. Trends Ecol. Evol, 17: 183-188

Anderson, T.J.C., Jaenike, J. 1997. Host specificity, evolutionary relationships and macrogeographic differentiation among *Ascaris* populations from humans and pigs. Parasitology, 115: 325-342

Arme, C. 1975. Tapeworm-host interactions. In: Jennings, D. & Lee, D.J. (Eds.) Symposium of the Society for Experimental Biology. No. 29 – Symbiosis. p. 505-532. Cambridge, University Press

Arme, C. 1997. Ligulosis in two cyprinid hosts: *Rutilus rutilus* and *Gobio gobio*. Helmithologia, 34: 191-196

Balloux, F., Lehmann L., De Meeus, T. , 2003. The population genetics of clonal and partially clonal diploids. Genetics, 164: 1635-1644

Barber, I., Hoare, D., Krause, J. 2000. Effects of parasites on fish behaviour: a review and evolutionary perspective. Rev. Fish Biol. Fish.,10: 131-165

Bargues, M.D., Morchon, R., Latorre, J.M., Cancrini, G., Mas-Coma, S., Simon, F. 2006. Ribosomal DNA second internal transcribed spacer sequence studies of Culicid vectors from an endemic area of *Dirofilaria immitis* in Spain. Parasitol. Res., 99: 205-213

Beerli, P., Felsenstein, J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. PNAS, 89: 4563-4568

Binz, T., Reusch, T.B.H., Wedekind, C., Schärer, L., Sternberg, J.M., Milinski, M., 2000. Isolation and characterization of microsatellite loci from the tapeworm *Schistocephalus solidus*. Mol. Ecol., 9: 1926-1927

Blair, D., Agatsuma, T., Watanobe, T. 1997. Molecular evidence for the synonymy of three species of *Paragonimus*, P-ohirai Miyazaki, 1939, P-iloktsuenensis Chen, 1940 and P-sadoensis Miyazaki et al., 1968. J. Helmintol., 71: 305-310

Brooks, D.R., Ferrao, A.L., 2005. The historical biogeography of co-evolution: emerging infectious diseases are evolutionary accidents waiting to happen. J. Biogeogr., 32(8): 1291-1299

Brooks, D.R., McLennan, D.A. 1993. Parascript, parasites and the language of evolution. Smithsonian Press, Washington DC

Brown, S.P., Loot, G., Grenfell, B.T., Guégan, J-F. 2001. Host manipulation by *Ligula intestinalis*: accident or adaptation? Parasitology, 123:519-529

Buckler, E.S., Ippolito, A., Holtsford, T.P. 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. Genetics, 145(3):821-32

Bush, G.L. 1969. Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). Evolution, 23: 237-251

Catteau, G. 2006. Contributión á la phylogeographie de mulot raye (*Apodemus agrarius*) sur l'ensemble de la region Palearctique. Master Thesis. Université de Liège

Charleston, M.A. 1998. Jungles: a new solution to the host/parasite phylogeny reconciliation problem. Math. Biosci., 149: 191-223

Clayton, D.H. and Moore, J. (eds.) 1997. Host-Parasite evolution - general principles & avian models. Oxford University Press, 473 p.

Clement, M., Posada, D., Crandall, K.A. 2000. TSC: a computer program to estimate gene genelogies. Mol. Ecol., 9: 1657-1659

Crandall, K.A., Templeton, A.R. 1996. Application of intraspecific phylogenetics. In: P.H. Harvey, A.J.L. Brown, J. Maynard Smith and S. Nee (eds.): New uses for new phylogenies. Oxford University Press, p. 81-99

Crandall, K.A., Templeton, A.R., Sing, C.F. 1994. Intraspecific cladogram estimation: problems and solutions. In: R. W. Scotland, D. J. Siebert, and D. M. Williams (eds.): Models in Phylogeny Reconstruction. Oxford: Clarendon Press

Criscione, C.D., Blouin, M.S., 2006. Minimal selfing, few clones, and no among-host genetic structure in a hermaphroditic parasite with asexual larval propagation. Evolution, 60: 553-562

Criscione, C.D., Blouin, M.S., 2007. Parasite phylogeographical congruence with salmon host evolutionarily significant units: implications for salmon conservation. Mol. Ecol., 16: 993-1005

Criscione, C.D., Poulin, R., Blouin, M.S. 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Mol. Ecol., 14: 2247-2257

Curry-Lindahl, K.1980. Les oiseaux migrateurs à travers mer et terre. Delachaux et Niestle, Neuchatel, Paris

De Graaf, M., Megens, H.J., Samallo, J., Sibbing, F.A. 2007. Evolutionary origin of Lake Tana's (Ethiopia) small *Barbus* species: indications of rapid ecological divergence and speciation. Animal Biol., 57: 39-48

Demanche, C., Berthelemy, M., Petit, T., Polack, B., Wakefield, A.E., Dei-Cas, E., Guillot, J. 2001. Phylogeny of *Pneumocystis carinii* from 18 primate species confirms host specificity and suggests coevolution. J. Clin. Microbiol., 39: 2126-2133

De Meeus, T., Michalakis, Y., Renaud, F. 1998. Santa Rosalia revisited: or why are there so many kinds of parasites in `The garden of earthly delights'? Parasitol. Today, 14: 10-13

Depaquit, J., Ferté, H., Léger, N., Lefranck, F., Alves-Pires, C., Hanafi, H., Maroli, M., Morillas-Marquez, F., Rioux, J.-A., Sobodová, M., Volf, P. 2002. ITS 2 sequences heterogeneity in *Phlebotomus sergenti* and *Phlebotomus similis* (Diptera, Psychodidae) possible consequences in their ability to transmit *Leishmania tropica*. Int. J. Parasitol., 32: 1123-1131

Drummond, A.J., Rambaut, A. 2003. BEAST v1.0, Available from http://evolve.zoo.ox.ac.uk/beast/

Dubinina, M.N. 1980. Tapeworms (Cestoda, Ligulidae) of the Fauna of the USSR. Amerind Publishing, New Delhi, India

Durden, L.A., Musser, G.G. 1994. The sucking lice (Insecta, Anoplura) of the World: a taxonomic checklist with records of mammal hosts and geographical distributions. B. Am. Mus. Nat. Hist., 218

Emerson, B.C., Paradis, E. and Thébaud, C. 2001. Revealing the demographic histories of species using DNA sequences. Trends Ecol. Evol, 16: 707-716

Falush, D., Stephens, M., Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes, 7: 574-578

Figuerola, J., Green, A.J., Michot, T.C. 2005. Invertabrate eggs can fly: evidence of waterfowl-mediated gene flow in aquatic invertebrates. Am. Nat., 165: 276-280

Filippucci, M.G., Macholán, M., Michaux, J.R. 2002. Genetic variation and evolution in the genus *Apodemus* (Muridae: Rodentia). Biol. J. Linn. Soc., 75: 395-419

Fleischer, R.C., Boarman, W.I., Gonzalez, E.G., Godinez, A., Omland, K.E., Young, S., Helgen, L., Syed, G., Mcintosh, C.E. 2008. As the raven flies: using genetic data to infer the history of invasive common raven (*Corvus corax*) populations in the Mojave Desert. Mol. Ecol., 17: 464-474

Fleischer, R.C., Loew, S. 1995. Construction and screening of microsatellite-enriched genomic libraries. P. 459-468 in Ferraris,J., Palumbi, S. (eds.): Molecular Zoology: Advances, Strategies and Protocols. Wiley-Liss, New York

Freeland, W.J., Boulton, W.J. 1992. Coevolution of food webs: parasites, predators and plant secondary compounds. Biotropica, 24: 309-327

Galazzo, D.E., Dayanandan, S., Marcogliese. D.J. and McLaughlin, J.D. 2002. Molecular systematics of some North American species of *Diplostomum* (Digenea) based on rDNA-sequence data and comparisons with European congeners. Can. J. Zool., 80: 2207-2217

Galvani, P.A., 2003. Epidemiology meets evolutionary ecology. Trends Ecol. Evol., 18: 132-139

Ganley, A.R.D., Kobayashi, T. 2007. Highly efficient concerted evolution in the ribosomal DNA repeats: Total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res., 17: 184-191

Gasser, R.B., Zhu, X.Q., Chilton, N.B., Newton, L.A., Nedergaard, T., Guldberg, P., 1998. Analysis of sequence homogenisation in rDNA arrays of *Haemonchus contortus* by denaturing gradient gel electrophoresis. Electrophoresis, 19: 2391-2395

Grillo, V., Jackson, F., Cabaret, J., Gilleard, J.S. 2007. Population genetic analysis of the ovine parasitic nematode *Teladorsagia circumcincta* and evidence for a cryptic species. Int. J. Parasitol., 435-447

Haag, K.L., Araujo, A.M., Gottstein, B., Siles-Lucas, M., Thompson, R.C.A., Zaha, A. 1999. Breeding systems in *Echinococcus granulosus* (Cestoda; Taeniidae): selfing or outcrossing. Parasitology, 118: 63-71

Hafner, M.S., Sudman, P.D, Villablanca, F.X., Spradling, T.A., Demastes, J.W., Nadler, S.A., 1994. Disparate rates of molecular evolution incospeciating hosts and parasites. Science, 265: 1087-1090

Harris, M.T., Wheeler, A. 1974. *Ligula* infections of bleak *Alburnus alburnus* (L.) in the tidal Thames. J. Fish Biol., 6:181-188

Hernandez, A.D., Sukhdeo M.V.K. 2008. Parasite effects on isopod feeding rates can alter the host's functional role in a natural stream ecosystem. Int. J. Parasitol., in press

Hudson, R.R. 1990. Gene genealogies and the coalescence process. Oxford surveys in evolutionary biology. 7: 1-14

Hudson, P.J., Greenman, J. 1998. Competition mediated by parasites: biological and theoretical progress. Trends. Ecol. Evol., 13: 387-390

Hunt, P.W., Knox, M.R., Le Jambre, L.F., McNally, J., Anderson, L.J. 2008. Genetic and phenotypic differences between isolates of *Haemonchus contortus* in Australia. Int. J. Parasitol., in press

Innal, D., Keskin, N., Erk'akan, F. 2007. Distribution of *Ligula intestinalis* (L.) in Turkey. Turk. J. Fish. Aquat. Sci., 7: 19-22

Janzen, D.H.. 1985. On ecological fitting. Oikos, 45:308-310

Johnson, K.P., Cruickshank, R.H., Adams, R.J., Smith, V.S., Page, R.D.M., Clayton, D.H. 2003. Damatically elevated rate of mitochondrial substitution in lice (Insecta: Phtiraptera). Mol. Phyl. Evol., 26: 231-242

Johnson, K.P., Drown, D.M. and Clayton, D.H. 2001a. A data based parsimony method of cophylogenetic analysis. Zool. Scr., 30: 79-87

Johnson, K.P., Moyle, R.G., Witt, C.C., Faucett, R.C, Weckstein, J.D. 2001b. Phylogenetic relationships in the louse genus Penenirmus based on nuclear (EF-1α) and mitochondrial (COI) DNA sequences. Syst. Entomol., 26: 491-497

Johnson, K.P., Williams, B.L., Drown, D.M., Adams, R.J., Clayton, D.H. 2002. The population genetics of host specificity: genetic differenciation in dove lice (Insecta: Phtiraptera). Mol. Ecol., 11: 25-38

Jousson, O., Bartoli, J., Pawlowski, J. 2000. Cryptic speciation among intestinal parasites (Trematoda:Digenea) infecting sympatric host fishes (Sparidae). J. Evol. Biol., 13: 778-785

Judson, O.P., Normark, B.B. 1996. Ancient asexual scandals. Trends Ecol. Evol., 11: 41-46

Kellogg, V.L., 1913. Distribution and species-forming of ectoparasites. Am. Nat., 47: 129-158

Kellogg, V.L. 1913. Ecto-parasites of the monkeys, apes and man Science, 38: 601-602

Kennedy, C.R. and Burrough, R.J. 1981 The establishment and subsequent history of population of *Ligula intestinalis* in roach *Rutilus rutilus* (L.). J. Fish Biol., 19: 105-126

Kingman, J.F.C. 1982a. On the genealogy of large populations. J. Appl. Prob., 19A: 27-43

Kingman, J.F.C. 1982b. The coalescent. Stoch. Proc. Appl., 13: 235-248

Klassen, G.J. (1992) Coevolution: A history of macroevolutionary aproach to studying host-parasite associations. J. Parasitol., 78: 573-587

Knowles, L.L. 2004. The burgeoning field of statistical phylogeography. J. Evol. Biol., 17: 1-10

Koblížková, A., Doležel, J., Macas, J. 1998. Subtraction with 3 ' modified oligonucleotides eliminates amplification artifacts in DNA libraries enriched for microsatellites. Biotechniques, 25: 32-38

Krištofík, J. 1999. Sucking lice (Phthiraptera) from Mongolian mammals. Biologia, 54: 143-149

Krištofík, J., Dudich, A. 2000. Sucking lice of the *Polyplax genus* (Phtiraptera) on small mammals (Insectivora, Rodentia) in Slovakia. Biologia, 55(2): 133-142

Kuchta, R., Scholz, T., Brabec, J., Bray, R. A. 2008. Suppression of the tapeworm order Pseudophyllidea (Platyhelminthes: Eucestoda) and the proposal of two new orders, Bothriocephalidea and Diphyllobothriidea. Int. J. Prasitol., 48: 49-55

Kuhner, M.K., 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics, 22: 768-770

Kuhner, M.K., Smith, L.P. 2007. Comparing likelihood and Bayesian coalescent estimation of population parameters. Genetics, 175: 155-165

Kuhner, M.K., Yamato, J., Felsenstein, J. 1998. Maximum likelihood estimation of population growth rates based on the coalescent. Genetics, 149: 429-434

Kunz, W. 2002. When is a parasite species a species? Trends in Parasitol., 18: 121-124.

Legendre, P., Desdevises, Y., Bazin, E. 2002. A statistical test for host-parasite coevolution. Syst. Biol., 51: 217-234

Leo, N.P., Barker, S.C., 2002. Intragenomic variation in ITS2 rDNA in the louse of humans, *Pediculus humanus*: ITS2 is not a suitable marker for population studies in this species. Insect Mol. Biol., 11: 651-657

Leo, N.P., Hughes, J.M., Yang, X., Poudel, S.K.S., Brogdon, W.G., Barker, S.C., 2005. The head and body lice of humans are genetically distinct (Insecta : Phthiraptera, Pediculidae): evidence from double infestations. Hereddity, 95: 34-40

Logan, F.J., Horák, A., Štefka, J., Aydogdu, A., Scholz, T. 2004. Phylogeny of diphyllobothriid tapeworms (Cestoda: Pseudophyllidea) based on sequences of the ITS-2 rDNA. Parasitol. Res., 94: 10-15

Loot, G., Aulagnier, S., Lek, S., Thomas F., Guégan, J-F. 2002a. Experimental demonstration of a behavioural modification in a cyprinid fish, *Rutilus rutilus* (L.), induced by a parasite, *Ligula intestinalis* (L.). Can. J. Zool., 80: 738-744

Loot, G., Francisco, P., Santoul, F., G., Lek, S., Guégan, J-F., 2001a. The three hosts of the *Ligula intestinalis* (Cestoda) life cycle in Lavernose-Lacasse gravel pit, France. Arch. Hydrobil., 152: 511-525

Loot, G., Lek, S., Dejean, D., Guégan, J-F. 2001b. Parasite-induced mortality in three host populations of the roach *Rutilus rutilus* (L.) by the tapeworm *Ligula intestinalis* (L.). *Ann. Limnol.*, 37: 151-159

Loot, G., Poulin, R., Lek, S., Guégan, J-F. 2002b. The differential effects of *Ligula intestinalis* (L.) plerocercoids on host growth in three natural populations of roach, *Rutilus rutilus* (L.). Ecol. Freshw. Fish, 11: 168-177

Luscher, A., Milinski, M. 2003. Simultaneous hermaphrodites reproducing in pairs self-fertilize some of their eggs: an experimental test of predictions of mixed-mating and Hermaphrodite's Dilemma theory. J. Evol. Biol., 16: 1030-1037

Luo, H.Y., Nie, P., Yao, W.J., Wang, G.T., Gao, Q. 2003. Is the genus *Diagramma* synonymus to the genus *Ligula* (Cestoda: Pseudophyllidea)? Parasitol. Res., 89: 419-421

Lymbery, A.J. and Thompson, R.C.A. 1996. Species of *Echinococcus*: Pattern and Process. Parasitol. Today, 12: 486-491

Lymbery, A.J., Constantine, C.C., Thompson, R.C.A., 1997. Self-fertilization without genomic or population structuring in a parasitic tapeworm. Evolution, 51: 289-294

Mailund, T., Schierup, M.H., Pedersen, C.N.S., Mechlenborg, P.J.M., Madsen, J.N., Schauser, L. 2005. CoaSim: A flexible environment for simulating genetic data under coalescent models. BMC: Informatics, 6: 252

Machordom, A., Doadrio, I., 2001. Evolutionary history and speciation modes in the cyprinid genus *Barbus.* P. Roy. Soc. B-Biol. Sci., 268: 1297-1306

Mallon, M., MacLeod, A., Wastling, J., Smith, H., Reilly, B., Tait, A., 2003. Population structures and the role of genetic exchange in the zoonotic pathogen *Cryptosporidium parvum*. J. Mol. Evol., 56: 407-417

Mas-Coma, S., Funatsu, I.R., Bargues, M.D. 2001. *Fasciola hepatica* and lymnaeld snails occurring at very high altitude in South America. Parasitology, 123(Suppl.): S115-S127

McCoy, K.D., Boulinier, T., Tirard, C. 2005. Comparative host-parasite population structures: disentangling prospecting and dispersal in the black-legged kittiwake *Rissa tridactyla*. Mol Ecol., 14: 2825-2838

McCoy, K.D., Boulinier, T., Tirard, C., Michalakis, Y. 2001. Host specifity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick *Ixodes uriae*. J. Evol. Biol., 14: 395-405

McGaughran, A., Hogg, I.D., Stevens, M.I. 2008. Patterns of population genetic structure for springtails and mites in southern Victoria Land, Antarctica. Mol. Phyl. Evol., 46: 606-618

McManus, D.P. 1985. Enzyme analyses of natural populations of *Schistocephalus solidus* and *Ligula intestinalis*. J. Helminthol., 59: 323-328

Michaux, J., Bellinvia, E., Lymberakis, P., 2005a. Taxonomy, evolutionary history and biogeography of the broad-toothed field mouse (*Apodemus mystacinus*) in the eastern Mediterranean area based on mitochondrial and nuclear genes. Biol. J. Linn. Soc., 85: 53-63

Michaux, J.R., Libois, R., Filippucci, M.G. 2005b. So close and so different: comparative phylogeography of two small mammal species, the Yellow-necked fieldmouse (*Apodemus flavicollis*) and the Woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. Heredity, 94: 52-63

Michaux, J.R., Libois, R., Paradis, E., Filippucci, M.G. 2004. Phylogeographic history of the yellow-necked fieldmouse (*Apodemus flavicollis*) in Europe and in the Near and Middle East. Mol. Phyl. Evol., 32: 788-798

Milinski, M. 2006. Fitness consequences of selfing and outcrossing in the cestode *Schistocephalus solidus*. Integr. Comp. Biol., 46:373-380

Morgan, D.L. 2003. Distribution and biology of *Galaxias truttaceus* (Galaxiidae) in south-western Australia, including first evidence of parasitism of fishes in Western Australia by *Ligula intestinalis* (Cestoda). Environ. Biol. of Fish., 66:155-167

Nadler, S.A. 1990. Molecular approaches to studying helminth population genetics and phylogeny. Int. J. Parasitol. 20: 11-29

Nadler, S.A., Adams, B.J., Lyons, E.T., Delong, R.L., Melin, S.R. 2000. Molecular and morphometric evidence for separate species of *Uncinaria* (Nematoda: Ancylostomatidae) in california sea lions and northern fur seals: hypothesis testing supplants verification. J. Parasitol., 85: 1099-1106

Nadler, S.A., Lindquist, R.L., Near, T. 1995 Genetic structure of midwestern *Ascarius suum* populations: a comparison of isoenzyme and RAPD markers. J. Parasitol., 81: 385-394

Nieberding, C.M., Olivieri, I. 2007. Parasites: proxies for host genealogy and ecology? Trends Ecol. Evol., 22: 156-165

Nee, S., Holmes, C., Rambaut, A. and Harvey, P.H. 1996. Inferring population history from molecular phylogenies. . In: P.H. Harvey, A.J.L. Brown, J. Maynard Smith and S. Nee (eds.): New uses for new phylogenies. Oxford University Press, p. 66-79

Nosil, P. 2002. Transition rates between specialization and generalization in phytophagous insects. Evolution, 56: 1701-1706

O'Donnell, K., Cigelnik, E. 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus Fusarium are nonorthologous Mol. Phyl. Evol., 7: 103-116

Olson, P.D., Littlewood, D.T.J., Griffiths, D., Kennedy, C.R. and Arme, C. 2002. Evidence for co-existence of separate strains or species of *Ligula* in Lough Neagh, Northern Ireland. J. Helminthol., 76: 171-174

Panchal, M. 2007. The automation of nested clade phylogeographic analysis. Bioinformatics, 23: 509-510

Page, R.D.M. 1994. Parallel phylogenies: reconstructing the history of host-parasite assemblages. Cladistics, 10: 155-173

Page, R.D.M. 2003. Tangled Trees: Phylogeny, Cospeciation, and Coevolution. The University of Chicago Press

Page, R.D.M., Charleston, M.A. 1997. From gene to organismal phylogeny: Reconciled trees and the gene tree species tree problem. Mol. Phyl. Evol., 7: 231-240

Page, R.D.M., Holmes, E.C. 1998. Molecular Evolution: A Phylogenetic Approach. Blackwell Science Inc. 352 p.

Paterson, S., Viney, M.E. 2000. The interface between epidemiolology and population genetics. Parasitol. Today, 16: 528-532

Perkins, S.L. 2000. Species concepts and malaria parasites: detecting a cryptic species of *Plasmodium*. Proc. R. Soc. London B Biol. Sci., 267:2345-2350

Perkins, S.L. 2001. Phylogeography of Caribbean lizard malaria: tracing the history of vector-borne parasites. J. Evol. Biol., 14: 34-35

Petit, R.J. 2008. The coup de grâce for the nested clade phylogeographic analysis? Mol. Ecol., 17: 516-518

Pinho, C., Harris, D.J., Ferrand, N. 2007. Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. Mol. Ecol., 16: 1191-1205

Posada, D., Crandall, K.A., Templeton, A.R. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Mol. Ecol., 9: 487-488.

Poulin, R. 1999. The functional importance of parasites in animal communities: many roles at many levels? Int. J. Parasitol., 29: 903-914

Poulin, R. 1998. Large-scale patterns of host use by parasites of freshwater fishes. Ecol. Lett., 1: 118-128

Poulin, R. 2007. Evolutionary Ecology of Parasites, 2nd ed., Princeton University Press, Princeton, 332 p.

Poulin, R., Krasnov, B.R., Shenbrot, G.I., Mouillot, D., Khokhlova, I.S. 2006. Evolution of host specificity in fleas: Is it directional and irreversible? Int. J. Parasitol. 36:185-191

Poulin, R., Mouillot, D. 2004. The relationship between specialization and local abundance: the case of helminth parasites of birds. Oecologia, 140: 372-378

Price, P.W. 1980. Evolutionary biology of parasites. Princeton University Press, Princeton, 237p.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A., Feldman, M.W. 2002. The genetic structure of human populations. Science, 298: 2381-2385

Reed, L.D., Smith, V.S., Hammond, S. L., Rogers, A.R., Clayton, D.H. 2004. Genetic analysis of lice supports direct contact between modern and archaic humans. PLoS Biol., 2: e340

Redman, E., Packard, E., Grillo, V., Smith, J., Jackson, F., Gilleard, J.S. 2008. Microsatellite analysis reveals marked genetic differentiation between *Haemonchus contortus* laboratory isolates and provides a rapid system of genetic fingerprinting. Int. J. Parasitol., 38: 111-122

Schjorring, S., Jäger, I. 2007. Incestuous mate preference by a simultaneous hermaphrodite with strong inbreeding depression. Evolution, 61: 423-430

Shields, B.A., Groves, K.L., Rombauch, C., Bellmore, R. 2002. Ligulosis associated with mortality in largescale suckers. J.Fish Biol., 61: 448-455

Šimková, A., Morand, S., Jobet, E., Gelnar, M., Verneau, O. 2004. Molecular phylogeny of congeneric monogenean parasites (*Dactylogyrus*): A case of intrahost speciation. Evolution, 58: 1001-1018

Slatkin, M. 2001. Simulating genealogies of selected alleles in a population of variable size. Genet. Res., 78: 49-57

Slatkin, M., Maddison, W.P. 1989. A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics, 123: 603-616

Smith, M.A., Woodley, N.E., Janzen, D.H., Hallwachs, W., Hebert, P.D.N. 2006. DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). PNAS, 103: 3657-3662

Stireman, J.O. 2005. The evolution of generalization? Parasitoid flies and the perils of inferring host range evolution from phylogenies. J. Evol. Biol., 18: 325-336

Stevanovitch, A., Gilles, A., Bouzaid, E., Kefi, R., Paris, F., Gayraud, R.R., Spadoni, J.L., El-Chenawi, F., Beraud-Colomb, E. 2004. Mitochondrial DNA sequence diversity in a sedentary population from Egypt. Ann. Hum. Genet., 68: 23-39

Szalazai, A,J,, Yang, X., Dick, T.A. 1989. Changes in number and growth of *Ligula intestinalis* in the spotted shiner (*Nothropis hudsonius*) and their roles in transmission. J.Parasitol., 75: 571-576

Tarjuelo, I., Posada, D., Crandall, K.A., Pascaul, M. and Turon, X. 2001. Cryptic species of *Clavellina* (Ascidiacea) in two different habitats: harbours and rocky littoral zones in northwestern Mediterranean. Mar. Biol., 139: 455-462

Tavaré, S. 1984. Line-of-descent and genealogical process, and their applications in population genetics models. Theor. Pop. Biol., 26:119-164

Templeton, A.R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Mol. Ecol., 7: 381-397

Templeton, A.R., Boerwinkle, E., Sing, C.F. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction andonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. Genetics, 117: 343-351

Templeton, A.R., Crandall, K.A., Sing, C.F. 1992 A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics, 132: 619-633

Templeton, A.R., Routman, E., Phillips, C.A. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics, 140: 767-782

Templeton, A.R., Sing, C.F. 1993. A Cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. Genetics, 134: 659-669

Thompson, J.N. 1994. The coevolutionary process. University of Chicago Press, Chicago, Il.

Thompson, R.C.A., Lymbery, A.J. 1990. Intraspecific variation in parsites-what is a strain? Parasitol. Today, 6: 345-348

Trouvé, S., Renaud, F., Durand, P., Jourdane, J. 1999. Reproductive and mate choice strategies in the hermaphroditic flatworm *Echinostoma caproni.* J. Hered., 90: 582-585

Vilá, C., Amorim, I.R., Leonard, J.A., Posada, D., Castroviejo, J., Petrucci-Etrucci-Fonseca, F., Crandall, K.A., Ellegren, H., Wayne, R.K. 1999. Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. Mol. Ecol., 8: 2089-2103

Wang, T. P., Shrivastava, J., Johansen, M. V., Zhang, S. Q., Wang, F. F., Webster, J.P. 2006. Does multiple hosts mean multiple parasites? Population genetic structure of *Schistosoma japonicum* between definitive host species. Int. J. Parasitol., 36: 1317-1325

Welch, M.D., Meselson, M. 2000. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Science, 288: 1211-1215

Wendel, J.F., Schnabel, A., Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). PNAS, 92: 280-284

Whiteman, N.K., Kimball, R.T., Parker P.G. 2007. Co-phylogeography and comparative population genetics of the threatened Galapagos hawk and three ectoparasite species: ecology shapes population histories within parasite communities. Mol. Ecol., 16: 4759-4773

Whiteman, N.K., Parker, P.G. 2005. Using parasites to infer host population history: A new rationale for parasite conservation. Animal Conservation, 8:175-181

Wilson, D.J., Falush, D., McVean, G. 2005. Germs, genomes and genealogies. Trends.Ecol. Evol., 420: 39

Wilson, G.A., Rannala, B., 2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics, 163: 1177-1191

Wordeheide, G., Hooper, J.N.A., Degnan, B.M. 2002. Phylogeography of western Pacific *Leucetta 'chagosensis'* (Porifera: Calcarea) from ribosomal DNA sequences: implications for population history and conservation of the Great Barrier Reef World Heritage Area (Australia). Mol. Ecol., 11: 1753-1768

Yan, G., 1996. Parasite-mediated competition: a model of directly transmitted macroparasites. Am. Nat., 148: 1089-1112

Manuscript No. 1:

Bouzid, W., Štefka, J., Hypša V., Lek, S., Scholz, T., Legal, L., Ben Hassine, O.K., Loot G. 2008. Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). Int. J. Parasitol., in press

Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae).

