Charles University in Prague
Faculty of Science

Study program: Biology

Study branch: Ecology

Bc. Nuria Viñuela Rodríguez

Phylogeny and phylogeography of the cyprinid fish genus Pelasgus (Teleostei: Cyprinidae)

Fylogeneze a fylogeografie kaprovitých ryb rodu Pelasgus (Teleostei: Cyprinidae)

Diploma Thesis

Supervisor: RNDr. Jasna Vukićová, Ph.D.

Prague, 2016

Declaration:

I hereby declare that this diploma thesis is entirely the result of my own work and I have acknowledged all the sources of information which I have been used in it. This thesis has not been submitted in order to obtain the same or any other academic degree.

In Prague, 15. 08. 2016

Signature

Acknowledgements

My deepest gratitude to my supervisor, RNDr. Jasna Vukićová, Ph.D., who has supported and helped me during my period as master student, providing me with wise advices and knowledge, and who first opened the door of Charles University to me. I have been very lucky to have such an encouraging and caring supervisor and it has been a great pleasure being a student of hers.

Secondly I would like to say thank you to Mgr. Radek Šanda, Ph.D. (National Museum, Prague) who firstly gave me the opportunity to learn the molecular methods in his laboratory and introduced me to the world of phylogeny and phylogenetic studies.

Last but not least, I would like to thank my beloved parents, for supporting me in the adventure of studying abroad, for always encouraging me to do what I want and what I like, and for helping me to heal my homesick by bringing a piece of home whenever they visited me. Thank you. Also the warmest thanks to my best friend and partner, for always being by my side, with infinite patience and care.

Abstract

The genus *Pelasgus* (Cyprinidae) is endemic to the southern part of the Balkan Peninsula and includes seven species. In this work, a multilocus approach has been applied to study phylogenetic relationships between the species and their populations and to revise their distribution areas. 180 specimens from 47 localities from 30 river drainages were analyzed, comprehensively covering the distribution range of the genus. Moreover, samples from type localities of all species were included in the analyses. Mitochondrial (cytochrome b) and nuclear markers (the first intron of ribosomal protein S7, recombination activating gene RAG1 and rhodopsin) were used. Existence of seven well supported lineages was revealed based on cytochrome b, which is the most variable marker. These lineages correspond to *P. laconicus*, *P. marathonicus*, *P. minutus*, *P. stymphalicus*, *P. thesproticus*, *P. prespensis* and *Pelasgus* sp. The most variable nuclear marker was first intron of S7, which provides almost the same results as cytochrome b, revealing six well supported lineages, whereas RAG1 and rhodopsin appear to be less informative, revealing only four well supported clades. These markers did not separate several species (*P. marathonicus*, *P. stymphalicus*, *P. thesproticus*, and *Pelasgus* sp.) due to low variability of the markers and common haplotype sharing between these species. *Pelasgus* sp. requires further research; it could be *P. epiroticus*. However, this issue could be resolved only by analyses of the specimens from Ioannina Lake, the population, which is possibly extinct. Molecular identification has allowed reshaping the knowledge of the distribution areas for the species of the genus. The introduction of *Pelasgus* sp. is suggested in two localities on the Peloponnesus Peninsula (Kandila springs and Stymphalia Lake) and possibly in one locality in the mainland (Zaravina Lake in W Greece). Sympatric occurrence of *P. stymphalicus* and *Pelasgus* sp. within the above mentioned Peloponnesus localities and hybridization between them was revealed.

Key words: Pelasgus, mitochondrial and nuclear markers, haplotype network, phylogeography, southern Balkans, freshwater fish.

Abstrakt

Rod *Pelasgus* je endemický pro jižní část Balkánu a je známo sedm druhů tohoto rodu. V této práci byl použit multilokusový přístup pro studium fylogenetických vztahů jednotlivých druhů a populací v rámci rodu *Pelasgus*. Dále bylo revidováno rozšíření jednotlivých druhů. Bylo analyzováno celkem 180 jedinců, pocházejících ze 47 lokalit náležejících k 30 povodím. Použitý materiál zcela pokrývá celou oblast rozšíření rodu. Navíc byl použit materiál z typové lokality každého druhu. Použity byly mitochondriální gen cytochrom b a nukleární markery první intron genu S7 (ribosomal protein gene), RAG1 (recombination activating gene) a rhodopsin. Analýzy cytochromu b, který byl z použitých markerů nejvariabilnější, odhalily existenci sedmi statisticky silně podpořených linií, které odpovídají těmto druhům: *P. laconicus*, *P. marathonicus*, *P. minutus*, *P. stymphalicus*, *P. thesproticus*, *P. prespensis* a *Pelasgus* sp. Nejvariabilnějším jaderným markerem byl první intron S7, jehož analýzy odhalily šest dobře podpořených linií, tedy téměř tolik, jako cytochrom b. Oproti tomu RAG1 a rhodopsin se ukázaly méně informativní, protože jejich analýzy odhalily existenci jen čtyř dobře podpořených linií. Tyto dva nukleární geny od sebe nedokázaly oddělit *P. marathonicus*, *P. stymphalicus*, *P. thesproticus* a *Pelasgus* sp. Tyto druhy mezi sebou sdílely stejné haplotypy a celková variabilita v rámci těchto druhů byla na obou genech velmi nízká.

Pelasgus sp. vyžaduje další výzkumy, neboť by to mohl být *P. epiroticus*. Avšak tato otázka může být vyřešena jednoznačně pouze zahrnutím jedinců z jezera Jánina do analýz. Tato populace je však pravděpodobně vyhynulá.

Díky molekulární identifikaci populací byl upřesněn areál výskytu několika druhů. Dále byla zjištěna pravděpodobná introdukce *Pelasgus* sp. na dvě lokality na Peloponésu (Kandila a jezero Stymphalia) a možná také na jednu lokalitu na pevnině, do jezera Zaravina v západním Řecku. Dále byl zjištěn sympatrický výskyt *P. stymphalicus* a *Pelasgus* sp. na obou výše zmíněných lokalitách na Peloponésu a byla odhalena hybridizace mezi oběma druhy.

Klíčová slova: Pelasgus, mitochondriální a jaderné markery, haplotypová síť, fylogeografie, jižní Balkán, sladkovodní ryby

Table of contents

1. Introduction 11

1.1 The Balkans as a biodiversity hotspot 11

1.2 The study area: Greece 11

1.2.1 Palaeogeography and development of ichthyofauna 11

1.2.2 Greek hydrological network and biodiversity 12

1.2.3 Greek ichthyofauna 13

1.3 The genus *Pelasgus* 14

1.3.1 Morphology 16

1.3.2 Distribution and ecology 16

Aims 19

2. Material and methods 20

2.1 Individuals and localities 20

2.2 Molecular methods 25

2.2.1 DNA extraction 25

2.2.2 DNA amplification 25

2.2.3 DNA purification 28

2.3 Phylogenetic analyses 29

3. Results 31

3.1 PCR amplification of Cyt b, S7, RAG and Rhodopsine 31

3.2 Mitochondrial and nuclear DNA phylogenetic analyses 31

4. Discussion 61

4.1 Distribution of Pelasgus species 62

4.2 Molecular identification of *Pelasgus* species: the incognita of *P. epiroticus*. 68

4.3 Stymphalia Lake, Kandila springs (Peloponnesus) and Zaravina Lake (mainland) as an example of introduction 69

5. Conclussions 71

6. References 73

7. Appendices 85

Appendix 1- Identification key to *Pelasgus* species. Taken and modified from Kottelat & Freyhof (2007a) 85

Appendix 2- Detailed information of the samples 86

Appendix 3 – Genomic DNA Mini Kit (Tissue) Protocol (Geneaid Biotech, Taiwan). Taken and modified from www.geneaid.com. 94

Appendix 4 – Details of haplotype network reconstruction for nuclear markers: visualization of alleles position in the network. 98

1. Introduction

1.1 The Balkans as a biodiversity hotspot

The Balkans belongs to the so-called hotspots of biodiversity and is a region known for a high degree of diversity and endemism of freshwater fish. The historical isolation of the Balkan Peninsula from the main continent and the recent post-glacial dispersal of the Danube fish fauna to the Balkan Peninsula are considered to have had a great influence on the composition of the Greek freshwater fish fauna (Economidis & Banarescu, 1991). As a part of the Balkans, Greece is one of the centers of biodiversity and endemism of freshwater fishes in Europe and one of the most important biogeographical regions (Zogaris et al., 2009; Oikonomou et al., 2014). However, especially the region of Western Greece appears to be so far insufficiently explored, in terms of taxonomy, phylogeography and phylogenetic relationships between populations of different species (e.g. Durand et al., 1999; Geiger et al., 2014).

As an evidence of still insufficient knowledge of freshwater fishes of Greece, several new species have been recently described, e.g. *Salaria economidisi* (Kottelat, 2004), *Phoxinus strymonicus* (Kottelat, 2007), *Squalius orpheus* (Kottelat & Economidis, 2006), *Pelasgus laconicus* (Kottelat & Barbieri, 2004), *Aphanius almiriensis* (Kottelat et al., 2007) or *Valencia robertae* (Freyhof et al., 2014). In addition, the results of several of the above-mentioned molecular genetic works suggest the existence of unknown species in the genera *Telestes* (Barbieri et al., 2015), *Barbus* (Marková et al., 2010), *Pelasgus* (Geiger et al., 2014), *Squalius* (Kottelat & Freyhof, 2007a) or *Alburnoides* (Stierandová et al., 2016).

1.2 The study area: Greece

1.2.1 Palaeogeography and development of ichthyofauna

In the development of Greece numerous climatic, geological and orogenic events played a role and gradually shaped this area, contributing to the high concentration of endemic species. Some of this events are the emergence of a land bridges in the early Oligocene and Miocene (Rögl, 1999), which connected the Balkans with Asia (Perea et al., 2010), the freshwater phase of Paratethys sea (Hsu et al., 1977), or the alpine orogeny, which culminated 11 MYA (million

years ago). All this events have gradually formed the Greek hydrological systems, leading to a long isolation of local freshwater ichthyofauna and subsequent speciation.

The beginning of the tectonic activity, volcanism and sea level changes have been estimated to have occurred during the Oligocene (Rögl, 1998), and the tectonic disturbances occurred in the Middle and Late-Miocene (12-8 MYA) initiated the development of the modern biogeographic history of this area (Poulakis et al., 2005). Some of the big paleo-geological events that authors have recognized as important in the evolutionary history of the freshwater ichthyofauna of Balkan Peninsula are the Messinian Salinity Crisis (MSC), which involved a great decline of the Mediterranean sea level, the subsequent filling with fresh water from Paratethys known as Lago Mare stage, and, according to some older works, the possible dispersion of freshwater fishes (Zardoya & Doadrio, 1999). However, more recent studies have demonstrated that older events most probably played even a more important role in dispersion and speciation of the freshwater fishes (Perea et al., 2010). Also the geologic isolation of Peloponnesus has been considered to play a role in the evolutionary history of organisms, but there are divergences on the timing: Sola et al. (2013) estimate the event of the isolation about 3-4 MYA ago, but Zardoya & Doadrio (1999) and Dubut et al. (2012) estimate it about 2.5 MYA. Pleistocene glaciations that might have allowed connections among land masses and ancient river beds are the last events proposed as involved in the shaping of the Greek fauna (Maurakis et al., 2001).

1.2.2 Greek hydrological network and biodiversity

Despite the fact that freshwater habitats cover a very small portion of the global surface (8% of the global surface, and only 0.01% of the world's water surface), about 10% of all animal species and 40% of fish species occur in freshwaters. This imbalance between the surface which it covers and the diversity that freshwater hosts, is what it is called the "paradox of freshwater biodiversity" (Martens, 2009).

Greek inland aquatic ecosystems are characterized by a rather stable geographical isolation, spatially and temporally speaking, by the existence of small river basin areas, and long-existing water bodies, respectively. Those water bodies acted as refugia during the Pleistocene glaciation and were the point of origin for subsequent dispersion (Economou et al., 2007).

Greek hydrological network consists of small rivers isolated from each other. In total, on the mainland and the neighboring islands of Greece, there are over 100 isolated hydrological systems, from large rivers (e.g. Marica, Vardar or Acheloos) and lakes (Prespa, Vegoritida or Pamvotis) to a high number of smaller streams (e.g. Sperchios, Alfios or Acheroon) and small wetland systems (Economou et al., 2007; Koutsikos et al., 2012).

Greek river basins usually comprise isolated hydrographic basins, characterized by short, steep fluvial systems that exhibit very erosive behavior, flashy irregular flow regimes and are influenced by varied geographical and climatic conditions. Most rivers run through narrow mountain valleys and descend abruptly to the coast, usually lacking extensive lowland sections and floodplains habitats (Economou et al. 2007).

1.2.3 Greek ichthyofauna

In terms of inland water biota, Greece is a crossroad between Mediterranean, temperate European Danubian-Black Sea, and Anatolian influences (Banarescu, 2004). Although the biotic influences due to the country's geography are unique, important biogeographic barriers cross the country and create even more heterogeneity.

Greek freshwater systems are inhabited by a very diverse ichthyofauna (160 species - Barbieri et al., 2015) characterized by the high number of endemic species. The fragmentation of aquatic habitats in the semi-arid and arid Mediterranean environments led to geographic isolation and speciation (allopatric speciation and vicariant events), resulting in the restriction of the species distribution to very small number of hydrographic basins or even to a single basin (Kottelat & Freyhof, 2007a). The average species richness in Greek waterbodies is low, which is explained, at least partly, by the small size of most basins (Economou et al. 2007).

This habitat fragmentation, the limited distribution of primary freshwater species and their low spreading potential from one watershed to another, makes them ideal as model organism for biogeographical studies. They are considered to be reliable mirrors of the biogeographical history of the river basins they inhabit (Zogaris et al., 2009).

Dominant component of Greek ichthyofauna are the members of the family Cyprinidae, which make up 52% of all native freshwater species (Barbieri et al., 2015), and represent a very significant endemism; not in vain the members of this family constitute 62% of all endemic fishes in Greece (Oikonomou et al., 2014). Recent phylogenetic and phylogeographic studies,

which included material from Greece, brought fundamental discoveries and confirmed the importance of this region as a center of biodiversity and biogeographically important area. Comprehensive phylogenetic studies of Cyprinidae in Greece were made within the broader scope of work (Zardoya & Doadrio, 1999; Perea et al., 2010; Geiger et al., 2014). These works have shown that some of Cyprinidae genera in this area are very old in evolutionary terms (*Pachychilon, Tropidophoxinellus* and *Pelasgus*).

The rate of endemism of Cyprinidae is the largest in southern and western Greece, endemic representatives belong to genera *Squalius*, *Barbus*, *Luciobarbus*, *Telestes*, *Pelasgus*, *Rutilus*, *Tropidophoxinellus* and *Scardinius*. Systematics and evolutionary history of the representatives of almost all of these genera from the territory of Greece remain unclear, even though some groups have been previously studied (e.g. Doadrio & Carmona, 1998; Durand et at., 1999; Gilles et al., 2010; Markova et al., 2010; Perea et al., 2010; Dubut et al., 2012; Geiger et al. 2014; Stierandova et al., 2015).

1.3 The genus *Pelasgus*

My study group is the genus *Pelasgus* (Teleostei, Cyprinidae), whose representatives are small fish endemic to the Southern Balkans. The genus is named after the Pelasgians, inhabitants of the lands around the Aegean Sea, predecessors of the proto-Greek people (Kottelat & Freyhof, 2007b).

Genus *Pelasgus* was first described in 2007 (Kottelat & Freyhof, 2007b) based on morphological characters. Previously, representatives of the genus, along with several other small cyprinid fishes from the Balkans and Turkey, were assigned into several different genera by different authors (*Pseudophoxinus, Paraphoxinus* or *Phoxinellus*) on the basis of their small size and morphological similarities (e.g. Bogutskaya & Zupančič, 2003).
Pelasgus species have been studied in detail morphologically (Kottelat & Barbieri 2004). Initially, only two species of *Pseudophoxinus* were recognized in Greece: *Pseudophoxinus epiroticus* and *Pseudophoxinus stymphalicus*, and Kottelat and Barbieri (2004) tentatively recognized two more species: *Pseudophoxinus marathonicus* and *Pseudophoxinus thesproticus*, grouping together the Balkan species of *Pseudophoxinus* and suggesting the existence of an

independent lineage within the genus. The further comparison of the Balkan group of *Pseudophoxinus* with the type species of related genera demonstrated that there were some morphological characteristics absent in the Balkan species of *Pseudophoxinus*, which instead showed more similarities to the genus *Delminichthys* (Kottelat & Frethoff, 2007). In 2006, Freyhof et al. (2006) published a work that showed that the southern Balkan *Pseudophoxinus* species form a monophyletic lineage, sister-group of *Delminichthys,* and supported their recognition as a distinct genus. Molecular genetic data provided in Zardoya et al. (1999), Perea et al. (2010) and Geiger et al. (2014) support the monophyly of the genus (see fig. 1).

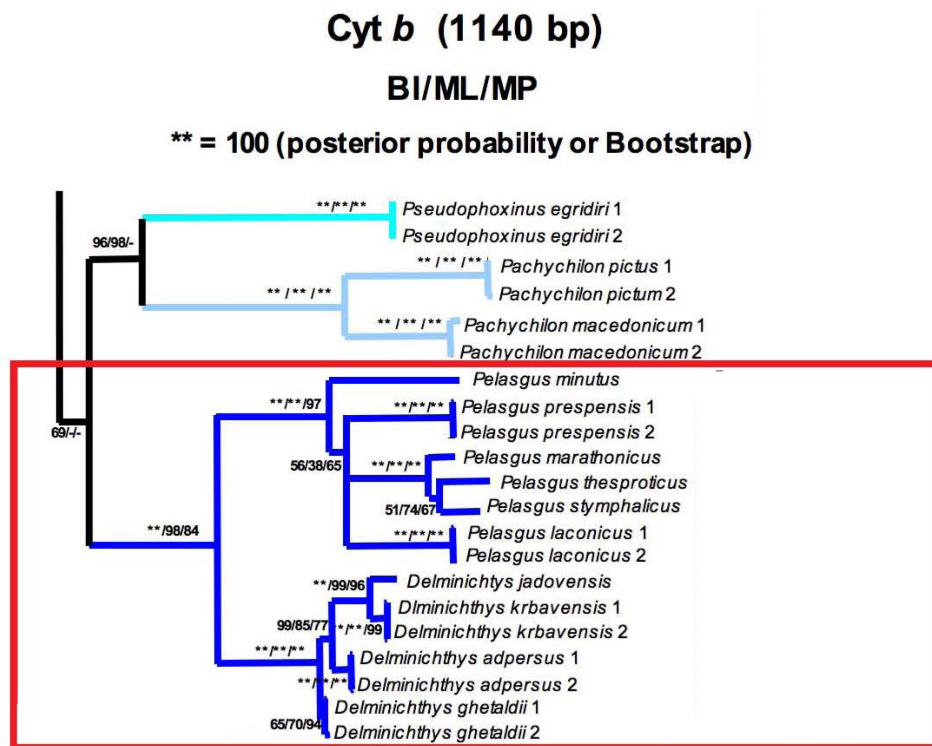

Fig. 1 – A part of the phylogenetic tree rendered by Bayesian analysis of the mitochondrial cytochrome b data set of Perea et al. (2010). Numbers above branches means posterior probabilities of BI/Bootstrap values of ML/Bootstrap values of MP (taken and modified from Perea et al., 2010).

The type species of the genus is *Pelasgus laconicus* (Kottletat & Barbieri, 2004) and the genus includes seven currently recognised species: *Pelasgus epiroticus* (Steindachner, 1896), *P. laconicus* (Kottetlat & Barbieri, 2004), *Pelasgus marathonicus* (Vinciguerra, 1921), *Pelasgus minutus* (Karaman, 1924), *Pelasgus stymphalicus* (Valenciennes, 1844), *Pelasgus thesproticus* (Stephanidis, 1939) and *Pelasgus prespensis* (Karaman, 1924).

Fig. 2 - *Pelasgus laconicus* (Evrotas, Greece); male, 42mm SL. Taken and modified from Kottelat & Freyhof (2007a).

1.3.1 Morphology

Species of *Pelasgus* are small sized fish (maximum size 130 mm SL), distinguished from all other genera of Eurasian and North African cyprinids (except *Delminichthys*) by a unique sexual dimorphism in the shape of the pectoral fin. It is distinguished from other genera by a unique combination of characters, which include the number of scales in the mid-lateral row, the shape of the head and snout, the number of scales on the caudal peduncle, the shape of the anal fin rays, the coloration and the scales disposition (Kottelat & Barbieri, 2004). Some of the morphological features used for distinguishing the different *Pelasgus* species between each other are: the number of scales in the lateral line (usually incomplete or missing), the shape and size of the scales, the shape of the pectoral fin and the characteristics of the caudal peduncle. For the complete key to species of *Pelasgus*, see Appendix 1 (Kottelat & Barbieri, 2004; Kottelat & Freyhof, 2007a, b).

1.3.2 Distribution and ecology

Pelasgus species are distributed in southern Greece, Albania, FYROM and Montenegro; species have non-overlapping occurrence area (Fig. 3; Kottelat & Freyhof, 2007a). In the Alfios River,

two species occur in the same river basin allopatrically, one of them in the upper course and the other in the lower course.

Fig. 3 - Distribution of *Pelasgus* species according to Kottelat & Freyhof (2007a). Map taken and modified from Google Maps.

Pelasgus species can be found in springs, small and shallow rills, backwaters and ditches, i.e. in narrow, shallow and slow flowing waters, and they are able to survive summer droughts in pools remaining in the river beds, in springs or in wells. They feed on a variety of aquatic invertebrates, algae and detritus (Kottelat & Freyhof, 2007a).

Not much is known about the biology and life cycle of *Pelasgus* species. *P. marathonicus* lives up to two years and spawns in May-September; *P. minutus* has been reported to live up to six years and probably matures at one year, with spawning period between June-July and deposition of eggs among algae or on stones; about *P. prespensis* only is know that it probably has an extended spawning season. *P. laconicus* reproduces between March and June, matures on the second year of life, and spawns its eggs on aquatic vegetation; *P. stymphalicus* lives up to three

years and spawns for the first time at one year, repeatedly between December and April, depositing the eggs on plants. Nothing is known about *P. epiroticus* neither about *P. thesproticus*.

Conservation status of *Pelasgus* species varies from least concern (LC, as the case of *P. stymphalicus*) to critically endangered (CE, as the case of *P. laconicus*), with *P. minutus* as data deficient (DD) (Kottelat & Barbieri, 2004; Kottelat & Freyhof, 2007a).

Aims

The aim of this project was to conduct a comprehensive phylogenetic and phylogeographic study of the endemic freshwater genus *Pelasgus* based on a multilocus approach. This study illustrates the evolutionary history of the species of the studied genus.

The main aims of the project are:

1. to study the phylogenetic relationships of the species belonging to the genus *Pelasgus* using a multilocus approach.
2. to determine the exact geographic distribution of each *Pelasgus* species in Greece.

2. Material and methods

2.1 Individuals and localities

I included in my analyses 180 individuals collected from 47 localities within 30 river basins, representing species of the genus *Pelasgus* occurring in Greece, populations of *P. thesproticus* and *P. minutus* occupying some localities of Albania and Macedonia, and populations of *P. prespensis* from the Prespa Lake basin (at the border between Macedonia, Albania and Greece) (Tab. 1, Figs 4 and 5).

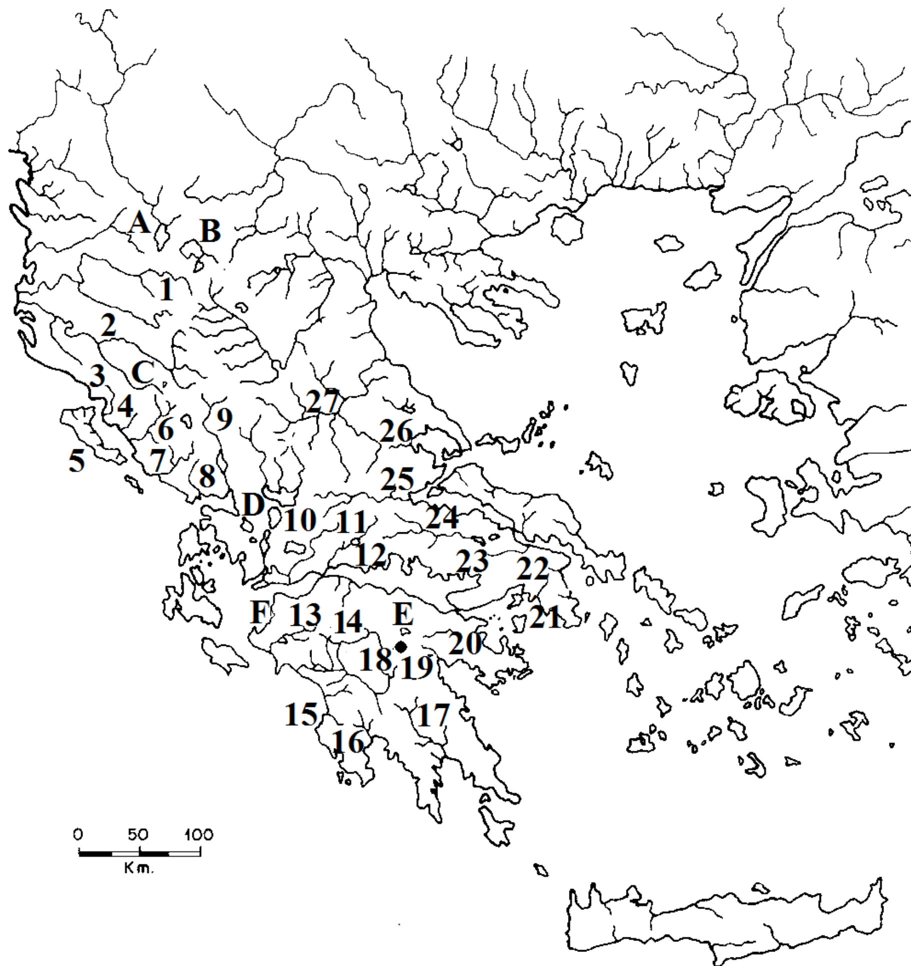

Fig. 4 - Map of Greece showing the main drainages sampled in this study. The river basins are: 1, Devoll; 2, Vjosa; 3, Bistrica; 4, Pavllo; 5, Messonghi; 6, Kalamas; 7, Acheroon; 8, Louros; 9, Arachthos; 10, Acheloos; 11, Evinos; 12, Mornos; 13, Pinios (Peloponnesus); 14, Alfios; 15, Peristeras; 16, Pamissos; 17, Evrotas; 18, Kandila springs; 19, Lerni spring; 20, Assopos (Peloponnesus); 21, Kifissos - Athens; 22, Vraonas; 23, Kato Souli (Marathon); 24, Kifissos - Beotia ; 25, Sperchios; 26, Xerias; 27, Pinios (Thessaly). The lakes are: A, Ohrid; B, Prespa; C, Zaravina; D, Trichonis; E, Stymphalia and F, Kotichi. Taken and modified from Zardoya et al., 1998.

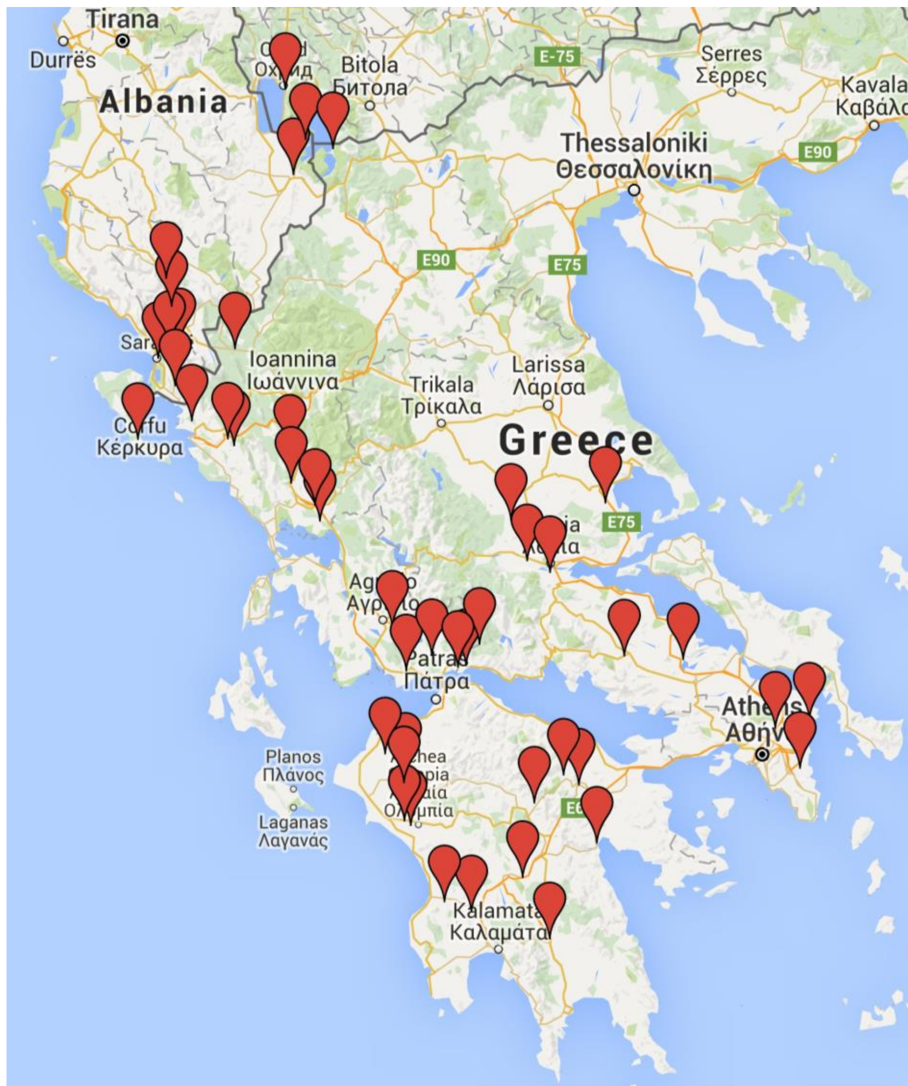

Fig. 5 - Map of the sampled localities (taken and modified from Google Maps).

From 2 to 13 individuals were analysed from each river basin (1-10 from each locality) for mitochondrial DNA and from 1 to 8 individuals for nuclear DNA (1-8 from each locality).

Table 1 - Sampled localities. The number of the river basin corresponds to the number in the Fig. 4.

Species according to Cyt b	Expected species	Locality	River basin		Coordinates
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Vergas stream	F	Kotichy Lake system	N 37°59'3.09" E 21°25'25.29"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Kalivakia, upstream of reservoir	13	Pinios – Pelloponesius	N 37°54'31.13" E 21°32'44.26"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	tributary of Pinios, Ag. Nicolaos			N 37°50'36.35" E 21°32'18.16"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	below Alfios dam, near Olympiada	14	Alfios	N 37°38'20.50" E 21°34'51.89"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Enipeas r., Salmoni			N 37°39'39.26" E 21°32'32.49"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	road to Artiki	15	Peristeras	N 37°16'9.45" E 21°47'6.37"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Meligalas	16	Pamissos	N 37°13'20.12" E 21°56'57.94"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Nafpaktos	12	Mornos	N 38°23'51.02" E 21°52'4.50"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Limnitsa			N 38°30'26.86" E 21°59'56.30"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Chiliadou spring			N 38°24'25.57" E 21°55'3.58"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Kokori	11	Evinos	N 38°22'46.11" E 21°33'13.45"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Vania			N 38°27'9.30" E 21°42'28.84"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	Trichonis Lake	10	Acheloos	N 38°35'20.19" E 21°28'2.68"
Pelasgus stymphalicus	*Pelasgus stymphalicus*		19	Milos – Lerni spring	N 37°33'10.08" E 22°43'2.89"
Pelasgus stymphalicus	*Pelasgus stymphalicus*		20	Assopos – Peloponnesius	N 37°50'4.86" E 22°36'21.33"
Pelasgus stymphalicus	*Pelasgus stymphalicus*	inflow of Stymphalia lake	E	Stymphalia lake	N 37°52'53.13" E 22°30'48.17"
Pelasgus thesproticus	*Pelasgus thesproticus*	Kokitos river, upper Acheroon	7	Acheroon	N 39°26'49.13" E 20°30'3.36"
Pelasgus thesproticus	*Pelasgus thesproticus*	reservoir Chaskovo, near Paramythia			N 39°28'45.74" E 20°27'41.72"
Pelasgus thesproticus	*Pelasgus thesproticus*	river Messonghi	5	Kerkyra island	N 39°28'36.65" E 19°54'42.68"
Pelasgus thesproticus	*Pelasgus thesproticus*	Anakoli spring	6	Kalamas	N 39°33'46.63" E 20°14'28.73"
Pelasgus thesproticus	*Pelasgus thesproticus*	Mesopotam	3	Bistrica	N 39°54'29.23" E 20°6'14.61"
Pelasgus thesproticus	*Pelasgus thesproticus*	Bistrica river			N 39°55'11.89" E 20°10'7.11"
Pelasgus thesproticus	*Pelasgus thesproticus*	Cuka			N 39°51'24.3" E 20°02'10.7"
Pelasgus thesproticus	*Pelasgus thesproticus*	Hormove, Drinos river	2	Vjosa basin	N 40°13'47.20" E 20° 5'7.60"

Pelasgus thesproticus	*Pelasgus thesproticus*	Burimi i Viroit spring, Drinos river			N 40° 5'53.13" E 20° 7'10.61"
Pelasgus thesproticus	*Pelasgus thesproticus*	Shalles	4	Pavllo	N 39°43'42.24" E 20° 8'30.00"
Pelasgus minutus	*Pelasgus minutus*	Ohrid Lake	A	Ohrid Lake	N 41° 6'8.23" E 20°48'50.69"
Pelasgus prespensis	*Pelasgus prespensis*	Beli Hill	B	Prespa lake	N 40°52'20.50" E 20°56'15.60"
Pelasgus prespensis	*Pelasgus prespensis*	Agios Germanos			N 40°49'46.48" E 21° 6'17.04"
Pelasgus prespensis	*Pelasgus prespensis*	Golo Brdo	1	Devoll river	N 40°42'37.45" E 20°51'49.21"
Pelasgus sp.	***Pelasgus epiroticus (?)***	Louros, springs, Terovo area	8	Louros	N 39°25'15.54" E 20°50'29.90"
Pelasgus sp.	***Pelasgus thesproticus***	Agios Georgios			N 39°16'14.56" E 20°50'55.31"
Pelasgus sp.	***Pelasgus thesproticus***		C	Zaravina lake	N39°54'03.0" E 20°30'24.6"
Pelasgus sp.	***Pelasgus sp.***		18	Kandila springs	N 37°44'45.21" E 22°20'2.79"
Pelasgus stymphalicus	***Pelasgus sp.***				
Pelasgus sp.	***Pelasgus stymphalicus***	inflow	E	Stymphalia lake	N 37°52'53.13" E 22°30'48.17"
Pelasgus sp.	***Pelasgus thesproticus***	Acropotamia	9	Arachthos	N 39° 5'29.64" E 21° 1'34.84"
Pelasgus sp.	***Pelasgus thesproticus***	Arta			N 39° 9'56.45" E 20°59'38.49"
Pelasgus marathonicus	*Pelasgus marathonicus*		25	Sperchios	N 38°54'14.33" E 22°17'30.22"
Pelasgus marathonicus	*Pelasgus marathonicus*	chanell near Lamia			N 38°50'54.60" E 22°25'54.46"
Pelasgus marathonicus	*Pelasgus marathonicus*		26	Xerias	N 39°10'40.44" E 22°46'11.24"
Pelasgus marathonicus	*Pelasgus marathonicus*	Akrefnio	24	Kifissos – Beotia	N 38°26'1.06" E 23°14'34.92"
Pelasgus marathonicus	*Pelasgus marathonicus*				N 38°27'2.12" E 22°53'3.02"
Pelasgus marathonicus	*Pelasgus marathonicus*		21	Kifissos – Athens	N 38° 6'26.76" E 23°48'25.53"
Pelasgus marathonicus	*Pelasgus marathonicus*	Kato Souli wetland	23	Kato Souli – Marathon	N 38° 8'56.76" E 24°0'59.44"
Pelasgus marathonicus	*Pelasgus marathonicus*		22	Vraonas	N 37°54'52.63" E 23°57'28.64"
Pelasgus marathonicus	*Pelasgus marathonicus*	former Xiniada lake basin	27	Pinios-Thessaly	N 39° 5'44.38" E 22°11'37.33"
Pelasgus laconicus	*Pelasgus laconicus*	Sparta	17	Evrotas	N 37° 5'28.71" E 22°25'38.74"
Pelasgus laconicus	*Pelasgus laconicus*	Paparis stream	14	Alfios – upper part	N 37°23'14.15" E22°15'48.73"

More detailed information about the samples is given in the appendix 2.

Fish were collected during fieldwork campaigns in 2009, 2011, 2012, 2013, 2014 and 2015 by electrofishing and samples for molecular analyses were preserved in ethanol. A piece of tissue from the right paired fins was used for DNA extraction, and in the case of too small individuals where fin tissue is not sufficient for DNA extraction, muscle tissue from the right side of the body was used (left side is left untouched for the morphological studies).

As the closest relative of *Pelasgus*, species of the genus *Delminichthys* were used as outgroup for phylogenetic analyses (Perea et al. 2010). Sequences for the outgroup *Delminichthys* were obtained from GenBank with the exception of Rhodopsin sequence (DEMA 8), which was provided by Dr. Ivana Buj from Zagreb University, due to the insufficient length of the sequences available in GenBank (table 2)

Table 2 - GenBank accession numbers of sequences of *Delminichthys* species used as outgroups.

Species	Gene	GenBank accession number
Delminichthys ghetaldii	Cyt b	HM560091.1
Delminichthys adspersus	Cyt b	HM560089.1
Delminichthys jadovensis	Cyt b	AY838924.1
Delminichthys krbavensis	Cyt b	HM560087.1
Delminichthys ghetaldii	S7	HM560491.1
Delminichthys ghetaldii	RAG	HM560404.1
Delminichthys adspersus	RAG	HM560403.1
Delminichthys krbavensis	RAG	HM560402.1

The sequences of cytochrome b of species of *Pelasgus* downloaded from GenBank are in Table 3.

Table 3 - GenBank accession numbers of *Pelasgus* sequences

Species	Gene	GenBank accession number
Pelasgus stymphalicus	Cyt b	AF090766
Pelasgus stymphalicus	Cyt b	AF090767
Pelasgus stymphalicus	Cyt b	HM560109
Pelasgus prespensis	Cyt b	AF090763
Pelasgus prespensis	Cyt b	HM560108
Pelasgus sp.	Cyt b	AF090769
Pelasgus marathonicus	Cyt b	AF090768

2.2 Molecular methods

2.2.1 DNA extraction

The extraction of DNA was carried out using *Geneaid™ DNA Isolation Kit (Tissue)* (Geneaid Biotech, Taiwan). Manufacturer instructions were followed with minor changes to optimize the results. Complete protocol and changes are available in Appendix 3. The DNA isolates were stored at -20 °C.

2.2.2 DNA amplification

A total of four genes were selected: Cytochrome b gene (Cyt b) as mitochondrial marker, and first intron of Ribosomal Protein gene S7 (S7), Recombination Activating Gene (RAG1) and Rhodopsin gene as nuclear markers.

Cytochrome b (cyt b) is a protein found in the mitochondria of eukaryotic cells as part of the electron transport chain and as part of respiratory chain. Cyt b gene has a length of 1140 bp and is commonly used for taxonomic and population genetic studies and to determine phylogenetic relationships between organisms due to its sequence variability (adequate mutation rate and large number of parsimony informative sites with a low consistency index) (Perea et al., 2010; Buj et at., 2014). It has been proven to reliably reflect the phylogenetic relationships between the fish species in numerous previous investigations (e.g. Bohlen et al., 2006; Perea et al. 2010; Yang et al. 2015).

All of the selected nuclear markers have been proven to be useful in molecular phylogenetics in different studies and specially RAG1 and S7 genes provide fundamental phylogenetic information (Buffalino et al. 2010) The first intron of Ribosomal Protein gene S7 is a complex marker whose use is being extended in the last years (Buffalino et al. 2009; Perea et al., 2010).

The Recombination Activating Gene (RAG) encodes components of the recombinase involved in recombination of immunoglobin and T-cell receptor genes and appears as conserved single copies in vertebrates (Hansen and Kaattari, 1996; Willett et al, 1997). The RAG gene has been widely used to evaluate intrageneric and intraspecific relationships (Baker et al., 2000; Lewis-Oritt et al., 2001; Lovejoy and Collette, 2001; Hardman, 2004) and it is also used to reveal higher-level phylogenetic relationships (Hoegg et al. 2004; Calcagnotto et al. 2005).

Rhodopsin gene is a nuclear visual pigment coding gene, member of the opsin gene family (Chen et al., 2003), often used in phylogenetic studies (e.g. Gvoždík et al., 2010; Picq et al., 2014; Galván-Quesada et al., 2016; Neves et al., 2016).

Polymerase chain reaction (PCR) was performed in 25 µl volume, for reaction information see the tables 4 to 7. Cyt b Gene was amplified using primers GluF and ThrR (Machordom & Doadrio, 2001). The specific internal primers CB4-Glu and PhoxThr were used in combination with external primers for amplification in the case of bad quality PCR products or no PCR products at all obtained with the external primers. S7 gene was amplified with the primers S7RPEX1F and S7RPEX2R (Chow and Hazama, 1998), from now on referred to as S7F and S7R. The specific internal primers PlgS7F and PlgS7R were used in combination with external primers in the case of bad quality PCR products or no PCR products at all with the external primers. RAG region was amplified using the primers RAG1F and RAG9R (Quenouille et al., 2004). Rhodopsin gene was amplified using the primers Rho193F and Rho1073R (Chen et al., 2003). PCR was performed in a thermocycler *GeneTouch* (Hangzhou bori Technology Co., BIOER). For the amplification, PPP Master Mix (Top-Bio, Czech Republic) was used (details in Table 4).

Table 4 – Composition of PCR master mix for amplification of Cyt b, RAG, S7 and Rhodopsin

Chemicals	**1 PCR reaction**
PPP Master Mix*	12.5 µl
PCR Ultra H_2O	9.2 µl
Forward primer (Table ?)	0.65 µl
Reverse primer (Table ?)	0.65 µl
Final volume	23 µl
+ **Template**	**2 µl**

* $MgCl_2$ was already added in the PPP Master Mix.

Table 5 - PCR program for the Cyt b amplification

Reaction	Temperature	Time	
Initial denaturation	94 °C	3 min	
Denaturation	**94 °C**	**45 s**	x 35
Annealing	**48 °C**	**1 min 30 s**	
Elongation	**78 °C**	**1 min 45 s**	
Final Elongation	72 °C	7 min	
Hold	4 °C		

Table 6 - PCR program for the S7 and RAG amplification

Reaction	Temperature	Time	
Initial denaturation	95 °C	5 min	
Denaturation	**94 °C**	**40 s**	x 5
Annealing	**60 °C**	**1 min**	
Elongation	**72 °C**	**2 min**	
Denaturation	**95 °C**	**30 s**	x 35
Annealing	**56 °C**	**1 min**	
Elongation	**72 °C**	**2 min**	
Final Elongation	72 °C	20 min	
Hold	4 °C		

Table 7 - PCR program for the rhodopsin amplification

Reaction	Temperature	Time	
Initial denaturation	92 °C	5 min	
Denaturation	**92 °C**	**1 min**	x 35
Annealing	**54 °C**	**1 min**	
Elongation	**72 °C**	**1 min 30 s**	
Final Elongation	72 °C	7 min	
Hold	4 °C		

Table 8 – Sequences of primers used for PCR and sequencing

Primer	Direction	Sequence 5' → 3'	Reaction
GluF	forward	AAC CAC CGT TGT ATT CAA CTA CAA	PCR, sequencing
ThrR	reverse	ACC TCC GAT CTT CGG ATT ACA AGA CCG	PCR, sequencing
CB4-Glu	forward	CCT GAA AYA TYG GYG TRG T	PCR, sequencing
PhoxThr	reverse	AGG AGG AAR TGR AAT GCG AA	PCR, sequencing
S7F	forward	TGG CCT CTT CCT TGG CCG TC	PCR, sequencing
S7R	reverse	AC TCG TCT GGC TTT TCG CC	PCR, sequencing
PlgS7F	forward	AGC YGG TGA ATT ACT TAG AWT CT	PCR, sequencing
PlgS7R	reverse	ACA GAT GCA AGT CMC GGG TT	PCR, sequencing
RAG1F	forward	AGC TGT AGT CAG TAY CAC AAR ATG	PCR
RAG9R	reverse	GTG TAG AGC CAG TGR TGY TT	PCR
RAG3F	forward	GGG AGA TGT CAG CGA GAA GCA	sequencing
RAG11F	forward	GCC TGT GAT GAG GAT GAA TGG	sequencing
RAG6R	reverse	ATG GCT TTC CGC TCT GCT AC	sequencing
Rho193F	forward	CNT ATG AAT AYC CTC AGT ACT ACC	PCR, sequencing
Rho1073R	reverse	CCR CAG CAC ARC GTG GTG ATC ATG	PCR, sequencing

The quality of the extracted genomic DNA and the amplified DNA after the PCR was checked via gel electrophoresis (150 V, 30 min) on a 1.5 % agarose gel in 50 ml of TBE buffer, with the use of the fluorescent nucleic acid gel stain GelRed for the visualization. 2 µl of PCR product were used for the electrophoresis. GeneRuler 100bp Plus DNA Ladder (ThermoFisher Scientific, USA) was used to estimate the size of the fragments. The concentration of purified DNA was checked with the use of Nanodrop spectrophotometer (Thermo Fisher Scientific, USA). PCR products were stored at 4 °C.

2.2.3 DNA purification

For purification of the PCR products, the ethanol precipitation method was applied: two volumes of 99.9% ethanol previously stored in freezer were added to the mixture of PCR product (23µl of PCR product, 77µl of water) and 1/10 of volume of sodium acetate (10 µl), and mixed thoroughly by vortexing and subsequently centrifuged for 10 min at 14 000 rpm at room

temperature. After centrifugation, supernatant was removed. In the next step, 70% ethanol (previously stored in freezer) was added and again mixed by vortexing and centrifuged for 10 min at 14 000 rpm. The supernatant was removed, the content of the tube was dried at 40 °C and 20µl of Millipore water was added.

The sequencing reactions were performed by Macrogen Inc. (South Korea) with an Applied Biosystems 3730XL DNA Analyzer. The primers used for sequencing were GluF, ThrR, CB4-Glu and PhoxThr for Cytochrome b, S7F, S7R, PlgS7F and PlgS7R for S7, RAG3F, RAG11F and RAG6R for RAG, and Rho193F and Rho1073R for Rhodopsin Gene (Table 8).

2.3 Phylogenetic analyses

The Cyt b DNA sequences were assembled and edited in the program *Geneious 7.0.6* (http://www.geneious.com, Kearse et al., 2012). Obtained RAG, S7 and rhodopsin DNA sequences were aligned manually and visually checked with the programs *Chromas version 2.4.4* and *BioEdit* (Hall, 1999).

The best-fit nucleotide substitution model was estimated with *jModelTest version 2.1.7* (Darriba et al., 2011). Model selection was conducted on the basis of Akaike information criterion (AIC) and Bayesian information criterion (BIC). Two types of phylogenetic analyses were used for both nuclear and mitochondrial markers: Bayesian Inference (BI) and Maximum Likelihood (ML). Bayesian analysis was conducted in *MrBayes version 3.2.6* (Ronquist et al., 2012). Markov Chain Monte Carlo (MCMC) sampling was run for two million generations (two runs, each consisting of four chains) with trees sampled every 100 generations. The first 5000 generations (25%) were discarded as burn-in and the remaining trees were used to construct a 50 % majority-rule consensus tree. I checked that standard deviation of split frequencies was < 0.01. Maximum Likelihood (ML) analysis was conducted with *MEGA6*. BioNJ method (an improved version of Neighbor-Joining algorithm, Gascuel, 1997) was used to compute an initial tree. Two thousand bootstrap replicates were completed.

Indels were present in S7 sequences of *P. prespensis*, so obtained sequences were not as long as those of the rest of the species. In order to include all species of *Pelasgus* in S7 analyses, two

parallel datasets were used for phylogenetic analyses: one with shorter sequences including *P. prespensis,* and another without *P. prespensis* sequences but with longer sequences.

MEGA6 was used to calculate intra- (*Within Group Mean Distance*) and interspecific (*Between Group Mean Distance*) distances within and between *Pelasgus* species (p-distances).

All computed trees were visualized using *FigTree version 1.4.2* (Rambaut, 2012) and modified with InkScape 0.91

DnaSP v5.10.1 (http://www.ub.edu/dnasp/, Rozas et al., 2010) was used for sorting haplotypes of mitochondrial and nuclear sequences and to unphase heterozygote positions in nuclear markers. In the case of well supported heterozygote positions, two different sequences corresponding to the two possible alleles were separated and included in the alignments as independent sequences. In the case of not enough support (phase probabilities lower than 85%), sequences were discarded. Nineteen individuals were discarded for S7, eight for RAG and two for Rhodopsin. Phased data were used for the sorting of haplotypes and constructing of phylogenetic trees and haplotype networks. For Bayesian and Maximum Likelihood analyses gaps in S7 alignments were treated as missing data.

Haplotype networks for nuclear and mitochondrial markers were constructed with *TCS* (Clement et al., 2000). Files were converted to PHYLIP format with *ALTER* (Alignment Transformation Environment; Glez-Peña et al., 2010). For haplotype network gaps in S7 alignments were treated as the fifth character. The network estimation implemented in TCS is Statistical Parsimony (Templeton et al., 1992). It was used under a 95% connection limit. Visualization and editing of the haplotype networks was done in *tcsBU* (TCS Beautifier; Santos et al. 2015), a web-based program that extends the capabilities of TCS.

3. Results

3.1 PCR amplification of Cyt b, S7, RAG and Rhodopsine

PCR products of the amplified genes displayed sizes from 500 to 2000 base pairs (bp) Out of a total of 4214 characters per individual obtained, 1100 bp corresponded to mitochondrial DNA and 3114 bp to nuclear DNA. The aligned sequences obtained from the sampled individuals ranged in size from 1473 bp for RAG gen to 495 bp for S7 gene, see Tab. 1.

Table 9 - Lengths of sequences for each individual gene.

Gene	Number of sequences	Length of sequences (bp)
Cytochrome b	182	1100
Ribosomal Protein Gene S7 (S7)	129 / 126	495 / 785*
Recombination Activating Gene (RAG1)	122	1473
Rhodopsin	126	856
Nuclear DNA	377 / 374	2824 / 3114*
Mitochondrial DNA	182	1100
TOTAL	559 /556	3924 / 4214*

* two datasets with different sequence length were analyzed, see methods.

3.2 Mitochondrial and nuclear DNA phylogenetic analyses

The cytochrome b dataset used in this study contains a total of 182 sequences, 55 of them correspond to *Pelasgus stymphalicus*, 36 to *P. thesproticus*, 41 to *P. marathonicus*, 10 to *P. minutus*, 5 to *P. prespensis*, 8 corresponding to *P. laconicus* and 24 tentatively named *Pelasgus* sp. (with 780 sites without polymorphism and 290 parsimony informative sites). Seventy seven unique haplotypes were identified.

GTR+G model (General Time Reversible model + Gamma distributed rates) was given as best model for all analysed markers and was used in all subsequent analyses.

The Bayesian inference and Maximum Likelihood phylogenetic analysis confirmed the existence of seven well supported lineages within the genus *Pelasgus* (see Figs 6 and 7), all of them with

high statistical support (1 posterior probability in Bayesian inference and 99% bootstrap values for Maximum Likelihood). The topology of the trees inferred by both methods are the same, with *P. minutus* placed in a sister position to all remaining clades, followed by *P. prespensis* and *P. laconicus*. Three big sister clades are identified in phylogenetic trees, placed in polytomy, corresponding to *Pelasgus* sp., *P. thesproticus*, and *P. stymphalicus* together with *P. marathonicus* as mutual sister groups as well.

P. marathonicus and *P. stymphalicus* show numerous haplotypes. In the case of *P. stymphalicus* Bayesian analysis distinguished two subgroups, one of which well supported and corresponding to populations from mainland and Northwestern Peloponnesus, and the other corresponding to the rest of Peloponnesus populations (see also haplotype network of P. stymphalicus, fig. 12). This subdivision is not is not well supported by ML.

Pelasgus sp. is divided in two well supported subgroups, that correspond to mainland populations (subgroup with 99% bootstrap support and 1 posterior probability), and Peloponnesus populations together with Zaravina Lake in the mainland (99% bootstrap and 1 posterior probability). This is also displayed in the haplotype network reconstruction, with two independent groups that fit with the tree structure (see figures 6 to 8). The only exception is the haplotype 60 from Zaravina Lake, which is very distant from all other haplotypes and in the trees, does not cluster with other haplotypes from the subgroup.