Wafa Bouzid [a,c,1], Jan Štefka [b,1], Václav Hypša [b,*], Sovan Lek [a], Tomáš Scholz [b], Luc Legal [d], Oum Kalthoum Ben Hassine [c], Géraldine Loot [a]

[a] *Laboratoire Evolution et Diversité Biologique, U.M.R CNRS-UPS 5174, Université Paul Sabatier 118 route de Narbonne, F-31062 Toulouse cedex 4, France*

[b] *Faculty of Science, University of South Bohemia and Biology Centre ASCR, Institute of Parasitology, Branišovská 31, 37005 České Budějovice, Czech Republic*

[c] *Unité de Recherche Biologie, Ecologie et Parasitologie des organismes aquatiques, Faculté des Sciences de Tunis, Tunisie*

[d] *ECOLAB, U.M.R. CNRS-UPS 5245, Université Paul Sabatier, Bâtiment IVR3, 118 Route de Narbonne, F-31062 Toulouse cedex 4, France*

* Corresponding author. Tel: +420 387 775 409; fax: + 420 385 310 388.

E-mail address: vacatko@paru.cas.cz (V. Hypša)

[1]These authors contributed equally.

Abstract

Parasite species with global distributions and complex life cycles offer a rare opportunity to study alternative mechanisms of speciation and evolution in a single model. Here, genealogy and genetic structure, with respect to geography and fish host preference, have been analyzed for *Ligula intestinalis*, a tapeworm affecting freshwater fish. The data analyzed consisted of 109 tapeworms sampled from 13 fish host species in 18 different localities on a macrogeographic scale. Two mitochondrial genes, cytochrome oxidase subunit I and cytochrome B, and the nuclear sequence of intergenic transcribed spacer 2 (ITS2) were used for the genetic reconstruction. Different evolutionary patterns were found at the local and at the global geographic scales. On a local scale, the flat genetic structure was mainly attributed to its contiguous range expansion. Migrating birds are the most likely cause of the homogenisation of the whole population, preventing the creation of significant genetic barriers. By contrast, on a global scale, genetically distant and well-separated clusters are present in different geographic areas. Reproductive isolation was found even between clades living in sympatry and infecting the same definitive host, suggesting the existence of efficient biologically determined genetic barriers, and thus possibly separate species. Although the ITS2 sequences were found to display considerable intragenomic variability, their relationships were generally in good agreement with the topology derived from mitochondrial genes.

Keywords: Phylogeography; Host specificity; Parasite evolution; Cryptic species; Intragenomic variability; *Ligula*

1. Introduction

Inferring the genetic structure of parasites provides great opportunities to elucidate questions about host specificity, co-speciation events and parasite evolution in general. The genetic structure of parasites is becoming accessible to rigorous analysis with the use of tools combining molecular phylogeny and population genetics. Several studies have demonstrated that the genetic structure and the evolution of parasites at the intraspecific level vary considerably among groups of parasites (e.g. Anderson, 2001; McCoy et al., 2001; Johnson et al., 2002; Brant and Orti, 2003; Wickström et al., 2003).

The large diversity of life-history traits displayed by parasites may explain such variation. For instance, diversity in modes of reproduction (e.g. hermaphroditic, parthenogenetic) and in dispersal capability of the free-living stage are major determinants of the genetic structure of parasites (Johnson et al., 2002; Criscione et al., 2005). In addition, host vagility may also predispose parasites to various genetic structures and ultimately dictate various patterns of speciation (Blouin et al., 1995; McCoy, 2003; Criscione and Blouin, 2004).

Most studies have focused on parasites with a simple life cycle. Examples include ectoparasitic arthropods such as ticks on nesting seabirds (McCoy et al., 2001; McCoy, 2003) and lice on doves (Page et al., 1998; Johnson et al., 2002). In parasitic platyhelminthes, nematodes and acanthocephalans, complex life cycles with one or several intermediate hosts are extremely common (Poulin, 1998). For such parasites, which present complex systems, the genetic structure is controlled by life-history traits in different parasite stages and by the mobility of the hosts (Jarne and Théron, 2003; Prugnolle et al., 2005). Criscione and Blouin (2004) demonstrated that parasites cycling in freshwater hosts only (an autogenic life cycle) had much more highly structured populations and much lower gene flow among sub-populations than parasites cycling through freshwater and terrestrial hosts (an allogenic life cycle).

The tapeworm *Ligula intestinalis* (L.) is the most common species of the genus *Ligula* (Bloch, 1782). It is widely distributed throughout the whole Holarctic region (Dubinina, 1980) and has recently been also reported from Australasia (Morgan, 2003; Chapman et al., 2006). This cestode presents a complex life cycle with a cyclopoid or diaptomid copepod as first intermediate host and planktivorous fish as second intermediate host. Fish-eating birds serve as the final host in which *L. intestinalis* quickly reaches sexual maturity and releases eggs into the water. The most conspicuous stage within the life cycle is the plerocercoid. It develops in the abdominal cavity of the second intermediate host and has a considerable effect on fish health, fecundity and behaviour. As a result it can cause heavy losses in freshwater pisciculture (Arme and Owen, 1968; Sweeting, 1977; Taylor and Hoole, 1989; Wyatt and Kennedy, 1989; Carter et al., 2005). Although typically reported from cyprinid fish, *L. intestinalis* has been shown to utilize a broad range of hosts, including other fish families such as Catostomidae, Salmonidae or Galaxiidae (e.g. Dubinina, 1980; Bean and Winfield, 1992; Groves and Shields, 2001; Museth, 2001; Barus and Prokes, 2002; Chapman et al., 2006). However, these lists of published records should be considered with caution, as they do not provide an accurate measure of host specificity in a given local population.

Records available on *L. intestinalis* indicate that the host spectra are not uniform across the investigated populations. For instance, analysis of host specificity in natural and experimental fish populations in south-west France has shown that silver bream, *Blicca bjoerkna*, is refractory to *Ligula* infection (Loot et al., 2002). This finding strongly contrasts with other studies identifying bream as a suitable host (Dubinina, 1980; Barus and Prokes,

1994). Loot et al. (2006) suggested that the abundance of potential hosts and the temporal dynamics of their occurrence strongly affect local host specificity. A similar role of host-abundance in the evolution of tapeworm specificity has been suggested for proteocephalids in South America (Hypša et al., 2005).

In addition to the differences in host spectra, Arme (1997) demonstrated that the pathology of *Ligula* infections in two cyprinid species, the roach (*Rutilus rutilus*) and the gudgeon (*Gobio gobio*) differs in several respects. In natural infections of roach, there is pronounced inflammation of host tissue which is never found in natural infections of gudgeon. There are also qualitative and quantitative differences in splenic and pronephric leucocyte counts and in spleen weights between the two hosts. It has been proposed that parasites from roach and gudgeon differ in their ability to stimulate a host response. This could be attributed to the existence of different *Ligula* strains or species (Arme, 1997; Kennedy and Burrough, 1981; McManus, 1985). Recently, Olson et al. (2002) studied *Ligula* tapeworms from three cyprinid fishes, namely gudgeon (*G. gobio*), minnow (*Phoxinus phoxinus*) and roach (*R rutilus*), and provided molecular evidence for separate *Ligula* strains. Luo et al. (2003) found relatively high intraspecific variability between *Ligula* specimens when comparing the genetic sequences of the entire internal transcribed spacer of the ribosomal DNA (ITS rDNA) and the 5' end of 28S rDNA. Logan et al. (2004), using nucleotide variation of ITS2, showed that samples from Turkey identified as *L. intestinalis* were genetically distinct from European and Chinese *Ligula* isolates from various fish hosts.

A further complication of the *L. intestinalis* phylogenetic picture stems from the indication that *L. intestinalis* may be paraphyletic with at least two species, *Ligula colymbi* and *Digramma interrupta*. In a previous study based on ITS2, these two species clustered within the assemblage of *L. intestinalis* samples (Logan et al., 2004). Whereas the validity of *L. colymbi* might be questioned due to the scarcity and ambiguity of discriminating morphological characters, *D. interrupta* is a morphologically well-defined species which can be reliably distinguished. Taken together, these facts invoke a suspicion that the taxon designated as *L. intestinalis* may in fact be composed of several lineages and/or cryptic species. This hypothesis, emerging from available observations and offering a good explanation of reported discrepancies, has not yet been addressed by rigorous analysis of the genealogy and genetic population structure. The relative significance of geography and host specificity in the speciation process and evolution within the *L. intestinalis* complex thus remains unclear.

The aim of this study is to examine the phylogenetic structure of the common cestode, *L. intestinalis* on a macroevolutionary scale using mtDNA and ITS2 regions. Sequence divergence was used to explore the possible presence of cryptic species and discuss the mechanisms of speciation. On a finer scale, we explored the spatial distribution of genetic variation within the major haplotype cluster (tapeworms from the Europe-Mediterranean region) to discriminate between historic versus geographic events.

2. Materials and methods

2.1. Parasite samples

Tapeworms were collected from 23 localities across a broad geographic range covering 18 countries (see Table 1). The samples were isolated from body cavities of fish intermediate hosts and preserved either by placing them in 96% ethanol or freezing them at -80°C. The

specimens were identified as *L. intestinalis*, *Ligula* sp. or *D. interrupta* using the characters suitable for species identification according to Dubinina (1980), i.e. the shape of the anterior end of the body, the presence/absence of external segmentation (metamerism) of the anterior part of the body and the number of genital complexes visible at fully developed plerocercoids. We also included 11 specimens isolated from definitive hosts (eight from the great crested grebe, *Podiceps cristatus*, two from *Gavia stellata* and one from *Mergus merganser*).

2.2. PCR amplification and DNA sequencing

Genomic DNA was extracted from small pieces of tapeworm tissue (10-20 mg) using a Qiagen Tissue extraction kit. Three DNA regions were amplified; the partial sequence of nuclear ribosomal ITS2, the 5.8S partial sequence rDNA and two mitochondrial genes, cytochrome oxidase subunit I (COI) and cytochrome B (COB). A new primer pair COIA2-COIB2 was designed on the basis of the primers COIA/COIB (Li et al., 2000) and cestode COI sequences available in GenBank. The COBA-COBB primers were constructed using the COB sequences available in the GenBank mitochondrion database (Accession No. **AF314223**). Primer sequences are given in Table 2. PCR reactions were carried out in a 20 µl vol. using 1 µl of extracted DNA solution (approximately 20 ng), 5 pM of each primer, 15 mM $MgCl_2$, 10 mM of each dNTP and 0.25 units of Taq polymerase. Amplifications were performed using methods modified from Logan et al. (2004), Li et al. (2000) or von Nickisch-Rosenegk et al. (2001). Purified DNA (20 ng/µl) was either sequenced directly with ABI BigDye chemistry using the amplification primers, or cloned into *Escherichia coli* (Invitrogen) using pGEM-Teasy vector (Promega). Plasmids were sequenced using the universal primers M13 or Sp6 and T7. Sequencing products were alcohol-precipitated and separated on ABI 3100 or 3130 automatic sequencers (Applied Biosciences).

2.3. Phylogenetic analysis

To prepare alignments, new sequences were supplemented with data from GenBank (Table 3). Mitochondrial genes were aligned with BioEdit 7.05 (Hall, 1999) and the program Collapse 1.2 (Posada, 2004, Collapse. A tool for collapsing sequences to haplotypes, Version 1.2, http://darwin. uvigo. es) was used to retrieve individual haplotypes. The ITS2 sequences could not be unequivocally aligned due to considerable length variation, particularly within the regions containing microsatellite repeats. The sequences were prealigned using Megalign (DNASTAR, Inc.) and the alignment was manually adjusted in Bioedit 7.05 (Hall, 1999). The following procedure was then applied to extract phylogenetic characters from the ITS2 sequences. The microsatellite regions were removed altogether since they could not be reliably aligned. Within the rest of the alignment, each indel (usualy 2 to 5 bp long) was treated as a single evolutionary event. To achieve this, the first position of each deletion was coded as a gap, while the rest was considered missing data. In this way, the weight of all indels was identical regardless of their actual lengths.

Sequences of mitochondrial genes were used in several subsequent steps. First, the relationships among the haplotypes on a global scale were inferred using the COI alignment. The COB sequences, displaying a higher degree of variability, were then explored to establish the intrapopulational relationships within the two European lineages (clades A and B; see Results). Finally, a concatenated matrix of both genes was used to evaluate the genealogical structure of clade A by Nested Clade Phylogeographic Analysis (NCPA) (Panchal, 2007).

Maximum parsimony (MP) using the matrix of COI haplotypes was performed in PAUP* (Swofford, 1998. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0. Sinauer Associates, Sunderland, MA) with Ts:Tv weights set to 1:1, 1:2 and 1:3 using the TBR algorithm with 50 replicates of random sequence addition. Calculation of bootstrap support was executed with 1,000 replications of TBR search with Ts:Tv set to 1:1. The maximum likelihood (ML) tree was constructed in Phyml (Guindon et al., 2005) with a GTR+γ+Inv model and the parameters estimated from data, bootstrap support was obtained by 1,000 replications. COI sequences of the diphyllobothridean tapeworm *Diphyllobothrium latum* (**NC008945**) and two cyclophyllidean tapeworms, *Taenia crassiceps* (**NC002547**) and *Hymenolepis diminuta* (**AF314223**), were used as outgroups. MP analysis of ITS sequences, together with *D. latum* (DQ768176.1) as an outgroup, was performed using the TNT program (Giribet, 2005) with a TBR algorithm with 500 replications.

Several methods were employed to evaluate the demographic information on the European and North African populations of *Ligula* samples (clades A and B, Fig. 3). Neutrality tests were performed separately for the COI and COB haplotypes; 95% and 99% confidence intervals were calculated by 10,000 coalescence simulations using DNASP 4.10 (Rozas et al., 2003). Mismatch distribution of pairwise substitutions between COB haplotypes were calculated and compared with the Poisson model using DNASP 4.10. The observed patterns of population growth/decline were tested using the Lamarc 2.0.2 software package (Kuhner, 2006) under a model of molecular evolution identified in Modeltest 3 (Posada and Crandall, 1998). Lamarc was run in a likelihood mode and the search strategy consisted of three replications of 10 short initial chains followed by two long final chains. The initial chains were run with 500 samples and a sampling interval of 20 (10,000 samples); burn-in was set to 1,000 samples for each chain. The same parameters were used for the final chains performed with 10,000 samples. Values of population size (*Theta*) and population growth (*g*) were estimated and confidence intervals calculated using the percentile approach.

The genealogical structure of clade A populations was constructed in program TCS 1.21 (Clement et al., 2000) using a concatenated matrix of COI and COB, which produced 37 haplotypes. Information about the biogeographical history of European populations (clade A) was inferred using software for NCPA. This program implements TCS and GeoDis algorithms (Clement et al., 2000; Posada et al., 2000) testing the congruence between network structure and geography, and an inference key of Templeton (2004) evaluating possible historical and geographical events. The probability of the null hypothesis (no association of genetic structure with geography) was estimated by 1 million permutations. The geographic coordinates of the localities were used to describe their distribution; the radius of the water body was used as a size estimate for each sampled area. To allow for possible long distance migration and/or population fragmentation between the European continent and its surroundings, the North Sea and the Mediterranean were entered as regions uninhabitable for *Ligula*. Following the suggestions of Posada et al. (2000) and Panchal (2007), three regions with excessively large gaps between the sampled populations were indicated to prevent a false inference of isolation by distance. These regions cover unsampled areas in southern Germany and Austria, two areas in Eastern Europe (corresponding approximately to Poland and Byelorussia), and the western part of Russia.

3. Results

3.1. Sequences and alignments

We obtained partial sequences of three DNA regions from 109 *L. intestinalis*, *Ligula* sp. and *D. interrupta* samples: two mitochondrial genes COI (396 bp) and COB (404 bp), and the non-coding nuclear sequence of ITS2 varying from 610 – 650 bp in size.

3.1.1. Mitochondrial genes

In the 101 samples sequenced for COI (including two sequences retrieved from GenBank) and 76 samples sequenced for COB, we identified 46 (COI) and 44 (COB) different haplotypes. Distribution of COI haplotypes among the samples was extremely uneven. While the majority of the haplotypes were only represented by a single *Ligula* specimen, the most frequent haplotype (H14) was retrieved from 39 European samples. The set of COB haplotypes was more evenly distributed, with the two most abundant haplotypes containing eight and seven samples. The complete list and distribution of COI haplotypes is provided in Tables 1 and 3. The levels of genetic variability were similar for both genes. Without outgroups, the 396 positions long COI matrix contained 275 (70%) constant and 89 (22%) parsimony informative characters. For COB, the proportion of constant sites was 278 out of 405 (67%) and the number of parsimony informative sites was 91 (22%). Alignments of mitochondrial genes were unequivocal and did not require any indel adjustments. The observed variability consisted almost exclusively of synonymous mutations at the third codon position.

3.1.2. ITS2

In total 188 ITS2 sequences from 131 specimens were obtained (Tables 1 and 3). Screening for ITS2 intragenomic variability confirmed that different forms of ITS2 can occur together in a single genome. Eight sequences contained 70 to 430 bp deletions extending into the 3' end of 5.8S rDNA. These sequences were considered pseudogenes and were excluded from further analyses. Among the presumably functional sequences, two types of intragenomic variability were observed. The first was only minor. In phylogenetic analyses these variants clustered as closely related sequences within the same COI-derived clade. This variability was mainly due to singletons, numbers of short repeats and several short indels. The second type was represented by highly divergent paralogues; in the phylogenetic tree, these paralogues clustered in different lineages. In contrast to the *Ligula* samples, a low level of variability was observed in *D. interrupta* (Ru Hl). The five haplotypes identified only differed by a few single-nucleotide mutations and in one case by a microsatellite region. The intragenomic variability is summarized in Table 4. The sequences were deposited in GenBank (NCBI) under the accession numbers **EU240978-EU241317**. The numbers of samples sequenced for particular genes are reported in Table 1.

3.2. Phylogenetic analysis

3.2.1. COI and COB

The MP analysis of COI performed under three different Ts/Tv ratios yielded 286 trees. Their strict consensus was entirely compatible with the ML tree and contained several well-supported clades (Fig. 1). Distribution of haplotypes among the clades was consistent with their geographic origin. Whereas some of the geographic areas were represented by a single

monophyletic clade (e.g. Canada, Mexico and Ethiopia), the samples from the Euro-Mediterranean area were split into two distinct lineages (the clades A and B). Clade A was comprised of samples from Europe and Tunisia, and clade B contained, apart from European and Algerian samples, the samples from China and Australia. The split of European populations into two clades reflected the difference in host range observed in these two groups of parasites. The European and Tunisian samples of Clade A were recovered exclusively from the phylogenetically derived cyprinid fish (*Abramis*, *Alburnus*, *Phoxinus*, *Rutilus*, *Scardinius*), whereas the European and Algerian samples of Clade B were found to parasitize the basal cyprinid species (*Barbus, Gobio*, *Rhodeus*). Further geographically dependent diversification was seen within Clade B. The Algerian lineage was composed of haplotypes 1 and 2, the mostly European lineage (Haps 25, 26, 37 and 40) also included a single Algerian haplotype, Hap 3. The Chinese and Australian samples again formed two distinct monophyletic branches, represented by haplotypes 10 - 11 and 4 – 5, respectively. The isolates collected from *P. cristatus* (Czech Republic) were scattered over clades A (haplotypes 14, 23, 24) and B (haplotypes 25, 26). The COB sequences provided less robust topology, but they supported the basic geographically-dependent pattern, including the split of European samples into two lineages (not shown). On the other hand, the higher degree of COB variability resulted in better resolution within clade A (26 haplotypes out of 51 samples compared with 22 haplotypes out of 66 COI haplotypes). This information was used for population structure analysis based on alignments of both mitochondrial genes (see below).

The samples determined as *D. interrupta* clustered at two different positions. The Russian and Chinese samples sequenced in this study (H9 and H43) formed a monophyletic clade separated from Chinese and European *Ligula* of clades A and B. By contrast, haplotypes H12 and H13, corresponding to the Chinese *Ligula* and *Digramma* samples sequenced by Li and Liao (2003) and Li et al. (2000), branched together with clades A and B.

3.2.2. ITS2

The ITS2 sequences provided a weaker phylogenetic signal resulting in lower topological resolution. Despite this limitation, the strict consensus of 286 MP trees was surprisingly compatible with the main geographically-specific lineages inferred from the mitochondrial genes (Fig. 2). Out of the 180 ITS2 sequences analysed, the positions of 170 sequences were consistent with COI topology. Only occasionally were some of the ITS2 paralogues placed into two or more different clades (Table 4).

3.3. Population structure and genetics

Neutrality tests, originally developed to examine selection neutrality of mutations, proved to be a useful tool for testing population expansion. For COI and COB matrices, the neutrality tests provided mainly insignificant results; the only exception was found for COI sequences in clade A (Table 5). This finding corresponds to the low variability of COI haplotypes observed within this clade and indicates possible changes in population size in clade A. A similar trend, although not significantly supported, was identified in the COB matrix (Table 5). Furthermore, for clade A the mismatch distribution of COB haplotypes created a bell-shaped pattern indicating rapid population growth (Fig. 3), while the multimodal pattern corresponding to a stable non-expanding population was found in clade B. Similar differences were detected by coalescence analysis; both *Theta* and *g* were much higher in clade A than in clade B (Table 6). *Theta* is a joint estimator of effective population size and mutation rate. Since there is no reason to expect any considerable difference in mutation rate between

the two lineages, we conclude that the higher value of *Theta* for the clade A populations reflects a larger and expanding population (larger *g* than in clade B). In concordance with these results, NCPA based on a concatenated alignment of both genes identified two cases of contiguous population expansion in European populations of clade A. The contiguous expansion was particularly significant for clade 3-2 and the tests were close to significant values for the entire cladogram (Fig. 4; Table 7). In addition, restricted gene flow with isolation by distance (IBD) was suggested for clade 2-3. In clade 3-3, due to insufficient genetic resolution, no decision was made to distinguish restricted gene flow with IBD from long distance range expansion. The general lack of geographically dependent structure was particularly well documented in several mixed populations (e.g., RuAb, TuRb and CZ1Rr), containing haplotypes separated by many mutational steps and distributed across distant clades of the network (Fig. 4).

4. Discussion

The data presented here contain several strong components of population structure. The most conspicuous is an apparent correspondence between genetics and geography on a broad geographic scale (Fig. 1). This finding is not entirely unexpected since genetic variability due to long-distance separation is a common phenomenon. However, despite this general correspondence, the genealogical/phylogenetic arrangement does not imply any simple succession of vicariance events. The mitochondrial clades, although geographically restricted, do not cluster in any obvious pattern corresponding to the evolutionary history or distances between localities. It is reasonable to assume that at least in some cases the phylogeographic pattern is affected by unique long-distance transmissions mediated by introduced fish or migrating birds and followed by local diversification. For example, our set contains, to our knowledge, the first reported molecular data on *Ligula* specimens from the southern hemisphere, namely haplotypes H4 and H5 from Australia. Fish infections with *Ligula* have previously been reported from various sites in Australia and New Zealand, and both salmonid fish and fish-eating birds have been suggested as factors responsible for their introduction (Pollard, 1974; Weekes and Penlington, 1986; McDowall, 1990; Morgan, 2003). This view corresponds with our DNA-based results: despite considerable geographic distance and isolation, the Australian samples were placed firmly within clade B. However, since Australian fauna was only represented by samples from two adjacent localities, its overall diversity could be seriously underestimated. Similarly, Canadian, Mexican and Ethiopian haplotypes formed well-supported distinct clades, although based on only a few samples. Thus, with knowledge as it stands and the data currently available, it would be difficult to infer any global phylogeographic pattern.

Compared with the geography-based structure, the host-specificity component is expressed in a weaker but intriguing way. Within the most abundant clades A and B, no obvious correspondence was found between tapeworm phylogeny and specificity to fish-hosts. This indicates that within the geographically delimited clades, tapeworms are able to utilize several intermediate hosts without suffering from any apparent genetic isolation. Such mixing is consistent with the current view of the life cycle of *L. intestinalis*, since the tapeworms from different fish are supposed to cross-breed in their definitive bird host. In contrast to this intraclade homogeneity, striking differences in composition of intermediate-host spectra can be found among the clades, even if they occur in sympatry. For example, out of 66 samples collected from the European species and clustering within clade A, 65 were collected from derived cyprinids. The single exception is sample Ua Cc from Ukraine, isolated from a basal

cyprinid fish *Carassius carassius*. No samples from derived cyprinids were placed into the related clade B, composed of tapeworms from basal cyprinids and other fish orders (Table 1).

The concept of cryptic species in which there are no discernible morphological characters despite genetic differences in mitochondrial loci is well documented. For example, sequence divergence between the chewing louse, *Columbicola passerinae*, parasitising the blue ground dove, *Claravis pretiosa,* differs from *C. passerinae* parasitising the common ground dove by 11.3% for a portion of the COI mitochondrial gene, indicating the two louse populations could be reclassified as two distinct species (Johnson et al., 2002). In our study, the genetic divergence seen in the mitochondrial COI gene tree between samples of haplotypes A and B was 8.1 ± 0.02%. By comparison, the genetic divergence between two cestode species, *Paranoplocephala buryatiensis* and *Paranoplocephala longivaginata* using COI sequence data was 7.05 ± 0.47% and intraspecific genetic diversity was 0.5 ± 0.5% and 0.2 ± 0.08%, respectively (Haukisalmi et al., 2007). Blouin et al. (1998) compared mtDNA sequence variation among individual nematodes of the same species and closely related species and showed that the typical intraspecific range was 2-6% and the interspecific range was 10-20%. Taking into account all these data, it is likely that genetic divergence alone cannot be straightforwardly used as proof of cryptic speciation. However, the analogy with closest known examples, i.e. the tapeworms of the genus *Paranoplocephala*, indicates that the clades A and B could in principle represent two different biological species.

Regardless of their taxonomic status, the European haplotypes of clades A and B provide a typical instance of two genetically distinct lineages/species maintained in sympatry. Such situations constitute potential models for studying the speciation processes in parasites. It has been emphasized that to analyze the speciation scenario correctly, an appropriate usage of the sympatric/allopatric concept is an important prerequisite, which requires a priori knowledge of the parasite's biology and transmission mode (McCoy, 2003; Giraud, 2006). For example, parasite lineages inhabiting different host species living in sympatry should only be considered sympatric if the hosts do not pose an extrinsic barrier to their genetic mixture (Giraud, 2006). In the case of *L. intestinalis*, adult worms develop and mate in the intestine of a fish-eating bird serving as a definitive host. Despite their specificity to different intermediate hosts, the members of both clades A and B were found in the same definitive host, *P. cristatus*, sampled from the same locality. Therefore, the genetic isolation of clades A and B is maintained in sympatry and requires an explanation based on biological traits rather than physical isolation.

Such sympatric occurrence of two phylogenetically related lineages can be explained either by sympatric speciation or by a secondary encounter of cryptic species originating from allopatry. In a context of parasitology, disruptive selection has regularly been invoked as a possible means of sympatric speciation dependent on host biology (Théron and Combes, 1995; McCoy et al., 2001). It has even been suggested that the fish feeding strategy can determine the composition of its parasitic fauna and thus affect the overall co-evolutionary scenario (Jousson et al., 2000; Hypša et al., 2005; Criscione et al., 2005). We may hypothesize that the feeding strategy of pelagic (clade A) versus benthic fish (clade B) represents a potential diversifying force in parasites. In Lake Tana in Ethiopia, host specificity of *L. intestinalis* was shown to be likely attributable to host food type. Dejen et al. (2006) reports that *L. intestinalis* shows host specificity to *Barbus tanapelagius,* a specialized zooplanktivore, compared with *B. humilis* which has both zooplanktivore and benthic invertebrate food types.

A secondary encounter of two or more cryptic species with different intermediate host spectra seems a more plausible explanation. We hypothesize the scenario of host-parasite co-speciation (Hafner and Nadler, 1990; Hafner and Page, 1995; Hoberg et al., 1997; Hughes, 2007). In such a scenario, the two genetically distinct European lineages would have originated

in allopatry and adapted to different groups of intermediate hosts. After a secondary encounter, these cryptic species retained genetic isolation and different host specificities. Indeed, the two groups of parasites reflected the divergence in host ranges with *Ligula* from Clade A recovered exclusively from derived cyprinid fish and *Ligula* from Clade B from basal cyprinid species. The phylogenetic analysis of European cyprinids showed the existence of two main fish clades corresponding to the sub-families Cyprininae and Leuciscinae (Zardoya and Doadrio, 1998; Gilles et al., 2001). The divergence of Cyprinidae based on fossil evidence took place approximately 27.7 million years ago (Zardoya and Doadrio, 1999) and we hypothesize that the two *Ligula* clades are the product of a co-evolution event. Unfortunately, calibrated rates are lacking for platyhelminths because they do not leave a fossil record (Haukisalmi et al., 2007).

This latter view (i.e. the encounter of already separate cryptic species) is also supported by several previous findings. First, Olson et al. (2002) described two sympatric lineages of *Ligula* sp. from Northern Ireland with different types of ITS rDNA. These lineages, characterized by their affinity to two different intermediate hosts, *Rutilus* and *Gobio*, obviously correspond to our clades A and B. As noted by Olson et al. (2002), *Rutilus* is not native to Ireland. Its introduction in the 1970s was later followed by an increased occurrence of the great crested grebe, the final host of *Ligula.* It is likely that grebes then secondarily introduced clade A *Ligula* into Ireland where clade B was already well established. We found a similar situation in *Ligula* populations from northern Africa. Whereas the Algerian samples collected from *Barbus* clustered within clade B, the samples recovered from *Rutilus rubilio* in a neighboring locality in Tunisia were placed into the “European” clade A. Since *R. rubilio* was introduced into Tunisia from Southern Europe in the 1960s (Losse, G.F., Nau, W., Winter, M., 1991. Le développement de la pêche en eau douce dans le nord de la Tunisie Projet de la coopération technique. Coopération technique Tuniso-Allemande, CGP/GTZ), we assume that *Ligula* clade A was either imported with the infected fish stock or immigrated later with the final bird host. Although both native and introduced fish species are nowadays common in the North African region, their *Ligula* fauna do not seem to undergo any substantial mixing.

We found that *D. interrupta* clustered within the *L. intestinalis* tree. This has already been reported by Logan et al. (2004) and Kuchta et al. (2007) who synonymized *Digramma* with *Ligula*. In this study, we encountered a conflict between the position of our *D. interrupta* samples isolated from the fish genus *Hemiculter* in Russia and China (haplotypes 9 and 43) and the Chinese samples collected and characterized previously (haplotypes 12 and 13) by Li and Liao (2003). This conflict cannot be unequivocally explained; however several facts indicate that the molecular data from the study of Li and Liao (2003) should be treated with caution. Analyzing COI sequences from several Chinese samples of *Digramma* and *Ligula*, these authors found extremely low variability, represented by only two haplotypes shared by both tapeworm species. Such genetic homogeneity contradicts not only the high genetic variation found in our COI set but also their own results obtained with a different mitochondrial gene; when using one sequence of NADH deshydrogenase subunit I, they detected clear differences between *Ligula* and *Digramma* samples. Since the authors worked with formalin-preserved tapeworms, the contamination originated during the process of DNA isolation might be an explanation. Still less clear is the taxonomic status of another species, *L. colymbi*. In our sample set, we found the suggested morphological characteristics unreliable and difficult to assess. Since the great crested grebe, *P. cristatus*, known to harbor both *L. intestinalis* and *L. colymbi*, is often considered a typical final host of the latter species (Dubinina, 1980), we paid special attention to the phylogenetic relationships of the adults isolated from *P. cristatus*. The tapeworms obtained from a single locality clustered among typical *L. intestinalis* samples within the clades A (five samples; haplotypes 14, 23, 24) and B (two samples; haplotypes 25,

26). This finding shows that species identification based on the final host only may be unreliable.

ITS has been the most frequently used marker in previous phylogenetic studies in the genus *Ligula* and in diphyllobothriid tapeworms in general (Olson et al., 2002; Li and Liao, 2003; Luo et al., 2003; Logan et al., 2004), thus knowledge of the occurrence of paralogues and the degree of intragenomic variability should be an important basis for evolutionary interpretations in such studies. In our study, we confirmed the presence of numerous ITS2 copies within a single genome and demonstrated that in some cases the copies may differ considerably in their phylogenetic position. In spite of this potentially serious disruption of the phylogenetic signal, the COI and ITS2 topologies displayed an unexpectedly high degree of congruence. Almost all of the clades revealed by COI were mirrored by corresponding monophyletic ITS2 clades (Fig. 2). The only exception was clade A, where the difference between COI and ITS2 was due to different levels of resolution rather than contradicting signals. We assume that the occurrence of divergent copies within a single genome is relatively rare and does not affect the overall topology in population studies where the sample is sufficiently large. It may, however, cause serious distortions in phylogenetic studies where few samples are collected and each is represented by a single ITS2 copy.

As many hundreds or even thousands of ITS2 copies are present in a single genome (e.g. Buckler-Iv et al., 1997; Lewis and Doyle, 2002; Pfeil et al., 2004; Bayly and Ladiges, 2006; Ganley and Kobayashi, 2007), it is unfeasible to thoroughly analyse the complete intragenomic variability. It is therefore impossible to distinguish the portion of the heterogeneity that is due to standard long-term variability within a single haploid genome, and how much of the variability should be attributed to occasional crossing among the clades. Considering the overall fit between the COI and ITS2 phylogenies, we assume that occasional crossings rather than a high degree of ancestral polymorphism are responsible for the presence of highly diverged forms within a single tapeworm. In this respect, it is interesting to note that no such group of highly divergent copies was observed in *D. interrupta*. This seems to support the status of *D. interrupta* as an isolated cryptic species which does not cross-breed with other lineages/species of the complex. Similarly, the pattern shown in Fig. 2 (i.e. the overall correspondence between COI and ITS2 with occasional "switches") indicates several well-formed and mostly isolated lineages/species of *Ligula* with rare crossings.

Besides their different host ranges, the two Euro-Mediterranean clades A and B differ in the parameters of their population genetics. Although statistics may be affected by the low number of samples in clade B, the tests of neutrality (Table 5), mismatch distribution (Fig. 3) and coalescence analysis (Table 6) provided identical patterns. Population sizes of the two main lineages were shown to differ: clade A was characterized as fast-expanding and having a wider distribution than clade B. This is in concordance with the host distribution of the clades: while clade A infects widespread fish taxa with high population densities, the European lineage of clade B was sampled exclusively from less common host taxa (information obtained from www.fishbase.org). As discussed above, the distribution of tapeworms from different fish species within the inner topology of clade A did not suggest any obvious co-speciation between the parasite and its intermediate host. Rather, the different tapeworm frequencies found in different fish species indicate that a certain degree of specificity might be expressed in the differential prevalence of various genotypes. Loot et al. (2001) described preferential infection of *Rutilus* by *L. intestinalis* at several localities in France. Even though both roach and bream are abundant species in France, among the 1,385 specimens of bream examined only two were infected (0.2%), whereas the prevalence of *Ligula* in roach at the same localities was 40% (G. Loot, unpublished data). Similarly, Griffiths and Bigsby (unpublished data ex Olson et al., 2002) did not find any *Ligula* infections in *Abramis* in Lough Neagh. By contrast,

L. intestinalis infections of *Abramis* are very common in middle and eastern European areas (eg. Dubinina, 1980; Baruš and Prokes, 1994). In the future, additional studies using fine-scale molecular markers (microsatellite sequence) and experimental cross-infection are needed to explore patterns of local parasite adaptation (see Greischar and Koskella, 2007) in clade A.

In the same way, no effect of the intermediate fish host on the tapeworm population structure was found using the haplotype network as a basis for NCPA. In this analysis, the flat population structure with respect to geography was mainly attributed to contiguous range expansion. Of the two equally probable hypotheses suggested for clade 3-3, we consider the range expansion with long distance dispersal more plausible since subclade 2-9 is formed by Tunisian samples. Indeed, these tapeworms are thought to have been recently imported into Tunisia and to have become genetically isolated from continental Europe by the Mediterranean Sea (W. Bouzid et al., unpublished data).

Within clade A, genealogical discordance between parasite and host phylogenies suggest extensive gene flow among parasites across the host species spectra. Among the factors reducing genetic variation within populations of parasites, an important role is usually attributed to the dissemination of eggs by highly mobile definitive hosts (Kennedy, 1998; Gittenberger et al., 2006). It has recently been demonstrated in a system of cladoceran crustaceans that waterfowl migrations can affect the genetic structure of aquatic invertebrates (Figuerola et al., 2005). The birds that act as potential definitive hosts of *Ligula* in Europe are mostly migrating and wintering species (Curry-Lindahl, 1980). For example, the Black-headed gull, *Larus ridibundus,* and the Herring gull, *Larus argentatus,* have a widespread dispersion and display wintering erraticism over central and southern Europe (Curry-Lindahl, 1980). Some populations even migrate to North African coasts and can reach as far as the African tropics during the winter (Curry-Lindahl, 1980). The Great Crested Grebe *P. cristatus*, from which several mature tapeworms were analyzed in this study, winters in several European countries and spends summer on Atlantic and North Sea coasts. Some of them migrate to North Africa and join the sedentary populations of the same species. Similar migration behaviour can be observed in many other definitive host species. Even though the duration of *Ligula* infection in a bird host is very short, usually lasting only 2 to 5 days (Dubinina, 1980), birds are probably able to disseminate the eggs over long distances within a short period of time during seasonal migration (Wyatt and Kennedy, 1989).

The origin and evolution of host specificity in parasites is one of the perpetual questions in parasitology. In host-parasite associations with strong co-evolutionary signals, co-speciation events are usually considered as the mechanisms behind host specificity. However, in parasites with free-living stages, mobile hosts and thus the capacity for population mixing, local adaptation rather than co-speciation processes might be a more important source of specificity (Gandon and Michalakis, 2002; Greischar and Koskella, 2007). Since most of the previous analyses on tapeworm phylogeny were performed at higher phylogenetic levels, it is difficult to make any strong generalization by comparing our results with other published analyses. It is, however, interesting to note that various phylogenetic studies on tapeworms indicate little sign of co-evolution or even completely uncoupled phylogenies of the parasites and their hosts (Caira and Jensen, 2001; Škeříková et al., 2001; Wickström et al., 2003). The phylogenetic patterns of the *Ligula* samples discussed above might provide an interesting insight into the origin of host specificity in this type of parasite. At low phylogenetic levels (i.e. within the clades) the homogenization guaranteed by mobile definitive hosts results in remarkably shallow population structures and prevents co-speciation between host and parasite. In genetically distinct lineages or cryptic species (clades A and B), we suggest a scenario of co-evolution between the tapeworm *Ligula* and its host range. In sufficiently divergent cryptic species, the different host specificities naturally persist even when the parasite populations

meet again and coexist in sympatry. Finally, at high phylogenetic levels, geographically isolated lineages diverge due to physical isolation and are thus able to adapt to local host fauna.

Acknowledgements

The following persons kindly provided samples of *Ligula* tapeworms or helped considerably in their sampling: J. Sitko, Czech Republic; C. Arme, United Kingdom; A. Aydogdu, Turkey; T. Boutorina, Russia; J.M. Caspeta-Mandujano, Mexico; A. de Chambrier, Switzerland; M. Davidová, Czech Republic; M. Borga Ergonul, Turkey; A. Kostadinova, Bulgaria; R. Kuchta, Czech Republic; S. Frank, Germany; A. Abdessalem, Algeria; A. Lymbery, Australia; P. Nie, China; A. Kangur and K. Kangur, Estonia. Samples of *Ligula* tapeworms from Lake Tana could not be obtained without the help of Abebe Getahun, Seyoum Mengistou, Eshete Dejen Dresilign, Moges Beleteu, and the staff of the ANRS Agricultural Research Institute, Bahir Dar, Ethiopia. The study was supported by grants LC06073 and MSM 60076605801 (Ministry of Education, Czech Republic), Barrande project (Egide, France), Grant Agency of the Czech Republic (projects Nos. 206/08/1019 and 524/04/0342) and Institute of Parasitology (research projects Z60220518 and LC 522).

References

Anderson, T.J.C., 2001. The dangers of using single locus markers in parasite epidemiology: Ascaris as a case study. Trends Parasitol. 17, 183-188.

Arme, C., 1997. Ligulosis in two cyprinid hosts: *Rutilus rutilus* and *Gobio gobio*. Helmintholgia 34, 191-196.

Arme, C., Owen, R.W., 1968. Occurrence and pathology of *Ligula intestinalis* infections in British fishes. J. Parasitol. 54, 272-280.

Barus, V., Prokes, M., 1994. Parasite load of *Ligula intestinalis* plerocercoids in adult silver bream, *Blicca bjoerkna.* Helminthologia 31, 91–94.

Barus, V., Prokes, M., 2002. Length and weight of *Ligula intestinalis* plerocercoids (Cestoda) parasitizing adult cyprinid fishes (Cyprinidae): a comparative analysis. Helminthologia 39, 29-34.

Bayly, M.J., Ladiges, P.Y., 2007. Divergent paralogues of ribosomal DNA in eucalypts (Myrtaceae). Mol. Phylogenet. Evol. 44, 346-356.

Bean, C.C., Winfield, I.J., 1992. Influences of the tapeworm *Ligula intestinalis* (L.) on the spatial distributions of juvenile roach *Rutilus rutilus* (L.) and gudgeon *Gobio gobio* (L.) in Lough Neagh, Northern Ireland. Neth. J. Zool. 42, 416-429.

Blouin, M.S., Yowell., C.A., Courtney, C.H., Dame, J.B., 1998. Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. Mol. Biol. Evol. 15, 1719-1727.

Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1995. Host movement and the genetic structure of populations of parasitic nematodes. Genetics 141, 1007-1014.

Brant, S.V., Ortí, G. 2003. Evidence for gene flow in parasitic nematodes between two host species of shrews. Mol. Ecol. 12, 2853-2859.

Buckler-Iv, E.S., Ippolito, A., Holtsford, T.P., 1997. The Evolution of Ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics 145, 821-832.

Caira, J.N., Jensen. K., 2001. An investigation of the co-evolutionary relationships between onchobothriid tapeworms and their elasmobranch hosts. Int. J. Parasitol. 31, 960-975.

Carter, V., Pierce, R., Dufour, S., Arme, C., Hoole, D., 2005. The tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea) inhibits LH expression and puberty in its teleost host, *Rutilus rutilus*. Reproduction 130, 939-945.

Chapman, A., Hobbs, R.P., Morgan, D.L., Gill, H.S., 2006. Helminth parasitism of *Galaxias maculatus* (Jenyns 1842) in southwestern Australia. Ecol. Freshw. Fish. 15, 559-564.
Criscione, C.D., Blouin, M.S., 2004. Life cycles shape parasite evolution: Comparative population genetics of salmon trematodes. Evolution 58, 198-202.
Criscione, C.D., Poulin, R., Blouin, M.S., 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Mol. Ecol. 14, 2247-2257.
Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657-1660.
Curry-Lindahl, K.,1980. Les oiseaux migrateurs à travers mer et terre. Delachaux et Niestle, Neuchatel, Paris.
Dejen, E., Vijverberg, J., Sibbing, F.A., 2006. Spatial and Temporal Variation of Cestode Infection and its Effects on Two Small Barbs (*Barbus humilis* and *B. tanapelagius*) in Lake Tana, Ethiopia. Hydrobiologia 556, 109-117.
Dover, G., 1982. Molecular drive: a cohesive mode of species evolution. Nature 299, 111-117.
Dubinina, M.N., 1980. Tapeworms (Cestoda, Ligulidae) of the Fauna of the USSR. Amerind Publishing, New Delhi, India.
Figuerola, J., Green, A.J., Michot, T.C., 2005. Invertebrate eggs can fly: Evidence of waterfowl-mediated gene flow in aquatic invertebrates. Am. Nat. 165, 274-280.
Gandon, S., Michalakis, Y., 2002. Local adaptation, evolutionary potential and host-parasite coevolution: interactions between migration, mutation, population size and generation time. J. Evolution Biol. 15, 451-462.
Ganley ARD, Kobayashi T (2007) Highly efficient concerted evolution in the ribosomal DNA repeats: Total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res. 17, 184-191.
Gilles, A., Lecointre, G., Miquelis, A., Loerstcher, M., Chappaz, R.M., Brun, G., 2001. Partial Combination Applied to Phylogeny of European Cyprinids Using the Mitochondrial Control Region. Mol. Phylogenet. Evol. 19, 22-33.
Giraud, T., 2006. Speciation in parasites: host switching does not automatically lead to allopatry. Trends Parasitol. 22, 151-152.
Giribet, G., 2005. TNT: Tree analysis using New Technology. Syst. Biol. 54, 176-178.
Gittenberger, E., Groenenberg, D.S.J., Kokshoorn, B., Preece, R.C., 2006. Molecular trails from hitch-hiking snails. Nature 439, 409-409.
Greischar, M.A., Koskella, B., 2007. A synthesis of experimental work on parasite local adaptation. Ecol. Lett. 10, 418-434.
Groves, K.L., Shields, B.A., 2001. Observations on the plerocercoid stage of the tapeworm *Ligula* in three Species of fish from the lower crooked river of central Oregon. J. Aquat. Anim. Health 13, 285-289.
Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online — a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 33, W557–W559.
Hafner, M.S., Nadler, S.A., 1990. Cospeciation in Host-Parasite Assemblages: Comparative Analysis of Rates of Evolution and Timing of Cospeciation Events. Syst. Zool. 39, 192-204.
Hafner, M.S., Page, R.D.M., 1995. Molecular phylogenies and host-parasite cospeciation: Gophers and Lice as a model system. Philos T. Roy. Soc. B. 349, 77-83.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95-98.
Haukisalmi, V., Hardman, L.M., Hardman, M., Laakkonen, J., Niemimaa, J., Henttonen, H., 2007. Morphological and molecular characterisation of *Paranoplocephala buryatiensis n.*

sp. and *P. longivaginata* Chechulin & Gulyaev, 1998 (Cestoda: Anoplocephalidae) in voles of the genus *Clethrionomys*. Systematic Parasitol. 66, 55-71.

Hoberg, E.P., Brooks, D.R., Siegel-Causey, D., 1997. Host-parasite cospeciation: History, principles, and prospects. In: Clayton, D., Moore, J. (Eds.), Host-parasite Evolution: General principles and avian models, Oxford University Press, pp. 212-235.

Hughes, J., 2007. Multiple Cophylogenetic Analyses Reveal Frequent Cospeciation between Pelecaniform Birds and *Pectinopygus* Lice. Syst. Biol. 56, 232-251.

Hypša, V., Škeříková, A., Scholz, T., 2005. Phylogeny, evolution and host–parasite relationships of the order Proteocephalidea (Eucestoda) as revealed by combined analysis and secondary structure characters. Parasitology 130, 359-371.

Jarne, P., Théron, A., 2003. Genetic structure in natural populations of flukes and snails: a practical approach and review. Parasitology 123, 27-40.

Johnson, K.P., Williams, B.L., Drown, D.M., Adams, R.J., Clayton, D.H., 2002. The population genetics of host specificity: genetic differentiation in dove lice (Insecta: Phthiraptera). Mol. Ecol. 11, 25-38.

Jousson, O., Bartoli, P., Pawlowski, J., 2000. Cryptic speciation among intestinal parasites (Trematoda: Digenea) infecting sympatric host fishes (Sparidae). J. Evol. Biol. 13, 778-785.

Kennedy, C.R., 1998. Aquatic birds as agents of parasite dispersal: a field test of the effectiveness of helminth colonisation strategies. Bull. Scand. Soc. Parasitol. 8, 23–28.

Kennedy, C.R., Burrough, R.J., 1981. The establishment and subsequent history of a population of *Ligula intestinalis* in roach *Rutilus rutilus* (L.). J. Fish Biol. 19, 105-126.

Kuchta, R., Scholz, T., Brabec, J., Bray, R.A., 2007. Suppression of the tapeworm order pseudophyillidea (Platyhelminthes: Eucestoda) and the proposal of two new orders, Bothricephalidea and Diphyllobothriidea. Int. J. Parasitol. 38, 49-55.

Kuhner, M.K., 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22, 768-770.

Lewis, C.E., Doyle, J.J., 2002. A phylogenetic analysis of tribe Areceae (Arecaceae) using two low-copy nuclear genes. Plant Syst. Evol. 236, 1-17.

Li, J., Liao, X., 2003. The taxonomic status of *Digramma* (Pseudophyllidae: Ligulidae) inferred from DNA sequences. J. Parasitol. 89, 792-799.

Li, J., Liao, X., Yang, H., 2000. Molecular characterization of a parasitic tapeworm (*Ligula*) based on DNA sequences from formalin-fixed specimens. Biochem. Genet. 38, 309-322.

Logan, F.J., Horák, A., Štefka, J., Aydogdu, A., Scholz, T., 2004. The phylogeny of diphyllobothriid tapeworms (Cestoda: Pseudophyllidea) based on ITS-2 rDNA sequences. Parasitol. Res. 94, 10-15.

Loot, G., Aulagnier, S., Lek, S., Thomas, F., Guegan, J.F., 2002. Experimental demonstration of a behavioural modification in a cyprinid fish, *Rutilus rutilus* (L.), induced by a parasite, *Ligula intestinalis* (L.). Can. J. Zool. 80, 738-744.

Loot, G., Francisco, P., Santoul, F., Lek, S., Guegan, J.F., 2001. The three hosts of the *Ligula intestinalis* (Cestoda) life cycle in Lavernose-Lacasse gravel pit, France. Arch. Hydrobiol. 152, 511-525.

Loot, G., Park, Y.S., Lek, S., Brosse, S., 2006. Encounter rate between local populations shapes host selection in complex parasite life cycle. Biol. J. Linn. Soc. 89, 99-106.

Luo, H.Y., Nie, P., Yao, W.J., Wang, G.T., Gao, Q., 2003. Is the genus *Digramma* synonymous to the genus *Ligula* (Cestoda: Pseudophyllidea)? Parasitol. Res. 89, 419-421.

McCoy, K.D., Boulinier, T., Tirard, C., Michalakis, Y., 2001. Host specificity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick *Ixodes uriae*. J. Evolution Biol. 14, 395-405.

McCoy, K.D., 2003. Sympatric speciation in parasites–what is sympatry? Trends Parasitol. 19, 400-404.

McDowall, R.M., 1990. New Zealand Freshwater Fishes: A Natural History and Guide, Heinemann Reed, Auckland, New Zealand

McManus, D.P., 1985. Enzyme analyses of natural populations of *Schistocephalus solidus* and *Ligula intestinalis*. J. Helminthol. 59, 323–332.

Möller, M., 2000. How universal are universal rdna primers? A cautionary note for plant systematists and phylogeneticists. Edinb. J. Bot. 57, 151-156.

Morgan, D.L., 2003. Distribution and biology of *Galaxias truttaceus* (Galaxiidae) in south-western Australia, including first evidence of parasitism of fishes in Western Australia by *Ligula intestinalis* (Cestoda). Environ. Biol. Fish. 66, 155-167.

Museth, J., 2001. Effects of *Ligula intestinalis* on habitat use, predation risk and catchability in European minnows. J. Fish Biol. 59, 1070-1080.

Olson, P.D., Littlewood, D.T.J., Griffiths, D., Kennedy, C.R., Arme, C., 2002. Evidence for the co-existence of separate strains or species of *Ligula* in Lough Neagh, Northern Ireland. J. Helminthol. 76, 171-174.

Page, R.D.M., Lee, P.L.M., Becher, S.A., Griffiths, R., Clayton, D.H., 1998. A different tempo of mitochondrial DNA evolution in birds and their parasitic lice. Mol. Phylogenet. Evol. 9, 276-293.

Panchal, M., 2007. The automation of nested clade phylogeographic analysis. Bioinformatics 23, 509-510.

Pfeil, B.E., Brubaker, C.L., Craven, L.A., Crisp, M.D., 2004. Paralogy and orthology in the Malvaceae rpb2 gene family: investigation of gene duplication in *Hibiscus*. Mol. Biol. Evol. 21, 1428-1437.

Pollard, D.A., 1974. The biology of a landlocked form of the normally catadromous salmoniform fish *Galaxias maculatus* (Jenyns) VI. Effects of cestode and nematide parasites. Aust. J. Mar. Freshwat. Res. 25: 105-120.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Posada D, Crandall, K.A., Templeton, A.R., 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Mol. Ecol. 9, 487-488.

Poulin, R., 1998. Evolutionary Ecology of Parasites: From Individuals to Communities. Chapman and Hall, London.

Prugnolle F, Liu H, de Meeûs T, Balloux F. 2005. Population genetics of complex life-cycle parasites: an illustration with trematodes. Int J. Parasitol. 35, 255-263.

Rozas, J., Sánchez-DelBarrio, J.C., Messeguer, X., Rozas, R., 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19, 2496-2497.

Škeri´ková, A., Hypša, V., Scholz, T., 2001. Phylogenetic analysis of European species of *Proteocephalus* (Cestoda: Proteocephalidea): compatibility of molecular and morphological data, and parasite-host coevolution. Int. J. Parasitol. 31, 1121-1128.

Sweeting, R.A., 1977. Studies on *Ligula intestinalis*. Some aspects of the pathology in the second intermediate host. J. Fish Biol. 10, 43-50.

Taylor, M., Hoole, D., 1989. *Ligula intestinalis* (L.) (Cestoda: Pseudophyllidea): An ultrastructural study of the cellular response of roach fry, *Rutilus rutilus* L., to an unusual intramuscular infection. J. Fish Dis. 12, 523-528.

Templeton, A.R., 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. Mol. Ecol. 13, 789-809.

Théron, A., Combes, C., 1995. Asynchrony of Infection Timing, Habitat Preference, and Sympatric Speciation of Schistosome Parasites. Evolution 49, 372-375.

von Nickisch-Rosenegk, M., Brown, W.M., Boore, J.L., 2001. Complete Sequence of the Mitochondrial Genome of the Tapeworm Hymenolepis diminuta: Gene Arrangements Indicate that Platyhelminths Are Eutrochozoans. Mol. Biol. Evol. 18, 721-730.

Weekes, P.J., Penlington, B., 1986. First records of *Ligula intestinalis* (Cestoda) in rainbow trout, *Salmo gairdneri*, and common bully *Gobiomorphus cotidianus* in New Zealand. J. fish biol. 28, 183-190.

Wickström, L.M., Haukisalmi, V., Varis, S., Hantula, J., Fedorov, V.B., Henttonen, H., 2003. Phylogeography of the circumpolar *Paranoplocephala arctica* species complex (Cestoda: Anoplocephalidae) parasitizing collared lemmings (*Dicrostonyx* spp.). Mol. Ecol. 12, 3359-3371.

Wyatt, R.J., Kennedy, C.R., 1989. Host-constrained epidemiology of the fish tapeworm *Ligula intestinalis* (L.). J. Fish Biol. 35, 215-227.

Zardoya, R., Doadrio, I., 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. Proc. Biol. Sci. 265, 1365-1372.

Zardoya, R., Doadrio, I., 1999. Molecular evidence on the evolutionary and biogeographical patterns of european cyprinids. J. Mol. Evol. 49, 227-237.