Within *P. thesproticus* there is a well-supported subdivision in two subgroups, that correspond to the population from Kerkyra Island (haplotypes 5 to 7), and the mainland populations.

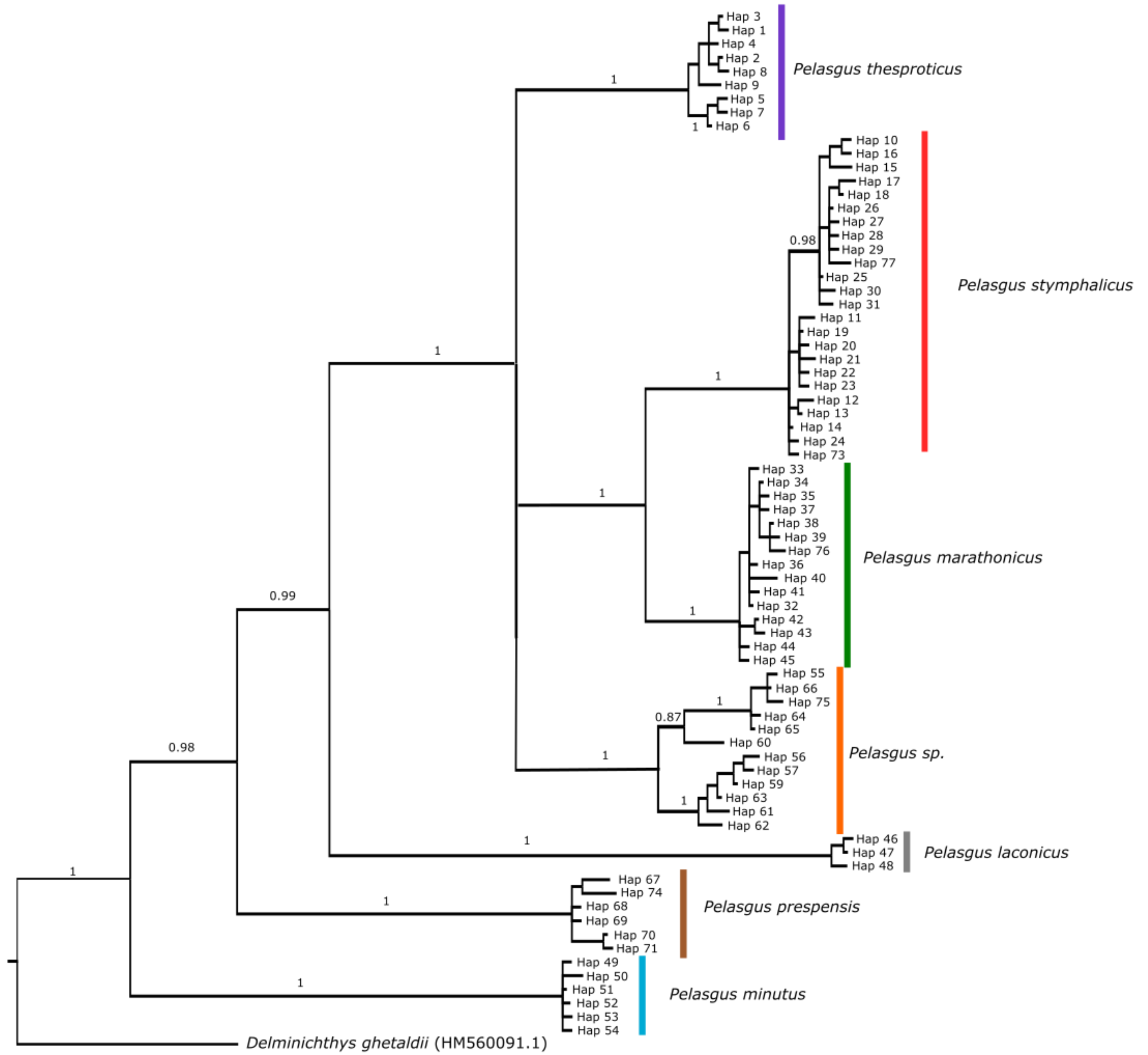

Fig. 6 - Phylogenetic tree for Cyt b rendered by Bayesian analysis. The numbers in the branches represent posterior probability. Only values higher than 0,87 are shown.

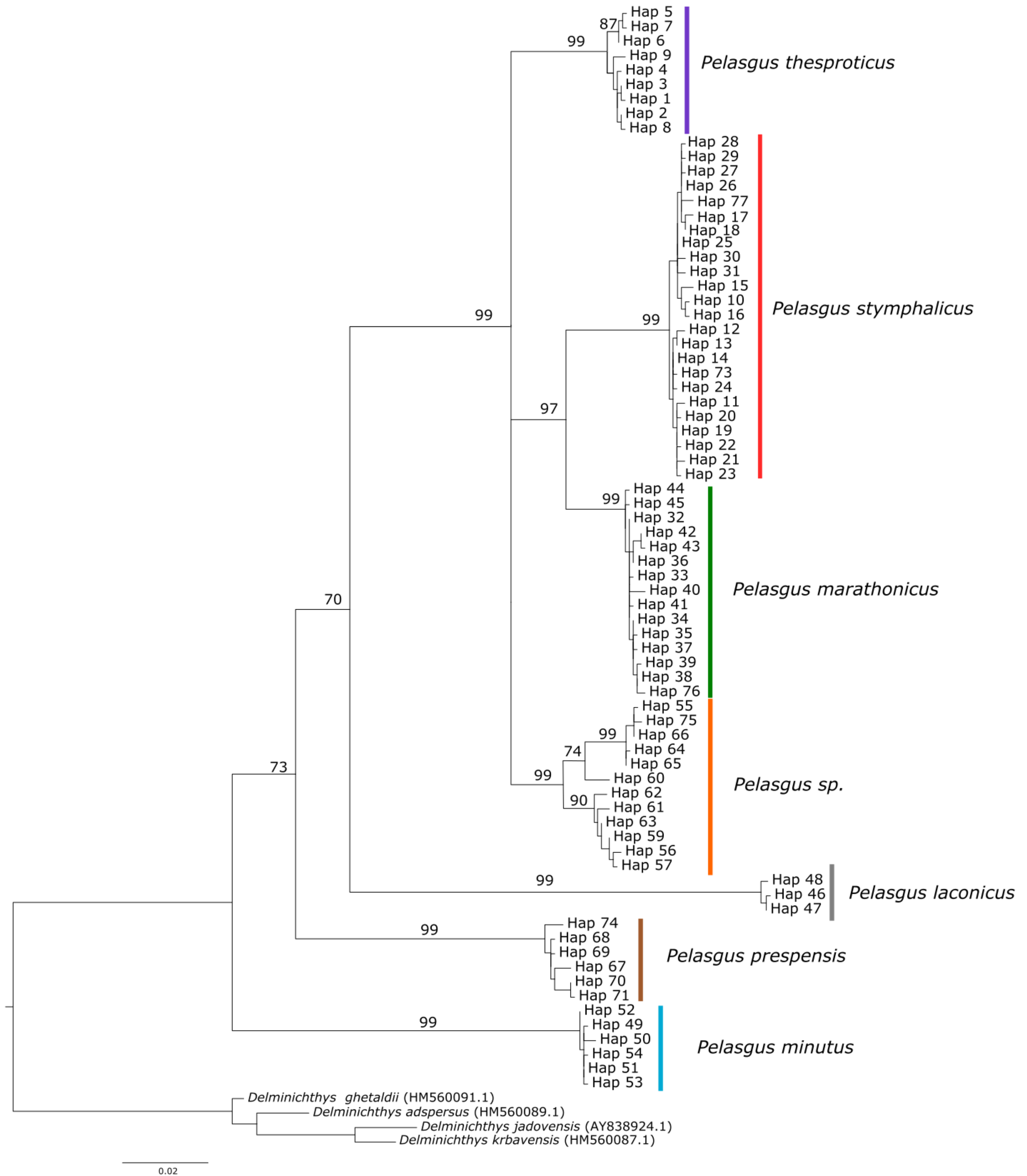

Fig. 7 - Phylogenetic tree for Cytochrome b from Maximum Likelihood. The numbers on the branches represent bootstrap values. Only values higher than 70% have been represented.

Assignation to species according to Cyt b sequences is not completely congruent with the expected occurrence of species according to distribution areas described by Kottelat & Freyhof (2007a). Extension of the distribution areas for *P. marathonicus* (Xerias River), *P. thesproticus* (Vjosa river basin) and *P. prespensis* (Devoll River), was revealed. The results suggest the occurrence in sympatry of two species (*P. stymphalicus* and *Pelasgus sp.*) in two localities in the Peloponnesus (Kandila springs and Stymphalia Lake). See table 1 in Methods section. For more details, see Appendix 2.

For cyt b, only the haplotype network of *Pelasgus* sp., *P. marathonicus*, *P. thesproticus* and *P. stymphalicus* was reconstructed since those were the most genetically diverse species, according to Bayesian and Maximum Likelihood analyses and have greatest distribution areas.

Pelasgus sp. haplotype network appears to be divided in two very distinct groups. These two groups correspond with the mainland (group 1, Louros and Arachthos populations) and the Peloponnesus populations (Kandila and Stymphalia Lake), with the exception of Zaravina Lake, that although is a locality from the mainland, lines up with Peloponnesus localities. Circle sizes in all haplotype network representations correlate to haplotype frequencies.

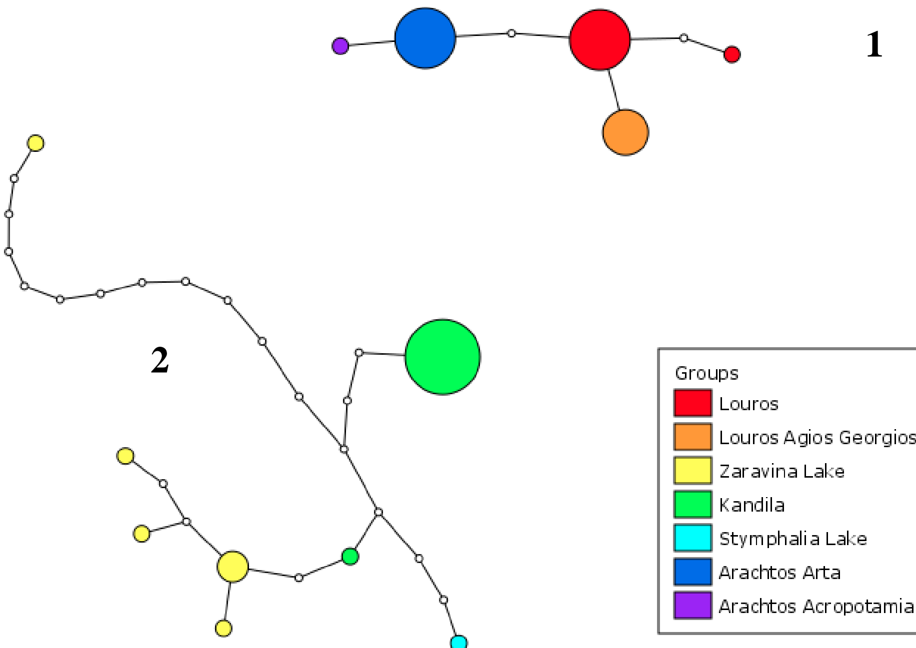

Fig. 8 - *Pelasgus sp*. haplotype network for Cyt b. Modified with tcsBU (TCS Beautifier).

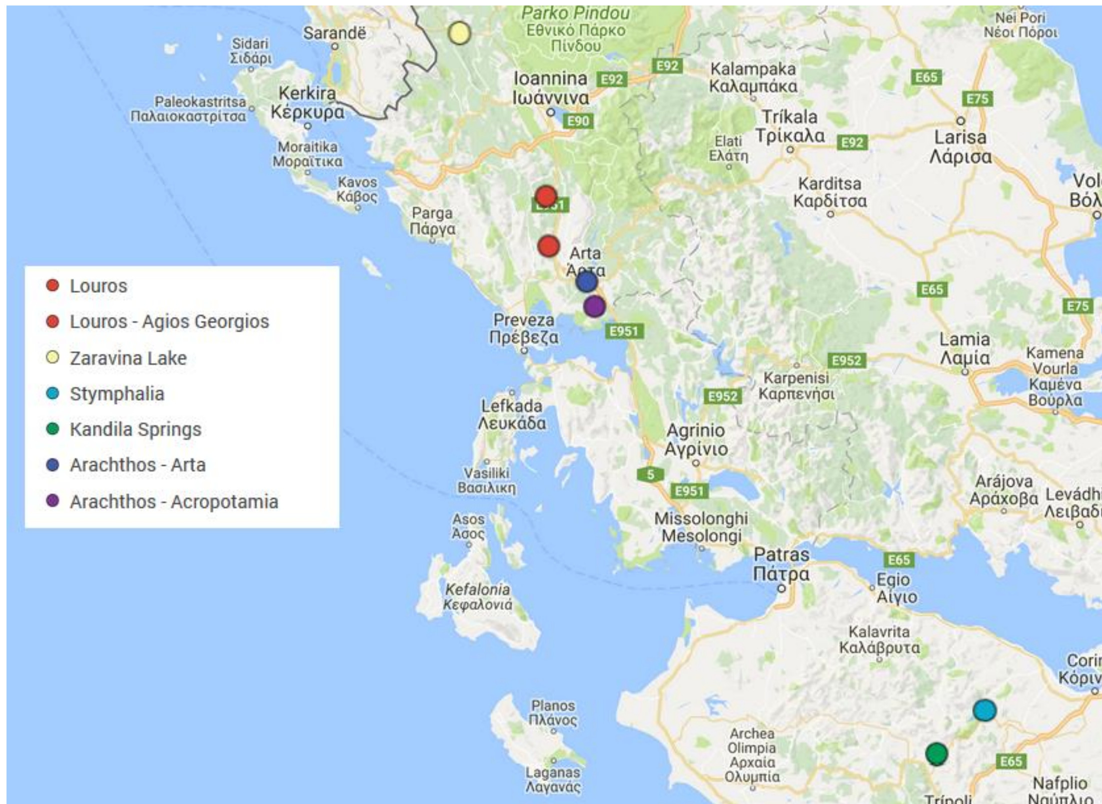

Fig. 9 - *Pelasgus sp.* Cyt b haplotype distribution. Taken and modified from Google Maps.

Even though in the phylogenetic trees P. *marathonicus* appears to be a very diverse group without a clear inner structure, haplotypes network reconstruction shows separated groups of haplotypes. Thirteen out of fifteen haplotypes are private, only two are shared between different localities. Kifissos River in Beotia shows the biggest haplotype diversity, with five different haplotypes. It is noteworthy the big separation between the two individuals from the locality of Sperchios River near Lamia (five mutations). There is not a clearly distinguishable geographical pattern in the haplotype network display. However, specimens from several drainages have specific haplotypes (e.g. Xerias, Vraonas or Kato Souli).

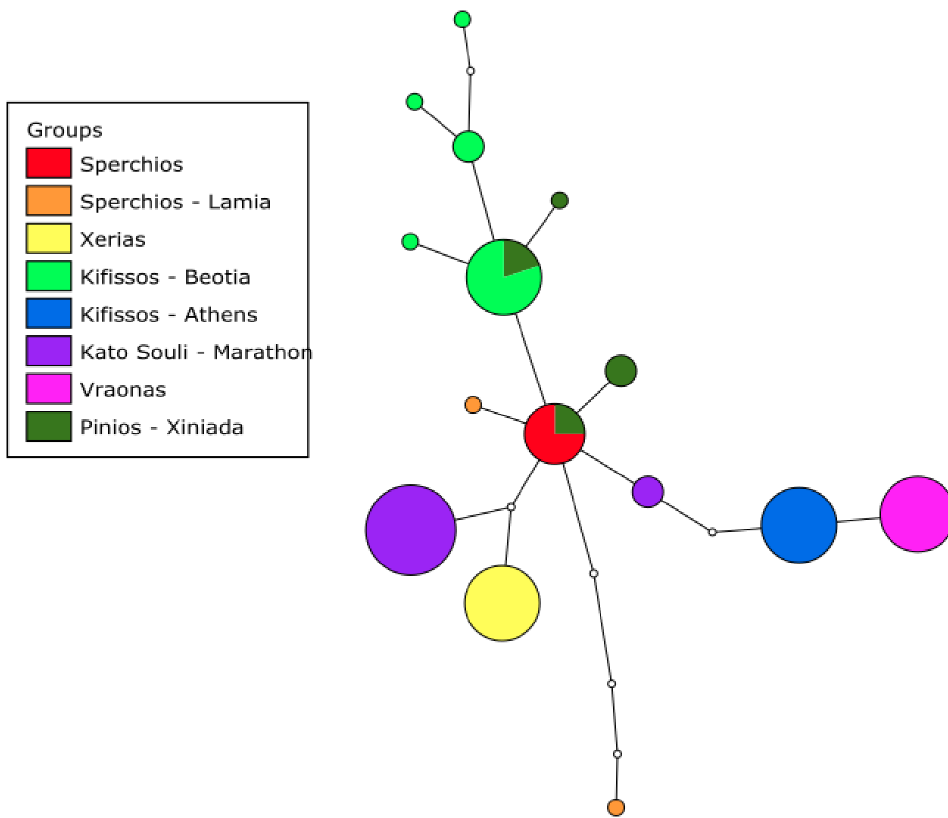

Fig. 10 - *Pelasgus marathonicus* haplotype network for Cyt b. Modified with tcsBU (TCS Beautifier).

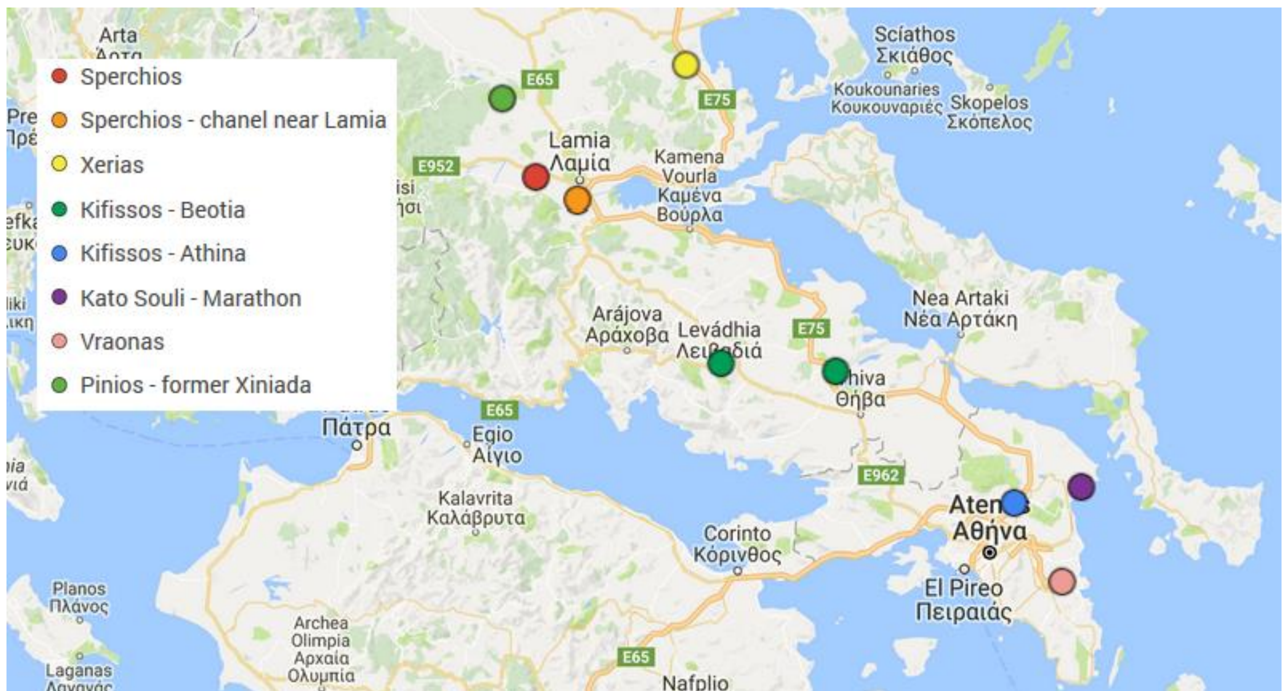

Fig. 11 - *Pelasgus marathonicus* Cyt b haplotype distribution. Map taken and modified from Google Maps.

P. stymphalicus haplotype network shows a division in two big groups, geographically the first one corresponding to the mainland populations (Mornos, Evinos and Acheloos) and the Northwestern populations in Peloponnesus peninsula (Kotichy Lake system, Pinios, and Alfios), and the second corresponding to the rest of Peloponnesus populations (Peristeras, Pamissos, Milos, Assopos, Kandila and Stymphalia Lake. It is important to note the presence of very divergent haplotypes within Alfios river basin (marked with arrows in the haplotype network).

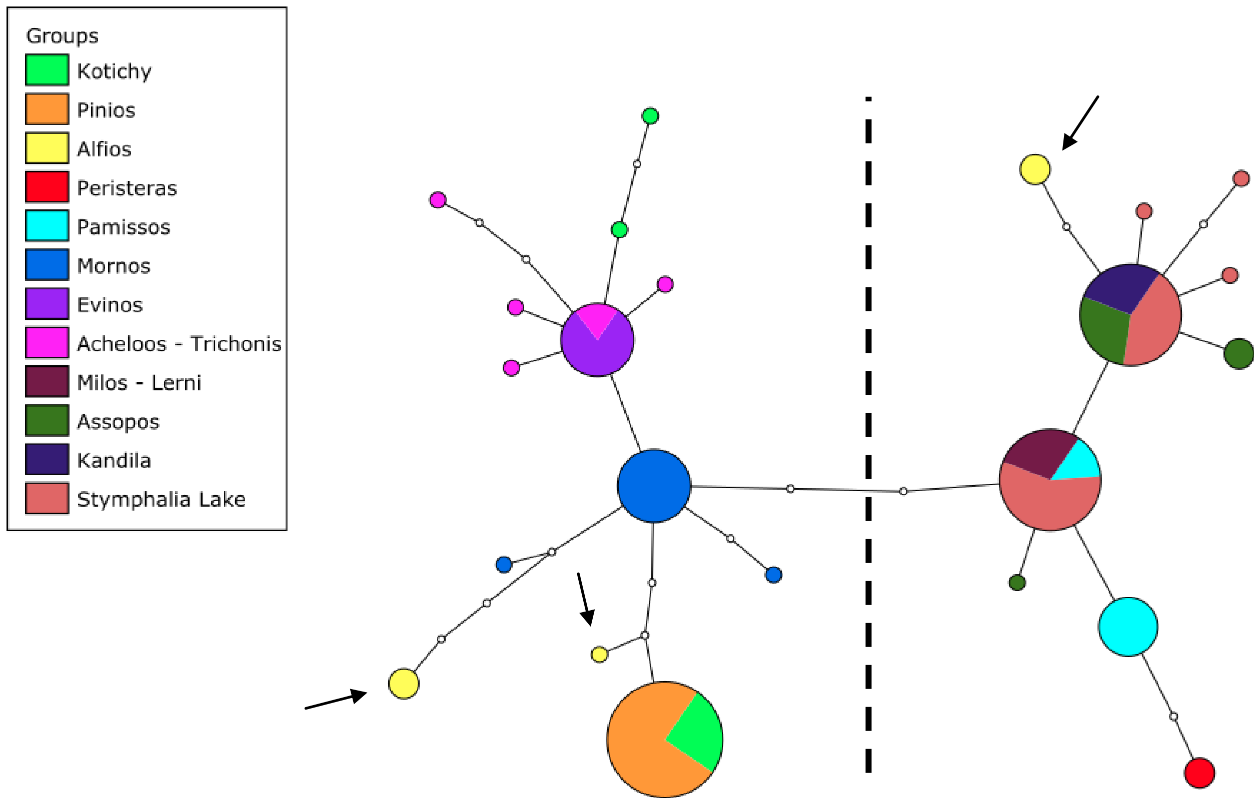

Fig. 12 - *Pelasgus stymphalicus* haplotype network for Cyt b. Modified with tcsBU (TCS Beautifier).

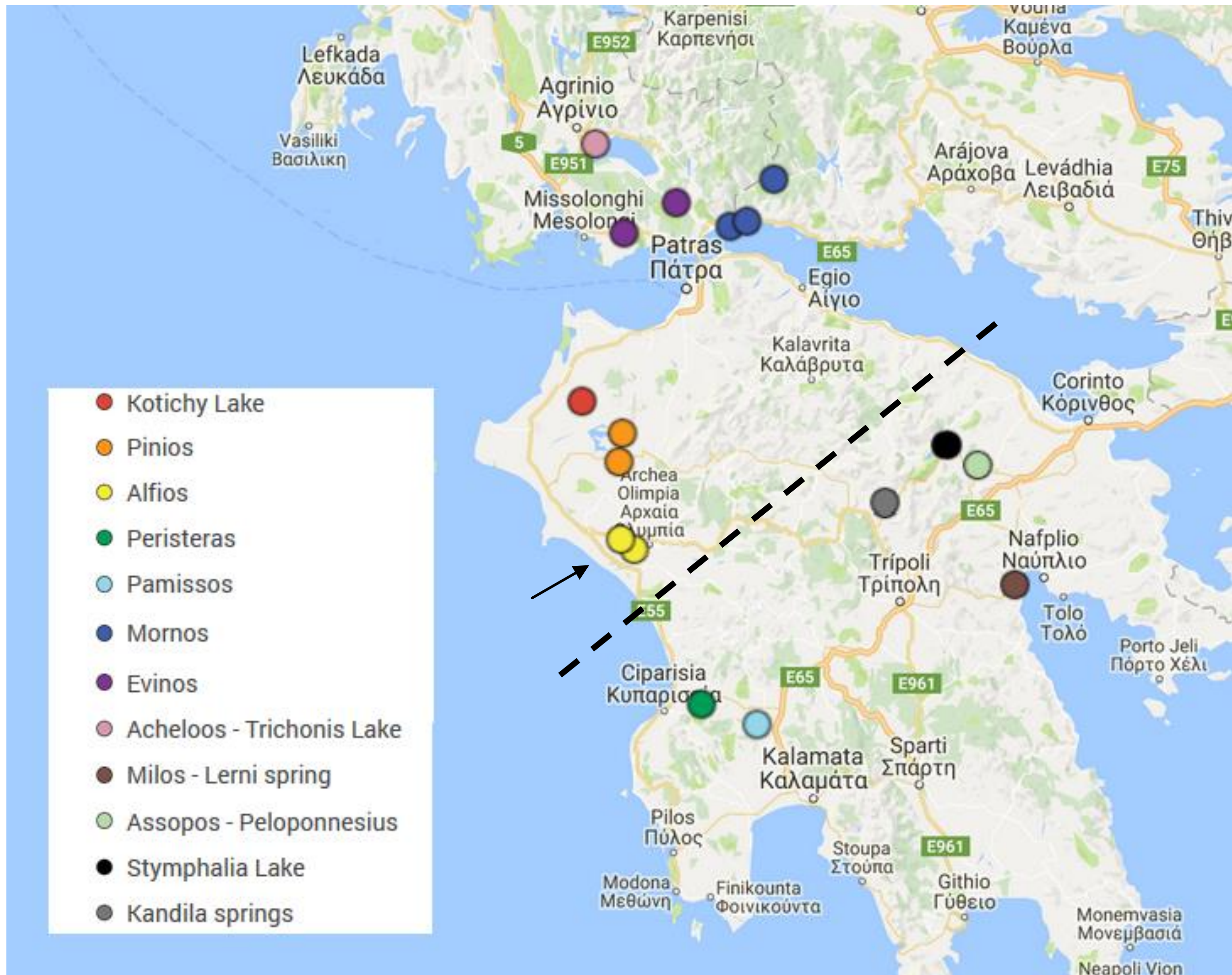

Fig. 13 - *Pelasgus stymphalicus* Cyt b haplotype distribution. Map taken and modified from Google Maps.