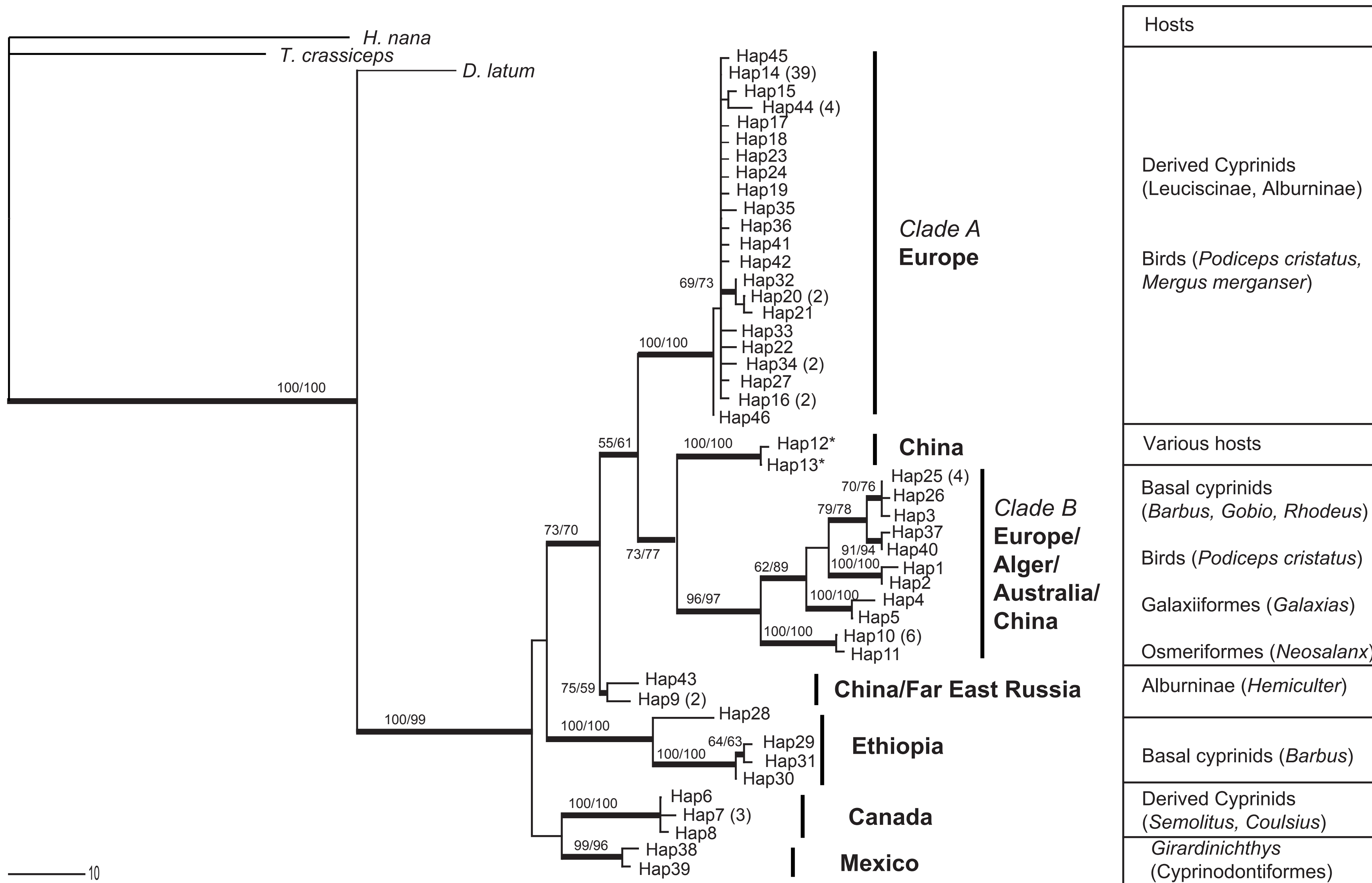


Fig. 1 One of the 288 MP trees provided assuming a Ts:Tv ratio of 1, 2 and 3; based on sequences of COI. The numbers at the nodes indicate bootstrap support values higher than 50% (MP/ML, 1000 replications). Strict consensus of all trees is depicted in bold lines. Haplotype numbers refer to the numbers listed in Tables 1 and 3. Asterisks indicate sequences obtained from GenBank. Numbers in parentheses are the number of samples grouped within a haplotype. Phylogenetic position of cyprinids (basal/derived) was based on Briolay et al. (1998); Gilles et al. (2001); Cunha et al. (2002) Durand et al. (2002); Liu & Chen (2003); Saitoh et al. (2006).

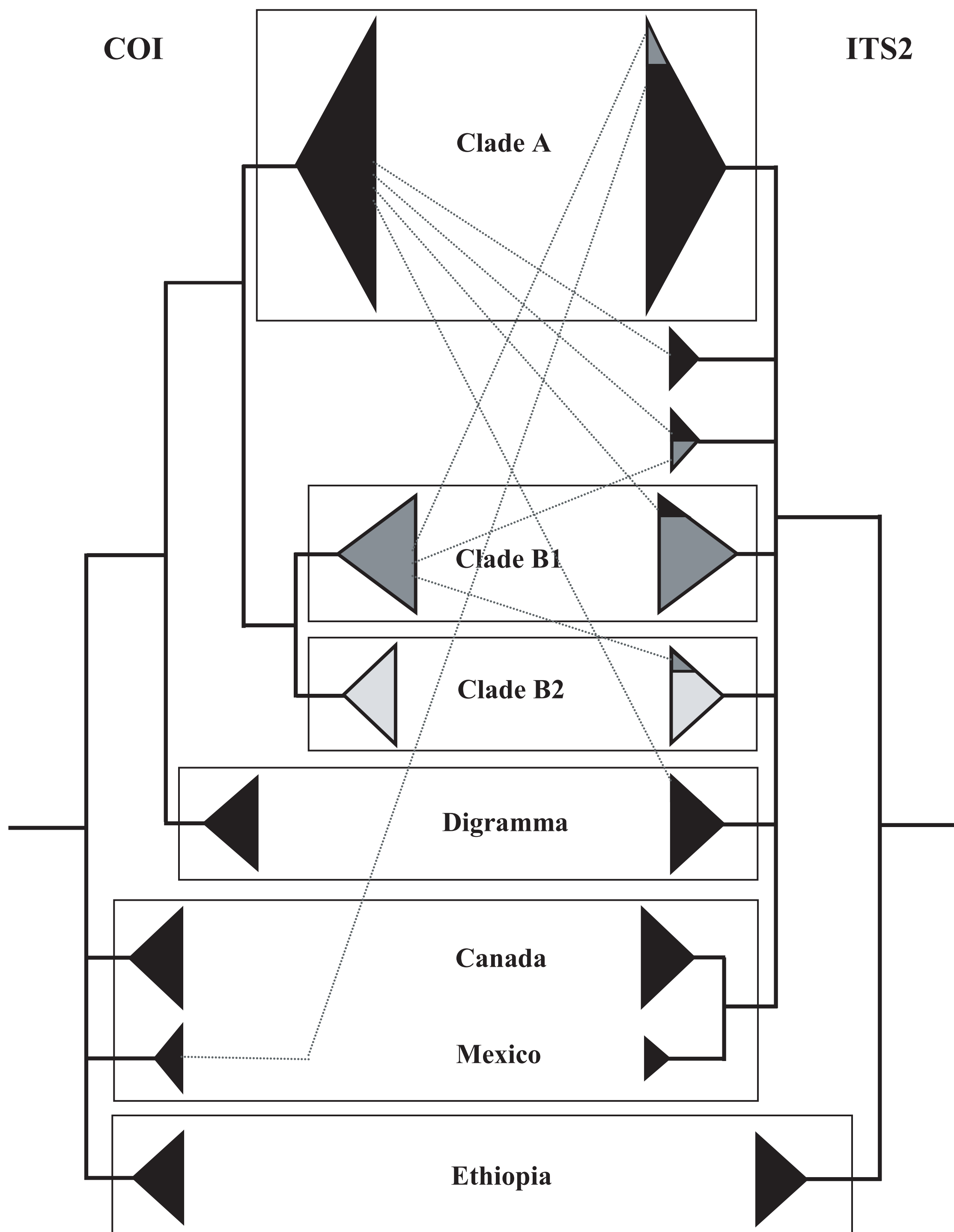


Fig. 2 Comparison of COI and ITS2 phylogenies. The COI tree corresponds to the topology shown in Fig. 1 with haplotypes 12 and 13 removed. The ITS2 tree is a strict consensus of 286 MP trees. Positions of putative paralogues and their counterparts in the COI topology are highlighted by dotted lines (see Table 4).

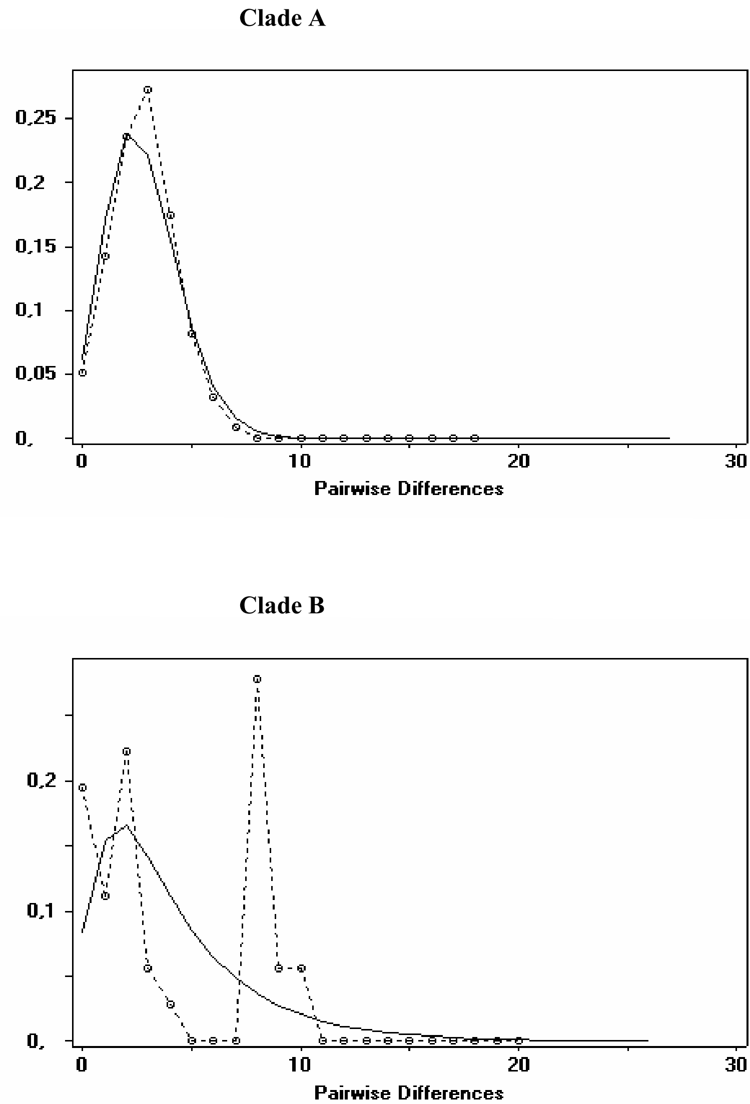


Fig. 3 Mismatch distribution of COB haplotypes for two European *Ligula* clades. The observed frequency of numbers of pairwise mismatches is represented by a dotted line. The expected frequency (solid line) was estimated under the population-growth model in DNASP 4.10 (Rozas *et al.* 2003).

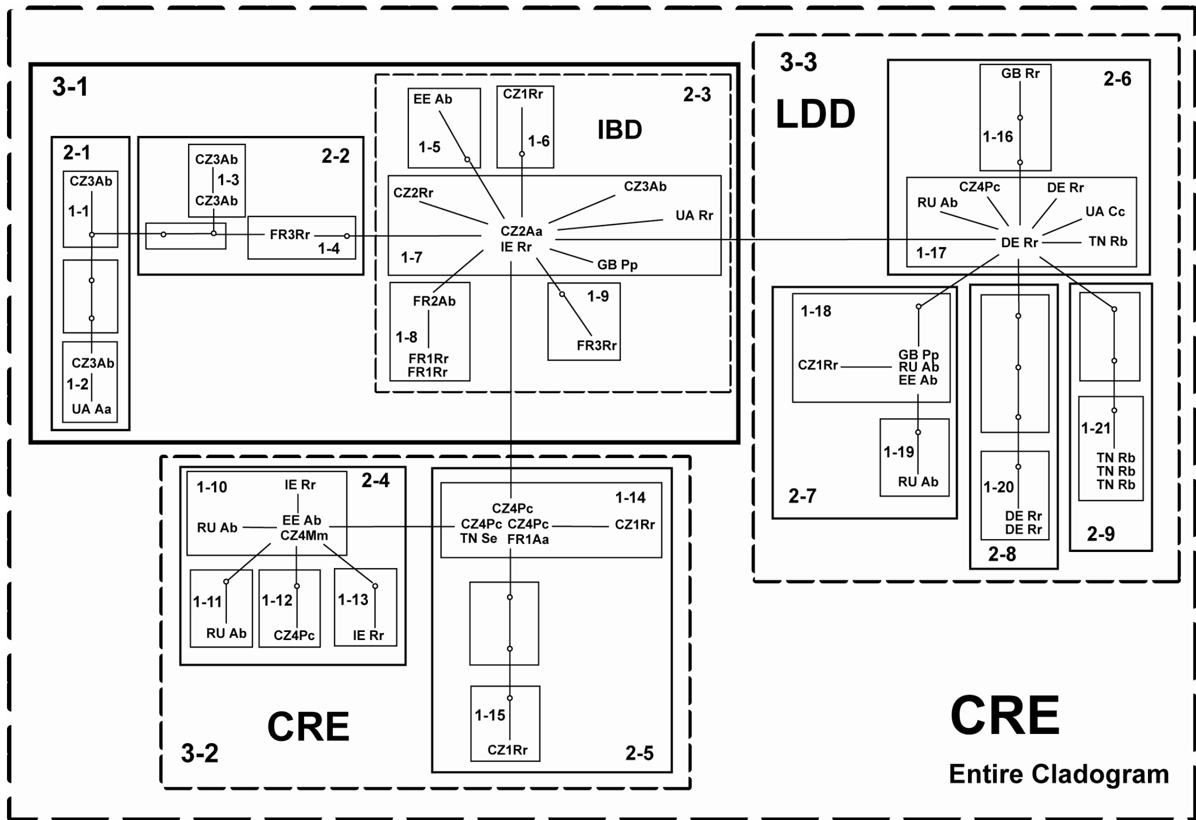


Fig. 4 Nested clade diagram obtained from concatenated matrix of the COI and COB haplotypes of the clade A. Neighbouring haplotypes are connected by single mutation, open circles represent missing haplotypes along mutational pathways. Sample codes are given in Table 1. Nested clade levels are shown by the numbers within a particular nested clade. Dashed lines highlight nesting categories where populational events were recognized using the NCPA. **IBD**, Isolation by distance; **CRE**, Contiguous range expansion; **LDD**, Long distance dispersal. For levels of significance see Table 7.

Table 1. Geographic origin of the samples sequenced here.

Geographic origin		Symbol	Host species	Species	No. of samples analysed per gene		
					ITS2	COI	COB
Algeria	Oued Hamiz Reservoir	Al BS	*Barbus* sp.	*L.* sp.	4	3 (1,2,3)	3
Australia	Goodga River	AU Gt	*Galaxias truttaceus* (Osmeriformes)	*L.* sp.	1	1 (4)	1
	Moates Lake	Gm	*Galaxias maculatus* (Osmeriformes)	*L.* sp.	1	1 (5)	1
Canada	Dalpec Lake	CA Cp	*Coulsius plumbeus*	*L.* sp.	1	1 (8)	1
	Lac Dumbo	Sa	*Semotilus atromaculatus*	*L.* sp.	8	4 (6,7)	4
China	Dong Tink Lake	CN Hb	*Hemiculter bleekeri*	*D. i.*	2	2 (9)	2
	Zhanghe reservoir	Nt	*Neosalanx taihuensis* (Osmeriformes)	*L.* sp.	7	7 (10,11)	3
Czech Rep.	Lipno Reservoir	CZ1Rr	*Rutilus rutilus*	*L. i.*	8	9 (14,15,16,17,18)	4
	Zelivka Reservoir	CZ2Rr	*Rutilus rutilus*	*L. i.*	-	1 (19)	1
		Aa	*Alburnus alburnus*	*L. i.*	1	1 (14)	1
	Nove Mlyny Reservoir	CZ3Ab	*Abramis brama*	*L. i.*	13	12 (14,20,21,22)	6
	Tlumacov ponds	CZ4Pc	*Podiceps cristatus* – bird host	*L.* sp.	4	7 (14,23,24,25,26)	8
		Mm	*Mergus merganser* – bird host	*L.*sp.	-	1 (14)	1
Estonia	Peipsi Lake	EE Ab	*Abramis brama*	*L. i.*	5	3 (14, 27)	3
Ethiopia	Tana Lake	ET Bh	*Barbus humilis*	*L.* sp.	2	1 (28)	1
		Bt	*Barbus tsanensis*	*L.* sp.	1	1 (29)	1
		Bi	*Barbus intermedius*	*L.* sp.	1	1 (30)	1
		Bb	*Barbus brevicephalus*	*L.* sp.	1	1 (31)	-
France	Pareloup Reservoir	FR1Rr	*Rutilus rutilus*	*L. i.*	7	1 (14)	1
		Aa	*Alburnus alburnus*	*L. i.*	1	1 (14)	1
	Lavernose and Muret	FR2Rr	*Rutilus rutilus*	*L. i.*	4	-	-
		Bb	*Blicca bjoerkna*	*L. i.*	2	1 (14)	1
	Creteil Lake	FR3Rr	*Rutilus rutilus*	*L. i.*	7	3 (14, 32, 33)	3
Germany	Müggelsee Lake	DE Rr	*Rutilus rutilus*	*L .i.*	4	4 (34)	4
Great Britain	Scotland, River Gryfe	GB Rr	*Rutilus rutilus*	*L .i.*	-	1 (35)	1
		Pp	*Phoxinus phoxinus*	*L. i.*	-	1 (36)	1
	Wales, Aberystwyth	Pp	*Phoxinus phoxinus*	*L.* sp.	-	1 (14)	1
Mexico	Tonatzahua lake	MXGm	*Girardinichthys multiradiatus* (Cyprinodontiformes)	*L.* sp.	2	2 (38,39)	[a]
N. Ireland	Lough Neagh	IE Rr	*Rutilus rutilus*	*L.* sp.	-	3 (14)	3
		Gg	*Gobio gobio*	*L.* sp.	-	3 (25, 37)	3
Poland	Gdansk	PL Gs	*Gavia stellata* – bird host	*L.* sp.	2	1 (40)	-
	Wloclawski reservoir	PL Ra	*Rhodeus amarus*	*L.*sp.	1	1 (25)	1
Russia	Khanka Lake, Far East	RU Hl	*Hemiculter lucidus*	*D. i.*	-	1 (43)	1
	Rybinsk Reservoir	RU Ab	*Abramis brama*	*L. i.*	9	9 (14, 41,42)	5
Tunisia	Sidi Salem Reservoir	TN Rb	*Rutilus rubilio*	*L. i.*	6	5 (14,44)	4
		Se	*Scardinius erythrophthalmus*	*L. i.*	1	1 (14)	1
Ukraine	Dniester River	UA Rr	*Rutilus rutilus*	*L. i.*	1	1 (46)	1
		Aa	*Alburnus alburnus*	*L. i.*	1	1 (45)	1
		Cc	*Carassius carassius*	*L. i.*	1	1 (14)	1
Totals					109	99 (46)	76

L.i: *Ligula intestinalis*; *L.* sp: sample could not be unequivocally determined (*L. intestinalis* versus *Ligula colymbi*; see text for details); *D.i.*: *Digramma interrupta*.
Numbers in parentheses refer to haplotype numbers used in Fig. 1.
[a] No PCR product was obtained for this gene. Sequences obtained from GenBank are listed in Table 3.

Table 2. Primers used for PCR

Target genes	Primers Forward Reverse	Sequence (5'-3'direction)	Reference
ITS2 (rDNA)	Flo1	CGGTGGATCACTCGGCTC	Logan et al., 2004
	ITS2	TCCTCCGCTTATTGATATGC	
COI (mtDNA)	COIA2	CATATGTTTTGATTTTTTGG	This study
	COIB2	AKAACATAATGAAAATGAGC	Modified from Li et al., 2000
COB (mtDNA)	COBA	GTATGTGGCTGATTCAGAGTTGAGC	This study
	COBB	TTCGAGCCCGAAGAATGCAAGTAG	

Table 3. Sequences downloaded from GenBank

Geographic origin		Symbol	Host species	Species	No. of sequences analysed per gene		
					ITS2	COI	Accession no.
China	Qinghai	CN Gp	*Gymnocypris przewalskii*	*L. sp.*	2	-	**AY121751/52** [a]
	Jiangxi	CN Hb2	*Hemiculter bleekeri*	*D. i.*	1	-	**AF354293** [a]
	Hubei	CN Cd	*Culter dabryi*	*D. i.*	2	-	**AF354291/92** [a]
	Hubei	CN Ca	*Carassius auratus*		1	-	**AY12175** [a]
	various localities	CN XL	various Cyprinidae and *Protosalanx hyalocranius* (Osmeriformes)	*L. sp.*	-	1 (12)	**AF524041** [b]
	various localities	CN XD	various cyprinidae	*D. i.*	-	1 (13)	**AF153910** [c]
Czech Rep.	Lipno Reservoir	CZ1Rr	*Rutilus rutilus*	*L. i.*	1	-	**AY549520** [d]
	Nove Mlyny Reservoir	CZ3Ab	*Abramis brama*	*L. i.*	1	-	**AY549519** [d]
	Tlumacov Ponds	CZ4Pc	*Podiceps cristatus*	*L. c.*	2	-	**AY549510/11** [d]
Great Britain	Scotland, River Gryfe	GB Rr	*Rutilus rutilus*	*L. i.*	1	-	**AY549512** [d]
		Pp	*Phoxinus phoxinus*	*L. i.*	1	-	**AY549513** [d]
	Wales, Aberystwyth	Pp	*Phoxinus phoxinus*	*L. sp.*	1	-	**AF385760** [e]
	N. Ireland, Lough Neagh	IE Rr	*Rutilus rutilus*	*L. sp.*	3	-	**AF385761/63/64** [e]
		Gg	*Gobio gobio*	*L. sp.*	3	-	**AF385765/67/69** [e]
Russia	Khanka Lake, Far East	RU Hl	*Hemiculter lucidus*	*L. i.*	1	-	**AY549506** [d]
Turkey	Iznik Lake	TR CS	*Chalcaburnus sp.*	*L. i.*	1	[f]	**AY549517** [d]
		Sg	*Silurus glanis* (Siluriformes)	*L. i.*	1	[f]	**AY549516** [d]
Totals					22	2 (2)	

Tapeworm species determination was adapted from the original studies: [a] Luo et al., 2003; [b] Li and Liao, 2003; [c] Li et al., 2000; [d] Logan et al., 2004; [e] Olson et al., 2002. Numbers in parentheses refer to haplotype numbers used in Fig. 1.

[f] No PCR product was obtained for this gene.

Table 4. Intragenomic variability of ITS2 sequences.

Sample-Voucher No.	COI Haplotype No.	No. of plasmids sequenced	No. of haplotypic sequences:		
			total	Clustering within the COI clade	Clustering outside the COI clade, or pseudogene
CHNt8	10	20	9	7 (Clade B2)	2 (pseudogenes)
CZ1Rr15	14	20	11	9 (Clade A)	2 (pesudogenes)
RUAb2	14	20	9	7 (Clade A)	2 (unknown ITS only lineage)
RUHl	43	15	5	5 (*Digramma*)	0
CNNt6	10	2	2	1 (China)	1 (pseudogene)
EEAb1	14	1	1	0 (Clade A)	1 (Digramma)
EEAb3	14	2	2	1 (Clade A)	1 (pseudogene)
CZ4Pc107	14	2	2	0 (Clade A)	2 (unknown ITS only lineage)
PLGs1	40	2	2	0 (Clade B1)	1 (Clade A) 1 (Clade B2)
CZ3Ab26	14	2	2	1 (Clade A)	1 (pseudogene)
CZ4Pc2	26	[a]	1	0 (Clade B1)	1[a] (unknown ITS only lineage)
IERr6	14	2	3	1 (Clade A) 1[a] (Clade A)	1 (Clade B1)
MXGm2	39	1	1	0 (Mexico)	1 (Clade A)

[a] Sequences obtained from GenBank.

“Clustering within/outside the COI clade” refers to the ITS sequence(s) position in Fig. 2.

Table 5. Molecular variability and neutrality tests.

Clades		Sample size	No. of haplotypes	P_i	Tajima's D	Fu and Li's D	Fu and Li's F
A:	COI	66	22	0.0035	-2.428 [a]	-2.6425 [b]	-3.055 [a]
	COB	51	26	0.0069	-0.326	-1.185	-1.589
B:	COI	10	7	0.0178	-0.241	0.900	0.727
	COB	9	5	0.0102	-0.326	0.020	-0.144

P_i, nucleotide diversity; levels of significance:[a] $P < 0.01$, [b] $P < 0.05$

Table 6. Demographic parameters of European samples of *Ligula* given by coalescence analysis.

Clade	*Theta*	*g*
A	0.118 (0.075-0.194)	953.1 (680.9-1208.6)
B	0.010 (0.005-0.023)	195.2 (163.3-214.1)

Values in parentheses represent 95% confidence intervals.

Table 7. Results of the Nested Clade Phylogeographic Analysis.

Nesting category	D_c		D_n		Chain of Inference/
	$P \le$	$P \ge$	$P \le$	$P \ge$	Inferred Pattern
Clade 2-3					1-2-3-4 NO/
Subclade: 1-5	1.000	1.000	1.000	0.079	Restricted gene flow with isolation by distance (IBD)
1-6	1.000	1.000	0.241	0.839	
1-7	0.485	0.517	0.691	0.311	
1-8	0.034 [a]	0.979	0.334	0.678	
1-9	1.000	1.000	0.170	1.000	
I-T	0.882	0.119	0.664	0.336	
Clade 3-2					1-19-20-2-11-12 NO/
Subclade: 2-4	0.032 [a]	0.970	0.015 [a]	0.987	Contiguous range expansion (CRE)
2-5	0.928	0.073	0.980	0.021 [a]	
I-T	0.032 [a]	0.970	0.018 [a]	0.984	
Clade 3-3					1-2-3-5-6*-7-8 YES/
Subclade: 2-6	0.282	0.727	0.257	0.751	Restricted gene flow/dispersal but with some long-distance dispersal (LDD) over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations.
2-7	0.568	0.433	0.788	0.213	Too few clades: Insufficient genetic resolution to discriminate between range expansion/colonization and restricted dispersal/gene flow.
2-8	0.114	1.000	0.078	0.965	
2-9	0.012 [a]	1.000	0.956	0.049 [a]	
I-T	0.855	0.115	0.234	0.766	
Entire cladogram					1-2-11-12 NO/
Subclade: 3-1	0.130	0.870	0.099	0.901	Contiguous range expansion (CRE)
3-2	0.753	0.247	0.671	0.329	
3-3	0.924	0.076	0.937	0.063 [b]	
I-T	0.072 [b]	0.928	0.062 [b]	0.938	

[a] Values of $P \le 0.05$

[b] Values of P close to 0.05 level, considered as significant when running the inference key.

Nesting categories are as in Fig. 4.

Manuscript No. 2:

Štefka, J., Gilleard, J.S., Grillo, V., Hypša, V. 2007. Isolation and characterization of microsatellite loci in the tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea). Mol. Ecol. Notes, 7: 794-796

Molecular Ecology Notes (2007) 7, 794–796 doi: 10.1111/j.1471-8286.2007.01704.x

PRIMER NOTE

Isolation and characterization of microsatellite loci in the tapeworm Ligula intestinalis (Cestoda: Pseudophyllidea)

J. STEFKA,* J. S. GILLEARD,† V. GRILLO† and V. HYPSA*
*Faculty of Biological Sciences, University of South Bohemia and Biology Centre ASCR, Institute of Parasitology, Brani&ovská 31, 37005 0eské Budejovice, Czech Republic, †Division of Infection and Immunity, Institute of Comparative Medicine, Faculty of Veterinary Medicine, Bearsden Road, University of Glasgow, G61 1QH, Scotland, UK

Abstract

Ligula intestinalis is a species with a complex biology. Several strains, distinct in the host specificity and biogeography, are supposed to exist; however, their morphological or molecular identification remains unresolved. We describe the isolation of 16 polymorphic microsatellite markers. The number of alleles ranged from four to 14 in a population of 21 L. intestinalis individuals. Loci showed amplification success across samples of worldwide distribution. A cross-amplification of 14 markers in Digramma interrupta (Cestoda: Pseudophyllidea) was observed.