P. thesproticus haplotype network shows that Kalamas and Acheroon river basins have private haplotypes, whereas populations from Butrint lagoon rivers (Pavllo and Bistrica) and Vjosa basin (Drinos) share haplotypes. Population from Kerkyra Island appears to be very well differentiated from the rest of populations, in concordance with the results displayed in the phylogenetic trees.

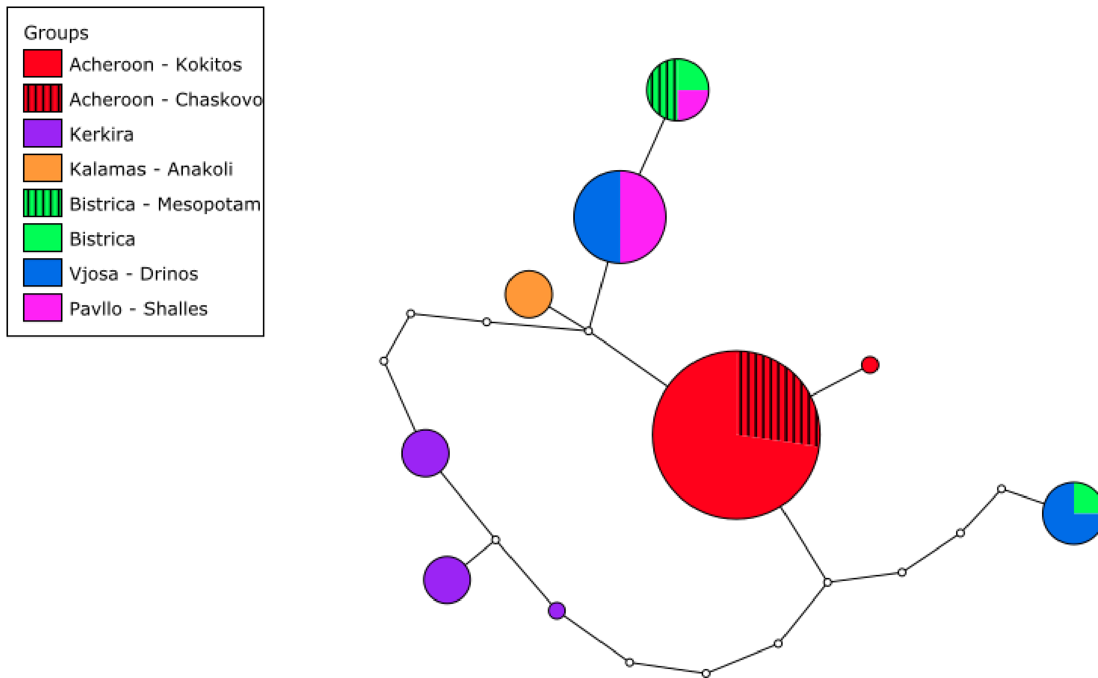

Fig. 14 - *Pelasgus thesproticus* haplotype network for Cyt b. Modified with tcsBU (TCS Beautifier).

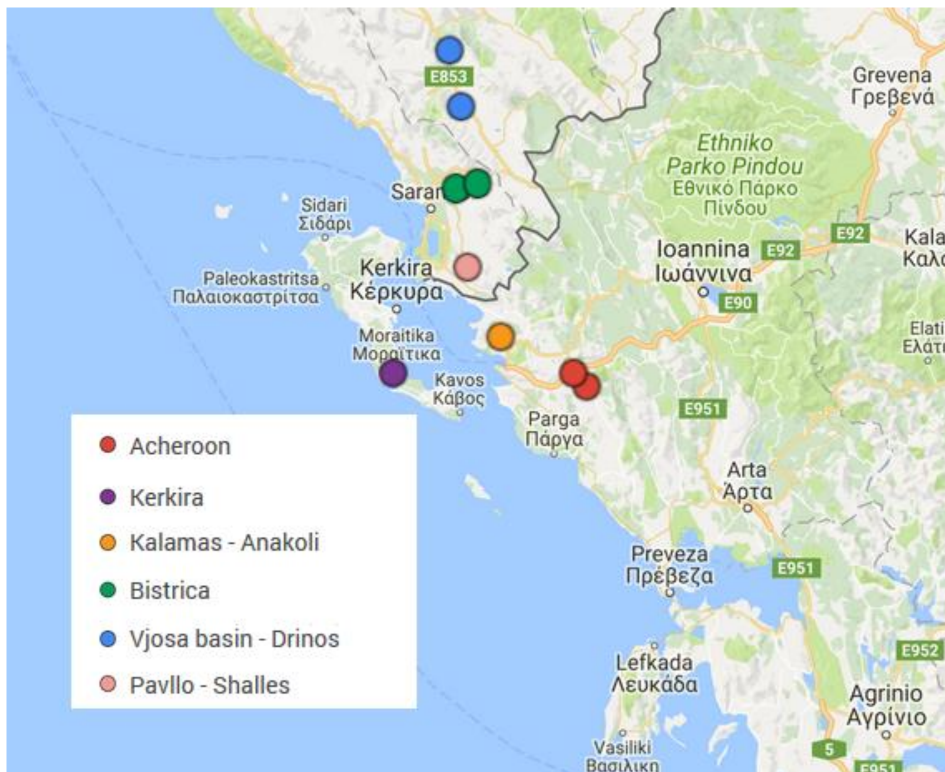

Fig. 15 - *Pelasgus thesproticus* Cyt b haplotype distribution. Map taken and modified from Google Maps.

As explained in Methods section, **S7** analyses were done in two different datasets, one including all species, but with less informative sites, and the other, excluding *P. prespensis* but having a better resolution for the remaining taxa, as the alignment contained longer sequences and therefore more informative sites (about one third of the characters more).

The dataset with shorter alignment including *P. prespensis* sequences contains a total of 129 sequences corresponding to 81 individuals. 37 of them correspond to *Pelasgus stymphalicus*, 16 to *P. thesproticus*, 34 to *P. marathonicus*, 10 to *P. minutus*, 3 to *P. prespensis*, 9 corresponding to *P. laconicus* and 19 tentatively identified as *Pelasgus sp.* The final length of the alignment was 495 base pairs with 641 sites without polymorphism and 74 parsimony informative sites. Thirty nine unique haplotypes were identified.

The dataset with longer alignment without *P. prespensis* sequences contains a total of 126 sequences corresponding to 79 individuals. The composition of the dataset is the same as the previously described with the absence of *P. prespensis* representatives. The final length of the alignment was 785 base pairs with 414 sites without polymorphism and 48 informative sites. Fifty two unique haplotypes were identified.

The significant difference of the number of identified haplotypes between the dataset with and without *P. prespensis* representatives is due to the high variability in the last section of the sequences which was not included in the shorter dataset. Shorter alignment included all sequences, but longer alignment showed better resolution.

Both analyses separated well *P. prespensis* (in the case of the shorter dataset) *P. minutus* and *P. laconicus* and the topology for these species is congruent with that for cyt b. The remaining taxa appear in two sister lineages. In the long dataset (figs 19 and 20), Bayesian analysis clearly separates *P. stymphalicus* and *P. marathonicus* as independent groups, but in Maximum Likelihood analyses the support was lower. Short dataset analysis does not show such a clear division (figs 16 and 17).

The sister cluster to the cluster *P. stymphalicus* + *P. marathonicus* contains haplotypes of *P. thesproticus, P. stymphalicus* from Kandila spring and *Pelasgus* sp. haplotypes from Kerkyra Island (hapl. 18-23) cluster together with a strong support. The same applies for the subclade of *Pelasgus* sp. individuals from Zaravina Lake (hapl. 24-25).

Fig. 16 - Phylogenetic tree for the S7 short sequences alignment, rendered by Bayesian analysis. The numbers in the branches represent posterior probability. Only values higher than 0.95 have been represented**.**

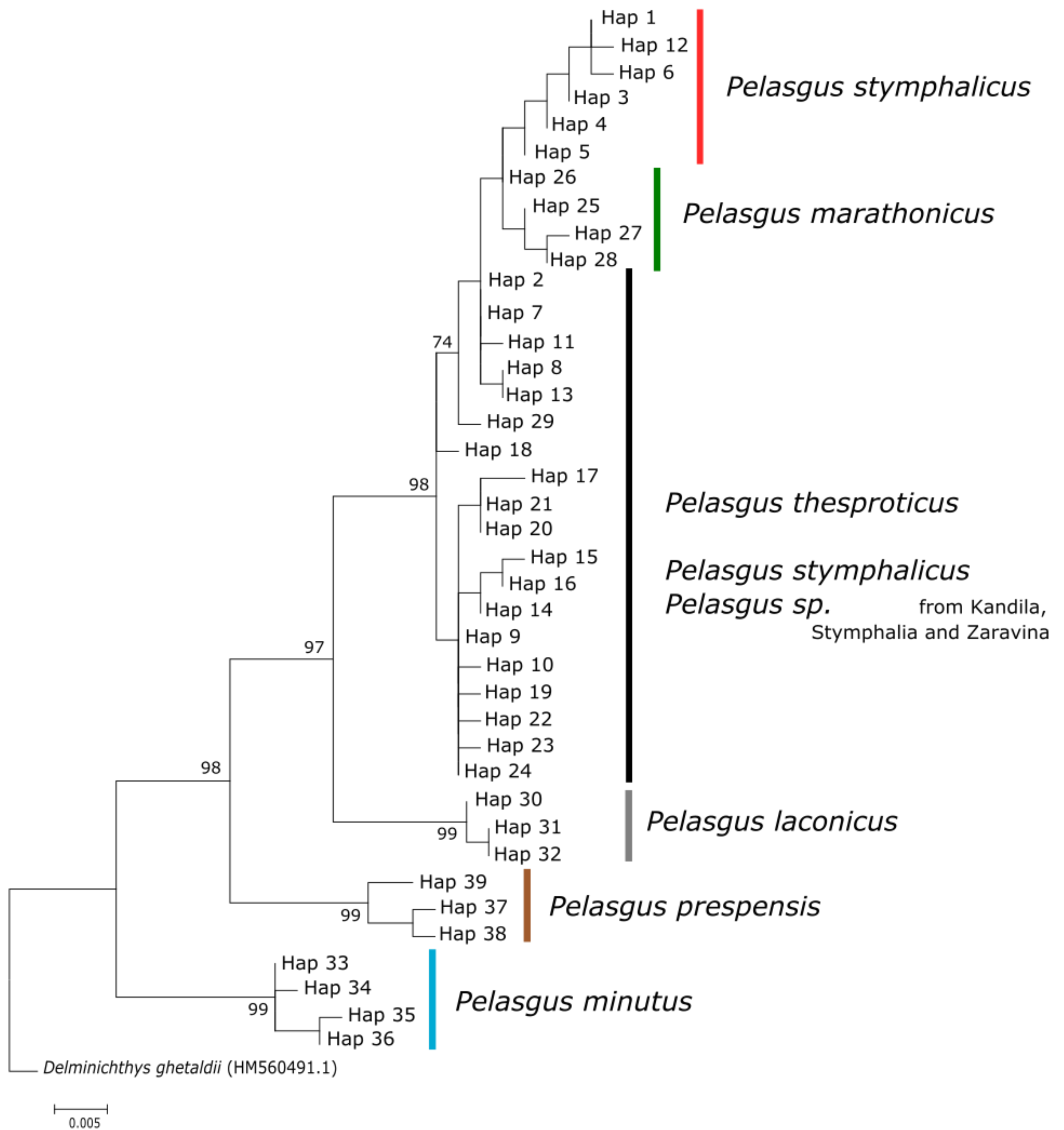

Fig. 17 - Phylogenetic tree for the S7 short sequences alignment, with Maximum Likelihood. The numbers in the branches represent bootstrap values. Only values higher than 75 have been represented.

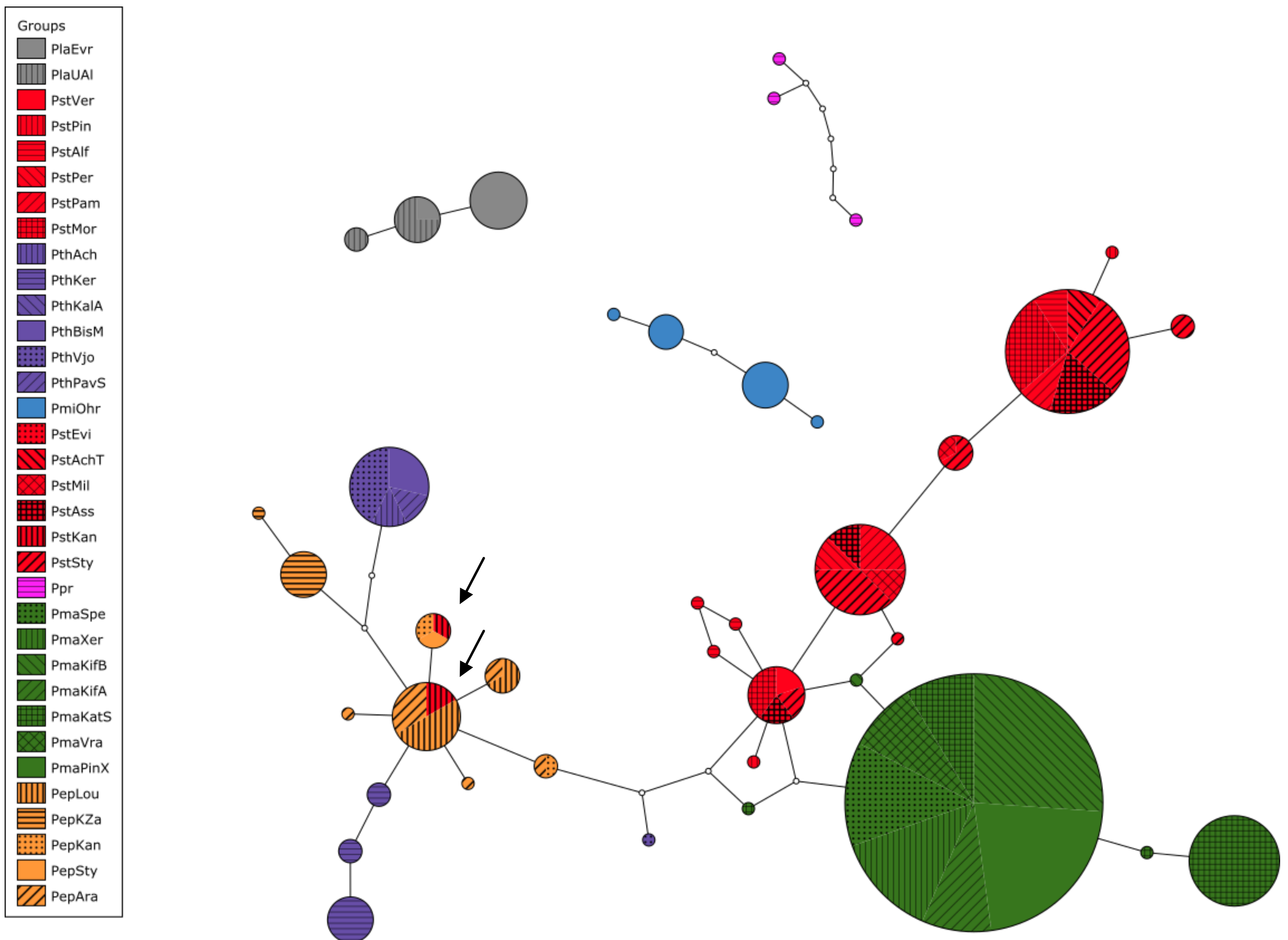

Fig. 18 - Haplotype network for S7 (short sequence). Modified with tcsBU (TCS Beautifier). Arrows show the occurrence of alleles from sample F1166.

Haplotype network (fig 18)lines up with the Bayesian and ML phylogenetic analyses. *P. prespensis*, *P. laconicus* and *P. minutus* appear as independent lineages, separated from the rest of the species. *P, thesproticus* appears to have two separate groups, one of them corresponding to Kerkyra Island and the second corresponding to mainland populations. We can observe that *P. marathonicus* populations are clustered together, as well as *P. stymphalicus*, even though there is a lot of variation within both groups and some of the haplotypes belonging to both species are connected to each other. Conflictive localities are Kandila springs and Stymphalia Lake, were the sympatric occurrence of two species has been observed in cyt b analyses. When observing on detail the occurrence of heterozygous individuals, it has been observed that one sample identified

by cyt b as *P. stymphalicus* from Kandila, here appears to be situated within *Pelasgus* sp. This sample is F1166, position of the alleles is marked in the haplotype network with arrows. Allele A is haplotype 9, together with other samples of *Pelasgus* sp. from Kandila, and allele B is haplotype 10, which is shared with individuals from Louros and Arachthos. *Pelasgus* sp. individuals from Zaravina Lake form a separate haplotype group.

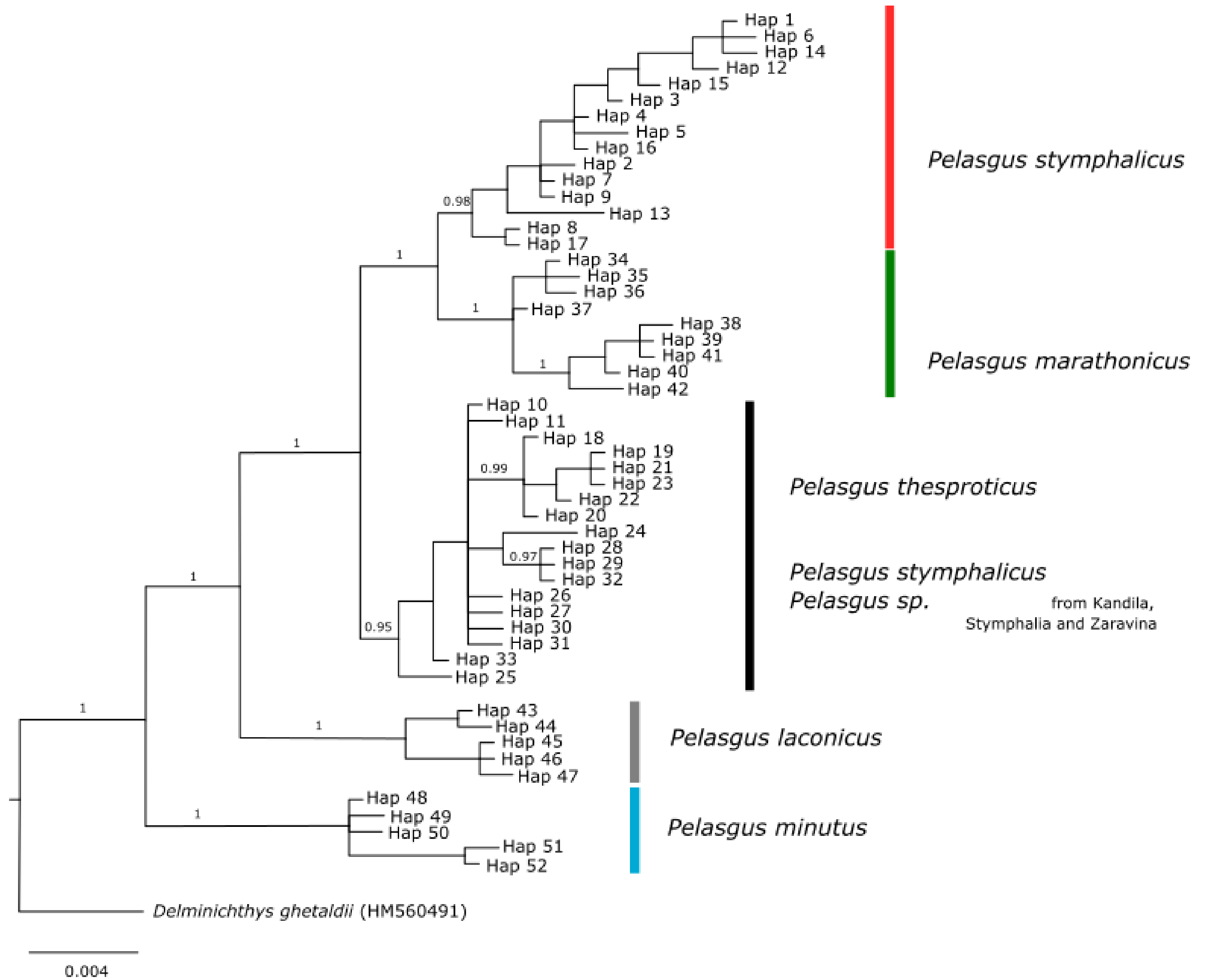

Fig 19. - Phylogenetic tree for the S7 long sequences alignment, rendered by Bayesian analysis. The numbers in the branches represent posterior prbability. Only values higher than 0.95 have been represented.

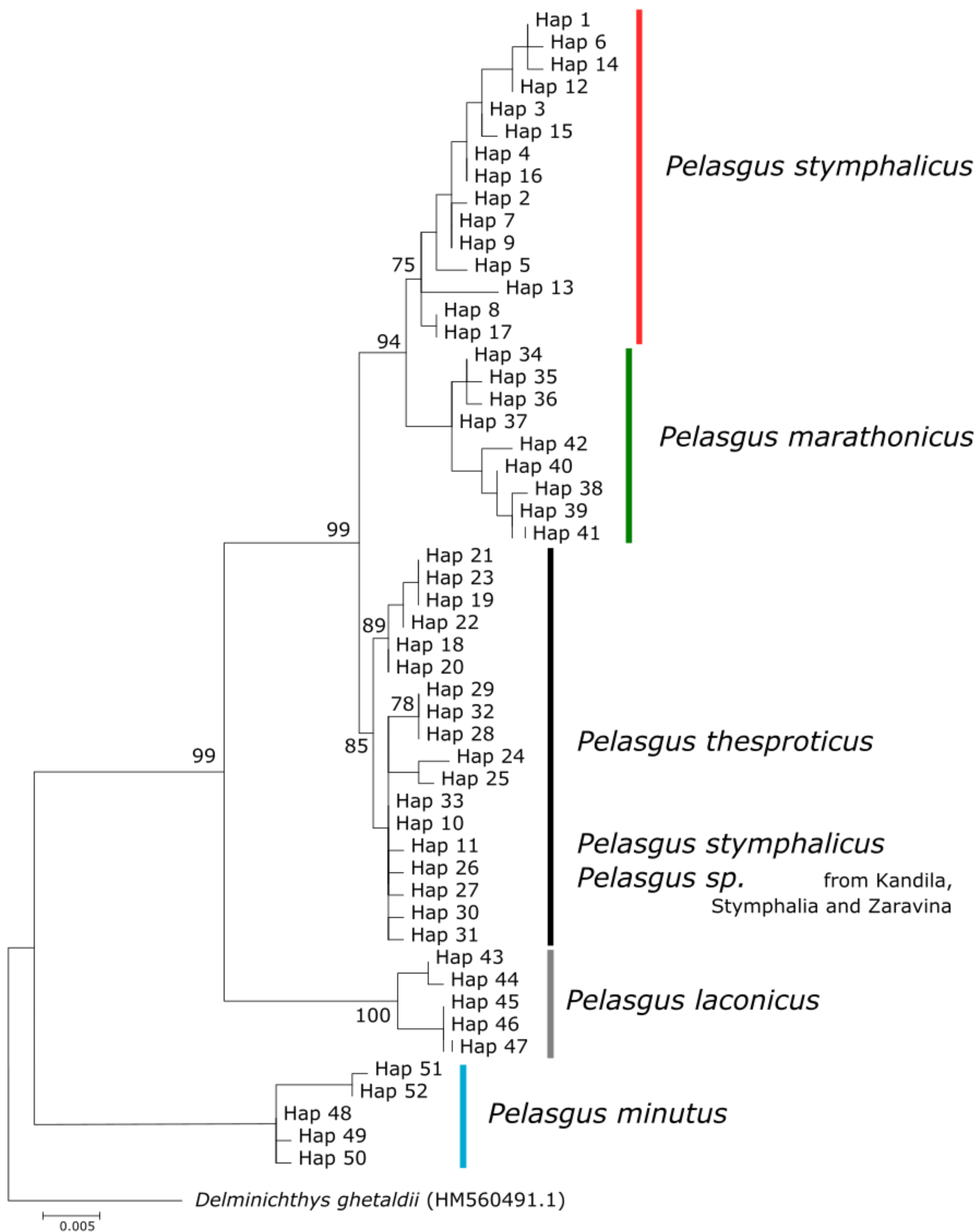

Fig. 20 - Phylogenetic tree for the S7 long sequences alignment from Maximum Likelihood. The numbers on the branches represent bootstrap values. Only values higher than 75% have been represented.

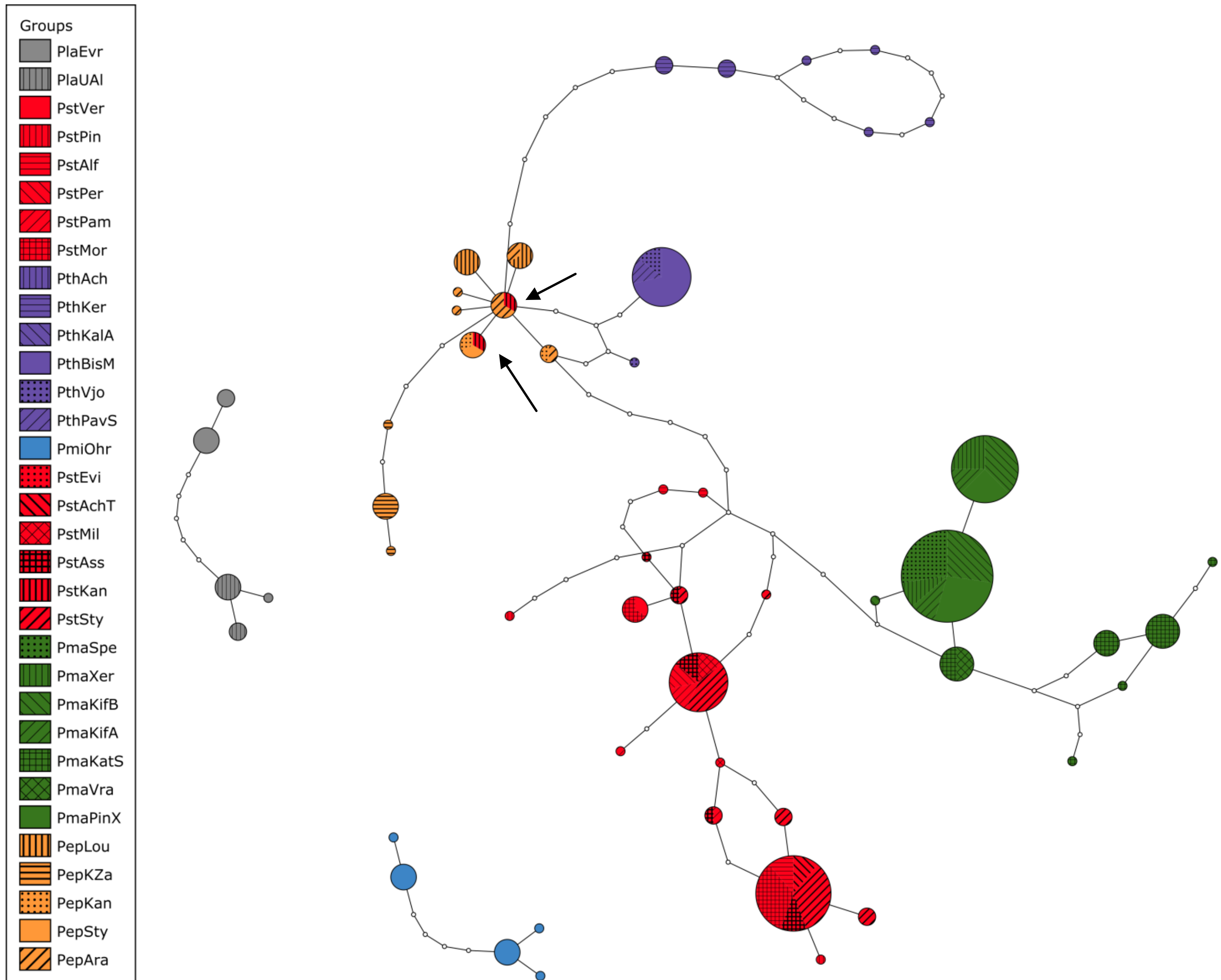

Fig. 21 - Haplotype network for S7 (long sequence). Modified with tcsBU (TCS Beautifier)

Haplotype network constructed on the bases of the long alignment (fig 21) shows much better separation of the groups. It can be commented in the same line the short was done. *P. laconicus* and *P. minutus* appear as independent lineages; *P. thesproticus* appears to have two separate groups corresponding to Kerkyra Island and mainland populations. *P. marathonicus* populations are clustered together, and *P. stymphalicus* and *Pelasgus* sp. perform the highest variability. Again, there is a sample of *P. stymphalicus* from Kandila springs that here appears to be situated with *Pelasgus* sp. (marked in the haplotype network with arrow). Allele A is the haplotype 10 shared with samples of *Pelasgus* sp. from Arachthos, and allele B is haplotype 11, which is

shared with *Pelasgus* sp. individuals from Kandila. *Pelasgus* sp. individuals from Zaravina Lake form a separate haplotype group.

The **RAG** dataset used contains a total of 122 sequences from 91 individuals, 30 of them correspond to *Pelasgus stymphalicus*, 18 to *P. thesproticus*, 30 to *P. marathonicus*, 3 to *P. minutus*, 2 to *P. prespensis*, 13 corresponding to *P. laconicus* and 24 tentatively identified as *P. epiroticus.* The final length of the alignment was 1473 base pairs with 1394 sites without polymorphism and 55 informative sites. Twenty eight unique haplotypes were identified, including one sequence obtained from GenBank (accession number HM560412.1).

Position of *P. prespensis* is different in RAG analysis; it appears to be more closely related to *Pelasgus* sp., *P. stymphalicus, P. thesproticus* and *P. marathonicus* than it was in Cyt b and S7 analyses. Polytomy is observed in Bayesian analysis, corresponding to *P. minutus*, *P. laconicus* and the group comprising rest of taxa, whereas four branches appear in Maximum Likelihood analysis, corresponding to *P. minutus*, *P. laconicus, P. prespensis* and the cluster with the rest of the taxa. *Pelasgus* sp., *P. stymphalicus, P. thesproticus* and *P. marathonicus* cluster together. In Bayesian analyses it is possible to see two groups within this cluster, but in Maximum Likelihood that division does not exist. In Bayesian analyses, haplotypes 9, 10, 27 and 28 appear as well supported clade, which corresponds to *Pelasgus* sp. and *P. stymphalicus* from Stymphalia Lake and some other *P. stymphalicus* populations (figs 22 and 23).

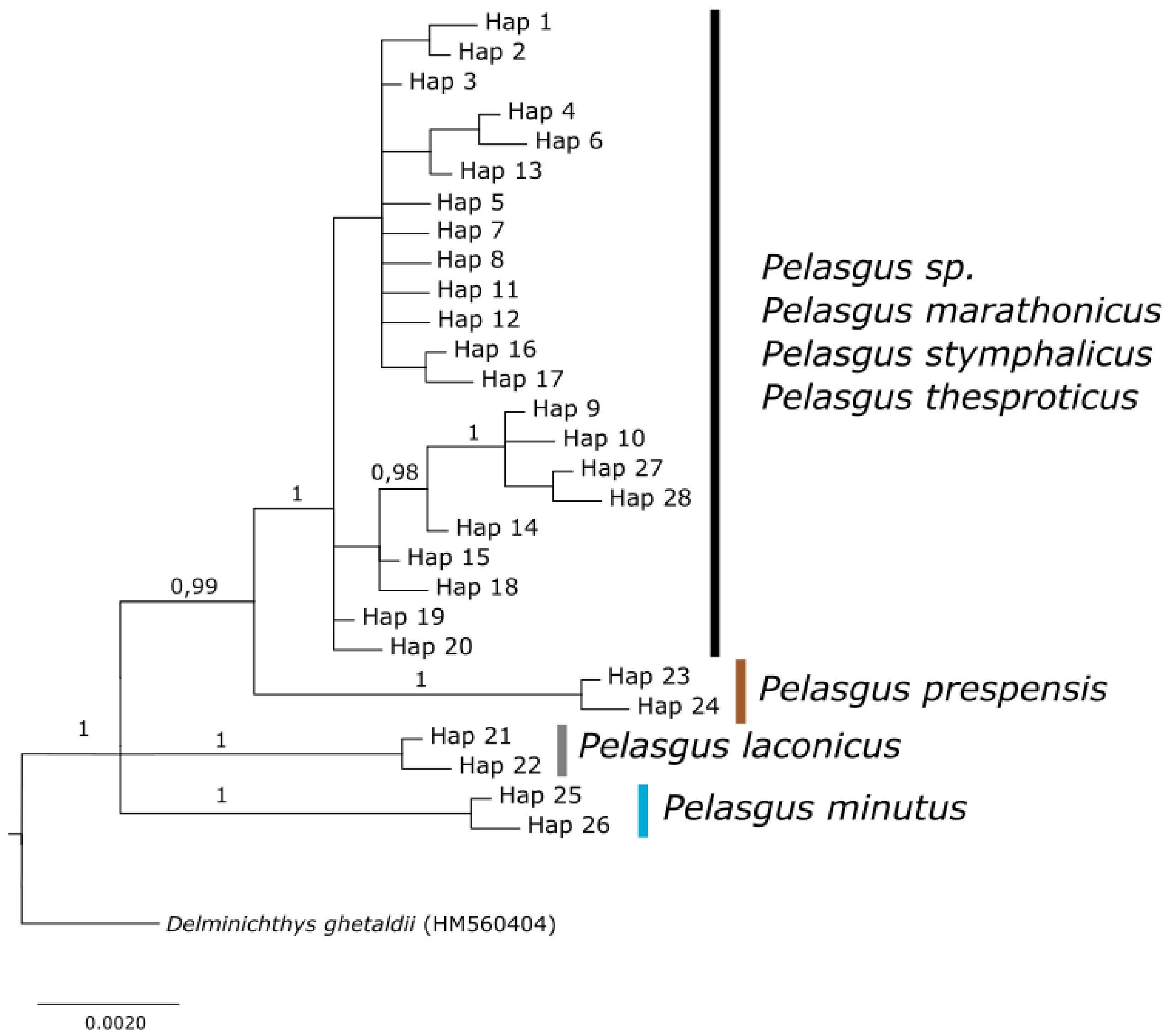

Fig. 22 - Phylogenetic tree for the RAG marker rendered by Bayesian analysis. The numbers in the branches represent posterior porbability. Only values higher than 0,95 have been represented.

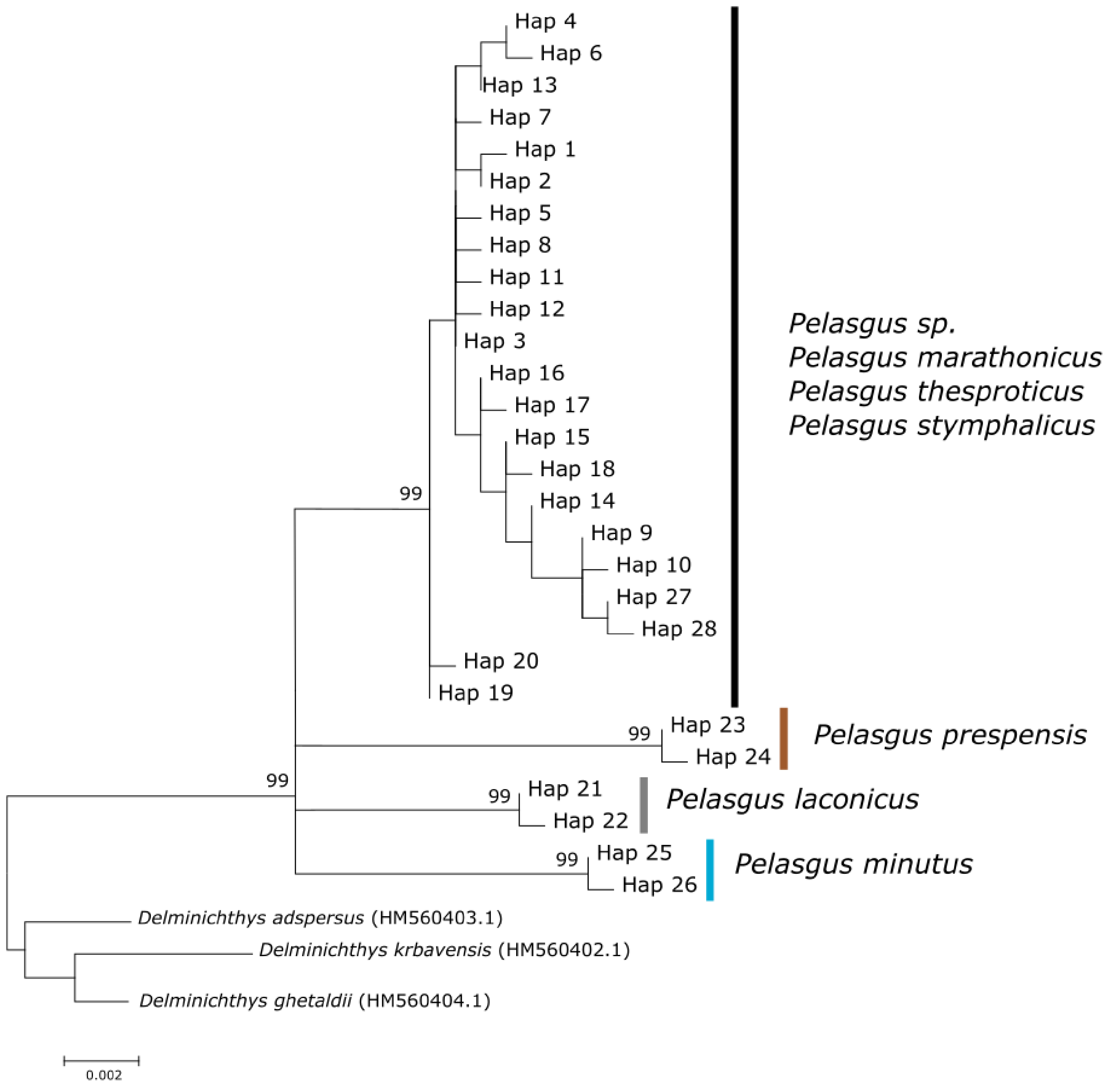

Fig. 23 - Phylogenetic tree for the RAG marker from Maximum Likelihood. The numbers on the branches represent bootstrap values. Only values higher than 75% have been represented.

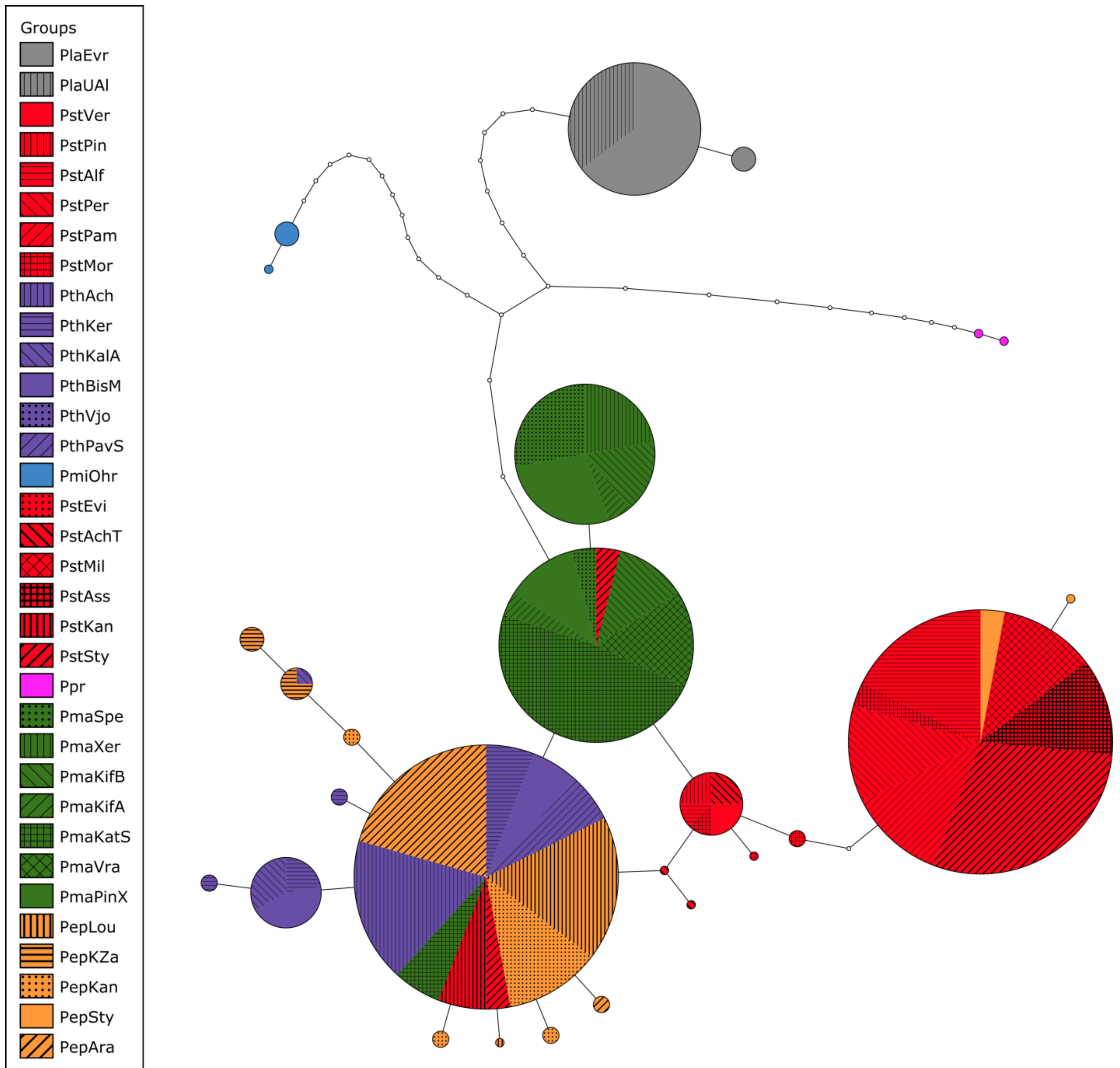

Fig. 24 - Haplotype network for RAG. Modified with tcsBU (TCS Beautifier)

RAG haplotype network reconstruction (fig 24) agrees with the structure of the phylogenetic trees. *P. minutus, P. laconicus and P. prespensis* appear to be well separated from the rest of the species by sixteen, twelve, and thirteen mutations, respectively. *Pelasgus marathonicus* individuals are grouped into two big haplotypes, and share another haplotype with *Pelasgus* sp.,

P. thesproticus and *P. stymphalicus* from Stymphalia and Kandila. *Pelasgus* sp. and *P. stymphalicus* appear to have the biggest haplotype variability. *Pelasgus* sp. from Zaravina has separated haplotypes from the rest of *Pelasgus* sp. individuals.

Rhodopsin dataset contains a total of 126 sequences from 108 individuals, 30 of them correspond to *Pelasgus stymphalicus*, 20 to *P. thesproticus*, 24 to *P. marathonicus*, 5 to *P. minutus*, 7 to *P. prespensis*, 10 corresponding to *P. laconicus* and 26 tentatively identified as *P. epiroticus.* The final length of the alignment was 856 base pairs with 797 sites without polymorphism and 19 informative sites. Eighteen unique haplotypes were identified.

P. prespensis together with *P. minutus* and *P. laconicus* occurs in polytomy with the group including the rest of taxa. Within that cluster *Pelasgus* sp. and *P. stymphalicus* from Stymphalia and Kandila form a separated group, involving haplotypes 5, 11 and 12. *Pelasgus thesproticus* and *P. marathonicus* cluster together. In Bayesian analyses it is possible to see two groups within *P. thesproticus*/*P. marathonicus* cluster, but in Maximum Likelihood that division was not revealed. Furthermore, in Bayesian analyses, haplotypes 9, 10, 27 and 28 appear as well supported clade, which corresponds with *P. epiroticus* and *P. stymphalicus* from Stymphalia Lake and some other *P. stymphalicus* populations (figs 25 and 26).

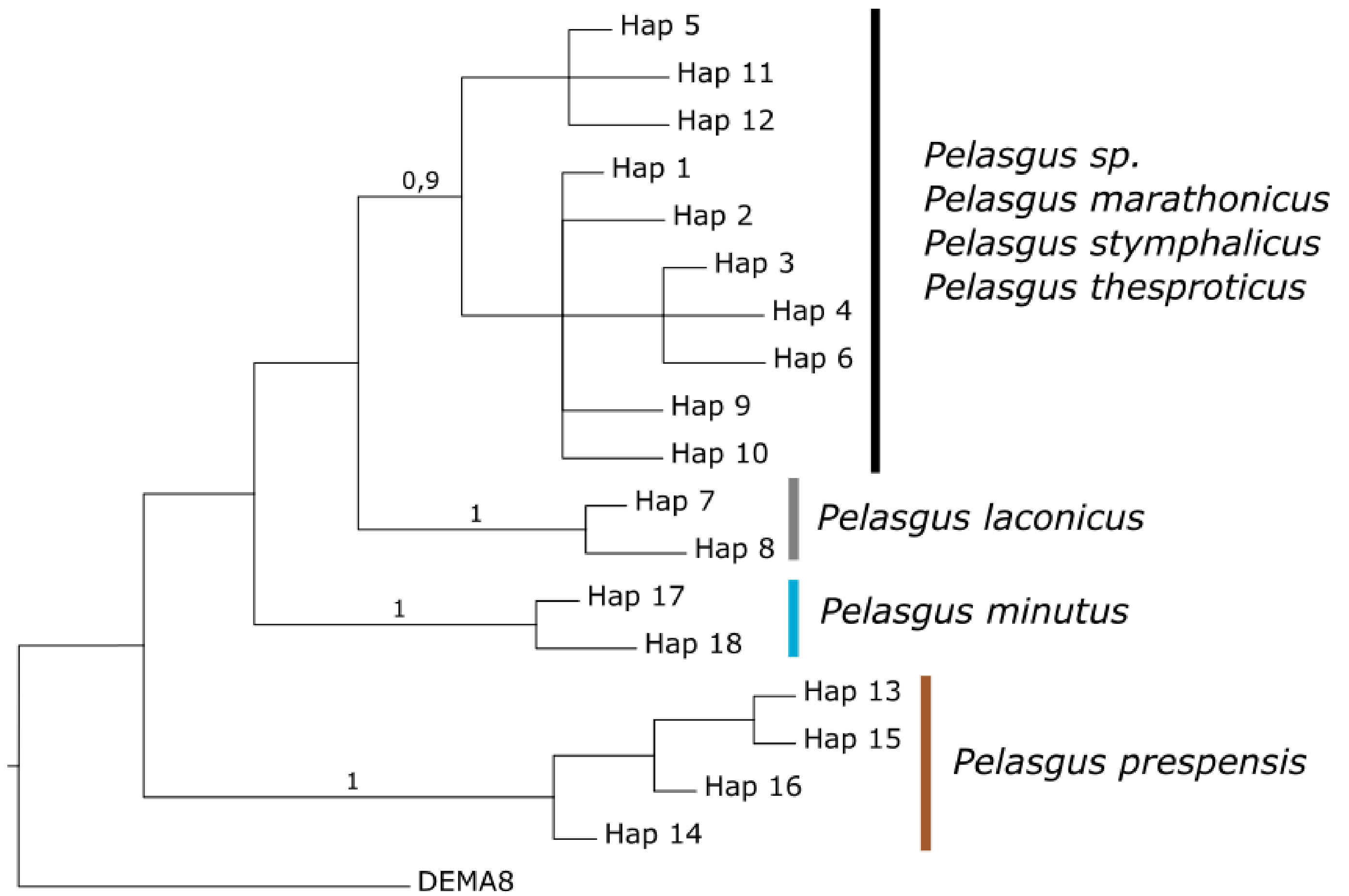

Fig. 25 - Phylogenetic tree for the Rhodopsin marker rendered by Bayesian analysis. The numbers in the branches represent posterior prbability. Only values higher than 0.95 have been represented.

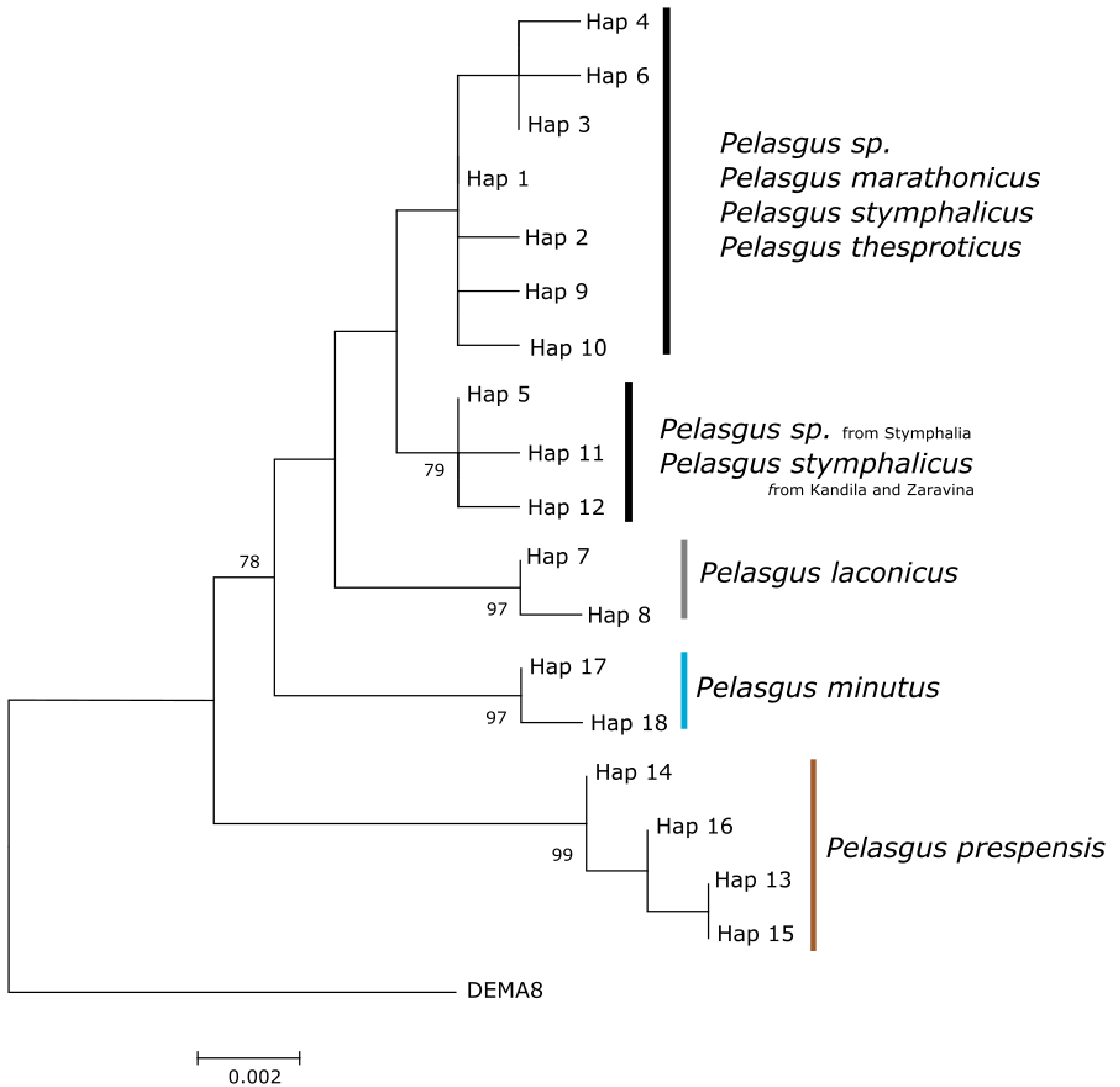

Fig. 26 - Phylogenetic tree for the Rhodopsin from Maximum Likelihood. The numbers on the branches represent bootstrap values. Only values higher than 75% have been represented.