Keywords: Cestoda, Ligula intestinalis, microsatellites

Received 24 November 2006; revision accepted 19 December 2006

Manuscript on pages 54 to 56 is a subject to copyright

Manuscript No. 3:
Microsatellite analysis of genetic structure of the endoparasite *Ligula intestinalis* in relation to biogeography and host specificity (to be submitted to Molecular Ecology)

Microsatellite analysis of genetic structure of the endoparasite *Ligula intestinalis* in relation to biogeography and host specificity

Introduction

Relation between phenotypic diversity and genetic structure is one of the fundamental evolutionary and ecological issues that became fully accessible only after introduction of modern molecular techniques. Number of studies demonstrated that a seemingly broad polymorphism of biological traits within some populations may often manifest a hidden genetic structure. Adaptation to different hydrographic conditions in foraminiferans (de Vargas, 1999), differencies in ontogenetic migration in fish (Hyde et al., 2008), shell morphology in limpets (Nakano and Spencer, 2007) or electrophysiological properties of electric fishes (Feulner et al., 2006) may serve as examples illustrating the broad diversity of biological characteristics explored in this context. Obviously, an important prerequisite of such investigations is an *a priori* determination and reliable delimitation of the traits to be tested against the genetic structure or mapped on the population genealogy. In parasitic organisms, one of the most frequently discussed properties in this respect is the affinity to different host taxa, called host-specificity. An emergence of host-mediated barriers to gene flow is even considered one of the major speciation modes (Brooks and McLennan, 1993). At the gross phylogenetic level, the genetic diversification of parasites due to co-speciation processess is an often studied and methodologically well explored problem (Clayton and Moore, 1997; Page and Holmes, 1998, Legendre et al., 2002). In contrast, very little is known about host-mediated genetic structure within populations. It is clear that at this level temporary fragmentations and reunions of various subpopulations may affect both genetic structure of the population and distribution of various biological properties. However, the scarcity of data on various ecological types and taxonomic groups of parasites do not allow for any consistant picture. As a result, the patterns observed in particular studies may be entirely unexpected. For example, the lack host-dependent genetic structure found in three species of chewing lice parasitizing the birds of genus *Calonestris* (Gomez-Diaz et al. 2007) is quite surprising and contradicts many other studies on these ectoparasites. Moreover, recent studies performed at the genealogical level demonstrate that the association between genetic diversification on one hand and various biological traits on the other hand can dramatically vary even among closely related lineages of the parasite (Reed et al., 2004 ; Štefka and Hypša 2008). For endoparasites with complex life cycle, the best illustration is provided by the study of Steinauer et al. (2007). Motivated by apparent phenotypic variability of the acanthocephalan species *Leptorhynchoides thecatus*, the authors investigated its population structure throughout North America and showed that it forms cluster of several genetic lineages (Steinauer et al. 2007). Whereas most of these lineages were represented by specialists with narrow host and microhabitat specificity, one lineage possessed wide specificity overlapping the host spectrum of all specialists.

In this study, we explore genetic structure of a different type of endoparasite with complex life cycle. The tapeworm *Ligula intestinalis* (Linnaeus, 1758) (Cestoda: Diphyllobothriidea) is broadly distributed but morphologically very uniform species (Dubinina, 1980). During its development, *L. intestinalis* utilizes three types of hosts (Fig. 1). Within the primary host, crustacean copepode, and the definitive host, fish eating bird, the infection is relatively short. In contrast, the secondary host, i.e. fish from various taxonomic groups, plays substantial role in the whole life cycle and harbours the *Ligula* plerocercoids for at least several months (Dubinina, 1980). Two biological features particularly predispose *L. intestinalis* to become a useful model for studying speciation, geographic variability and significance of host-specificity. First, *L. intestinalis* is a species with wide geographic distribution and has been

reported from numerous localities (e.g., Weekes and Penlington, 1986; Dejen et al., 2006, Innal, 2007). Second, during its secondary larval stage, it infects high diversity of fish hosts. Apart from various species of cyprinids, which are the most common hosts, the plerocercoids were recovered from three other fish families (Shields et al., 2002; Morgan, 2003; Innal, 2007).

Recently, this parasite has become an object of several populationally oriented studies (Olson, et al., 2002; Logan et al., 2004; Bouzid et al., 2008a, b) that found high degree of hidden genetic variability, indicating possible occurrence of cryptic lineages. Based on various types of molecular markers, these studies determined deep level of structuring associated with large geographic distances. However, they failed to find any substantial differences among the populations within the European area sampled from several host species across wide geographical range. This was unexpected, since Loot et al. (2001a) detected a host-specificity by performing experimental infections of different hosts (*Rutilus* and *Blicca*) with French isolates of *Ligula*. It is generally known that various types of molecular markers are prone to different types of artifacts. Out of the markers used in the previous studies, the mtDNA often fails to detect recent diversifications due to incomplete lineage sorting (e.g., Anderson, 2001; McGuire et al., 2007), while the ITS2 sequences suffer from presence of many paralogs within a single genome (Buckler et al., 1997). To obtain more complex and consistant picture of diversification processes within *L. intestinalis*, we utilize data from 15 microsatellite loci for exploring broad geographic sample from variety of host species.

Methods

Sample and markers

We used plerocercoid sample from several previous studies on the *Ligula* ecology and evolution (Loot et al., 2001a, 2002; Bouzid et al., 2008b) and thus considerably extended the number of specimens geneticaly studied to date. All plerocercoid samples were either preserved in 96% ethanol or stored at -20 ºC prior the DNA isolation. Adults worms, stored in denatured ethanol, were excluded from the microsatellite analysis since they failed to amplify at multiple loci. Altogether, we analyzed 338 plerocercoids from 24 localities with worlwide distribution (Fig. 2, Table 1). To analyze population structure, we used a set of 15 out of 16 microsatellite loci characterized by Štefka et al. (2007). Locus Li.15 was excluded due to elevated presence of null alleles (misamplification) in many samples. Isolation of DNA and PCR conditions were adopted from Štefka et al. (2007). PCR products were run on 1.5% agarose gels and analysed on ABI PRISM 3170 Automatic Sequencer (PerkinElmer). Null homozygotes (i.e. loci that failed to amplify) were checked by additional PCR. The programs GENEMAPPER 3.7 (Applied Biosystems) and MICRO-CHECKER 2.2.3 (Van Oosterhout et al., 2004) were used to score individual alleles and to search for typing errors, respectively. One sample which returned an unexpected microsatellite profile was further sequenced for cytochrome oxidase subunit I (COI) using PCR conditions and primers from Bouzid et al. (2008b). This sequence was deposited in Genbank under Acc. number EU 636655.

Data analysis

Since hidden population structure within the samples collected from different hosts at single locality could not be a priori excluded, the analyses with a priori defined populations were run in two different modifications. In the first set of data, the individuals were pooled according to the hosts and localities, whereas input of the second batch was less structured, treating all individuals from a given locality as single population. The designations of the data sets are explained in Table 1.

Hardy-Weinberg Equilibrium (HWE), Linkage Disequilibrium (LD), numbers of private alleles and their frequences in populations (or subclades) were calculated using GENEPOP 3.4 (Raymond and Rousset, 1995). This program was also used to characterize the basic level of interpopulation or interclade variability by test of genic differentiation (exact G test) and values of pairwise F_{ST} coefficients. The tests were run using 1 000 Markov chain batches, with 10 000 iterations per batch. Obtained p-values were adjusted by sequential Bonferroni correction. Fixation indexes (F_{IS}) were calculated using FSTAT (Goudet, 1995) and p-values were estimated by 10 200 randomizations. Values of per population self-fertilization (selfing) rates, s(F_{IS}), were counted following the equatation: $F_{IS} = s/(2\text{-}s)$. Since the F_{IS} statistics is built on among locus calculations of heterozygote deficiency, the unrecognised null alleles tend to produce overestimates of the level of selfing. It has been shown by David et al. (2007) that calculation based on multilocus comparing of linkage disequilibria is less vulnerable to bias caused by occurrence of null alleles than the traditional F_{IS} based estimates. Therefore, we calculated alternative estimates of *s* using the program RMES (David et al., 2007) which implements the latter procedure.

MICROSATELLITE ANALYSER (MSA) (Dieringer and Schlötterer, 2003) was used to estimate allelic richness of populations and to construct distance matrices based on proportion of shared alleles (D_A) and Nei's chord distances (D_N) (Nei, 1983). Two types of matrices were prepared, one containing the distances among all 338 individuals and the other with distances among the populations. Bootstrap support was calculated by 1 000 replicates. Neighbour-Joining trees were constructed in PHYLIP 3.06 package (Felsenstein, 2007).

Analysis of Molecular Variance (AMOVA) was performed in software package ARLEQUIN 3.11 (Excoffier, 2007). The calculations based on F_{ST} (Weir and Cockerham, 1984) and R_{ST} (Slatkin, 1995) coefficients were run using 10 000 Markov chain replications. Several sets of data were explored. First, the level of structuring of populations at the the global scale was explored by arranging the populations into 5 groups (clade A, clade B, DIGR, ETH, CAN) according to the clustering obtained by mtDNA analyses of Bouzid et al. (2008b) (definition of the groups is provided in Table 1). Second, finer structure was modelled, where clade B populations were split according to their origin resulting in 8 different groups (cladeA, Ie Gg+AL, CN, AU, DIGR, ETH, CAN). Finally, possible structuring inside the clade A was explored by grouping the populations in three different sets: (1) according to the geographical origin ([FR1+FR2+FR3+FR4] [DE+CZ1+CZ2+CZ3] [IE+GB] [EE+RU] [UA] [TU]); (2) according the host spectra, where plerocercoids infecting the same host species were grouped together. The third set was prepared as a compromise between the former two combinations, ([FR1+FR2+FR3+FR4+D+CZ1+CZ2] [CZ3+EE+RU] [IE+GB] [UA] [TU]).

To obtain population structure independent on *a priori* assignement of the individuals to particular populations, we used Bayesian clustering algorithm implemented in the program STRUCTURE 2.2 (Falush et al., 2003, 2007). Since the algorithm is able to accomodate null alleles and ensures that they do not affect the clustering procedure, the null alleles were allowed for all loci in the analysis. Computations were run under the admixture ancestry model and the assumption of correlated allele frequencies as suggested by Falush et al. (2003). Because of large number of populations analysed in the study, the following strategy of the search was used: at first dataset containing all samples was used for modelling K values from 1 to 35 in short computation cycles (Burning length 2 000/5 000 Markov Chain replications). Because the values of log likelihood did not rise when K reached 15 and the variance of log likelihood increased steeply beyond this point, the upper bound for K was lowered to 15 in the subsequent analysis. This more thorough analysis with length of the cycles increased to 50 000/100 000 was performed in 20 replications. After obtaining the first clustering, the analysis was continued in hirerarchical manner where each cluster was used as an input set for subsequent analysis. The process was terminated when the program failed to divide the inputs into separate clusters and began assigning individual genotypes to several different clusters

(i.e. “horizontal” instead of “vertical” separations were observed in the plot). Based on the NJ topologies, we paid particular attention to genetic assignment of Ethiopian samples. To evaluate thorougly possible subdivision within the ET population collected from 4 fish species, we run the assignment analysis separately for the ET samples using three models to fit the data. Model 1 was the same as specified above. In model 2 a non-admixed ancestry of populations was fitted, following the suggestion of the STRUCTURE authors (Falush et al., 2003), that it may provide better resolution at detecting subtle structure. Model 3 was the most relaxed model; the non-admixed ancestry was fitted under the assumption of independent allele frequences with lambda parameter inferred from the data. K from 1 to 6 was modelled in the three tests using the same number of runs, burning length and MCMC replications as above. For finding the optimal number of clusters, we used the ΔK statistics of Evanno et al. (2005). Graphical visualization of assignment analysis was prepared in DISTRUCT (Rosenberg, 2004).

BOTTLENECK 1.2.02 (Cornuet and Luickart, 1996) was used to test populations for size contractions. Wilcoxon’s sign test with 10 000 permutations at each population, using strict stepwise mutation model (SMM) and two-phase model (TPM), was employed. The TPM was run modelling 90% of SMM mutations and 10% infinite allele model (IAM) mutations.

Results

Genetic character of populations

Almost all populations exhibited HW disequilibrium at least in one locus (Table 1), but the occurrence of HW disequilibiria was not distributed randomly across the studied loci. When exluding the populations with small numbers of samples, where fewer HW disequlibria were found due to weak significance of the tests, it can be seen that HW disequilibria are centered around 5 loci (Li. 7, 8, 18, 21, and 30; see Table 2). Correspondingly, the frequencies of null alleles estimated by comparing the expected values of homo- and heterozygosity (Brookfield, 1996) were generally low with the exception of some of the loci which were out of HWE. On the other hand, despite the relatively frequent occurrence of HW disequilibria, the data were not distorted in any substantial way, only 6 cases of null homozygotes were found among 243 samples from European populations. Furthermore, after sequential Bonferroni correction, significant values of LD were found only in two populations, one pair of loci was in LD for *TN RbSe*, 12 pair-locus combinations were found in LD for CZ3Ab. The pattern of HW disequilibria across loci and scarcity of LD shows that most localities represent populations without inner substructuring.

Estimates of the inbreeding coefficient (F_{IS}) were moderate in most populations, significant values were observed in several populations (Table 1). In two cases the s(F_{IS}) values exceeded 0.4 (populations AL BS and CZ3Ab), suggesting that strong inbreeding or frequent selfing is common at least in some *Ligula* populations. In contrast, the RMES method found significant values of *s* only in 4 populations (CZ3Ab, FR1Aa, FR2Rr and FR4Rr) and the estimated frequencies of selfing were low, reaching 0.16 at maximum (FR1Aa).

Results of the Wilcoxon’s sign test under the SMM model indicated bottlenecks in populations CA Sa, *CASaCp*, *CZ2AaRr* and Fr4Rr ($P \leq 0.001$); and in CZ3Ab, Fr1Rr and TN Rb ($P \geq 0.05$). Under the more relaxed hypothesis of the TPM model only pooled data for *CZ2AaRr* were significant ($P=0.035$).

Genetic variability and structure

The 15 analysed loci exhibited different levels of polymorphism over the 24 populations. In non-European populations some of the loci failed to amplify (Table 1) and some were

monomorphic (Table 2). No monomorphic loci were found within the European populations. Consistently with this finding, the allelic richness (vaules adjusted for n=9) for 10 loci was highest in the European samples (Table 2). This is most likely due to the fact that development of the microsatelite markers was originally based on plerocercoids from two European localities (Štefka et al., 2007). Despite this, 5 loci showed highest allelic richness in 2 of the non-European populations (Table 2). Such distribution of variability among the loci and populations proved applicability of the markers to study all sampled populations.

NJ trees show that both individuals and populations cluster mainly in accordance with their geographical origin (Figs. 3, 4). Particularly, the geographically isolated samples from Canada and Australia, but also the populations from Ethiopia and the *Digramma* samples, produced long branches with high bootstrap supports. Values of pairwise F_{ST} (Appendices 1, 2, 3) and results of analysis of molecular variance (Table 3) have shown that the structuring between different geographical areas is very deep. The major clade comprised samples from derived cyprinids collected across a broad European area and in North Africa (Tunisia). It corresponded to the mitochondrial clade A according to Bouzid et al. (2008b) and will be designated as clade A hereafter. On the other hand, the samples isolated from basal cyprinids in Europe (Ireland sample) and North Africa (Algerian sample) clustered outside the Europe/Tunisian clade. Unexpected position was observed for two samples from Great Britain that clustered at positions distant from all other European and North African samples (designated by asterisks in the Fig. 4). A conspicuous element of the NJ topologies was separation of Ethiopian population from Lake Tana into two groups (Figs. 3, 4) supported by high bootstrap values. The separation was also reflected in high values of pairwise F_{ST}, differences in HW disequilibria and considerable numbers of private alleles specific for the two groups (Appendix 4). This separation was highly host specific in respect to two host species, *Barbus humilis* and *B. tsanensis*.

Clustering of individual samples was further specified by Bayesian assignment algorithm in the program STRUCTURE. When performed on complete data set, the analysis found a largest indent of ΔK for division of samples into two clusters (Fig. 5). First cluster corresponded to the clade A (i.e. the European and Tunisian samples from derived cyprinids). The second cluster encompassed all remaining populations without any further differentiation (Fig. 6A). However, the sequential hierarchical procedure (as described in the MM section) led to much more complex structure with several well defined clusters (Fig. 5). The complete assignment pattern of these subclusters is provided in Fig. 6 B; here we only summarize the most important associations. Unlike the distance-based clustering, the Bayesian assignment indicated a split between European and Tunisian samples within the clade A. This separation was not strict; the assignment of some FR3Rr and FR4Rr individuals was equivocal and some were even assigned to the Tunisian cluster. Further difference against the NJ topologies are assignments of the Great Britain sample from *Rutilus* (GB Rr) and Ireland sample from *Gobio* (IE Gg). The sample GB Rr was clearly assigned to the clade A, which corresponds to taxonomic position of its host. The sample IE Gg clustered together with ALG BS, but considerable probability was also put on its assignment to the clade A. Surprisingly, the assignment of specimen FR4Rr_076 was similar to that of IE Gg. This was unexpected considering the host (*Rutilus*) of the former specimen. Position of this specimen was therefore further tested by sequencing and analysing COI gene. The result revealed only one bp difference between the two samples and thus confirmed their high similarity. For the Ethiopian population, the models 1 and 2 did not yield any further structuring. However, clear host-dependent structure, consistent with the NJ topology, was found under the model no. 3 (Fig. 6 C).

Discussion

Genetic variability and population structure

The analysis based on 15 microsatellite loci revealed several distinct components of genetic structure within the worldwide population of *Ligula intestinalis*. The most conspicuous is the high level of polymorphism and strong genetic structure detected among populations inhabiting different geographical areas. This fragmented pattern, conditioned by geographical distribution at long distances, is generally compatible with the results obtained from mitochondrial data (Bouzid et al., 2008b). The populations from distant localities, such as Canada, Australia or China, form entirely consistent clades even in the individual-based tree (Fig. 4), with no indication of mixing with other populations. This indicates that the *L. intestinalis s lat.* is likely to be composed of several/many highly isolated lineages or even cryptic species. Good illustration of such strong inner differentiation is demonstrated by recent reclassification of the genus *Digramma* (Kuchta et al., 2008). This tapeworm, originally distinguished by morphological differencies, forms one subclade of the *L. intestinalis* complex.

Differences between the mtDNA and microsatellite data were, however, found at the shorter genetic distances. The most noticeable is the separation of European and Tunisian populations from derived cyprinids. In mtDNA trees the Tunisian samples were mixed with various European populations and shared two of the three detected haplotypes. Using the Bayesian approach, distinct population subdivision was found between Europe and the population from Tunisia. This contrast is a typical illustration of inadequacy of mitochondrial data to disentangle recent events due to incomplete lineage sorting. Population of *Rutilus rubilio* is not native in northern Africa but was introduced into Tunisia during late 1960' from Italy (Losse et al., 1991). The Tunisian samples, collected from this host, are therefore recent offshot(s) of the European population and they could not be differentiated by the mitochondrial haplotypes. The same lack of divergence across the Mediterranean Sea was found by Bouzid et al. (2008a) using the ISSR analysis. Even within our microsatellite data, the individual-based NJ tree did not cluster the Tunisian samples into a single monophyletic clade. This demonstrates the shallow genetic signal joining the Tunisian samples and surprisingly good performance of STRUCTURE program when assigning samples to individual populations; with only one exception all Tunisian individuals were assigned back to their original population (Fig. 6). In correspondence with this view (i.e., recent import of the Tunisian population) the genetic data indicated that Tunisian population has been a subject to genetic drift or size bottleneck; 8 loci were found out of HWE accompanied by LD between two loci and positive test for bottleneck under the SMM model.

The recent formation of genetic barrier between the North African and European populations can be most plausibly explained as an effect of the Mediterranean Sea. The dispersal capabilities of most parasites are strongly limited and it was hypothesized that gene flow in parasites with complex life cycles is governed by the most mobile host (Prugnole et al., 2005). Birds represent such host in the life-cyle of *Ligula*, and even though the Mediterranean Sea is crossed by birds during migrations, it forms a barrier sufficiently strong to prevent populations of *Ligula* from substantial mixing. The duration of the *Ligula* infections in the definitive host is strongly limited. The adults mature and produce eggs within 48 hours after infection and the infection is terminated within a week (Dubinina, 1980). An effect of the Mediterranean Sea as a strong genetic barrier even for organisms with advanced dispersal abilities is well known. It was for example shown that populations of a bat *Myotis myotis* are stronly isolated at both sides of the Gibraltar Strait, whereas no population structure was observed either within the Iberian peninnsula or within Morocco (Castella et al., 2000).

Despite this apparently reduced gene flow, the isolation is probably not absolute and occasional crossings of *Ligula* between the two areas are possible. This is indicated by unequivoal assignements of several European samples.

Similar, but less transparent situation was identified in the Algerian population. Within the haplotype tree, the Algerian sample was polyphyletic with individual haplotypes clustering in two different and well supported lineages. This indicates that the current Algerian fauna associated with native hosts of the genus *Barbus* is a mixture of several (at least two) independent sources. The microsatellite data yielded clearly formed Algerian population, either as monophyletic clade in the NJ analysis or as highly homogeneous cluster within the Bayesian analysis (Figs. 4, 6).

Host-mediated structure

For two populations, European and Ethiopian, the host spectrum was sufficiently broad to allow for assessing the relation between host-distribution and genetic structure. In both of these samples the host records were at leasts partially associated with the observed genetic structure. Majority of the large European population was formed by samples collected from basal cyprinids. This population was highly homogeneous with no apparent association between genetics and intermediate host (Fig. 6). Distribution of individual intermediate hosts on the NJ tree (Fig. 4) shows that even phylogenetically distant host species, such as *Rutilus* and *Alburnus* or *Phoxinus*, can be invaded by close tapeworm genotypes. This result was further supported by the AMOVA analysis. While according to the Fst estimator combination between geographical and host-dependent grouping explained slightly more variability than each of the two factors separately, the Rst statistics attributed equal power to geographical and combined explanation.

In contrast to this situation, a clear distinction was found between the European samples from basal and derived cyprinids (Figs. 4, 6) and this specificity is also reflected in exchanges between the European and North African fauna. According to the haplotype-derived genealogy, the European/North African region is inhabitted by two different lineages of *Ligula*. One lineage is restricted to derived taxa of cyprinid hosts whereas the other was collected from the basal taxa. Within our data, the former lineage is represented by the major group of European samples. As discussed above, this lineage has recently been introduced to North Africa (Tunisian samples) where it rapidly constituted a genetically separated population retaining the specificity to introduced derived cyprinids. Situation of the latter lineage is more complicated. In Africa, it was collected from the hosts representing the native ichthyofauna, namely the genus *Barbus*, and the few European samples were isolated from another basal cyprinid, the genus *Gobio*. This host-based delimitation was almost absolute; only one (UA Cc) of the 243 samples of the first lineage (clade A) was found in basal cyprinids and vice versa only one sample (FR4 Rr_076) of the second lineage was found in a derived cyprinid. Interestingly, substantial differences between *Gobio* and *Rutilus* immune responses were observed during establishment of *Ligula* in the fish body cavity (Arme, 1997). This indicates that immunity may be at least one of the factors determining host-specificity of the *Ligula* lineages with different genotypes.

Considering evolutionary and genetic complexity of the whole *Ligula* system, the obtained patterns and topologies should be interpreted with maximal caution. However, several facts indicate that these samples represent a unique lineage adopted to basal cyprinids, which underwent several exchanges between North Africa and Europe. The most conspicuous is the position of Algerian/European samples outside the European/Tunisian clade and their apparent relatedness within Bayesian clusters. With current sample, it is impossible to estimate the number and directions of the exchanges between North Africa and Europe.

Even more pronounced dependence between the host taxonomy and tapeworm genetics was identified in the Ethiopian population collected from Lake Tana. Out of the three sufficiently sampled hosts, the species *B. humilis* and *B. tsanensis* yielded two genetically separated *Ligula* populations and the tapeworms from *B. intermedius* were scattered in both populations (Fig. 4).

Although this result suggests a strong host-dependent structure, it cannot be more thoroughly interpreted with the current data. The genus *Barbus* possesses a complicated taxonomy and evolutionary history. Till now, more than 800 nominal species have been described and found to form several lineages with different ploidity and geographical distribution (e.g. Machordom and Doadrio, 2001; Tsigenopoulos et al., 2002).

Members of two lineages are present in Lake Tana. *B. humilis* is a representative of diploid lineage, common elsewhere in Ethiopia and South Africa (Machordom and Doadrio, 2001; Tsigenopoulos et al., 2002; de Graaf et al., 2007). The 3 other species of barbels belong to a flock of cryptic species, originally described as *Barbus intermedius*. They are endemic to Lake Tana and belong to a clade of hexaploid barbels which is distributed throughout most of the African continent (Machordom and Doadrio, 2001). *B. intermedius* flock species are larger in size than *B. humilis* and posses piscivorous feeding strategy. Particular species of the flock are supposed to have evolved through adaptive radiation after riverine *B. intermedius* invaded Lake Tana (Sibbing et al; 1998). Although estimates of the age of the two barbel lineages are not available, the depth of branching in their phylogeny and the width of their distribution suggests that they are of an ancient origin (Machordom et al., 2001). The split between the two *Ligula* lineages thus may be an ancient event which resulted in establishment of two cryptic species living in sympatry.

Origin of the intrapopulation relationships: selfing vs. outcrossing

Deviations from HWE were detected in most populations. This fact becomes particularly important within the best sampled clade A from derived cyprinids. The complete lack of host-dependent structure within this population is rather surprising. Loot et al. (2001a, b) demonstrated that populations of *Ligula* from France, which were most frequently harboured by roach, were unable to infect breams, although both fish species occurred in sympatry. Although the genealogical relationships did not reveal any conspicuous host-dependence, the deviation from HWE arises question whether the mating process might be affected by some degree of host specificity. For example, indications of host-dependent substructure, although much more subtle than that found in Lake Tana, were detected among the sympatric samples from *Rutilus* and *Alburnus* at localities FR1 and CZ2. Possible structure was implied mainly by the differences in distribution of HW disequilibria between the host-specific data and the pooled populations (Table 2). Even more suggestive was finding that despite the low value of F_{ST} between the two FR1 samples (Appendix 1), the test of genic differenciation (G test) provided significant result. Similar trend was seen in CZ2, but in this case the values of F_{ST} and G test were weaker (P=0.0023) and thus insignificant after the sequential Bonferroni correction. These low levels of differentiation do not seem to reflect an occurrence of entirely separated host-specific strains. Rather, they may indicate host-dependent fluctuations in the gene pool within few generations of parasite and thus pose an interesting question on the mechanisms underlying such host-mediated separation. Differences in hosts feeding-strategies, suggested to cause differences of parasites distribution in other host-parasite systems (Jousson et al., 2000; Criscione et al, 2002; Hypša et al., 2005) cannot be assumed here, because the diet of the two species does not differ (Vašek et al., 2003 and 2004). Perhaps, immunological aspect mentioned above in connection to *Rutilus* and *Gobio* (Arme, 1997) could be also involved in this more subtle structuring. Under such view, the slight differences in the distribution of alleles detected by the G test within FR1 and CZ2 subpopulations could thus be a result of differential survival of plerocercoids originated from one generation of eggs.

However, it is important to realize that at least two additional processes might be responsible for HW deviations, hermaphroditic selfing and occurrence of null alleles (Hartl and Clark, 1997). Recognition of these factors is not trivial. Since the estimates of null alleles frequency (Table 2) are based on the calculation of homo- and heterozygosity, they are not independent

of the non-random mating (i.e. selfing) explanation. The RMES etimates of selfing may thus be a more reliable approach.

The data on the *Ligula* occurrence in its definitive hosts shows that multiple infections are common. The intensity of infestation reaches very high numbers at least in some of the bird hosts (Dubinina, 1980). Also, multiple infestations of the intermediate hosts play important role (Dubinina, 1980), since they provide the definitive host with a load of parasites in a single batch. Thus, *Ligula* is probably not forced to self frequently, due to a lack of sexual counterpart inside the bird host. This view is supported by insignificant or very low levels of selfing inferred by the RMES, usually not exceeding 10% (Table 1). Furthermore, under the clonal scenario of evolution, *Ligula* would have to reproduce almost entirely by selfing for several generations, which would have resulted in an extremely fragmented population structure and strong LD between loci across populations. The pattern of genetic variability found in our data does not conform to such view. ANOVA results revealed flat population structure within the European lineage across a large geographical area (Table 3) and low values of pairwise F_{ST} (Appendix 1). The pattern of populational structure together with scarcity of the cases, where selfing or LD were detected, show that outcrossing dominates the mating system of *Ligula*. Even in populations where selfing was found its levels were low, and thus insufficient to pose any significant force against gene flow inside the populations. Furthermore, the contrast between F_{IS} and the multi-locus based statistics, and the fact that only some loci are affected by HW disequilibria, suggests that null alleles are responsible for most of the observed HW disequilibria. On the other hand, the finding of only a few null homozygotes indicates that the frequency of null alleles in the populations is generally low.

The identified patterns of geography- and host-dependent genetic structure pose interesting questions to the future research. Considering the strong host-mediated constraints in some parts of the structure and lack of pronounced host-specificity in other parts, it would be particularly interesting to see whether these patterns are merely unique “accidental” outcomes of genealogical processess or whether they are determined by some biological (e.g., physiological, immunological or behavioral) traits of the host species.

Acknowledgement

The following persons kindly provided samples of *Ligula* tapeworms: W. Bouzid and G. Loot, France; C. Arme, United Kingdom. Samples of *Ligula* tapeworms from Lake Tana could not be obtained without help of A. de Chambrier, Switzerland; Abebe Getahun, Seyoum Mengistou, Eshete Dejen Dresilign, Moges Beleteu, and the staff of the ANRS Agricultural Research Institute, Bahir Dar, Ethiopia. The study was supported by grants LC06073 and MSM 60076605801 (Ministry of Education, Czech Republic), Grant Agency of the Czech Republic (projects Nos. 206/08/1019 and 524/04/0342) and Institute of Parasitology (research projects Z60220518 and LC 522).

References

Anderson, T.J.C. 2001. The dangers of using single locus markers in parasite epidemiology: Ascaris as a case study. Trends Parasitol., 17: 183-188

Aarme, C. 1997. Ligulosis in two cyprinid hosts: *Rutilus rutilus* and *Gobio gobio*. Helmithologia, 34:191-196

Bouzid, W., Lek., S., Mace, M., Ben Hassine, O.K., Etiene, R., Legal, L., Loot, G. 2008a Genetic diversity of *Ligula intestinalis* (Cestoda: Pseudophyllidea) based on analysis of inter-simple sequence repeat (ISSR) markers. J. Zool. Syst. Evol. Res., in press

Bouzid, W., Štefka, J., Hypša V., Lek, S., Scholz, T., Legal, L., Ben Hassine, O.K., Loot G. 2008b. Geography and host specificity: two forces behind the genetic structure of the

freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). Int. J. Parasitol., in press

Brookfield, J.F.Y. 1996. A simple new method for estimating null allele frequency from heterozygote deficiency. Mol. Ecol., 5: 453-455

Brooks, D.R., McLennan, D.A. 1993. Parascript, parasites and the language of evolution. Smithsonian Press, Washington DC

Buckler, E.S., Ippolito, A., Holtsford, T.P. 1997. The Evolution of Ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics, 145: 821-832

Castella, V., Ruedi, M., Excoffier, L. (2000) Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera : Vespertilionidae)? Mol. Ecol., 9: 1761-1772

Clayton, D.H., Moore, J. 1997. Host-Parasite evolution - general principles & avian models. Oxford University Press, 473 p.

Cornuet, J.M., Luikart, G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics, 144: 2001-2014

Criscione, C.D., Watral, V., Whipps, C.M., Blouin, M.S., Jones, S.R.M., Kent, M.L. 2002. Ribosomal DNA sequences indicate isolated populations of *Ichthyophonus hoferi* in geographic sympatry in the north-eastern Pacific Ocean. J. Fish Dis., 25: 575-582

David, P., Pujol, B., Viard, F., Castella, V., Goudet, J. 2007. Reliable selfing rate estimates from imperfect population genetic data. Mol. Ecol., 16: 2474-2487

de Graaf, M., Megens, H.J.W.C., Samallo, J., Sibbing, F.A. 2007. Evolutionary origin of Lake Tana's (Ethiopia) small Barbus species: indications of rapid ecological divergence and speciation. Animal Biol., 57: 39-48

Dejen, E., Vijverberg, J., Sibbing, F.A. 2006. Spatial and temporal variation of cestode infection and its effects on two small barbs (*Barbus humilis* and *Barbus tanapelagius*) in Lake Tana, Ethiopia. Hydrobiologia, 556: 109-117

de Vargas, C., Norris, R., Zaninetti, L., Gibb, S.W., Pawlowski, J. 1999. Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces. PNAS, 96: 2864-2868

Dieringer, D., Schlötterer, C. 2003. Microsatellite Analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes, 3: 167-169

Dubinina, M.N. 1980. Tapeworms (Cestoda, Ligulidae) of the Fauna of the USSR. Amerind Publishing, New Delhi, India

Evanno, G., Regnaut, S., Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol., 14: 2611-2620

Excoffier, L., Laval, G., Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online, 1:47-50

Fallush, D., Stephens, M., Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 164: 1567-1587

Falush, D., Stephens, M., Pritchard, J.K. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes, 7: 574-578

Felsenstein, J. 2007. PHYLIP (Phylogeny Inference Package) version 3.67. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle. `http://evolution.gs.washington.edu/phylip.html`

Feulner, P.G.D., Kirschbaum, F., Schugardt, C., Ketmaier, V., Tiedemann, R. 2006. Electrophysiological and molecular genetic evidence for sympatrically occuring cryptic species in African weakly electric fishes (Teleostei : Mormyridae : *Campylomormyrus*). Mol. Phyl. Evol., 39: 198-208

Gomez-Diaz, E., Gonzalez-Solis, J., Peinado, M.A., Page, R.D.M. 2007. Lack of host-dependent genetic structure in ectoparasites of Calonectris shearwaters. Mol. Ecol. 16: 5204-5215

Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. J. Heredity, 86: 485-486

Hartl, D.L., Clark, A.G. 2007. Principles of population genetics. Sunderland: Sinauer Associates; 4th edition, 652 p.

Hyde, J.R., Kimbrell, C.A., Budrick, J.E., Lynn, E.A.,Vetter, R.D. 2008. Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. Mol. Ecol., 17: 1122-1136

Hypša, V., Škeříková, A., Scholz, T. 2005. Phylogeny, evolution and host-parasite relationships of the order Proteocephalidea (Eucestoda) as revealed by combined analysis and secondary structure characters. Parasitology, 130: 359-371

Innal, D., Keskin, N., Erk'akan, F. 2007. Distribution of *Ligula intestinalis* (L.) in Turkey. Turk. J. Fish. Aquat. Sci., 7: 19-22

Jousson, O., Bartoli, J., Pawlowski, J. 2000. Cryptic speciation among intestinal parasites (Trematoda:Digenea) infecting sympatric host fishes (Sparidae). J. Evol. Biol., 13: 778-785

Kuchta, R., Scholz, T., Brabec, J., Bray, R. A. 2008. Suppression of the tapeworm order Pseudophyllidea (Platyhelminthes : Eucestoda) and the proposal of two new orders, Bothriocephalidea and Diphyllobothriidea. Int. J. Prasitol., 48: 49-55

Legendre, P., Desdevises, Y., Bazin, E. 2002. A statistical test for host-parasite coevolution. Syst. Biol., 51: 217-234

Logan, F.J., Horák, A., Štefka, J., Aydogdu, A., Scholz,T. 2004. Phylogeny of diphyllobothriid tapeworms (Cestoda: Pseudophyllidea) based on sequences of the ITS-2 rDNA. Parasitol. Res., 94: 10-15

Loot, G., Francisco, P., Santoul, F., G., Lek, S., Guégan, J-F. 2001a. The three hosts of the *Ligula intestinalis* (Cestoda) life cycle in Lavernose-Lacasse gravel pit, France. Arch. Hydrobil., 152: 511-525

Loot, G., Poulin, R., Lek, S., Guégan, J-F. 2002. The differential effects of *Ligula intestinalis* (L.) plerocercoids on host growth in three natural populations of roach, *Rutilus rutilus* (L.). Ecol. Freshw. Fish, 11: 168-177

Loot, G., Lek, S., Dejean, D., Guégan, J-F. 2001b. Parasite-induced mortality in three host populations of the roach *Rutilus rutilus* (L.) by the tapeworm *Ligula intestinalis* (L.). Ann. Limnol., 37: 151-159

Losse, G.F., Nau, W., Winter, M. 1991. Le développement de la pêche en eau douce dans le nord de la Tunisie Projet de la coopération technique. Coopération technique Tuniso-Allemande, CGP/GTZ

Machordom, A., Doadrio, I. 2001. Evolutionary history and speciation modes in the cyprinid genus *Barbus*. P. Roy. Soc. B-Biol. Sci., 268: 1297-1306

McGuire, J.A., Linkem, C.W., Koo, M.S., Hutchison, D.W. , Lappin, A.K., Orange, D.I., Lemos-Espinal, J., Riddle, B.R., Jaeger, J.R. 2007. Mitochondrial introgression and incomplete lineage sorting through space and time: Phylogenetics of crotaphytid lizards. Evolution, 61: 2879-2897

Morgan, D.L. 2003. Distribution and biology of *Galaxias truttaceus* (Galaxiidae) in south-western Australia, including first evidence of parasitism of fishes in Western Australia by *Ligula intestinalis* (Cestoda). Environ. Biol. Fish., 66:155-167

Nakano, T., Spencer, H.G. 2007. Simultaneous polyphenism and cryptic species in an intertidal limpet from New Zealand. Mol. Phyl. Evol., 45: 470-479

Nei, M., Tajima, F. Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. J. Mol. Evol. 19: 153-170

Olson, P.D., Littlewood, D.T.J., Griffiths, D., Kennedy, C.R., Aarme, C. 2002. Evidence for the co-existence of separate strains of *Ligula* in Lough Neagh, Northern Ireland. J. of Helmintology, 76: 171-174

Page, R.D.M., Holmes, E.C. 1998. Molecular Evolution: A Phylogenetic Approach. Blackwell Science Inc. 352 p.

Prugnolle, F., Theron, A., Pointier, J.P., Jabbour-Zahab, R., Jarne, P., Durand, P., De Meeus, T. 2005. Dispersal in a parasitic worm and its two hosts: Consequence for local adaptation. Evolution, 59: 296-303

Raymond, M., Rousset, F. 1995. An exact test for population differentiation. Evolution, 49: 1280-1283

Reed, L.D., Smith, V.S., Hammond, S. L., Rogers, A.R., Clayton, D.H. 2004. Genetic analysis of lice supports direct contact between modern and archaic humans. PLoS Biol., 2: e340

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Mol. Ecol. Notes, 4: 137-138

Shields, B.A., Groves, K.L., Rombauch, C., Bellmore, R. 2002. Ligulosis associated with mortality in largescale suckers. J. Fish Biol., 61: 448-455

Sibbing, F.A., Nagelkerke, L.A.J., Stet, R.J.M. 1998. Speciation of endemic Lake Tana barbs (Cyprinidae, Ethiopia) driven by trophic resource partitioning; a molecular and ecomorphological approach. Aquat. Ecol., 32: 217-227

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics, 139: 457-462

Štefka, J., Gilleard, J.S., Grillo, V., Hypša, V. 2007. Isolation and characterization of microsatellite loci in the tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea). Mol. Ecol. Notes, 7: 794-796

Štefka, J., Hypša, V. 2008. Host specificity and genealogy of the louse *Polyplax serrata* on field mice, *Apodemus* species: A case of parasite duplication or colonisation? Int. J. Parasitol., 38: 731-741

Steinauer, M.L., Nickol, B.B., Orti, G. 2007. Cryptic speciation and patterns of phenotypic variation of a highly variable acanthocephalan parasite. Mol. Ecol., 16: 4097-4109

Trouve, S., Renaud, F., Durand, P., Jourdane, J. 1996. Selfing and outcrossing in a parasitic hermaphrodite helminth (Trematoda, Echinostomatidae). Heredity, 77: 1-8

Tsigenopoulos, C.S., Ráb, P., Naran, D., Berrebi, P. 2002. Multiple origins of polyploidy in the phylogeny of southern African barbs (Cyprinidae) as inferred from mtDNA markers. Heredity, 88: 466-473

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P. 2004. Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes, 4: 535-538

Vašek, M., Kubečka, J., Sed'a, J. 2003. Cyprinid predation on zooplankton along the longitudinal profile of a canyon-shaped reservoir. Arch. Hydrobiol., 156:535-550

Vašek, M., Kubečka, J. 2004. In situ diel patterns of zooplankton consumption by subadult/adult roach *Rutilus rutilus*, bream *Abramis brama*, and bleak *Alburnus alburnus*. Folia Zool., 53: 203-214

Weekes, P.J., Penlington, B. 1986. First record of *Ligula intestinalis* (Cestoda) in eainbow trout , *Salmo gairgnerii*, and common bully, *Gobiomorphus cotidianus*, in New Zealand. J. Fish Biol., 28: 183-190

Weir, B.S., Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. Evolution, 38: 1358-1370

Table 1. List of samples (legend on the next page).

Geographic origin		Symbol	Host species	Species	mtDNA/ ITS2	*N* (total=338)	*L*	*HW*	F_{IS}	$s(F_{IS})$	*s(g2)*	*s(ML)*	*CI(95)*
Algeria	Oued Hamiz Reservoir	**AL BS**	*Barbus* sp.	*L.* sp.	B	10	15(0)	3	0.262*	0.415	×	×	×
Australia	Goodga River	**AU Gt**	*Galaxias truttaceus* (Osmeriformes)	*L.* sp.	B	4	13 (6)	n.d.	0.130	0.230	×	×	×
	Moates Lake	**Gm**	*Galaxias maculatus* (Osmeriformes)	*L.* sp.	B	5	13 (6)	n.d.	0.000	0.000	×	×	×
	pooled population	***AuGmGt***		*L. sp.*	*B*		*13 (6)*	*0*	*0.048**	*0.092*	×	×	×
Canada	Dalpec Lake	**CA Cp**	*Coulsius plumbeus*	*L.* sp.	CAN	1	14(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	Lac Dumbo	**Sa**	*Semotilus atromaculatus*	*L.* sp.	CAN	26	14 (2)	0	0.072	0.134	×	×	×
	pooled population	***CaSaCp***		*L. sp.*	*CAN*		*14 (2)*	*0*	*0.070*	*0.131*	×	×	×
China	Dong Tink Lake	**CN Hb**	*Hemiculter bleekeri*	*D. i.*	DIGR	2	14(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	Zhanghe Reservoir	**Nt**	*Neosalanx taihuensis* (Osmeriformes)	*L.* sp.	B	21	14 (1)	4	0.166	0.285	×	×	×
Czech Rep.	Lipno Reservoir	**CZ1Rr**	*Rutilus rutilus*	*L. i.*	A	9	15(0)	0	0.146	0.255	×	×	×
	Zelivka Reservoir	**CZ2Rr**	*Rutilus rutilus*	*L. i.*	A	18	15(0)	4	0.113	0.203	×	×	×
		Aa	*Alburnus alburnus*	*L. i.*	A	17	15(0)	2	0.132*	0.233	×	×	×
	pooled population	***Cz2AaRr***		*L. i.*	*A*		*15(0)*	*3*	*0.126**	*0.224*	×	×	×
	Nove Mlyny Reservoir	**CZ3Ab**	*Abramis brama*	*L. i.*	A	14	15(0)	**2**	0.253*	0.404	0.122	0.089	[0.009 , 0.203]
Estonia	Peipsi Lake	**EE Ab**	*Abramis brama*	*L. i.*	A	5	15(0)	0	0.101	0.183	×	×	×
Ethiopia	Tana Lake	**ET Bh**	*Barbus humilis*	*L.* sp.	ETH	7	12 (3)	n.d.	0.287	0.446	×	×	×
		Bt	*Barbus tsanensis*	*L.* sp.	ETH	7	12 (3)	n.d.	0.137	0.241	×	×	×
		Bi	*Barbus intermedius*	*L.* sp.	ETH	7	12 (3)	n.d.	0.250	0.400	×	×	×
		Bb	*Barbus brevicephalus*	*L.* sp.	ETH	1	12 (3)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	pooled population	***EtBhBtBiBb***		*L. sp.*	*ETH*		*13 (3)*	*3*	*0.303*	*0.465*	×	×	×
France	Pareloup Reservoir	**FR1Rr**	*Rutilus rutilus*	*L. i.*	A	23	15(0)	3	0.108*	0.195	×	×	×
		Aa	*Alburnus alburnus*	*L. i.*	A	5	15(0)	0	0.029	0.056	0.160	2,05E-5	[0 , 0.204]
	pooled population	***Fr1RrAa***		*L. i.*	*A*		*15(0)*	*1*	*0.110**	*0.198*	×	×	×
	Lavernose	**FR2Rr**	*Rutilus rutilus*	*L. i.*	A	26	15(0)	3	0.134*	0.236	0.094	0.039	[0 , 0.132]
		Bb	*Blicca bjoerkna*	*L. i.*	A	2	15(0)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	pooled population	***Fr2RrAb***		*L. i.*	*A*		*15(0)*	*5*	*0.144**	*0.252*	*0.101*	*0.060*	*[0 , 0.167]*
	Muret	**FR3Rr**	*Rutilus rutilus*	*L. i.*	A	11	15(0)	2	0.093	0.170	×	×	×
	Creteil Lake	**FR4Rr**	*Rutilus rutilus*	*L. i.*	A	35+1	15(0)	4	0.122*	0.217	0.055	0.047	[0.003, 0.121]
Germany	Müggelsee Lake	**DE Rr**	*Rutilus rutilus*	*L .i.*	A	6	15(0)	0	0.105	0.190	×	×	×
Great Britain	Scotland, River Gryfe	**GB Rr**	*Rutilus rutilus*	*L .i.*	A	1	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		Pp	*Phoxinus phoxinus*	*L. i.*	A	1	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	Wales, Aberystwyth	**Pp**	*Phoxinus phoxinus*	*L.* sp.	A	1	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
N. Ireland	Lough Neagh	**IE Rr**	*Rutilus rutilus*	*L.* sp.	A	3	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		Gg	*Gobio gobio*	*L.* sp.	B	3	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Russia	Khanka Lake, Far East	**RU Hl**	*Hemiculter lucidus*	*D. i.*	DIGR	1	14(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	Rybinsk Reservoir	**RU Ab**	*Abramis brama*	*L. i.*	A	9	15(0)	2	0.107	0.193	×	×	×
Tunisia	Sidi Salem Reservoir	**TN Rb**	*Rutilus rubilio*	*L. i.*	A	48	15(0)	7	0.155*	0.268	×	×	×
		Se	*Scardinius erythrophthalmus*	*L. i.*	A	5	15(0)	0	0.179	0.304	×	×	×
	pooled population	***TnRbSe***		*L. i.*	*A*		*15(0)*	***6***	*0.155**	*0.268*	×	×	×
Ukraine	Dniester River	**UA Rr**	*Rutilus rutilus*	*L. i.*	A	1	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		Aa	*Alburnus alburnus*	*L. i.*	A	1	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		Cc	*Carassius carassius*	*L. i.*	A	1	15(n.d.)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

mtDNA/ITS, lineage characterization according to Bouzid et al., (in press.): A, clade A (containing European samples isolated from phylogenetically derived cyrinids); B, clade B (containing samples from phylogenetically basal cyprinids from Europe and Alger, and samples from China and Australia; ETH, Ethiopia; CAN, Canada; DIGR, *Digramma*. *N*, nuber of samples in each population; *L*, number of loci analysed, in parentheses number of monomorphic loci; *HW*, number of loci out of Hardy-Weinberg equilibrium (see Table 2 for details), values in bold mark populations with significant Linkage Disequilibria; F_{IS}, Fixation index (* values significant at adjusted nominal level 0.0001); $s(F_{IS})$, estimate of selfing rate based on F_{IS} ; *s(g2)* and *s(ML)* estimates of selfing based on multilocus comparisons (see text for details); *CI(95)*, 95% confidence interval for *s(ML);* x, *s(g2)* results not significant; n.d., not diagnosed (when *N*<5). Symbols and results for populations pooled for all host species of a given locality are in italics.

Table 2. Summary statistics of genetic diversity by population and locus. Only data for populations of size N≥5 are given.

Population		Li.1	Li.2	Li.5	Li.6	Li.7	Li.8	Li.11	Li.14	Li.17	Li.18	Li.21	Li.23	Li.24	Li.25	Li.30	Mean
AL BS	H_O	1.00	0.50	0.20	0.70	0.80	0.00	0.40	0.40	0.10	0.30	0.60	0.40	0.60	0.80	0.60	0.49
	H_E	0.82	0.52	**0.72***	0.83	0.84	0.19	0.50	**0.77***	0.39	**0.85***	0.48	0.59	0.76	0.86	0.75	**0.66***
	A	6.69	2.00	3.90	6.79	7.68	1.99	3.89	6.69	2.00	6.97	2.99	3.89	4.80	6.79	4.89	4.80
	A_N	-0.10	0.01	0.30	0.07	0.02	0.16	0.07	0.21	0.21	0.30	-0.08	0.12	0.09	0.03	0.09	0.10
AU GmGt	H_O	0.89	n.d.	0.33	n.d.	0.44	n.d.	n.d.	n.a.	n.a.	0.11	n.d.	n.d.	0.22	0.00	0.00	0.29
	H_E	0.75	n.d.	0.29	n.d.	0.54	n.d.	n.d.	n.a.	n.a.	0.31	n.d.	n.d.	0.21	n.d.	n.d.	0.42
	A	4.00	2.00	2.00	1.00	4.00	1.00	2.00	n.a.	n.a.	3.00	1.00	1.00	2.00	1.00	1.00	1.92
	A_N	-0.08	n.d.	-0.03	n.d.	0.06	n.d.	n.d.	n.a.	n.a.	0.15	n.d.	n.d.	-0.01	n.d.	n.d.	0.02
CA Sa	H_O	0.81	n.d.	0.54	n.d.	0.92	n.d.	0.54	0.08	0.73	n.d.	n.a.	0.23	0.85	0.73	0.54	0.60
	H_E	**0.89**	n.d.	0.58	n.d.	0.82	n.d.	0.56	0.08	0.79	n.d.	n.a.	0.22	0.91	0.76	0.61	**0.62**
	A	9.18	1.35	5.17	1.00	6.60	1.00	2.92	1.63	5.45	1.35	n.a.	2.85	9.53	6.16	3.32	4.11
	A_N	0.05	n.d.	0.03	n.d.	-0.06	n.d.	0.01	0.00	0.03	n.d.	n.a.	-0.01	0.04	0.02	0.05	0.01
CA CpSa	H_O	0.81	n.d.	0.52	n.d.	0.93	n.d.	0.56	0.11	0.74	n.d.	n.a.	0.22	0.85	0.74	0.56	0.60
	H_E	**0.89**	n.d.	0.57	n.d.	0.82	n.d.	0.57	0.11	0.78	n.d.	n.a.	0.27	0.91	0.76	0.62	**0.63**
	A	9.05	1.33	5.08	1.00	6.61	1.00	3.12	1.77	5.39	1.33	n.a.	3.35	9.41	6.08	3.53	4.15
	A_N	0.04	n.d.	0.03	n.d.	-0.06	n.d.	0.01	0.00	0.02	n.d.	n.a.	0.04	0.03	0.01	0.04	0.02
CN Nt	H_O	0.95	0.67	0.33	0.86	0.86	0.67	n.d.	0.29	0.38	0.86	n.a.	0.48	0.38	0.00	0.33	0.54
	H_E	0.93	0.75	0.42	0.74	0.90	0.76	n.d.	**0.61***	**0.48***	0.78	n.a.	0.54	**0.54***	**0.22***	0.34	**0.62***
	A	11.32	7.15	2.82	4.92	9.42	5.88	1.00	3.67	3.25	5.08	n.a.	3.96	2.90	2.00	2.00	4.67
	A_N	-0.01	0.05	0.06	-0.07	0.02	0.05	n.d.	0.20	0.07	-0.05	n.a.	0.04	0.10	0.18	0.01	0.05
CZ1Rr	H_O	0.67	0.67	0.89	0.44	0.89	0.78	0.89	0.67	0.78	0.67	0.11	0.56	0.78	0.44	0.89	0.67
	H_E	0.87	0.82	0.61	**0.82**	0.94	0.82	0.85	**0.87**	0.66	0.81	**0.57**	**0.86**	0.61	**0.74**	0.75	**0.77***
	A	6.00	8.00	4.00	6.00	11.00	7.00	7.00	9.00	4.00	6.00	n.d.	7.00	5.00	5.00	5.00	6.43
	A_N	0.11	0.08	-0.17	0.21	0.03	0.03	-0.02	0.11	-0.07	0.08	0.29	0.16	-0.10	0.17	-0.08	0.05
CZ2Aa	H_O	0.88	0.94	0.65	0.41	0.88	0.59	0.59	0.71	0.53	0.47	0.41	0.88	0.65	0.59	0.76	0.66
	H_E	0.85	0.91	0.69	**0.78***	0.82	0.73	**0.84**	0.91	**0.69**	**0.86***	0.51	0.84	0.71	0.48	0.79	**0.76***
	A	6.83	10.76	4.47	6.08	7.80	5.11	6.72	9.66	3.94	7.32	3.48	6.81	4.06	3.68	5.38	6.14
	A_N	-0.02	-0.02	0.03	0.20	-0.03	0.08	0.14	0.10	0.10	0.21	0.07	-0.02	0.04	-0.07	0.01	0.05
CZ2Rr	H_O	0.67	0.94	0.67	0.94	0.78	0.56	0.72	0.89	0.72	0.67	0.44	0.89	0.61	0.33	0.44	0.69
	H_E	**0.84**	0.91	0.70	0.92	**0.94***	**0.86***	0.81	0.93	0.63	**0.85***	0.47	0.80	0.66	0.45	**0.77***	**0.77***
	A	5.96	9.94	4.75	9.82	11.77	8.19	5.66	11.68	3.64	6.69	2.50	7.34	4.64	2.88	4.49	6.66
	A_N	0.09	-0.02	0.02	-0.01	0.09	0.16	0.05	0.02	-0.06	0.10	0.02	-0.05	0.03	0.08	0.18	0.05
CZ2AaRr	H_O	0.77	0.94	0.66	0.69	0.83	0.57	0.66	0.80	0.63	0.57	0.43	0.89	0.63	0.46	0.60	0.67
	H_E	0.84	0.91	0.70	0.87	0.91	**0.80**	**0.82***	0.92	0.67	**0.86***	0.50	0.83	0.69	0.46	**0.78**	**0.77***
	A	6.45	10.35	4.84	8.32	9.97	6.93	6.10	10.46	3.76	7.36	3.35	7.27	4.39	3.28	5.16	6.53
	A_N	0.04	-0.02	0.03	0.10	0.04	0.13	0.09	0.06	0.02	0.15	0.05	-0.03	0.04	0.00	0.10	0.05
CZ3Ab	H_O	0.64	0.64	0.93	0.57	0.86	0.14	0.50	0.36	0.64	0.64	0.14	0.57	0.57	0.43	0.36	0.53
	H_E	**0.84**	**0.86**	0.68	**0.83**	0.89	**0.68***	**0.57**	**0.76**	0.63	**0.85**	**0.65***	0.85	0.46	**0.54**	0.53	**0.71***
	A	7.00	8.05	4.28	6.92	8.58	3.00	4.81	5.17	3.93	7.41	3.76	7.77	2.88	2.99	4.39	5.40
	A_N	0.11	0.12	-0.15	0.14	0.02	0.32	0.04	0.23	-0.01	0.11	0.31	0.15	-0.08	0.07	0.11	0.10
EEAb	H_O	1.00	0.80	0.40	0.40	0.80	0.80	0.80	0.80	0.60	0.60	0.80	0.60	0.60	n.d.	0.80	0.70
	H_E	0.89	0.78	0.78	**0.89**	0.91	0.80	0.71	0.64	0.51	0.93	0.80	0.80	0.47	n.d.	0.89	**0.77**
	A_N	-0.06	-0.01	0.21	0.26	0.06	0.00	-0.05	-0.09	-0.06	0.17	0.00	0.11	-0.09	n.d.	0.05	0.04
ET BhBt BiBb	H_O	n.d.	n.d.	0.41	0.50	0.55	0.00	0.45	0.64	n.a.	n.d.	n.a.	0.50	0.59	n.d.	0.09	0.41
	H_E	n.d.	n.d.	**0.65**	0.50	**0.82***	**0.09**	0.51	**0.92***	n.a.	n.d.	n.a.	**0.93***	**0.79**	n.d.	0.09	**0.59***
	A	n.d.	1.00	3.88	2.80	7.33	1.66	2.00	10.67	n.a.	1.00	n.a.	11.09	5.41	1.00	1.66	4.12
	A_N	n.d.	n.d.	0.15	0.00	0.15	0.08	0.04	0.15	n.a.	n.d.	n.a.	0.22	0.11	n.d.	0.00	0.10
FR1Rr	H_O	0.91	0.91	0.61	0.91	0.78	0.52	0.65	0.74	0.48	0.78	0.26	0.78	0.57	0.35	0.70	0.66
	H_E	0.87	0.86	0.70	0.90	**0.91***	**0.70***	0.79	0.82	0.48	**0.89***	0.38	0.82	0.65	0.57	0.78	**0.74***
	A	7.65	8.17	3.91	9.42	10.04	5.32	5.68	8.38	2.78	8.46	3.50	6.76	4.49	4.21	5.14	6.26
	A_N	-0.02	-0.03	0.05	-0.01	0.07	0.11	0.08	0.05	0.00	0.06	0.09	0.02	0.05	0.14	0.05	0.05

Table 2. Continued.