When observing haplotype network for rhodopsin (fig 27), samples F1160, F1157 and F1159 (*P. stymphalicus* from Stymphalia based on cytochrome b) have an allele grouped with *P. stymphalicus* (hap 3) and another allele lined up with *Pelasgus* sp. (hap 6). However, the variability within Rhodopsin sequences is small, so it is difficult to make any conclusions. There is a common shared haplotype between *Pelasgus* sp., *P. thesproticus* and *P. marathonicus* (hap. 1). *P. marathonicus* populations are divided into two haplotypes (hap. 1 and hap. 2). Again, *P. minutus*, *P. laconicus* and *P. prespensis* appear to be the most divergent taxa, separated by seven, five and eleven mutation points, respectively.

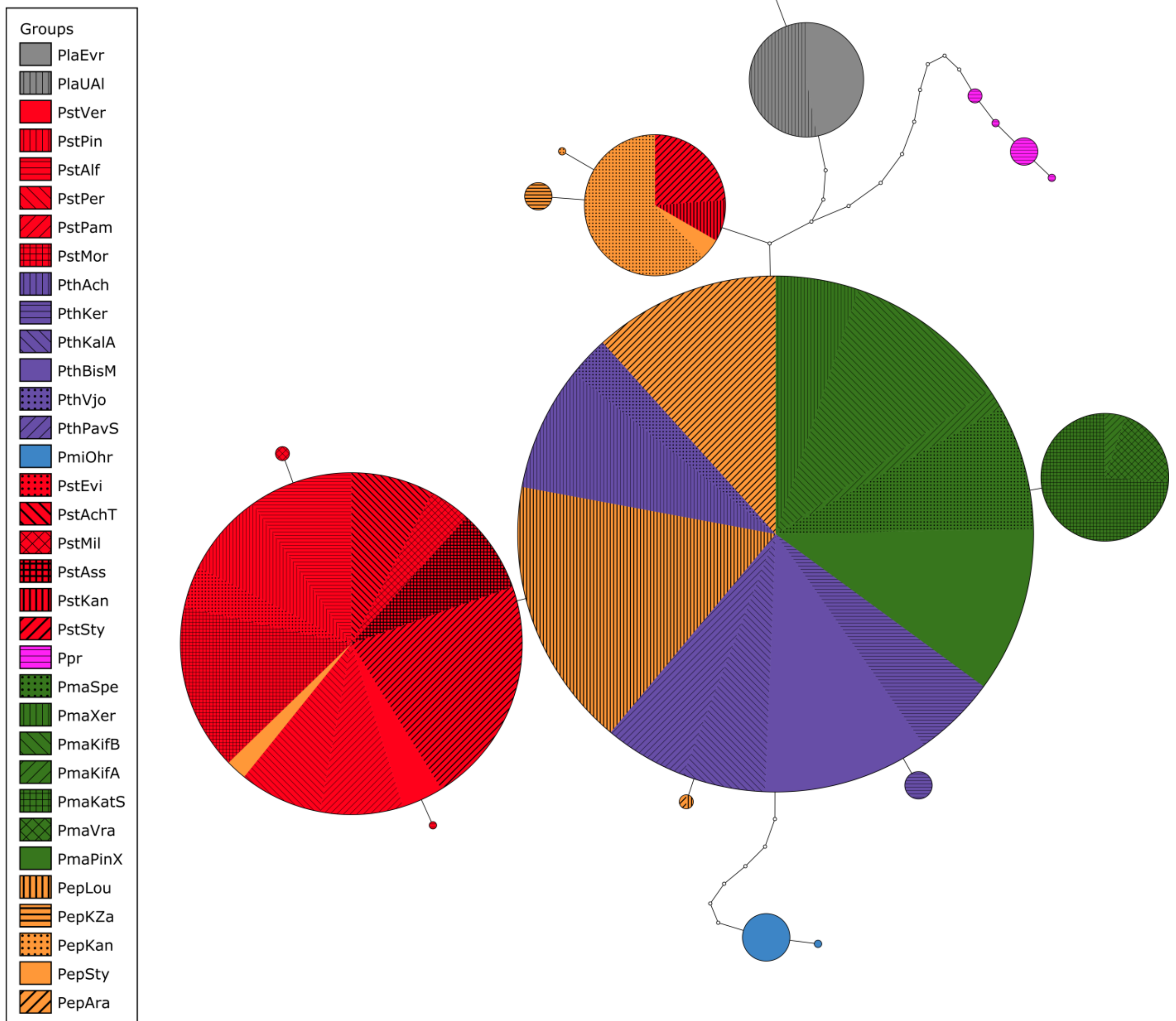

Figure 27 - Haplotype network for Rhodopsin. Modified with tcsBU (TCS Beautifier)

Maximum Likelihood and Bayesian analysis from RAG1 and Rhodopsin separated well only Pelasgus minutus, P. prespensis and P. laconicus. The position of P. minutus and P. prespensis in case of Rhodopsin inferred from both approaches was reversed.

Out of nine individuals analyzed from Kandila springs, seven had cyt b sequences of Pelasgus sp. and two of Pelasgus stymphalicus, while in Stymphalia out of ten individuals one has the cyt b of Pelasgus sp. Based on S7 sequences, hybridization occurred in both localities. Although there are signs of hybridization also from other nuclear markers, the variability within the dataset is too low to be able to draw conclusions from these data.

Details about heterozygous samples and its allele's positions in are in Appendix 1 and 4.

Mean genetic distances within and between *Pelasgus* species are compiled below in tables 10 to 15. Interespecific genetic distances (p-distances) are higher between *P. minutus, P. prespensis* and *P. laconicus* and the rest of the species, and lower between the other species (Table 10).

Table 10 - Mean intraspecific genetic distances (p-distance) within *Pelasgus* species for all markers (%)

	Cyt b	S7 (S)	S7 (L)	RAG	Rhodopsin
P. thesproticus	0.4	0.7	0.6	0.1	-
P. stymphalicus	0.5	0.5	0.5	0.1	0.3
P. marathonicus	0.3	0.2	0.3	-	0.1
P. laconicus	0.2	0.2	0.4	0.1	0.1
P. minutus	0.2	0.3	0.5	0.1	0.1
Pelasgus sp.	1.3	0.3	0.3	0.3	0.4
P. prespensis	0.5	1	-	0.1	0.1

Distances calculated for cyt b confirm the clades separation already observed in the phylogenetic analysis. *P. laconicus* and *P. minutus* appear to be the most divergent species, according to cyt b p-distance, and *Pelasgus* sp. and *P. marathonicus* appear to be the most genetically similar ones.

Table 11 - Mean genetic distances (p-distance) between *Pelasgus* species for Cyt b (%)

	P. th	P. st	P. ma	P. la	P. mi	Pelasgus sp.
P. thesproticus						
P. stymphalicus	5.8					
P. marathonicus	4.9	3.9				
P. laconicus	10.2	10.2	10.2			
P. minutus	10.9	10.4	10.5	**11**		
Pelasgus sp.	4.4	4.9	**4.3**	10	10.7	
P. prespensis	9.1	10.3	9.8	10.2	10.3	9.3

The greatest genetic distances between species are observed for S7 in *P. minutus*, which appears to be the most distant species compared with all other *Pelasgus* species, followed by *P. prespensis*. Distance of *Pelasgus* sp. to the rest of species differs in the nuclear markers. According to S7 (table 12), it is closely related to *P. thesproticus*, but according to RAG and

Rhodopsin (tables 13 and 14), it also very close to *P. stymphalicus* and *P. marathonicus*, with similar values in all cases.

Table 12 - Mean genetic distances (p-distance) between *Pelasgus* species for short alignment of S7 (%)

	P. th	P. st	P. ma	P. la	P. mi	Pelasgus sp.
P. thesproticus						
P. stymphalicus	1.6					
P. marathonicus	1.8	0.9				
P. laconicus	3.2	3	3.2			
P. minutus	5.5	5.7	**5.9**	5.2		
Pelasgus sp.	**0.6**	1.2	1.4	2.8	5.2	
P. prespensis	4.9	4.7	5	4.3	4.7	4.5

Table 13 - Mean genetic distances (p-distance) between Pelasgus species for long alignment of S7 (%)

	P. th	P. st	P. ma	P. la	P. mi
P. thesproticus					
P. stymphalicus	1.7				
P. marathonicus	1.7	1.3			
P. laconicus	3.8	4	3.9		
P. minutus	5.7	**5.9**	5.8	5.5	
Pelasgus sp.	**0.6**	1.5	1.5	3.5	5.5

According to p-distances calculated for RAG (table 14), *P. minutus* and *P. prespensis* appear to be the more distant species of the genus, and *Pelasgus* sp. seems to be closely related to *P. thesproticus*, *P. marathonicus* and *P. stymphalicus*, as it was already described in the phylogenetic trees.

Table 14 - Mean genetic distances (p-distance) between *Pelasgus* species for RAG (%)

	P. th	P. st	P. ma	P. la	P. mi	Pelasgus sp.
P. thesproticus						
P. stymphalicus	0.2					
P. marathonicus	0.2	0.2				
P. laconicus	1	1	0.9			
P. minutus	1.3	1.2	1.2	1.4		
Pelasgus sp.	0.2	0.2	0.3	1	1.3	
P. prespensis	1.1	1	1	1.2	1.6	1.1

According to Rhodopsin gene p-distances (table 15), *P. prespensis* is the most divergent species of the genus *Pelasgus*, and the most similar species are *P. marathonicus* and *P. thesproticus*

Table 15 - Mean genetic distances (p-distance) between *Pelasgus* species for Rhodopsin (%)

	P. th	P. st	P. ma	P. la	P. mi	Pelasgus sp.
P. thesproticus						
P. stymphalicus	0.3					
P. marathonicus	**0.2**	0.3				
P. laconicus	0.8	0.8	0.7			
P. minutus	0.8	0.9	0.7	1.1		
Pelasgus sp.	0.4	0.4	0.3	0.8	0.9	
P. prespensis	**1.5**	**1.5**	**1.4**	**1.5**	**1.4**	**1.5**

4. Discussion

The molecular study of genus *Pelasgus* confirmed some of the previous resuls, as for example the existence of several strongly supported lineages within the genus (Perea et al. 2010, Geiger et al. 2014), but it has also uncovered some interesting facts that should be pointed, each of them with different particularities that deserve additional attention.

The genetic analyses confirmed the validity of six species defined on the bases of morphology, namely *Pelasgus laconicus, Pelasgus stymphalicus, Pelasgus thesproticus, Pelasgus marathonicus, Pelasgus prepensis* and *Pelasgus minutus*. In addition, the study revealed the existence of a phylogenetic lineage within the genus, that requires more detailed studies to determine if it corresponds to the described species *Pelasgus epiroticus* or if it actually is a non described separate species.

Analyses of nuclear DNA markers are not as informative as the mitochondrial one. Nuclear genes effectively differentiate *P. minutus, P. laconicus* and *P. prespensis*. The long S7 intron fragment appears to distinguish the rest of species as well, with the exception of *Pelasgus* sp. and *P. thesproticus*. The remaining nuclear genes failed to distinguish *P. thesproticus*, *P. marathonicus*, *P stymphalicus* and *Pelasgus* sp. due to haplotpye sharing between these species. Even though this situation has been reported in some studies (e.g. Gauble et al., 2008), since the time to reach the monophyly in nuclear genes takes about four times longer than it does for mitochondrial DNA, an incomplete lineage sorting should not be considered as an evidence against the status of species (Markova et al., 2010).

Although Rhodopsin is not considered a neutral phylogenetic marker, all sampled populations of *Pelasgus* were form the similar habitat (small streams; in the case of the lake basins, small streams were samples; the only exception was Zaravina Lake, where samples were taken from the littoral), so the adaptive difference of vision in different populations was not expected.

The various molecular analyses used have shown a similar schema: *P. minutus, P. prespensis* and *P. laconicus* always appear as basal groups, with greater genetic distances to the

rest of the *Pelasgus* species. It has been suggested by Durand et al. (1999) that the presence of species of the genus *Squalius* from Western Greece as basal groups on phylogenetic analyses is a proof of ancient colonization events on that side of the Balkan Peninsula, that might have occurred during the mid-Pleistocene from Albania southwards. The position of *P. minutus, P. prespensis* and *P. laconicus* in the phylogenetic trees agrees with this proposed dispersion theory., thought the dating of the evolutionary events leading to separation of the species is obviously much older. Perea et al (2010) estimated the origin of the genus Pelasus to aproxximately 14 MY. There are no other available data for further separation of different linegaes/species within the genus, but based on the observed differences on cytochrome b (1.9 to 11%) and estimated mutation rate for this gene 0,4% per MY for Leuciscinae (Perea et al. 2010), none of the haplotype clusters is of Pleistocene origin. -.

The results of cytochrome b analyses revealed that within *P. thesproticus* and *P. marathonicus* there was a complicated haplotype structure (Fig 6 and 7), but the reconstruction of the haplotype network for both the mitochondrial and nuclear DNA showed that despite the apparent diversity within both species, they cluster together in very well defined and supported groups, except in the case of RAG and Rhodopsin. This can be explained by to the lower variability of the latter markers, in comparison with S7, which has more informative sites. In the haplotype network reconstructions of nuclear markers (Figs18, 21, 24 and 27), besides having their own haplotypes, both species tend to share haplotypes.

It has been observed a great diversity between Ohrid and Prespa lake species, not only within *Pelasgus* but also in other species (e.g. *Barbus,* Markova et al., 2010; *Alburnoides*, Stierandová et al., 2016), which can be explained by the mountains that separate both basins, which probably played a role as physical barrier, promoting genetic divergence (Tsoumani et al., 2014). This diversity has been assessed by both mitochondrial and nuclear data analyses (Figs. 6 to 27).

4.1 Distribution of Pelasgus species

My results confirm the distribution areas described by Kottelat and Freyhof (2007a) in the case of three *Pelasgus* species, but have also shown that the rest of the species occurs on larger area,

because I analyzed localities not sampled before (*P. thesproticus, P. marathonicus, P. prespensis* and possibly *P. epiroticus*). In contrast, I revealed restriction of the distribution area in the case of *P. thesproticus*.

Pelasgus laconicus (common name Evrotas minnow), was described from Evrotas river and upper Alfios river basins, located in the Southern Peloponnesus, and so far it is not known from any other locality. My data are inagreement with this described distribution. Sampled individuals, initially indentified as *P. laconicus,* from the Evrotas River near Sparta (which is just ten km downstream of the type locality of this species, Vivari spring by the Evrotas River) and upper part of the Alfios River, has been confirmed by phylogenetic analyses to form a monophyletic cluster (Figs 6 and 7).

The type locality of *Pelasgus marathonicus* (common name Marathon minnow) is a spring near Kato Souli, a few km NE of Athens. The expected distribution area comprises the rivers from vicinity of Athens up to Cholorema in Volos bay (Kottelat and Barbieri 2005, Economou et al. 2007), includind the former Xiniada lake basin in Thessalian Pinios river drainage (Koutsikos et al. 2012). I analyzed individuals from Xerias, Sperchios, Kifissos-Beotia, Kifissos-Athens (different river basin from the Kifissos in Beotia) and Vraonas rivers, as well as from Kato Souli wetland (type locality), and former Xiniada Lake basin in Thessalian Pinios river drainage. Haplotypes of cyt b of all samples from these populations cluster together (Figs 6 and 7). Xerias is a new locality of distribution of this species, located more northernly from the initially described area. Therefore, it can be said that *P. marathonicus* distribution area is actually wider than was previously thought.

Pelasgus minutus (common name Ohrid minnow) occurs in Ohrid and Skadar lakes basins, in Albania, Macedonia and Montenegro. Since the main objective of the study were the Greek species of *Pelasgu*s, only individuals from Ohrid Lake (which is the type locality of this species) were analyzed, confirming the identification and presence of *Pelasgus minutus* in Ohrid Lake.

Pelasgus presensis (commonly named Prespa minnow) has been only reported in Great Prespa Lake and Small Prespa Lake in north-western Greece, Albania and Macedonia, with a

continuous decline of the population, monitored since 1980s until present (Crivelli, 2006b; Kottelat&Frehof, 2007a). Samples analyzed in this study originate from Prespa Lakes and Devoll (Devolli) River, situated west of Prespa Lakes. All phylogenetic analyzes confirmed very close similarity of samples from Prespa lake and Devoll River. Thus, *P. prespensis* is not endemic to Prespa Lake system. A possible explanation of the presence of *P. prespensis* in the Devoll River is the existence of a connection between Small Prespa Lake and Devoll River (Tziritis, 2014). Similar situation has been described for other genera, such as the case of *Barbus prespensis* (Markova et al., 2010) and *Alburnoides prespensis* (Stierandova et al., 2015): in both studies it has been demostrated that *B. prespensis* and *A. prespensis* are not endemic species of Prespa Lake as initialy expected, but populate a large region of southern Albania, both species occurring in the same area. It would be interesting to do more sampling in the area to see if *P. prespensis* shows the same distribution pattern and if it is present in the same river basins of *Barbus* and *Alburnoides* species.

Pelasgus stymphalicus´s (commonly named Stymphalia minnow) type locality is Stymphalia Lake, in the Peloponnesus, which is an endorheic basin (without surface connection with other river drainages). All known populations from Peloponnesuse (with exception of the distribution range of *P. laconicus* and Kandila springs endorheic basin), together with rivers in Etolia-Acarnania up to the Lefkas island, were tentatively included as conspecific by Kottelat and Barbieri (2004), Kottelat and Freyhof (2007a), Economou et al. (2007). Molecular identification of samples in my work from the Peloponnesus and Etolia-Acarnania suggested that they belong to the same species, including individuals from lower Alfios. It is important to note that in the upper part of Alfios *P. laconicus* occurs, and the limit of the distribution of both species within Alfios river has not been identified yet (Kottelat and Barbieri 2004).

When analyzing the position of the sampled localities in the phylogenetic tree and in the haplotype network of cyt b, it was possible to differenciate two big groups of haplotypes (Figs 6, 7, 12 and 13),. One of the groups corresponds to the mainland populations (Mornos, Evinos and Acheloos) and the northwestern populations in Peloponnesus peninsula (Kotichy Lake system, Pinios, and Alfios), and the second corresponds to the rest of Peloponnesus populations (Alfios, Peristeras, Pamissos, Milos, Assopos, Kandila and Stymphalia Lake). Interestingly, two Alfios haplotypes appear in both haplogroups, and are thus very divergent from each other. The

similarity of populations from Evinos, Kotichy and Pinios, and its divergence from Alfios River population, can be explained by a river connection during the Pleistocene glacial sea regresions. This would agree with a very similar pattern that has been observed in the distribution of other species from this region (e.g. *Luciobarbus albanicus* or *Tropidophoxinellus hellenicus*, Economou et al., 2007). Acheloos and Evinos populations appear to be closely related, even sharing the same haplotype. Guinand et al. (2001) suggested migrations between both drainages during the Pleistocene, and this would agree with my observations.
No new populations have been discovered for this species, but molecular analyses of cytochrome b showed the sympatric occurrence of two species (*Pelasgus stymphalicus* and *Pelasgus* sp.) in two localities within the distribution area of *P. stymphalicus* in the Peloponnesus Peninsula (Stymphalia Lake and Kandila Springs, see Figs 8, 9, 12 and 13), the later confirmed also with the analyses of nuclear markers (Figs. 16 to 27). It was already suspected by Kottelat and Barbieri (2004) for Kandila Spring, that a different species of *Pelasgus* can be present there. Based on phylogenetic trees and the haplotype network reconstructions (see Figs 6 to 27), it is likely that an introduction occurred in that localities, as previously supposed by Kottelat and Barbieri (2004).

The type material of *Pelasgus thesproticus* (common name Epiros minnow) comes from the vicinity of Paramythya village, in the upper Kokitos river drainage, which is the inflow of the Acheroon River. Several populations from Greece (Kerkyra island and from Kalamas to Arachthos river basins) and Albania (Butrint lagoon drainage) were expected to belong to this species (Kottelat and Barbieri, 2004; Kottelat and Freyhof, 2007a; Economou et al., 2007).

My result are not concordant with the above mentioned distribution range for this species. Sampled populations covered all the river basin of expected range. Only some of the sampled localities initially asigned to *P. thesproticus* could be confirmed as occurrence localities of this species, based on the inclusion of samples from the type locality: Acheroon and Kalamas river basins in Grecee and Pavllo and Vjosa basins in Albania, as well as Kerkyra Island population. Arachthos river basin was originally included in the distriburion area of *P. thesproticus* (Kottelat & Barbieri, 2004; Kottelat & Freyhoff, 2007a), but the molecular analyses retrieved this population together with Louros river populations. Both Louros and Arachthos rivers drain to Amvrakikos Bay. Extension of the distribution range of *P. thesproticus* was, on the other hand,

discovered in Albania. Samples from two localities within Drinos river, which is an inflow of Vjosa (called Aoos in Greece) basin belong to *P. thesproticus* (Figs 6, 7, 14 and 15). Observed populations in the Drinos River are not far from the populations in Butrit drainage (Bistrica river) but they are separated by a chain of mountains. Its presence in both river basins might be due to river capture events or the fish could migrate undergound, the same as the species of the genus *Delminichthys* (Palandačić et al., 2012), which is the closest relative to *Pelasgus.* There are numerous springs on both sides of the karstic mountains (J. Vukić, pers. comm). Populations from both river drainages bear the same haplotypes (Fig 14). This suggests recent connection between both river basins, and the northwards extension of *P. thesproticus* distribution in Albania. As summary, known distribution area of *P. thesproticus* has been reduced southwards but extended northwards (Fig 15, 28).

Markova et al. (2014) distinguished in their study of *Barbus* from Albanian region a clade formed by populations from Pavllo river basin and some western Greek rivers. The comparison of my *P. thesproticus* populations also agrees with that distribution pattern, grouping together populations from Bistrica, Pavllo, Kalamas, and Acheroon. However, there is a discrepancy on the situation of Vjosa populations of both genera: in the case of *P. thesproticus* they are lined up with the previously mentioned populations from southernly isoolated rivers, but in the case of *Barbu*s, they were related to the populations from northern rivers in Albania, including Prespa Lake population, suggesting a differente distribution pattern and dispersion processes for both genera.

Figure 28 - Map summarizing the results about *Pelasgus* distribution areas.

4.2 Molecular identification of *Pelasgus* species: the incognita of *P. epiroticus*.

The identification of individuals as *Pelasgus epiroticus* is controverted and their study shows a big incognita. Type locality of *P. epiroticus* is Ioannina Lake and the upper Louros. Since the early 1990s, the population has dramatically diminished in the Ioannina Lake due to several factors, such as the introduction of alien species, high levels of water pollution and overfishing (Crivelli, 2006). This led to the loss of an estimated 90% of the population since 1995, and the inclusion of the species in the IUCN Red List of Threatened Species as CR (Critically Endangered) (Crivelli, 2006). During nearly the last twenty five years, no individual of this species has been found in Ioannina Lake (Perdikaris et al., 2003).

I analyzed three individuals from the uppermost Louros River, i.e. from a part of the type locality of *P. epiroticus*. Their cytochrome b sequences were considered the reference to tentatively identify other individuals as possibly belonging to *P. epiroticus*. Nonetheless, since the presence of *P. epiroticus* in Louros River has not been confirmed yet morphologically, and due to the complete absence of individuals from Ioannina Lake, the species which can be *P. epiroticus* is named *Pelasgus* sp. until its conspecificity is confirmed or disproved. *Pelagus epiroticus* from Ioannina Lake has a very distinctive morphology (Kottelat and Freyhof, 2007a) and to resolve the status of the *Pelasgus* sp. individuals, morphological studies and comparisons between my specimens and the type material stored in the Natural History Museum in Vienna will be needed.

According to sequences analyzes, a total of 27 individuals from six different river basins were included in the *Pelasgus* sp. cluster (Fig 6 and 7), some of them from localities previously considered to belong to the distribution area of *P. thesproticus* described by Kottelat and Freyhof (2007a), see Fig 8, 9 and 28.

I found *Pelasgus* sp. in the Louros and Arachthos river basins and Zaravina Lake on the mainland, and Kandila springs and Stymphalia Lake in the Peloponnesus, occurring in these two latter localities with *Pelasgus stymphalicus*. Zaravina Lake is small endorheic lake located between the river basins of Vjosa, Kalamas and Ioannina lake basins. The individuals sampled from the basins around the lake (Kalamas and Vjose) belong to *P. thesproticus* (Figs6, 7, 14 and 15). There are two possible explanations for the presence of *Pelasgus sp.* in Zaravina Lake: either it is natural distribution range of this evolutionary lineage, or they were introduced.

Zaravina Lake haplotypes are not identical to any of the other haplotypes of *Pelasgus sp*., but they are most similar to those from Kandila and Stymphalia on Peloponnesus. Kandila was suggested to be populated by *Pelasgus epiroticus* from Ioannina Lake (Kottelat and Barbieri 2004). If this is true, then most likely *Pelasgus* sp. could be really *Pelasgus epiroticus*. However, as already mentioned, morphological comparison of adult specimens (my material include only small juveniles) and molecular analysis of the old museum specimens from Ioannina lake are necessary to confirm this hypothesis.

4.3 Stymphalia Lake, Kandila springs (Peloponnesus) and Zaravina Lake (mainland) as an example of introduction.

Stymphalia Lake and Kandila springs are two small water bodies from Peloponnesus, geographically included in the distribution area of *P. stymphalicus* (Economou et al., 2007; Kottelat & Freyhof, 2007a). Kandila springs are located in a small endorheic basin, and even though it is considered to be within *P. stymphalicus* distribution area, individuals collected there did not belong to this species and were tentatively named *Pelasgus* sp. In the case of Stymphalia Lake, which is also endorheic basin, most collected individuals were assigned to *P. stymphalicus*.

Molecular identification of all individuals based on cyt b from both basins uncovered the sympatric occurrence of two species: *P. stymphalicus*, as it was expected, and a second species that was phylogenetically clustering with individuals of *Pelasgus* sp. collected from Louros, Arachthos and Zaravina Lake on the Greek mainland (Figs 6 and 7).

In 1988 and 1989, local fisheries administration stocked *Ctenopharyngodon idella* (grass carp), coming from a fisheries station of Ioannina Lake to Kandila Spring. Since Ioannina Lake is one of the localities from which *Pelasgus epiroticus* type material comes (together with upper Louros), individuals of this species could have been mixed with the stocked *C. idella* individuals (Kottelat & Barbieri, 2004). Due to the occurrence of these species in sympatry, hybridization could occur. To uncover possible cases of hybridization and/or introgression, both mitochondrial and nuclear marker were analyzed for most individuals available from both basins.

Gilles et al. (1998) define in their work the best way to test the presence of introgression and hybridization. In the case of introgression, they consider that it is valid when the introgressed mitochondrial haplotype is clustered with others of different geographic origin with a high

bootstrap support, combined with the presence of hybrid morphology. In our case is complicated to fulfill these two requirements since no morphological studies have been done for the collected individuals (they are mostly small juveniles). Therefore our consideration of introgression events sticks to the first condition. In the case of hybridization, Gilles et al. (1998) state that the presence of hybridization among lineages can only be resolved with the combination of mitochondrial and nuclear markers, as we have done in this study. Also the addition of haplotype networks to classical phylogenetic analysis is a useful tool to decipher cases of hybridization.

Based on the results of S7 sequence analyses, hybridization occurred in both Kandila spring and Stymphalia Lake (Figs 16 to 21). Although there are signs of hybridization also from other nuclear markers, the variability within the dataset for these markers is too low to be able to confirm it with certainty without further analyses. One of the samples from Kandila springs (F1166) had the haplotype 19 for cytochrome b, together with other *P. stymphalicus* individuals from other localities (Stymphalia and Assopos, see figs 6 and 7 and Appendix 2 for more detailed information), but the obtained nuclear sequences for S7 and Rhodopsin gene situated the sample with *Pelasgus* sp. individuals, which can be explained as a case of introgression (Figs 18, 21 and 27). This issue deserves more attention. Further research, combining the use of other more sensitive nuclear markers and morphological analysis, and on larger sample size is needed.

The case of Zaravina Lake is slightly different. It is located near the Kalamas river basin and therefore presence of *P. thesproticus* was expected at that site (Economou et al., 2007). Geiger et al. (2014) analysed samples from this lake and labelled them as *Pelasgus thesproticus*. In their tree, based on COI, this population is related to Arachthos samples, which agrees with my analysis of cyt b (Fig 6, 7). My molecular identification of the collected individuals showed that their tentative identification as *P. thesproticus* was wrong, and they lined up with *Pelasgus* sp. individuals. Zaravina is an endorheic basin, and the comparison of the individuals collected with the ones from Kandila and Stymphalia did not show any close relationship between them. Zaravina appears to be a differentiated group within *Pelasgus* sp. lineage.

5. Conclussions

As a summary of the results obtained in this work, we can conclude that:

- Genus *Pelasgus* comprises seven well supported monophyletic lineages of cytochrome b haplotypes.
- Six of these lineages correspond to already described species.
- The last lineage might correspond to the seventh described species, *Pelasgus epiroticus*, or corresponds to a non-described species, but this cannot be confirmed until material is compared with the type material of *Pelasgus epiroticus*.
- Markers showed to be useful for this phylogenetic study are mitochondrial gene for cytochrome b and intron of the nuclear ribosomal protein gene S7. Other nuclear markers were not variable enough.
- The distribution of *Pelasgus* species described by Kottelat & Freyhof (2007a) does not fully correspond with the distribution observed in this study.
- Occurrence of *P. prespensis*, *P. thesproticus* and *P. marathonicus* was confirmed for a few localities out of the known/expected range, thus changing the knowledge of their distribution areas.
- Occurrence of both *Pelasgus* sp. and *P. stymphalicus* in two localities on the Peloponnese localities Stymphalia Lake and Kandila springs was revealed. This is most probably the consequence of introduction of *Pelasgus* sp. to the native range of *P. stymphalicus.*
- Hybridization was observed between these two species occurring at same localities, but this issue deserves more attention and investigation.
- Existence of a high intraspecific diversity on cytochrome b, partly correlating with geographic distribution, was observed within the species with wider distribution areas (*P. thesproticus*, *P. marathonicus*, *P. stymphalicus* and *Pelasgus* sp.).

6. References

Baker, R. J., Porter, C. A., Patton, J. C., Van Den Bussche, R. A. (2000). Systematics of bats of the family Phyllostomidae based on RAG2 DNA sequences. Occasional Papers, Museum of Texas Technology University 202:1–16.

Banarescu, P. M. (2004): Distribution pattern of the aquatic fauna of the Balkan Peninsula, in: Balkan biodiversity. Pattern and process in the European hotspot, edited by: Griffiths, H. I., Krystufek, B., and Reed, J. M., Kluwer Academic Publishers, Dordrecht, Boston, London, 203–217.

Barbieri, R., Zogaris, S., Kalogianni, E., Stoumboudi, M., Chatzinikolaou, Y., Giakoumi, S., Kapakos, Y., Kommatas, D., Koutsikos, N., Tachos, V., Vardakas, L., Economou, A. N. (2015): Freshwater fishes and lampreys of Greece: An annotated checklist. Hellenic Centre for Marine Research, 130 pp.

Bogutskaya & Zupancic (2003). *Phoxinellus pseudalepidotus* (Teleostei: Cyprinidae), a new species from the Neretva basin with an overview of the morphology of *Phoxinellus* species of Croatia and Bosnia-Herzegovina. Ichthyological Exploration of Freshwaters 14(4):369-383.

Bohlen J, Perdices A, Doadrio I & Economidisi PS. 2006. Vicariance, colonisation, and fast local speciation in Asia Minor and the Balkans as revealed from the phylogeny of spined loaches (Osteichthyes; Cobitidae). Molecular Phylogenetics and Evolution, 39: 552-561.

Bufalino, A.P., Mayden, R.L. (2010). Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. Molecular Phylogenetics and Evolution, 55:274-283.

Buj, I., Šanda, R., Marčić, Z., Ćaleta, M., Mrakovčić, M. (2014): Combining Morphology and Genetics in Resolving Taxonomy–A Systematic Revision of Spined Loaches (Genus *Cobitis*;

Cypriniformes, Actinopterygii) in the Adriatic Watershed. PLoS ONE 9(6): e99833. doi:10.1371/journal.pone.0099833

Calcagnotto, D., Schaefer, S.A., DeSalle, R. (2005): Relationships among characiform fishes inferred from analysis of nuclear and mitochondrial gene sequences. Molecular Phylogenetics and Evolution, 36(1):135-153.

Chen, W.J., Bonillo, C., Lecointre, G. (2003): Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. Molecular Phylogenetics and Evolution, 26(2):262-288

Chow, S., Hazama, K. (1998). Universal primer for S7 ribosomal protein gene intron in fish. Molecular Ecology, 7:1255-1256.

Clement, M., Posada, D. and Crandall, K. (2000). TCS: a computer program to estimate gene genealogies. Molecular Ecology 9(10): 1657-1660

Crivelli, A.J. (2006a). *Pelasgus epiroticus*. The IUCN Red List of Threatened Species 2006: e.T61256A12454831

Crivelli, A.J. (2006b). *Pelasgus prespensis*. The IUCN Red List of Threatened Species 2006: e.T61259A12455146.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9(8), 772.

Doadrio, I. & Carmona, J. A. (1998). Genetic divergence in Greek populations of the genus Leuciscus and its evolutionary and biogeographical implications. Journal of Fish Biology 53(3):591-613.

Dubut V., Fouquet A., Voisin A., Costedoat C., Chappaz R., Gilles A. (2012): From Late Miocene to Holocene: Processes of Differentiation within the *Telestes* Genus (Actinopterygii: Cyprinidae). PlosOne 7/3, e34423.

Durand, J.D., Tsigenopoulos, C.S., Ünlü, E., Berrebi, P. (1999). Phylogeny and Biogeography of the family Cyprinidae in the Middle East inferred from cytochrome b DNA-Evolutionary significance of this region. Molecular Phylogenetics and Evolution, 22(1):91-100.

Economidis, P. S., Banarescu, P. M. (1991). The Distribution and Origins of Freshwater Fishes in the Balkan Peninsula, Especially in Greece. Internationale Revue der gesamten Hydrobiologie und Hydrographie, 76: 257–284. doi:10.1002/iroh.19910760209

Economou A. N., Giakoumi S., Vardakas L., Barbieri R., Stoumboudi M., Zogaris S. (2007): The freshwater ichthyofauna of Greece - an update based on a hydrographic basin survey. Mediterranean Marine Science 8: 91-166

Farias, I., Ortí, G., Sampaio, I., Schneider, H., Meyer, A. (2001). The Cytochrome b Gene as a Phylogenetic Marker: The Limits of Resolution for Analyzing Relationships Among Cichlid Fishes. Journal of Molecular Evolution 53: 89

Felsenstein J (1985). Confidence-limits on phylogenies: An approach using the bootstrap. Evolution, 39:783-791. 148.

Freyhof, J., Kärst, H., Geiger, M. (2014): *Valencia robertae*, a new killifish from southern Greece (Cyprinodontiformes: Valenciidae). Ichthyological Exploration of Freshwaters 24(4): 289-298.

Freyhof, J., Lieckfeldt, D., Bogutskaya, N.G., Ludwig, A. (2006): Phylogenetic position of the Dalmatian genus *Phoxinellus* and description of *Delminichthys*, new genus, (Teleostei: Cyprinidae). Molecular Phylogenetics and Evolution, 38:416-425.

Gamble, T., Berendzen, P.B., Shaffer, H.B., Starkey, D.E., Simons, A.M
Species limits and phylogeography of North American cricket frogs (Acris: Hylidae) Molecular Phylogenetics and Evolution, 48:112–125

Gascuel, O. (1997): BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. Molecular Biology and Evolution, 14(7):685-695.

Galván-Quesada, S., Doadrio, I., Alda, F., Perdices, A., Reina, R.G., García Varela, M., Hernandez, N., Mendoza, A.C., Bermingham, E., Dominguez-Dominguez, O. (2016): Molecular phylogeny and biogeography of the amphidromous fish genus *Dormitator* Gill 1861 (Teleostei: Eleotridae). PLoS ONE 11(4): e0153538.

Geiger, M. F., Herder, F., Monaghan, M. T., Almanda, V., Barbieri, R., Bariche, M., Berrebi, P., Bohlen, J., Casal-Lopez, M., Delmastro, G. B., Denys, G. P. J., Dettai, A., Doadrio, I., Kalogianni, E., Kärst, H., Kottelat, M., Kovačić, M., Laporte, M., Lorenzoni, M., Marčić, Z., Özuluğ, M., Perdices, A., Perea, S., Persat, H., Porcelotti, S., Puzzi, C., Robalo, J., Šanda, R., Schneider, M., Šlechtová, V., Stoumboudi, M., Walter, S., Freyhof, J. (2014): Spatial heterogeneity in the Mediterranean biodiversity hotspot affect barcoding accuracy of its freshwater fishes. Molecular Ecology Resources 14: 1210-1221.

Gilles, A., Costedoat, C., Barascud, B., Voisin, A., Banarescu, P. et al. (2010). Speciation pattern of *Telestes souffia* complex (Teleostei, Cyprinidae) in Europe using morphological and molecular markers. Zoologica Scripta 39: 225-242.

Glez-Peña, D; Gómez-Blanco, D; Reboiro-Jato, M; Fdez-Riverola, F; Posada, D. (2010) ALTER: program-oriented format conversion of DNA and protein alignments. Nucleic Acids Research. Web Server issue. ISSN: 0305-1048

Guinand, B., Durand, J.D., Laroche, J. (2001): Identifying main evolutionary mechanisms shaping genetic variation of Leuciscus cephalus L. 1758 (Cyprinidae) in Western Greece:

discordance between methods. Comptes Rendus De L'Academie Des Sciences Serie III-Sciences De La Vie-Life Sciences, 324(11):1045-1060

Guindon, S., Gascuel, O. (2003): A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology, 52(5):696-704. 145.

Gvozdik, V., Jandzik, D., Lymberakis, P., Jablonski, D., Moravec, J. (2010): Slow worm, *Anguis fragilis* (Reptilia: Anguidae) as a species complex: Genetic structure reveals deep divergences. Molecular Phylogenetics and Evolution 55(2):460-472.

Hall, T. (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium 41: 95-98.

Hansen, J.D., Kaattari, S.L. (1996): The recombination activating gene 2 (RAG2) of the rainbow trout *Oncorhynchus mykiss*. Immunogenetics 44(3): 203-211

Hardman, M. (2004): The phylogenetic relationships among Noturus catfishes (Siluriformes: Ictaluridae) as inferred from mitochondrial gene cytochrome b and nuclear recombination activating gene 2. Molecular Phylogenetics and Evolution (30) 2:395–408

He, S., Mayden, R.L., Wang, X., Wang, W., Tang, K.L., Chen, W.-J., Chen, Y. (2008): Molecular phylogenetics of the family Cyprinidae (Actinpoterygii: Cypriniformes) as evidenced by sequence variation in the first intron of S7 ribosomal protein-coding gene: further evidence from a nuclear gene of the systematic chaos in the family. Molecular Phylogenetics and Evolution, 46(3):818-829.

Hoegg, S.,Vences, M., Brinkmann, H., Meyer, A. (2004): Phylogeny and comparative substitution rates of frogs inferred from sequences of three nuclear genes. Molecular Biology and Evolution, 21(7):1188-1200.

Hsü K. J., Montadert L., Bernoulli D., Cita M. B., Erickson A., Garrison R. E., Kidd R. B., Mélières F., Müller C., Wright R. (1977): History of the Mediterranean salinity crisis. Nature 267: 399-403.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012): Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.

Kottelat, M. (2004): *Salaria economidisi*, a new species of freshwater fish from Lake Trichonis, Greece, with comments on variation in S. fluviatilis (Teleostei: Blenniidae). Revue Suisse De Zoologie, 111:121-137.

Kottelat, M. (2007): Three new species of *Phoxinu*s from Greece and southern France (Teleostei: Cyprinidae). Ichthyolical Exploration of Freshwaters. 18(2):145-162.

Kottelat, M., Barbieri, R. (2004): *Pseudophoxinus laconicus*, a new species of minnow from Peloponnese, Greece, with comments on the West Balkan *Pseudophoxinus* species (Teleostei: Cyprinidae). Ichthyological Exploration of Freshwaters, 15(2), 147–160.

Kottelat, M., Economidis, P.S. (2006): *Squalius orpheus*, a new species of cyprinid fish from Evros drainage, Greece (Teleostei: Cyprinidae). Ichthyological Exploration of Freshwaters 17(2):181-186.

Kottelat, M., Freyhof, J. (2007a): Handbook of European freshwater fishes. Publications Kottelat, 647 pp.

Kottelat, M., Freyhof, J. (2007b): *Pelasgus*, a new genus name for the Balkan species of *Pseudophoxinus* (Teleostei: Cyprinidae). Ichthyological Exploration of Freshwaters, 18: 103-108

Kottelat, M., Barbieri, R., Stoumboudi, M.Th. (2007): *Aphanius almiriensis*, a new species of toothcarp from Greece (Teleostei: Cyprinodontidae). Revue suisse de zoologie, 114(1):13-31

Koutsikos, N., Zogaris, S., Vardakas, L., Tachos, V., Kalogianni, E., Šanda, R., Chatzinikolaou, Y., Giakoumi, S., Economidis, P.S., Economou, A.N. (2012): Recent distributional contributions to the freshwater ichthyofaunal of Greece. Mediterranean Marine Science 13: 268-277.

Librado, P. & J. Rozas, (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452

Lewis-Oritt, N., Van den Bussche, R.A., Baker, J.R. (2001): Molecular evidence for evolution of piscivory in *Noctilio* (Chiroptera: Noctilionidae). Journal of Mammalogy, 82(3):748–759.

Lovejoy, N.R., Collette, B.B. (2001): Phylogenetic relationships of new world needlefishes (Teleostei : Belonidae) and the biogeography of transitions between marine and freshwater habitats Copeia 2001(2):324-338

Machordom, A. & Doadrio, I. (2001). Evidence of a Cenozoic Betic-Kabilian connection based on freshwater fish phylogeography (Luciobarbus, Cyprinidae). Molecular Phylogenetics and Evolution 18: 252-263.

Marková S., Šanda R., Crivelli A., Shumka S., Wilson I.F., Vukić J., Berrebi P., Kotlík P. (2010): Nuclear and mitochondrial DNA sequence data reveal the evolutionary history of Barbus (Cyprinidae) in the ancient lake systems of the Balkans. Molecular Phylogenetics and Evolution 55: 488-500.

Martens, K. (2010): The International Year of Biodiversity. Hydrobiologia, 637:1-2.

Maurakis, E. G., Pritchard, M. K., Economidis, P. S. (2001). Historical relationships of mainland river drainages in Greece. BIOS 6: 109-124.

Neves, J.M.M., Lima, S.M.Q., Mendes, L.F., Torres, R.A., Pereira, R.J., Mott, T. (2016): Population Structure of the Rockpool Blenny *Entomacrodus vomerinus* Shows Source-Sink Dynamics among Ecoregions in the Tropical Southwestern Atlantic. PLoS ONE 11(6): e0157472.

Oikonomou, A., Leprieur, F., Leonardos, I.D. (2014) Biogeography of freshwater fishes of the Balkan Peninsula. Hydrobiologia 738:205-220.

Page, R.D.M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12: 357-358.

Palandačić, A., Matschiner, M., Zupančič, P. (2012). Fish migrate underground: the example of *Delminichthys adspersus* (Cyprinidae). Molecular Ecology, 5, 1658-1671.

Perea, S., Böhme, M., Zupančič, P., Freyhof, J., Šanda, R., Özuluğ, M., Abdoli, A., Doadrio, I. (2010). Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. BMC Evolutionary Biology 10: 265.

Picq, S., Alda, F., Krahe, R., Bermingham, E. (2014): Miocene and Pliocene colonization of the Central American Isthmus by the weakly electric fish *Brachyhypopomus occidentalis* (Hypopomidae, Gymnotiformes). Journal of Biogeography 41(8):1520-1532.

Poulakakis, N., Lymberakis, P., Valakos, E., Pafilis, P., Zouros, E., Mylonas, M., (2005). Phylogeography of Balkan wall lizard (*Podarcis taurica*) and its relatives inferred from mitochondrial DNA sequences. Molecular Ecology 14: 2433-2443.

Quenouille, B., Bermingham, E., Planes, S. (2004). Molecular systematics of the damselfishes (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA sequences. Molecular Phylogenetics and Evolution, 31: 66-88.

Rambaut, A. (2012). Figtree 1.4.2. Distributed by the author via http://tree.bio.ed.ac.uk/software/figtree/

Robalo, J., Almada, V.C., Levy, A., Doadrio, I. (2007). Re-examination and phylogeny of the genus *Chondrostoma* based on mitochondrial and nuclear data and the definition of 5 new genera. Molecular Phylogenetics and Evolution, 42:362-372.

Rögl F. (1999): Mediterranean and Paratethys fact and hypotheses of an Oligocene to Miocene paleogeography (short overview). Geologica Carpathica 50: 339-349.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Systematic Biology, 61(3), 539–542.

Rozas, J. and Rozas, R. (1995). DnaSP, DNA sequence polymorphism: an interactive program for estimating Population Genetics parameters from DNA sequence data. Computer Applications in the Biosciences. 11: 621-625.

Santos, A.M., Cabezas, M.P., Tavares, A.I., Xavier, R., Branco, M. (2015). tcsBU: a tool to extend TCS network layout and visualization. Bioinformatics, btv636

Stierandová, S., Vukić, J., Vasil'evac, E.D., Zogaris S., Shumkae, S., Halačka, K., Vetešník, L., Švátora, M., Nowak, M., Stefanov, T., Koščoh, J., Mendel, J. (2016): A multilocus assessment of nuclear and mitochondrial sequence data elucidates phylogenetic relationships among European spirlins (Alburnoides, Cyprinidae). Molecular Phylogenetics and Evolution 94B: 479–491

Solà, E., Sluys, R., Gritzalis, K., Riutort, M. (2013). Fluvial basin history in the northeastern Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelmintes, Tricladida, Dugesiidae). Molecular Phylogenetics and Evolution 66: 877-888.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729.

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.

Towner, P. (1991). Purification of DNA. In Essential molecular biology, A practical approach. Edited by: Brown T. Oxford: Oxford University Press; 47-68. 138.

Tsoumani, M., Georgiadis, A., Giantsis, I.A., Leonardos, I., Apostolidis, A.P. (2014): Phylogenetic relationships among Southern Balkan *Rutilus* species inferred from cytochrome b sequence analysis: Micro-geographic resolution and taxonomic implications. Biochemical Systematics and Ecology, 54:172-178

Tziritis, E.P. (2014): Environmental monitoring of Micro Prespa Lake basin (Western Macedonia, Greece): hydrogeochemical characteristics of water resources and quality trends. Environmental Monitoring and Assessment, 186(7):4553-4568.

Wang, X., Li, J., He, S. (2007) Molecular evidence for the monophyly of East Asian groups of Cyprinidae (Teleostei: Cypriniformes) derived from the nuclear recombination activating gene 2 sequences. Molecular Phylogenetics and Evolution, 42(1):157-170.

Willet, C.E., Zapata, A.G, Hopkins, N., Steiner, L.A. (1997): Expression of Zebrafish rag Genes during Early Development Identifies the Thymus. Developmental Biology 182(2):331-341

Yang L., Sado T., Hirt M. V., Pasco-Viel E., Arunachalam M., Li J., Wang X., Freyhof J., Saitoh K., Simons A. M., Miya M., He S., Mayden R. L. (2015): Phylogeny and polyploidy:

Resolving the classification of cyprinine fishes (Teleostei: Cypriniformes). Molecular Phylogenetics and Evolution 85: 97-116.

Zardoya R., Doadrio I. (1999): Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. Journal of Molecular Evolution 49: 227-237.

Zardoya, R., Economidis, P.S., Doadrio, I. (1999): Phylogenetic Relationships of Greek Cyprinidae: Molecular Evidence for at Least Two Origins of the Greek Cyprinid Fauna. Molecular Phylogenetics and Evolution 13: 122-131.

Zogaris, S., Economou, A., Dimopoulos, P. (2009) Ecoregions in the Southern Balkans: should they be revised? Journal of Environmental Management, 43:682-697.

7. Appendices

Appendix 1- Identification key to *Pelasgus* species. Taken and modified from Kottelat & Freyhof (2007a)

Keys to species of *Pelasgus*

1a – 58–73 + 2 scales in midlateral row; scales very small, not overlapping, deeply embedded in skin (Lakes Prespa).
.... *P. prespensis*

1b – 35–58 + 1–2 scales in midlateral row; scales overlapping, distinct, with or without scale pockets covering part of them (Fig. 45).
.... 2

2a – 53–58 + 2 scales in midlateral row; 16–23 pored lateral line scales (Lake Pamvotis).
.... *P. epiroticus*

2b – 35–48 + 1–2 scales in midlateral row; 0–14 pored lateral line scales.
.... 3

3a – 6–6½ branched anal rays (Peloponnese).
.... *P. laconicus*

3b – 7–7½ branched anal rays.
.... 4

4a – Depth of caudal peduncle about 1.8–2.0 times in its length (Lake Ohrid and Drin drainage).
.... *P. minutus*

4b – Depth of caudal peduncle 1.4–1.8 times in its length.
.... 5

5a – ½7–8½ scale rows on caudal peduncle; 36–41 + 2 scales in midlateral row (eastern Greece).
.... *P. marathonicus*

5b – ½8–9½ (mode ½9½) scale rows on caudal peduncle; 38–43 + 1–3 scales in midlateral row.
.... 6

6a – Body ovoid in section; scale pockets exposed in anterior half of body; depth of caudal peduncle 11–13 % SL; in both sexes, belly rounded, ventral profile from pelvic to caudal base clearly concave (Peloponnese, western Greece).
.... *P. stymphalicus*

6b – Body compressed; scale pockets not or only very weakly exposed; depth of caudal peduncle 12–16 % SL; ventral profile from pelvic to caudal base almost straight in male, concave in female (northwestern Greece).
.... *P. thesproticus*

Appendix 2- Detailed information of the samples

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		Short		Allele A	Allele B	Allele A	Allele B
							Allele A	Allele B	Allele A	Allele B				
F1227	*Pelasgus stymphalicus*	Vergas stream	Kotichy Lake system	N 37°59'3.09" E 21°25'25.29"	PstVer	Hap 18	Hap 2		Hap 2		Hap 15		Hap 3	
F1228	*Pelasgus stymphalicus*					Hap 17	X (*)		x					
F1213	*Pelasgus stymphalicus*					Hap 16								
F1230	*Pelasgus stymphalicus*					Hap 16								
F1318	*Pelasgus stymphalicus*	Kalivakia, upstream of reservoir	Pinios - Pelloponesius	N 37°54'31.13" E 21°32'44.26"	PstPin	Hap 16	x		x		Hap 15		Hap 3	
F1319	*Pelasgus stymphalicus*					Hap 16								
F1320	*Pelasgus stymphalicus*					Hap 16								
F1321	*Pelasgus stymphalicus*			N 37°50'36.35" E 21°32'18.16"		Hap 16								
F1350	*Pelasgus stymphalicus*	tributary of Pinios, Ag. Nicolaos				Hap 16	Hap 13 / Hap 14		Hap 11 / Hap 12		x		Hap 3	
F1351	*Pelasgus stymphalicus*					Hap 16								
F1401	*Pelasgus stymphalicus*	below Alfios dam, near Olympiada	Alfios	N 37°38'20.50" E 21°34'51.89"	PstAlf	Hap 15	Hap 7 / Hap 8		Hap 7 / Hap 8		Hap 9		Hap 3	
F1402	*Pelasgus stymphalicus*					Hap 15								
F1400	*Pelasgus stymphalicus*					Hap 11								
G331	*Pelasgus stymphalicus*	Enipeas r., Salmoni		N 37°39'39.26" E 21°32'32.49"		Hap 11	Hap 1		Hap 1		Hap 9		Hap 3	
G332	*Pelasgus stymphalicus*					Hap 10	Hap 17		Hap 13		Hap 9		Hap 3 / Hap 6	
G238	*Pelasgus stymphalicus*	road to Artiki	Peristeras	N 37°16'9.45" E 21°47'6.37"	PstPer	Hap 12	x		x		Hap 9		Hap 3	
G239	*Pelasgus stymphalicus*					Hap 12	Hap 4		Hap 4		Hap 9		Hap 3	
F1542	*Pelasgus stymphalicus*	Meligalas	Pamissos	N 37°13'20.12" E 21°56'57.94"	PstPam	Hap 13	Hap 4 / Hap 15		Hap 4 / Hap 1		Hap 9		Hap 3	
F1543	*Pelasgus stymphalicus*					Hap 13	Hap 16		Hap 4		Hap 9		Hap 3	
F1544	*Pelasgus stymphalicus*					Hap 13								
F1545	*Pelasgus stymphalicus*					Hap 14								
F1546	*Pelasgus stymphalicus*					Hap 13								
C2039	*Pelasgus stymphalicus*	Nafpaktos	Mornos	N 38°23'51.02" 21°52'4.50"	PstMor	Hap 31	Hap 2 / Hap 1		Hap 2 / Hap 1		Hap 15 / Hap 9		Hap 3	
C2040	*Pelasgus stymphalicus*					Hap 25								