Polulation		Li.1	Li.2	Li.5	Li.6	Li.7	Li.8	Li.11	Li.14	Li.17	Li.18	Li.21	Li.23	Li.24	Li.25	Li.30	Mean
FR1Aa	H_O	1.00	0.80	1.00	0.80	0.80	0.80	0.40	0.40	0.20	0.60	0.40	1.00	0.60	0.20	1.00	0.67
	H_E	0.84	0.93	0.73	0.62	0.73	0.78	0.73	0.76	0.56	0.64	0.38	0.89	0.73	0.20	0.73	0.68
	A_N	-0.08	0.07	-0.15	-0.11	-0.04	-0.01	0.19	0.20	0.23	0.03	-0.02	-0.06	0.08	0.00	-0.15	0.01
FR1RrAa	H_O	0.93	0.89	0.68	0.89	0.79	0.57	0.61	0.68	0.43	0.75	0.29	0.82	0.57	0.32	0.75	0.66
	H_E	0.88	0.89	0.71	0.88	**0.92**	**0.71***	**0.79**	0.81	0.50	**0.87**	**0.37**	0.83	0.68	**0.52**	0.79	**0.74***
	A	7.93	9.68	4.18	9.37	10.33	5.34	5.63	8.24	2.70	7.99	3.57	7.04	4.44	4.00	5.22	6.38
	A_N	-0.03	0.00	0.02	-0.01	0.07	0.08	0.10	0.07	0.05	0.06	0.06	0.00	0.07	0.13	0.02	0.05
FR2Rr	H_O	0.96	0.77	0.58	0.77	0.73	0.58	0.88	0.69	0.65	0.73	0.31	0.81	0.65	0.42	0.54	0.67
	H_E	0.85	0.85	0.69	**0.87***	**0.91**	**0.82**	0.84	**0.90***	0.64	**0.86**	**0.46**	0.88	**0.69**	0.58	**0.77**	**0.77***
	A	7.50	8.68	3.35	8.99	10.10	6.28	6.14	9.81	3.87	7.26	2.94	8.33	5.37	3.81	5.36	6.52
	A_N	-0.06	0.04	0.07	0.05	0.09	0.13	-0.02	0.11	-0.01	0.07	0.10	0.04	0.02	0.10	0.13	0.06
FR2RrAb	H_O	0.96	0.79	0.54	0.75	0.68	0.57	0.86	0.64	0.64	0.75	0.32	0.79	0.68	0.39	0.57	0.66
	H_E	0.85	0.85	0.69	**0.88***	**0.90***	**0.83***	0.84	**0.89***	0.63	0.86	0.46	0.87	0.70	0.56	**0.77**	**0.77***
	A	7.40	8.63	3.32	9.08	9.86	6.36	6.10	9.66	3.86	7.35	2.92	8.22	5.29	3.73	5.34	6.47
	A_N	-0.06	0.03	0.09	0.07	0.12	0.14	-0.01	0.13	-0.01	0.06	0.09	0.05	0.01	0.10	0.11	0.06
FR3Rr	H_O	0.73	0.82	0.82	0.82	0.64	0.55	0.82	0.82	0.64	0.64	0.09	1.00	0.82	0.36	0.45	0.67
	H_E	0.91	0.79	0.62	0.94	**0.89***	**0.77**	0.77	0.91	0.45	0.85	**0.40***	0.83	0.63	0.58	0.62	**0.73***
	A	9.21	7.22	3.00	11.49	9.06	5.74	4.97	9.21	2.00	6.61	3.77	6.58	4.77	3.97	3.79	6.09
	A_N	0.10	-0.01	-0.12	0.06	0.14	0.13	-0.03	0.05	-0.13	0.12	0.22	-0.09	-0.12	0.13	0.10	0.04
FR4Rr	H_O	0.86	0.86	0.74	0.77	0.63	0.43	0.89	0.60	0.71	0.86	0.40	0.66	0.51	0.57	0.65	0.68
	H_E	0.90	0.87	0.71	**0.88**	**0.92***	**0.78***	0.82	**0.89***	0.61	0.88	0.44	**0.84***	0.56	0.64	0.78	**0.77***
	A	9.26	8.52	4.53	8.23	10.55	6.66	6.65	9.32	4.16	7.93	4.05	7.53	4.54	3.70	5.86	6.81
	A_N	0.02	0.01	-0.02	0.06	0.15	0.20	-0.04	0.15	-0.06	0.01	0.03	0.10	0.03	0.04	0.07	0.05
D Rr	H_O	0.83	0.83	0.50	0.83	0.83	0.50	1.00	0.33	0.50	0.33	0.50	0.67	0.67	0.33	0.67	0.62
	H_E	0.76	0.89	0.62	0.86	0.80	0.53	0.85	0.56	0.64	0.68	0.68	0.68	0.68	0.32	0.77	0.69
	A_N	0.45	0.45	0.33	0.45	0.45	0.33	0.50	0.25	0.33	0.25	0.33	0.40	0.40	0.25	0.40	0.37
RU Ab	H_O	1.00	0.56	0.56	1.00	0.56	0.67	0.78	0.78	0.67	0.56	0.44	0.56	0.67	0.33	0.78	0.66
	H_E	0.89	**0.87***	0.53	0.91	**0.93***	0.66	0.76	0.84	0.50	0.64	0.57	0.75	0.54	0.67	0.80	**0.72**
	A	7.00	9.00	3.00	9.00	10.00	4.00	5.00	10.00	3.00	8.00	n.d.	7.00	4.00	4.00	6.00	6.36
	A_N	-0.06	0.17	-0.02	-0.05	0.20	0.00	-0.01	0.04	-0.11	0.05	0.08	0.11	-0.08	0.20	0.01	0.04
TN Rb	H_O	0.83	0.73	0.67	0.79	0.63	0.40	0.77	0.77	0.54	0.58	0.31	0.67	0.54	0.65	0.42	0.62
	H_E	0.83	0.76	0.71	0.87	**0.87***	**0.59***	0.74	0.82	0.60	**0.79***	**0.49***	**0.81***	**0.58**	**0.73***	**0.73***	**0.73***
	A	6.53	5.36	3.99	8.30	8.34	4.65	4.53	9.00	2.99	7.11	3.83	6.97	3.41	4.57	4.15	5.58
	A_N	0.00	0.02	0.03	0.04	0.13	0.12	-0.02	0.02	0.04	0.12	0.12	0.08	0.03	0.05	0.18	0.06
TN Se	H_O	1.00	0.60	0.60	0.80	0.40	0.40	0.60	1.00	0.80	0.80	0.40	0.20	0.60	1.00	0.40	0.64
	H_E	0.93	0.73	0.78	0.87	**0.89**	**0.84**	0.82	0.96	0.64	0.76	0.62	0.51	0.73	0.73	0.64	**0.76***
	A_N	-0.03	0.08	0.10	0.04	0.26	0.24	0.12	-0.02	-0.09	-0.03	0.14	0.21	0.08	-0.15	0.15	0.07
TN RbSe	H_O	0.85	0.72	0.66	0.79	0.60	0.40	0.75	0.79	0.57	0.60	0.32	0.62	0.55	0.68	0.42	0.62
	H_E	0.84	0.76	0.71	**0.86**	**0.87***	**0.62***	0.75	0.83	0.60	**0.79***	**0.50***	**0.79***	**0.60**	**0.73**	**0.72***	**0.73***
	A	6.64	5.24	4.01	8.11	8.27	5.07	4.57	9.11	2.99	7.04	3.84	6.71	3.47	4.52	4.08	5.58
	A_N	-0.01	0.02	0.03	0.04	0.14	0.14	0.00	0.02	0.02	0.11	0.12	0.09	0.03	0.03	0.18	0.06

H_O, observed heterozygosity; H_E, expected heterozygosity; A, allelic richness adjusted for n=9; A_N, frequency of null alleles[$(H_E - H_O)/(H_E + 1)$] (Brookfield, 1996); n.a., nonamplified locus; n.d., not diagnosed – too few data. In bold, values of H_E where significant departure from Hardy-Weinberg equilibrium was observed ($P \leq 0.05$; *, values significant after sequential Bonferroni correction). Values of mean H_E in bold when significant departure from HWE observed in global test (across loci in population).

Table 3. AMOVA results.

analysed set			Fst		Rst	
All poulations	Grouping criterion	d.f.	Variance components	%	Variance components	%
1	*mtDNA*					
	Among groups	4	1.076	**22.26**	157.477	**33.8**
	Among populations within groups	29	0.322	**6.65**	21.229	**4.56**
	Among samples within populations	304	0.471	**9.74**	74.169	**15.92**
	Within samples	338	2.966	**61.35**	212.998	**45.72**
2	*NJ-msat*					
	Among groups	7	1.396	**28**	159.957	**34.37**
	Among populations within groups	26	0.152	**3.05**	18.318	**3.94**
	Among samples within populations	304	0.471	**9.45**	74.169	**15.94**
	Within samples	338	2.966	**59.51**	212.998	**45.76**
Clade A	Grouping criterion	d.f.	Variance components	%	Variance components	%
1	*Geography*					
	Among groups	5	0.117	**2**	19.035	4.04*
	Among populations within groups	15	0.135	**2.29**	5.744	*1.22*
	Among samples within populations	222	0.783	**13.32**	113.15	**24.02**
	Within samples	243	4.846	**82.39**	333.15	**70.72**
2	*Hosts*					
	Among groups	5	0.07	1.18*	14.904	3.14*
	Among populations within groups	15	0.194	**3.29**	13.691	**2.88**
	Among samples within populations	222	0.783	**13.29**	113.15	**23.82**
	Within samples	243	4.846	**82.23**	333.177	**70.15**
3	*Compromise*					
	Among groups	4	0.156	**2.63**	18.679	3.94*
	Among populations within groups	16	0.132	**2.24**	9.3	1.96*
	Among samples within populations	222	0.783	**13.24**	113.15	**23.86**
	Within samples	243	4.846	**81.89**	333.147	**70.25**

*, level of significance P=(0.001-0.05), results significant at P<0.001 are in bold, nonsignificant in italics

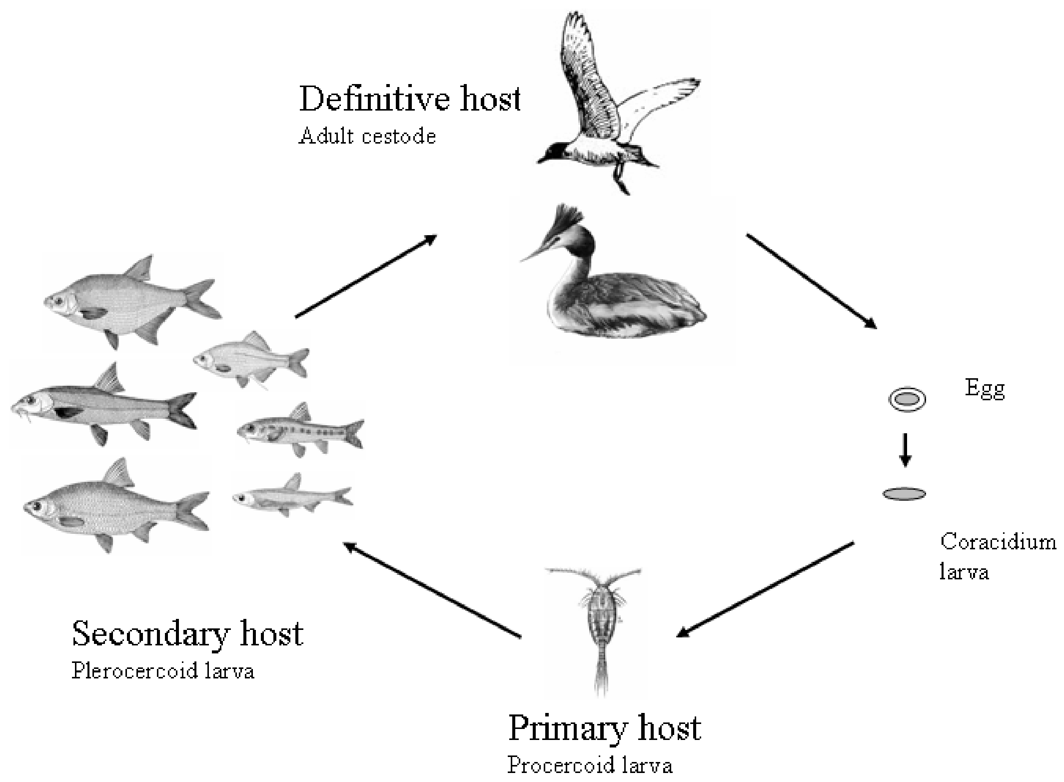


Fig. 1. Three-host life-cycle of *Ligula intestinalis*.

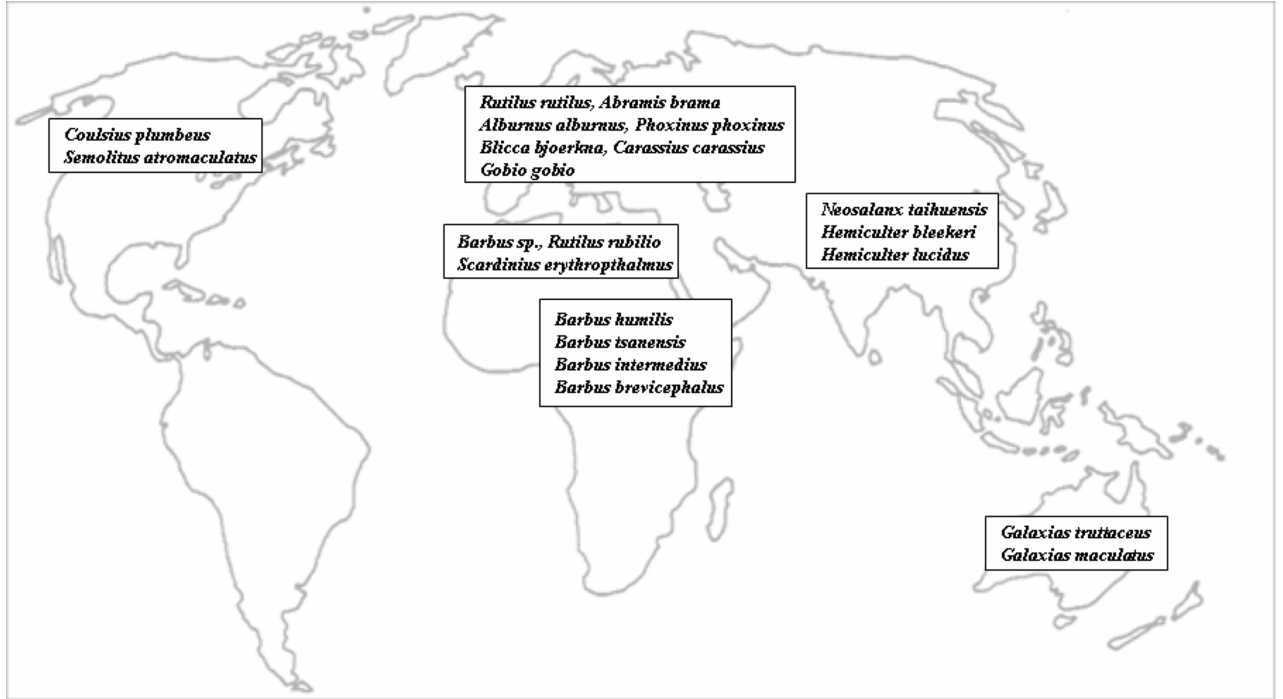


Fig 2. Map of distribution of *Ligula* sp. isolates analysed in this study. Detailed information on localities and sample sizes is provided in Table 1.

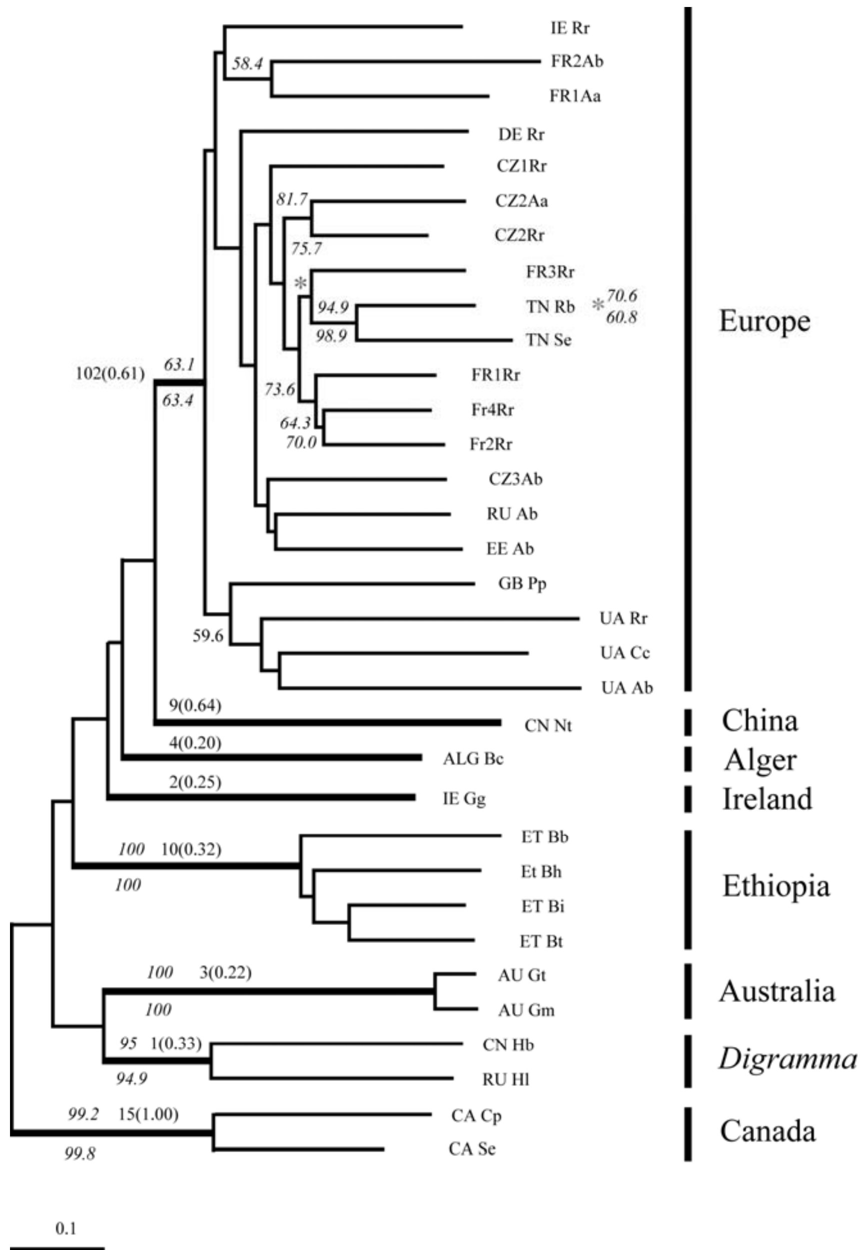
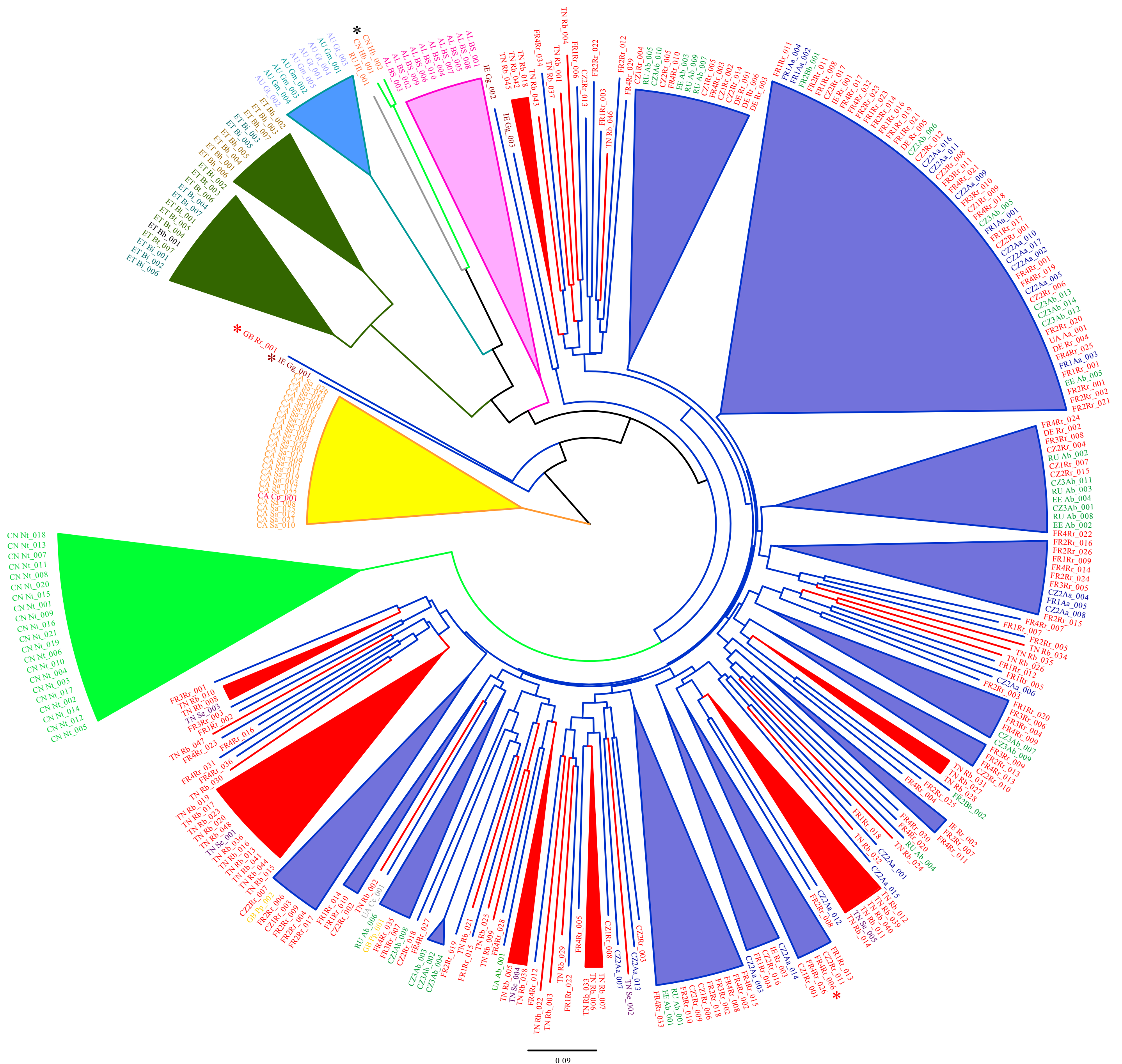


Fig 3. Unrooted NJ tree of *Ligula* populations. The values at nodes printed in italics show bootstrap supports higher than 50% for clustering according to D_A distances (above the branches) and D_N distances (below the branches). Numbers above the clades printed in bold indicate numbers of private alleles specific for the given clade; values in parentheses stand for the frequency of the most frequent private allele .



Region	Host fish
Europe, Tunisia	*Rutilus rutilus* and *R. rubilio*, *Abramis brama* and *Blicca bjoerkna*, *Alburnus alburnus*, *Scardinius erythrophthalmus*, *Phoxinus phoxinus*, *Carassius carassius*, *Gobio gobio*
Alger	*Barbus sp.*
China	*Neosalanx tainuensis*
Australia	*Galaxias maculatus*, *Galaxias truttaceus*
China, Far East Russia	*Hemiculter bleekeri* and *H. lucidus* *
Ethiopia	*Barbus humilis*, *Barbus tsanensis*, *Barbus intermedius*, *Barbus brevicephalus*
Canada	*Semolitus atromaculatus*, *Coulsius plumbeus*

* samples from genus *Hemiculter* belong to *Ligula interrupta* (formerly *Digramma interrupta*)

* * check for cluster affiliation of specimens in Fig. 6

Fig. 4. Unrooted NJ tree based on pairwise D_A distances between 338 *Ligula* sp. individuals. The sample labels are coloured according to the host fish, the clades‘ coloures refer to the region of the sample origin. Monophyletic subclades containing three or more specimens of the same geographic origin are joined into common clusters. The two sympatric Ethiopian lineages differing in their host spectra are presented as separate clades.

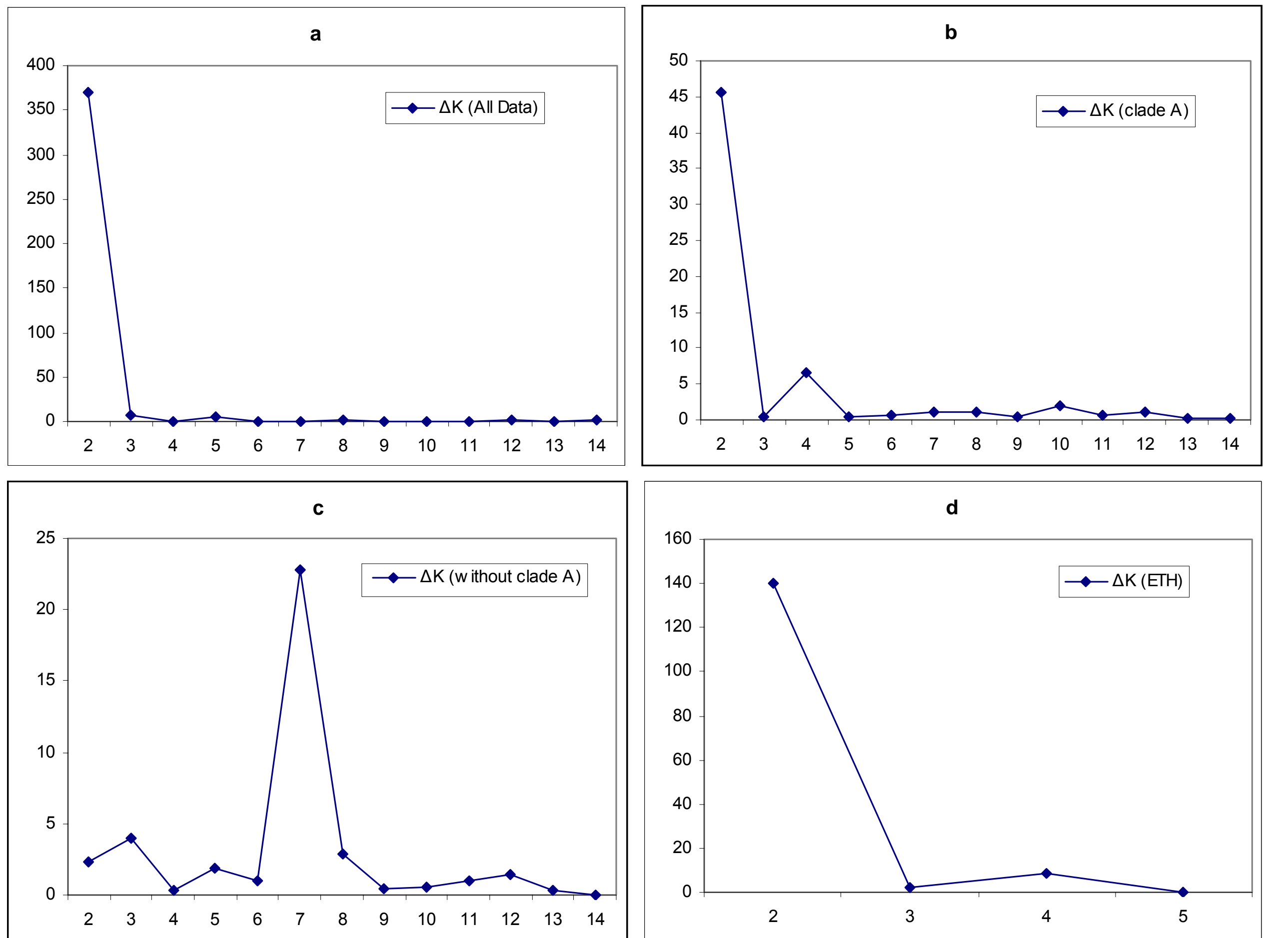


Fig. 5. Values of ΔK (Evanno et al., 2005), calculated for each number of clusters (y axis) estimated for datasets analysed using STRUCTURE 2.2. a, whole dataset; b, set of Clade A populations; c, set without Clade A populations; d, population Ethiopia under model no. 3.

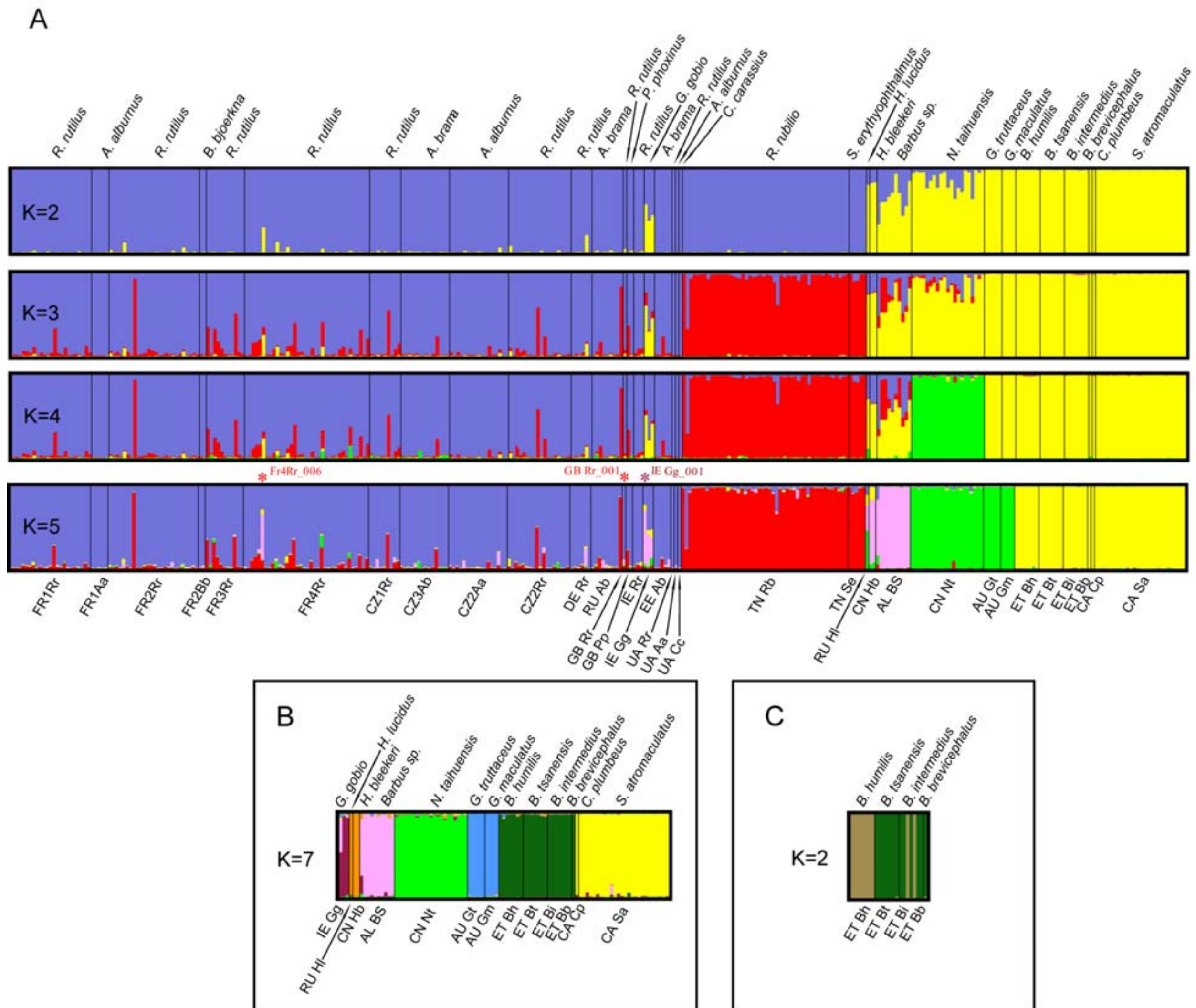


Fig.6. Results of STRUCTURE clustering analysis. A, results for the complete dataset. Bar plots for numbers of clusters (K) form 2 to 5 are shown, where each of the 388 individuals is represented by a vertical bar. Estimated proportion of membership to each cluster is indicated by different colours. Labels on top of the plots indicate host spectra of samples, labels below indicate population origin of samples. B, results for a dataset without Clade A populations. C, results for Ethiopian populations under model no. 3.

App.1. Population differentiation and pairwise Fst for all populations

pop	FR1 Rr	FR1 Aa	FR2 Rr	FR2 Bb*	FR3 Rr	FR4 Rr	CZ1 Rr	CZ3 Ab	CZ2 Aa	CZ2 Rr	D Rr	RU Ab	GB1 Rr*	GB2 Pp*	IE Rr*	IE Gg*	EE Ab	UA Aa*	UA Cc*	UA Rr*	TN Rb	TN Se*	RU Hl*	CN Hb*	AL BS	CN Nt	ET Bh	ET Bt	ET Bi	Et Bb*	AU Gm	AU Gt*	CA Cp*
FR1Aa	**0.05**																																
FR2Rr	**0.01**	**0.05**																															
FR2Ab*	0.03	0.06	0.00																														
FR3Rr	**0.01**	**0.07**	0.01	0.04																													
FR4Rr	0.01	**0.07**	**0.01**	0.05	0.01																												
CZ1Rr	**0.03**	**0.11**	**0.02**	0.06	**0.03**	**0.02**																											
CZ3Ab	**0.06**	**0.11**	**0.06**	**0.11**	**0.06**	**0.05**	**0.04**																										
CZ2Aa	**0.02**	**0.05**	**0.02**	0.05	**0.04**	**0.03**	**0.04**	**0.08**																									
CZ2Rr	**0.02**	**0.05**	**0.02**	0.05	**0.03**	**0.01**	0.00	**0.05**	0.01																								
D Rr	**0.06**	**0.11**	**0.07**	**0.11**	**0.10**	**0.06**	**0.05**	**0.06**	**0.07**	**0.03**																							
RU Ab	**0.05**	**0.12**	**0.04**	**0.12**	**0.06**	**0.03**	0.02	**0.03**	**0.05**	**0.02**	**0.07**																						
GB1Rr*	0.12	0.18	0.09	0.01	0.11	0.09	0.09	**0.14**	0.12	0.09	**0.22**	0.14																					
GB2Pp*	0.02	0.07	0.01	0.03	0.01	-0.01	-0.02	**0.04**	**0.05**	0.00	0.06	0.03	0.01																				
IE Rr*	0.04	**0.12**	0.03	-0.01	0.05	0.03	0.02	**0.12**	0.02	0.03	**0.11**	0.08	0.13	-0.03																			
IE Gg*	**0.14**	**0.16**	**0.13**	**0.14**	0.13	**0.12**	**0.13**	**0.20**	**0.13**	**0.13**	**0.21**	**0.16**	0.08	0.03	**0.09**																		
EE Ab	**0.03**	0.07	**0.04**	0.07	0.03	0.03	0.01	**0.02**	**0.05**	0.01	**0.05**	0.00	0.13	0.00	0.05	**0.14**																	
UA Aa*	0.05	0.13	0.01	0.02	0.05	0.01	0.00	0.06	0.05	0.03	0.11	0.05	0.63	-0.11	-0.04	0.13	0.06																
UA Cc*	-0.02	0.04	-0.05	-0.07	-0.01	-0.03	-0.09	0.02	-0.03	-0.07	0.05	0.01	0.57	-0.25	-0.08	0.06	-0.03	0.51															
UA Rr*	0.05	0.09	0.03	-0.02	0.07	0.03	-0.01	0.00	0.04	0.01	0.00	0.08	0.64	-0.13	-0.04	0.15	0.08	0.59	0.51														
TN Rb	**0.04**	**0.10**	**0.04**	**0.08**	**0.04**	**0.03**	**0.05**	**0.09**	**0.07**	**0.05**	**0.11**	**0.07**	**0.12**	**0.03**	**0.06**	**0.15**	**0.06**	0.08	0.01	**0.12**													
TN Se*	**0.03**	**0.10**	**0.02**	0.04	0.00	0.01	0.01	**0.06**	**0.04**	0.02	**0.10**	**0.06**	0.07	0.00	0.01	**0.13**	0.03	0.02	-0.05	0.07	-0.01												
RU Hl*	**0.25**	0.30	**0.24**	0.23	**0.27**	**0.23**	**0.23**	**0.26**	**0.24**	**0.22**	0.22	**0.27**	0.72	0.17	0.22	0.20	**0.28**	0.75	**0.71**	0.71	**0.28**	**0.26**											
CN Hb*	**0.36**	**0.45**	**0.34**	**0.47**	**0.38**	**0.34**	**0.35**	**0.39**	**0.35**	**0.35**	**0.43**	**0.40**	**0.75**	**0.51**	**0.45**	**0.48**	**0.44**	**0.77**	0.73	**0.74**	0.38	**0.41**	0.70										
AL BS	**0.18**	**0.24**	**0.17**	**0.20**	**0.19**	**0.17**	**0.15**	**0.22**	**0.18**	**0.17**	**0.22**	**0.18**	**0.18**	**0.17**	**0.19**	**0.16**	**0.18**	**0.18**	0.10	**0.24**	**0.19**	**0.18**	**0.26**	**0.41**									
CN Nt	**0.21**	**0.29**	**0.21**	**0.29**	**0.22**	**0.19**	**0.22**	**0.24**	**0.23**	**0.21**	**0.26**	**0.24**	**0.31**	**0.23**	**0.29**	**0.30**	**0.26**	**0.34**	**0.25**	**0.31**	**0.22**	**0.25**	**0.36**	**0.46**	**0.29**								
ET Bh	**0.32**	**0.44**	**0.32**	**0.48**	**0.35**	**0.30**	**0.34**	**0.40**	**0.31**	**0.31**	**0.43**	**0.37**	**0.52**	**0.45**	**0.39**	**0.45**	**0.40**	**0.57**	**0.49**	**0.53**	**0.32**	**0.35**	**0.49**	**0.65**	**0.36**	**0.44**							
ET Bt	**0.31**	**0.44**	**0.31**	**0.45**	**0.33**	**0.28**	**0.32**	**0.38**	**0.31**	**0.30**	**0.42**	**0.35**	**0.50**	**0.43**	**0.37**	**0.41**	**0.37**	**0.54**	**0.48**	**0.51**	**0.30**	**0.32**	**0.49**	**0.63**	**0.32**	**0.41**	**0.19**						
ET Bi	**0.30**	**0.42**	**0.30**	**0.43**	**0.32**	**0.27**	**0.30**	**0.37**	**0.29**	**0.28**	**0.40**	**0.33**	**0.46**	**0.40**	**0.33**	**0.39**	**0.35**	**0.50**	**0.44**	**0.47**	**0.29**	**0.29**	**0.42**	**0.61**	**0.32**	**0.40**	**0.07**	0.02					
Et Bb*	**0.27**	**0.39**	**0.25**	0.25	**0.28**	**0.24**	0.24	**0.34**	**0.26**	**0.25**	**0.37**	**0.30**	0.80	0.21	0.10	0.22	0.29	0.83	0.77	0.82	**0.26**	0.19	0.86	**0.98**	**0.28**	**0.42**	0.25	0.17	0.09				
AU Gm	**0.40**	**0.52**	**0.39**	**0.66**	**0.46**	**0.39**	**0.46**	**0.46**	**0.43**	**0.42**	**0.53**	**0.48**	**0.71**	**0.61**	**0.63**	**0.52**	**0.51**	**0.78**	**0.74**	**0.78**	**0.41**	**0.52**	**0.75**	**0.83**	**0.43**	**0.46**	**0.74**	**0.71**	**0.71**	**0.87**			
AU Gt*	**0.37**	**0.46**	**0.36**	**0.58**	**0.42**	**0.36**	**0.42**	**0.43**	**0.40**	**0.39**	**0.47**	**0.44**	**0.61**	**0.52**	**0.56**	**0.44**	**0.46**	**0.70**	**0.66**	**0.70**	**0.39**	**0.46**	**0.66**	**0.78**	**0.39**	**0.44**	**0.70**	**0.67**	**0.67**	**0.81**	-0.04		
CA Cp*	**0.22**	**0.29**	**0.19**	0.15	**0.23**	**0.19**	**0.20**	**0.23**	**0.19**	**0.20**	**0.28**	**0.24**	0.65	0.15	0.17	0.08	0.23	0.68	0.65	0.71	**0.23**	0.19	0.76	**0.79**	**0.20**	**0.38**	**0.58**	**0.53**	**0.50**	0.85	**0.76**	**0.68**	
CA Se	**0.34**	**0.42**	**0.33**	**0.46**	**0.37**	**0.32**	**0.37**	**0.37**	**0.35**	**0.34**	**0.42**	**0.38**	**0.39**	**0.42**	**0.46**	**0.38**	**0.40**	**0.50**	**0.49**	**0.50**	**0.35**	**0.41**	**0.51**	**0.59**	**0.38**	**0.46**	**0.60**	**0.58**	**0.58**	**0.64**	**0.55**	**0.52**	0.08

Bold Fst values mark population pairs where significant level of genic differentiation was observed (P-level was adjusted by sequential Bonferroni correction); populations comprising less than 5 samples are designated with an asterisk.

App.2. Population differentiation and pairwise Fst for pooled populations.

pop	***FR1Rr Aa***	***FR2 RrAb***	**FR3Rr**	**FR4Rr**	**CZ1Rr**	**CZ3Ab**	***CZ2 AaRr***	**D Rr**	**RU Ab**	***GB RrPp****	**IE Rr***	**IE Gg***	**EE Ab**	***UARr CcAa****	***TN RbSe***	**RU Hl***	**CN Hb***	**AL BS**	**CN Nt**	***ET Bh BtBiBb***	***AU GtGm***
FR2RrAb	**0.01**																				
FR3Rr	**0.02**	**0.01**																			
FR4Rr	**0.01**	**0.01**	0.01																		
CZ1Rr	**0.04**	**0.03**	**0.03**	0.02																	
CZ3Ab	**0.06**	**0.06**	**0.06**	**0.05**	**0.04**																
CZ2AaRr	**0.02**	**0.02**	**0.03**	**0.02**	**0.02**	**0.06**															
D Rr	**0.06**	**0.07**	**0.10**	**0.06**	**0.05**	**0.06**	**0.05**														
RU Ab	**0.05**	**0.05**	**0.06**	**0.03**	**0.02**	**0.03**	**0.03**	**0.07**													
GB RrPp*	**0.02**	**0.01**	0.01	0.00	-0.01	**0.05**	0.02	**0.07**	**0.04**												
IE Rr*	0.05	0.03	0.05	0.03	0.02	**0.12**	0.03	**0.11**	0.08	0.02											
IE Gg*	**0.13**	**0.13**	**0.13**	**0.12**	**0.13**	**0.20**	**0.13**	**0.21**	**0.16**	0.03	**0.09**										
EE Ab	0.03	**0.04**	0.03	0.03	0.01	**0.02**	**0.03**	**0.05**	0.00	0.01	0.05	**0.14**									
UARrCcAa*	0.03	0.02	0.04	0.02	-0.01	**0.05**	0.02	0.05	0.05	-0.01	0.04	**0.17**	**0.04**								
TN RbSe	**0.04**	**0.04**	**0.04**	**0.03**	**0.05**	**0.08**	**0.05**	**0.11**	**0.07**	**0.03**	**0.06**	**0.15**	0.06	**0.08**							
RU Hl*	**0.24**	**0.24**	**0.27**	**0.23**	**0.23**	**0.26**	**0.22**	0.22	**0.27**	0.15	0.22	0.20	**0.28**	**0.22**	**0.28**						
CN Hb*	**0.36**	**0.33**	**0.38**	**0.34**	**0.35**	**0.39**	**0.34**	**0.43**	**0.40**	**0.39**	**0.45**	**0.48**	**0.44**	0.43	**0.37**	0.70					
AL BS	**0.19**	**0.17**	**0.19**	**0.17**	**0.15**	**0.22**	**0.17**	**0.22**	**0.18**	**0.15**	**0.19**	**0.16**	**0.18**	**0.18**	**0.19**	**0.26**	**0.41**				
CN Nt	**0.21**	**0.21**	**0.22**	**0.19**	**0.22**	**0.24**	**0.21**	**0.26**	**0.24**	**0.21**	**0.29**	**0.30**	**0.26**	**0.27**	**0.22**	**0.36**	**0.46**	**0.29**			
ET BhBtBiBb	**0.33**	**0.33**	**0.36**	**0.30**	**0.35**	**0.41**	**0.30**	**0.43**	**0.38**	**0.39**	**0.37**	**0.42**	**0.39**	**0.43**	**0.31**	**0.42**	**0.57**	**0.35**	**0.42**		
AU GtGm	**0.41**	**0.41**	**0.49**	**0.41**	**0.51**	**0.50**	**0.42**	**0.57**	**0.52**	**0.55**	**0.67**	**0.57**	**0.56**	**0.63**	**0.43**	**0.72**	**0.80**	**0.47**	**0.49**	0.67	
CA CpSe	**0.33**	**0.33**	**0.37**	**0.32**	**0.37**	**0.36**	**0.32**	**0.42**	**0.38**	**0.35**	**0.46**	**0.37**	**0.39**	**0.44**	**0.34**	**0.51**	**0.59**	**0.38**	**0.46**	0.56	**0.56**

Bold Fst values mark population pairs where significant level of genic differentiation was observed (P-level was adjusted by sequential Bonferroni correction); populations comprising less than 5 samples are designated with an asterisk.

App.3. Population differentiation and pairwise Fst for NJ clades.

clade	Eu cladeA	IEGg+ Fr4Rr_006	Digramma	AL BS	CN NT	ET1 BhBi	ET2 BtBiBb	AU
IEGg+Fr4Rr_006	0.10							
Digramma	0.29	0.36						
AL BS	0.16	0.14	0.36					
CN NT	0.18	0.28	0.41	0.29				
ET1 BhBi	0.29	0.41	0.58	0.38	0.45			
ET2 BtBiBb	0.27	0.37	0.54	0.35	0.42	0.18		
AU	0.37	0.52	0.72	0.47	0.49	0.75	0.70	
CA	0.28	0.37	0.54	0.38	0.46	0.60	0.57	0.56

Significant level of genic differentiation was observed between all pairs of clades (P-level was adjusted by sequential Bonferroni correction).

Appendix 4. HW equilibrium for Ethiopian clades.

Population	N		Li.1	Li.2	Li.5	Li.6	Li.7	Li.8	Li.11	Li.14	Li.17	Li.18	Li.21	Li.23	Li.24	Li.25	Li.30	Mean
EtBhBtBiBb	22	H_O	n.d.	n.d.	0.41	0.5	0.55	0	0.45	0.64	n.a.	n.d.	n.a.	0.5	0.59	n.d.	0.09	0.41
		H_E	n.d.	n.d.	**0.65**	0.5	**0.82***	**0.09**	0.51	**0.92***	n.a.	n.d.	n.a.	**0.93***	**0.79**	n.d.	0.09	**0.59***
ET1 BhBi	9	H_O	n.d.	n.d.	0.11	0.44	0.44	n.d.	0.56	0.44	n.a.	n.d.	n.a.	0.67	0.56	n.d.	0.222	0.43
		H_E	n.d.	n.d.	0.11	0.39	0.63	n.d.	0.5	**0.87***	n.a.	n.d.	n.a.	**0.94**	0.75	n.d.	0.209	**0.55***
ET2 BtBiBb	12	H_O	n.d.	n.d.	0.67	0.58	0.67	0	0.42	0.83	n.a.	n.d.	n.a.	0.42	0.67	n.d.	n.d.	0.53
		H_E	n.d.	n.d.	0.8	0.55	0.71	**0.16**	0.55	0.96	n.a.	n.d.	n.a.	**1.00***	0.68	n.d.	n.d.	**0.68***

	No. of private alleles
ET1 BhBi	18 (0.556)
ET2 BtBiBb	26 (0.557)

H_O, observed heterozygosity; H_E, expected heterozygosity; n.a., nonamplified locus; n.d., not diagnosed – too few data. In bold, values of H_E where significant departure from Hardy-Weinberg equilibrium was observed ($P \leq 0.05$; *, values significant after sequential Bonferroni correction). Values of mean H_E in bold when significant departure from HWE observed in a global test (across loci in population). In the bottom part - numbers of private alleles differing between the two populations, in brackets - frequency of a private allele occurring with the highest frequency.

Manuscript No. 4:

Štefka, J., Hypša, V. 2008. Host specificity and genealogy of the louse *Polyplax serrata* on field mice, *Apodemus* species: A case of parasite duplication or colonisation? Int. J. Parasitol., 38: 731-741

Available online at www.sciencedirect.com

ScienceDirect

International Journal for Parasitolo gy 38 (2008) 731–741

www.elsevier.com/locate/ijpara

Host specifcity and genealogy of the louse Polyplax serrata on feld mice, Apodemus species: A case of parasite duplication or colonisation? [q]

Jan Stefka *, V áclav Hyps a

Faculty of Sciences and Biology Centre, Institute of Parasitology, Branisovská 31, Ceské Budejovice 37005, Czech Republic

Received 26 July 2007; received in revised form 20 September 2007; accepted 21 September 2007

Abstract

The genealogy, population structure and population dynamics of the sucking louse Polyplax serrata were analysed across four host species of the genus Apodemus An analysis of 126 sequences of cytochrome c oxidase subunit I using phylogenetic approaches and haplotype networking revealed a clear structure of European samples, forming three distinct and genetically distant clades with di?erent host speci?cities. Although a clear connection was detected between the host and parasite genealogies/phylogenies, a uniform pattern of co-speciation was not found. For example, a dramatic shift in the degree of host speci?city was demonstrated for two related louse lineages living in sympatry and sharing one of their host species. While one of the louse lineages frequently parasitised two di?erent host taxa (Apodemus sylvaticus and Apodemus ?avicollis), the other louse lineage was strictly speci?c to A. ?avicollis . The estimate of divergence time between the two louse lineages indicates that they may have arisen due to parasite duplication on A. ?avicollis .

2007 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: Parasite duplication; Host speci?city; Genealogy; Speciation; Polypla x; Apodemus

Manuscript on pages 86 to 96 is a subject to copyright

Curriculum vitae

Jan Štefka, MSc.

University of South Bohemia, Faculty of Science and Laboratory of Molecular Phylogeny and Evolution of Parasites, Institute of Parasitology, Biology Centre ASCR. Branišovská 31, 370 05 České Budějovice, Czech Republic.
tel.: +420-38-777-5409, e-mail: stefka@paru.cas.cz, web: www.jstefka.ic.cz

Date of birth: April 25, 1979
Place of birth: Přílepy, Czech Republic

Education:
2001: Bachelor degree: Introductory Biology Programme, Faculty of Biological Sciences, University of South Bohemia. Thesis: Analýza rodu *Bombina* v Předšumaví. [Analysis of the genus *Bombina* in region Předšumaví, 20 pp., in Czech with English summary].
2003: Master degree: Zoology, Faculty of Biological Sciences, University of South Bohemia. Thesis: Ecological aspects of hybridisation between fire-bellied toads *Bombina bombina* and *Bombina variegata*. 27 pp.
2003-present: Student of PhD programme in Parasitology, Faculty of Science, University of South Bohemia.
PhD thesis topic: Host specificity, genetic variability and genealogy in populations of model parasite species.

Appointments:
2003-present: Research assistant, Institute of Parasitology, Biology Centre, ASCR.
2005-present: Research assistant, Faculty of Science, University of South Bohemia.

Abroad stays and fellowships:
2006 (April-June): University of Glasgow, UK, Faculty of Veterinary Medicine. Research theme: Microsatellite characterization in *Ligula intestinalis* (Cestoda: Pseudophyllidea), Socrates/Erasmus fellowship.
2006-2007 (several short-term stays): Université Paul Sabatier Toulouse, France, Laboratoire Evolution Diversité Biologique, Research theme: Genetic variability of European populations of *Ligula intestinalis*. Barrande fellowship.
2007 (Nov.-Dec.): Argentina, Bolivia, Brasil. Field sampling with the focus on helminth fauna of endemic fishes.

Publications:
Logan, F.J., Horák, A., **Štefka, J.**, Aydogdu, A., Scholz, T. 2004. Phylogeny of diphyllobothriid tapeworms (Cestoda: Pseudophyllidea) based on sequences of the ITS-2 rDNA. *Parasitol. Res.*, 94: 10-15.

Štefka, J., Gilleard, J.S., Grillo, V., Hypša, V. 2007. Isolation and characterization of microsatellite loci in the tapeworm *Ligula intestinalis* (Cestoda: Pseudophyllidea). *Mol. Ecol. Notes.*, 7: 794-796.

Štefka, J., Hypša, V. 2008. Host specificity and genealogy of the louse *Polyplax serrata* on field mice, *Apodemus* species: A case of parasite duplication or colonisation? *Int. J. Parasitol.*, 38: 731-741.

Bouzid, W., **Štefka, J.**, Hypša V., Lek, S., Scholz, T., Legal, L., Ben Hassine, O.K., Loot G. 2008. Geography and host specificity: two forces behind the genetic structure of the freshwater fish parasite *Ligula intestinalis* (Cestoda: Diphyllobothriidae). *Int. J. Parasitol.*, in press.

Conference presentations:

Štefka, J., Škeříková A., Hypša V. 2005. ITS-2 and phylogenetic studies. Vth International Workshop on Cestode Systematics and Phylogeny, České Budějovice, Czech Republic.

Štefka, J., Hypša, V., Fričová, J., Stanko, M. 2006. Host-parasite associations, geography and population subdivision in *Polyplax serrata* (Anoplura: Polyplacidae) and wood mice of genus *Apodemus*. 3rd International Congress on Phthiraptera, Oct. 16 - 20, 2006, Buenos Aires, Argentina.

Štefka, J., Hypša, V., Scholz, T. 2007. Genetic structure of *Ligula sp.* (Cestoda: Pseudophyllidea): true population history vs. rDNA paralogues. 15th helminthological days held at Ředkovec (Czech Republic) May 14 – 18, 2007.

Invited lectures:

Horák, A., **Štefka, J.** 2005. Inter- and intraspecific genetic variability of diphyllobothriid tapeworms, Proceedings of the 13th helminthological days held at Ředkovec (Czech Republic) May 9 – 13, 2005, *Helminthologia*, 42(3): 171-186.

Štefka, J., Bouzid W., Hypša, V., Scholz, T. 2007. Ribosomal DNA spacers: Eligible tool to resolve low-level phylogenetics in helminths? 10th International Helminthological Symposium, Sept. 9-14, 2007, Stará Lesná, Slovakia.