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		Short		Allele A	Allele B	Allele A	Allele B
							Allele A	Allele B	Allele A	Allele B				
F1010	*Pelasgus stymphalicus*	Limnitsa		N 38°30'26.86" E 21°59'56.30"		Hap 25	x		x				Hap 3	
F1011	*Pelasgus stymphalicus*					Hap 25	Hap 1		Hap 1		Hap 14		Hap 3	
D2107	*Pelasgus stymphalicus*	Chiliadou spring		N 38°24'25.57" E 21°55'3.58"		Hap 25								
D2109	*Pelasgus stymphalicus*					Hap 25								
D2108	*Pelasgus stymphalicus*					Hap 30	Hap 2 / Hap 1		Hap 2 / Hap 1				Hap 3	
C2066	*Pelasgus stymphalicus*	Kokori	Evinos	N 38°22'46.11" E 21°33'13.45"	PstEvi	Hap 26	x		x		Hap 15 / Hap 18		Hap 3	
F983	*Pelasgus stymphalicus*	Vania		N 38°27'9.30" E 21°42'28.84"		Hap 26								
F984	*Pelasgus stymphalicus*					Hap 26								
F985	*Pelasgus stymphalicus*					Hap 26								
G803	*Pelasgus stymphalicus*	Trichonis Lake	Acheloos	N 38°35'20.19" E 21°28'2.68"	PstAchT	Hap 26								
G801	*Pelasgus stymphalicus*					Hap 27	x		x		Hap 15		Hap 3	
D2152	*Pelasgus stymphalicus*					Hap 29	Hap 1		Hap 1		Hap 16 / Hap 17		Hap 3	
D2153	*Pelasgus stymphalicus*					Hap 28								
F1126	*Pelasgus stymphalicus*		Milos - Lerni spring	N 37°33'10.08" E 22°43'2.89"	PstMil	Hap 14	Hap 3 / Hap 4		Hap 3 / Hap 4		Hap 9		Hap 4	
F1127	*Pelasgus stymphalicus*					Hap 14	x		x		Hap 9		Hap 3	
F1142	*Pelasgus stymphalicus*		Assopos - Peloponnesius	N 37°50'4.86" E 22°36'21.33	PstAss	Hap 24	Hap 1 / Hap 15		Hap 1		Hap 9		Hap 3	
F1143	*Pelasgus stymphalicus*					Hap 23	Hap 9 / Hap 4		Hap 2 / Hap 4		Hap 9		Hap 3	
F1144	*Pelasgus stymphalicus*					Hap 23								
F1145	*Pelasgus stymphalicus*					Hap 19								
F1146	*Pelasgus stymphalicus*					Hap 19								
F1165	***Pelasgus stymphalicus***		Kandila springs	N 37°44'45.21" E 22°20'2.79"	PstKan	Hap 19	x		x		x		Hap 5	
F1166	***Pelasgus stymphalicus***					Hap 19	Hap 10 / Hap 11		Hap 9 / Hap 10		Hap 3		x	
F1154	*Pelasgus stymphalicus*	inflow of Stymphalia lake	Stymphalia lake	N 37°52'53.13" E 22°30'48.17"	PstSty	Hap 19	Hap 11 / Hap 1		Hap 3 / Hap 1		Hap 9		Hap 3	
F1156	*Pelasgus stymphalicus*					Hap 19	Hap 11 / Hap 1		Hap 3 / Hap 1		x		Hap 3	
F1159	*Pelasgus stymphalicus*					Hap 19	Hap 9 / Hap 4		Hap 2 / Hap 4		Hap 19 / Hap 9		Hap 3 / Hap 5	

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		Short		Allele A	Allele B	Allele A	Allele B
							Allele A	Allele B	Allele A	Allele B				
F1155	*Pelasgus stymphalicus*					Hap 14	x		x				x	
F1158	*Pelasgus stymphalicus*					Hap 14	x		x		Hap 9		Hap 5	
F1157	*Pelasgus stymphalicus*					Hap 21	Hap 5 / Hap 6		Hap 5 / Hap 6		Hap 9		Hap 3 / Hap 5	
F1161	*Pelasgus stymphalicus*					Hap 20	Hap 4 / Hap 1		Hap 4 / Hap 1		Hap 9		Hap 3	
H194	*Pelasgus stymphalicus*					Hap 22	x		x		x		Hap 3 / Hap 5	
H196	*Pelasgus stymphalicus*					Hap 14	Hap 4 / Hap 6		Hap 4 / Hap 6		Hap 3 / Hap 9		Hap 3	
G353	*Pelasgus thesproticus*	Kokitos r., upper Acheroon drainage	Acheroon	N 39°26'49.13" E 20°30'3.36"	PthAch	Hap 2								
G355	*Pelasgus thesproticus*					Hap 2								
F586	*Pelasgus thesproticus*					Hap 2								
F587	*Pelasgus thesproticus*					Hap 8	x		x		Hap 3		Hap 1	
F588	*Pelasgus thesproticus*					Hap 2	Hap 24		Hap 17		Hap 3		Hap 1	
F589	*Pelasgus thesproticus*					Hap 2								
F590	*Pelasgus thesproticus*					Hap 2								
F591	*Pelasgus thesproticus*					Hap 2								
F592	*Pelasgus thesproticus*					Hap 2								
Pth1	*Pelasgus thesproticus*	reservoir Chaskovo, near Paramythia		N 39°28'45.74" E 20°27'41.72"		Hap 2								
Pth2	*Pelasgus thesproticus*					Hap 2								
G811	*Pelasgus thesproticus*					Hap 2					Hap 3		Hap 1	
G411	*Pelasgus thesproticus*	river Messonghi	Kerkyra island	N 39°28'36.65" E 19°54'42.68"	PthKer	Hap 6	Hap 18 / Hap 19		Hap 14 / Hap 15		Hap 2 / Hap 3		Hap 1 / Hap 9	
G412	*Pelasgus thesproticus*					Hap 6	Hap 20 / Hap 21		Hap 14 / Hap 15		Hap 1		Hap 1 / Hap 9	
G413	*Pelasgus thesproticus*					Hap 5								
G415	*Pelasgus thesproticus*					Hap 7	Hap 22 / Hap 23		Hap 16 / Hap 15		Hap 5		Hap 1 / Hap 9	
G416	*Pelasgus thesproticus*					Hap 5	Hap 22 / Hap 23		Hap 16 / Hap 15		Hap 2 / Hap 3		Hap 1 / Hap 9	
G417	*Pelasgus thesproticus*					Hap 6								
Kor42	*Pelasgus thesproticus*					Hap 5								
F756	*Pelasgus thesproticus*	Anakoli spring	Kalamas	N 39°33'46.63"	PthKalA	Hap 4	x		x		Hap 2 / Hap 4		Hap 1	

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7) Long (Allele A / Allele B)	Ribosomal Protein Gene S7 (S7) Short (Allele A / Allele B)	Recombination Activating Gene (RAG) (Allele A / Allele B)	Rhodopsin (Allele A / Allele B)
F757	*Pelasgus thesproticus*			E 20°14'28.73"		Hap 4	x	x	Hap 2	Hap 1
G806	*Pelasgus thesproticus*					Hap 4				
F310	*Pelasgus thesproticus*	Mesopotam	Bistrica	N 39°54'29.23" E 20° 6'14.61"	PthBisM	Hap 1	Hap 24	Hap 17	Hap 2 / Hap 3	Hap 1
F311	*Pelasgus thesproticus*					Hap 1				
Sc276	*Pelasgus thesproticus*	Bistrica river		N 39°55'11.89" E 20°10'7.11"		Hap 1				
A279	*Pelasgus thesproticus*					Hap 9				
D411	*Pelasgus thesproticus*	Cuka		N 39°51'24.3" E 20°02'10.7"		Hap 9	Hap 24	Hap 17	x	Hap 1
D392	*Pelasgus thesproticus*	Hormove Drinos River	Vjosa basin	N 40°13'47.20" E 20° 5'7.60"	PthVjo	Hap 9	Hap 25 / Hap 24	Hap 18 / Hap 17	x	Hap 1
F262	*Pelasgus thesproticus*	Burimi i Viroit spring, Drinos river		N 40° 5'53.13" E 20° 7'10.61"		Hap 3	Hap 24	Hap 17	Hap 2	Hap 1
F263	*Pelasgus thesproticus*					Hap 9	Hap 24	Hap 17	Hap 2 / Hap 3	Hap 1
F264	*Pelasgus thesproticus*					Hap 3				
F265	*Pelasgus thesproticus*					Hap 3				
G819	*Pelasgus thesproticus*	Shalles	Pavllo	N 39°43'42.24" E 20° 8'30.00"	PthPavS	Hap 3				
G820	*Pelasgus thesproticus*					Hap 1	Hap 24	Hap 17	x	Hap 1
G821	*Pelasgus thesproticus*					Hap 3	x	x	Hap 3	Hap 1
G823	*Pelasgus thesproticus*					Hap 3				
D1883	*Pelasgus minutus*	Ohrid Lake	Ohrid Lake	N 41° 6'8.23" E 20°48'50.69"	PmiOhr	Hap 51	Hap 48	Hap 33	x	Hap 17 / Hap 18
D1884	*Pelasgus minutus*					Hap 49	Hap 48 / Hap 49	Hap 33 / Hap 34	Hap 25 / Hap 26	Hap 17
D1885	*Pelasgus minutus*					Hap 52	Hap 48 / Hap 50	Hap 33	Hap 25	x
D1886	*Pelasgus minutus*					Hap 51				
D1887	*Pelasgus minutus*					Hap 51				
D1888	*Pelasgus minutus*					Hap 51				
D1890	*Pelasgus minutus*					Hap 53	Hap 51 / Hap 52	Hap 35 / Hap 36	x	Hap 17
D1891	*Pelasgus minutus*					Hap 51				
D1893	*Pelasgus minutus*					Hap 54	Hap 52	Hap 36	x	x

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		**Short**		**Allele A**	**Allele B**	**Allele A**	**Allele B**
							Allele A	**Allele B**	**Allele A**	**Allele B**				
D1894	*Pelasgus minutus*					Hap 50	Hap 52		Hap 36		x		Hap 17	
D300	*Pelasgus prespensis*	Beli Hill	Prespa lake	N 40°52'20.50" E 20°56'15.60"	Ppr	Hap 68	x		x		Hap 23 / Hap 24		x	
Ppres1	*Pelasgus prespensis*	Agios Germanos		N 40°49'46.48" E 21° 6'17.04"		Hap 67	x		x		x		Hap 13	
D305	*Pelasgus prespensis*	Golo Brdo, Devoll river	Devoll river	N 40°42'37.45" E 20°51'49.21"		Hap 70	x		x				Hap 14 / Hap 16	
H984	*Pelasgus prespensis*					Hap 69	x		x		x		Hap 14 / Hap 13	
H985	*Pelasgus prespensis*					Hap 71	x		Hap 37 / Hap 38				Hap 13 / Hap 15	
Pep1	***Pelasgus* sp. (?)**	Louros, springs, Terovo area	Louros	N 39°25'15.54" E 20°50'29.90”	PepLou	Hap 66	Hap 26		Hap 9		x		Hap 1	
Pep2	***Pelasgus* sp. (?)**					Hap 66	x		x				Hap 1	
Pep3	***Pelasgus* sp. (?)**					Hap 66	Hap 26 / Hap 27		Hap 9 / Hap 19		Hap 3		Hap 1	
F480	***Pelasgus* sp. (?)**	Agios Georgios		N 39°16'14.56" E 20°50'55.31"		Hap 66	x		x		Hap 3 / Hap 7		Hap 1 / Hap 10	
F486	***Pelasgus* sp. (?)**					Hap 55	x		x		Hap 3		Hap 1	
F487	***Pelasgus* sp. (?)**					Hap 55	x		x		Hap 3		Hap 1	
F488	***Pelasgus* sp. (?)**					Hap 55	Hap 26 / Hap 27		Hap 9 / Hap 19		x		Hap 1	
G824	***Pelasgus* sp. (?)**		Zaravina lake	N39°54'03.0" E20°30'24.6"	PepKZa	Hap 59	Hap 32		Hap 20		Hap 4		Hap 12	
G827	***Pelasgus* sp. (?)**					Hap 59								
G826	***Pelasgus* sp. (?)**					Hap 60								
G828	***Pelasgus* sp. (?)**					Hap 58								
H1362	***Pelasgus* sp. (?)**					Hap 56	Hap 28		Hap 20		Hap 4 / Hap 6		Hap 12	
H1366	***Pelasgus* sp. (?)**					Hap 57	Hap 28		Hap 20		Hap 6		x	
H1368	***Pelasgus* sp. (?)**					x	Hap 29 / Hap 28		Hap 21 / Hap 20		x		x	
F1162	***Pelasgus* sp. (?)**		Kandila springs	N 37°44'45.21" E 22°20'2.79"	PepKan	Hap 62					Hap 3 / Hap 13		Hap 5	
F1163	***Pelasgus* sp. (?)**					Hap 62							Hap 5	
F1164	***Pelasgus* sp. (?)**					Hap 62					Hap 12		Hap 5	
F1167	***Pelasgus* sp. (?)**					Hap 62	Hap 11		Hap 10		Hap 3 / Hap 11		Hap 5	

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		Short		Allele A	Allele B	Allele A	Allele B
							Allele A	Allele B	Allele A	Allele B				
H1345	***Pelasgus* sp. (?)**					Hap 62							Hap 5 / Hap 11	
H1346	***Pelasgus* sp. (?)**					Hap 63	Hap 33 / Hap 11		Hap 24 / Hap 10		Hap 3 / Hap 11		Hap 5	
H1347	***Pelasgus* sp. (?)**					x					Hap 3 / Hap 13		Hap 5	
F1160	***Pelasgus* sp. (?)**	inflow	Stymphalia lake	N 37°52'53.13" E 22°30'48.17"	PepSty	Hap 61	x		x		Hap 9 / Hap 10		Hap 3 / Hap 5	
C2100	***Pelasgus* sp. (?)**	Acropotamia	Arachthos	N 39° 5'29.64" E 21° 1'34.84"		Hap 64	Hap 33 / Hap 27		Hap 24 / Hap 19		Hap 3		Hap 1	
F800	***Pelasgus* sp. (?)**	Arta		N 39° 9'56.45" E 20°59'38.49"		Hap 65	Hap 10 / Hap 30		Hap 9 / Hap 22		Hap 8		Hap 1 / Hap 10	
F801	***Pelasgus* sp. (?)**					Hap 65	Hap 10 / Hap 31		Hap 9 / Hap 23		Hap 3		Hap 1	
F802	***Pelasgus* sp. (?)**					Hap 65					Hap 3		Hap 1	
F803	***Pelasgus* sp. (?)**					Hap 65							Hap 1	
G90	*Pelasgus marathonicus*		Sperchios	N 38°54'14.33" E 22°17'30.22"	PmaSpe	Hap 32	Hap 34		Hap 25		Hap 20		Hap 1	
G91	*Pelasgus marathonicus*					Hap 32								
G92	*Pelasgus marathonicus*					Hap 32								
G799	*Pelasgus marathonicus*	chanell near Lamia		N 38°50'54.60" E 22°25'54.46"		Hap 41	Hap 34		Hap 25		Hap 20		Hap 1	
G800	*Pelasgus marathonicus*					Hap 40	Hap 34		Hap 25		Hap 19 / Hap 20		Hap 1	
C2224	*Pelasgus marathonicus*		Xerias	N 39°10'40.44" E 22°46'11.24"	PmaXer	Hap 45								
C2225	*Pelasgus marathonicus*					Hap 45	Hap 34 / Hap 35		Hap 25		Hap 20		Hap 1	
C2226	*Pelasgus marathonicus*					Hap 45								
C2227	*Pelasgus marathonicus*					Hap 45								
C2228	*Pelasgus marathonicus*					Hap 45	Hap 35		Hap 25		Hap 20		Hap 1	
D2088	*Pelasgus marathonicus*	Beotia - Akrefnio	Kifissos – Beotia	N 38°26'1.06" E 23°14'34.92"	PmaKifB	Hap 39	Hap 36 / Hap 34		Hap 26 / Hap 25		x		Hap 1	
D2089	*Pelasgus marathonicus*					Hap 34								
D2090	*Pelasgus marathonicus*					Hap 34								
F741	*Pelasgus marathonicus*	Beotia		N 38°27'2.12"		Hap 38								

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		Short					
							Allele A	Allele B	Allele A	Allele B	Allele A	Allele B	Allele A	Allele B
F742	*Pelasgus marathonicus*			E 22°53'3.02"		Hap 38	Hap 35		Hap 25		Hap 19 / Hap 20		Hap 1	
F743	*Pelasgus marathonicus*					Hap 37	Hap 34 / Hap 35		Hap 25		Hap 19 / Hap 20		Hap 1	
F739	*Pelasgus marathonicus*					Hap 34								
F740	*Pelasgus marathonicus*					Hap 34	Hap 34 / Hap 35		Hap 25		Hap 19 / Hap 20		Hap 1	
F1062	*Pelasgus marathonicus*					Hap 42	Hap 34 / Hap 35		Hap 25		Hap 19 / Hap 20		Hap 1 / Hap 2	
F1063	*Pelasgus marathonicus*					Hap 42								
F1064	*Pelasgus marathonicus*		Kifissos – Athens	N 38° 6'26.76" E 23°48'25.53"	PmaKifA	Hap 42								
F1065	*Pelasgus marathonicus*					Hap 42								
F1066	*Pelasgus marathonicus*					Hap 42								
F1049	*Pelasgus marathonicus*					Hap 36	Hap 40 / Hap 41		Hap 28 / Hap 27		Hap 19		Hap 2	
F1051	*Pelasgus marathonicus*					Hap 36	Hap 37 / Hap 39		Hap 25 / Hap 27		Hap 3		Hap 2	
F1050	*Pelasgus marathonicus*					Hap 44	x		x		Hap 19		Hap 2	
F1053	*Pelasgus marathonicus*	Kato Souli wetland	Kato Souli – Marathon	N 38° 8'56.76" E 24° 0'59.44"	PmaKatS	Hap 44	Hap 37 / Hap 38		Hap 25 / Hap 27		Hap 19		Hap 2	
F1054	*Pelasgus marathonicus*					Hap 44	Hap 38 / Hap 39		Hap 27		Hap 19		Hap 2	
F1055	*Pelasgus marathonicus*					Hap 44	Hap 38 / Hap 39		Hap 27					
F1056	*Pelasgus marathonicus*					Hap 44	Hap 42 / Hap 39		Hap 29 / Hap 27		Hap 19		Hap 2	
F1057	*Pelasgus marathonicus*					Hap 44	x		x		Hap 19		Hap 2	
F1042	*Pelasgus marathonicus*					Hap 43								
F1043	*Pelasgus marathonicus*					Hap 43	Hap 37		Hap 25		Hap 19		Hap 2	
F1044	*Pelasgus marathonicus*		Vraonas	N 37°54'52.63" E 23°57'28.64"	PmaVra	Hap 43	Hap 37		Hap 25		Hap 19		Hap 2	
F1045	*Pelasgus marathonicus*					Hap 43								
F1046	*Pelasgus marathonicus*					Hap 43								

Sample code	Species according to cytochrome b	Locality	River basin	Coordinates	Haplotype codes	Cytochrome b (Cyt b)	Ribosomal Protein Gene S7 (S7)				Recombination Activating Gene (RAG)		Rhodopsin	
							Long		Short		Allele A	Allele B	Allele A	Allele B
							Allele A	Allele B	Allele A	Allele B				
G814	*Pelasgus marathonicus*	former Xiniada lake basin	Pinios – Thessaly	N 39° 5'44.38" E 22°11'37.33"	PmaPin	Hap 35	Hap 35		Hap 25		Hap 19 / Hap 20		Hap 1	
G815	*Pelasgus marathonicus*					Hap 34	Hap 34		Hap 25		Hap 19 / Hap 20		Hap 1	
G816	*Pelasgus marathonicus*					Hap 33								
G817	*Pelasgus marathonicus*					Hap 33	Hap 34 / Hap 35		Hap 25		Hap 19 / Hap 20		Hap 1	
G818	*Pelasgus marathonicus*					Hap 32	Hap 34		Hap 25		Hap 20		Hap 1	
F1592	*Pelasgus laconicus*	Sparta	Evrotas	N 37° 5'28.71" E 22°25'38.74"	PlaEvr	Hap 46	Hap 43		Hap 30		Hap 21 / Hap 22		Hap 7	
F1593	*Pelasgus laconicus*					Hap 48	Hap 43 / Hap 44		Hap 30		Hap 21 / Hap 22		Hap 7	
F1594	*Pelasgus laconicus*					Hap 48	Hap 45		Hap 31		Hap 21		Hap 7	
F1597	*Pelasgus laconicus*					Hap 46	Hap 43 / Hap 44		Hap 30		Hap 21 / Hap 22		Hap 7	
F1598	*Pelasgus laconicus*					x	x		x		Hap 21		x	
F1621	*Pelasgus cf. laconicus*	Paparis stream	Alfios – upper part	N 37°23'14.15" E 22°15'48.73"	PlaUAl	Hap 47	Hap 46 / Hap 47		Hap 31 / Hap 32		Hap 21		Hap 7	
F1622	*Pelasgus cf. laconicus*					Hap 46	x		x		Hap 21		Hap 7 / Hap 8	
F1623	*Pelasgus cf. laconicus*					Hap 46	Hap 46		Hap 31		Hap 21		Hap 7	
F1624	*Pelasgus cf. laconicus*					x	Hap 46 / Hap 47		Hap 31 / Hap 32		Hap 21		Hap 7	
F1625	*Pelasgus cf. laconicus*					Hap 46	x		x		Hap 21		Hap 7	

*Samples marked with (x) are samples sequenced but whose sequences were not used due to low quality.

Appendix 3 – Genomic DNA Mini Kit (Tissue) Protocol (Geneaid Biotech, Taiwan). Taken and modified from www.geneaid.com.

Genomic DNA Mini Kit (Tissue)

For research use only

Sample: up to 30 mg of tissue (tailsnips, liver, kidney, brain, adipose tissue, earpunches, insects etc.)
Yield: 10-20 µg (0.5 cm of mouse tail, 20 mg of mouse liver), 20-50 µg (20 mg of mouse kidney)
Format: spin column
Time: within 30 minutes
Elution volume: 30-200 µl
Storage: dry at room temperature (15-25ºC) for up to 1 year without showing any reduction in performance

Introduction

The Genomic DNA Mini Kit (Tissue) was designed specifically for purifying total DNA (including genomic, mitochondrial and viral DNA) from a variety of tissue and insect samples. The provided micropestle can efficiently homogenize tissue samples to shorten the time in the Lysis Step. Proteinase K and chaotropic salt are used to lyse cells and degrade protein, allowing DNA to be easily bound by the glass fiber matrix of the spin column. Once any contaminants have been removed, using a Wash Buffer (containing ethanol), the purified DNA is eluted by a low salt Elution Buffer, TE or water. The entire procedure can be completed without phenol/chloroform extraction or alcohol precipitation. The purified DNA (approximately 20-30 kb) is suitable for use in PCR or other enzymatic reactions.

Quality Control

The quality of the Genomic DNA Mini Kit (Tissue) is tested on a lot-to-lot basis by isolating genomic DNA from a 20 mg mouse liver sample. The purified DNA (more than 10 µg with an A260/A280 ratio of 1.8-2.0) is quantified with a spectrophotometer and analyzed by electrophoresis.

Kit Contents

Component	GT004	GT050	GT100	GT300
GT Buffer	3 ml	30 ml	30 ml	75 ml
GBT Buffer	4 ml	40 ml	40 ml	75 ml
W1 Buffer	2 ml	45 ml	45 ml	130 ml
Wash Buffer* (Add Ethanol)	1 ml (4 ml)	25 ml (100 ml)	25 ml (100 ml)	50 ml (200 ml)
Proteinase K** (Add ddH_2O)	1 mg (0.1 ml)	11 mg (1.1 ml)	11 mg x 2 (1.1 ml x 2)	65 mg (6.5 ml)
Elution Buffer	1 ml	30 ml	30 ml	75 ml
GS Columns	4	50	100	300
2 ml Collection Tubes	8	100	200	600
Micropestle	4	50	100	300

*Add absolute ethanol (see the bottle label for volume) to the Wash Buffer prior to initial use
**Add ddH_2O (see the bottle label for volume) to prepare Proteinase K (vortex to dissolve and spin down) and store at 4ºC

Order Information

Product	Package Size	Catalogue Number
Genomic DNA Mini Kit (Blood/Cultured Cell)	100/300 preps	GB100/300
Genomic DNA Maxi Kit (Blood/Cultured Cell)	10/25 preps	GDM010/25
Genomic DNA Mini Kit (Tissue)	50/100/300 preps	GT050/100/300
gSYNC™ DNA Extraction Kit	50/100/300 preps	GS050/100/300
Genomic DNA Mini Kit (Plant)	100 preps	GP100
Genomic DNA Maxi Kit (Plant)	10/25 preps	GPM010/25
GENEzol™ DNA Reagent Plant	100/200 rxns	GR100/200
Presto™ Mini gDNA Yeast Kit	100/300 preps	GBY100/300
Presto™ Mini gDNA Bacteria Kit	100/300 preps	GBB100/101/300/301
Geneius™ Micro DNA Extraction Kit	100/300 preps	GMB100/300
Presto™ Buccal Swab gDNA Extraction Kit	100/300 preps	GSK100/300
Presto™ 96 Well Blood gDNA Extraction Kit	4/10 x 96 preps	96GBP04/10
Presto™ 96 Well Plant gDNA Extraction Kit	4/10 x 96 preps	96GPP04/10

Caution

GBT Buffer contains guanidine hydrochloride. During operation, always wear a lab coat, disposable gloves, and protective goggles.

Genomic DNA Mini Kit (Tissue) Functional Test Data

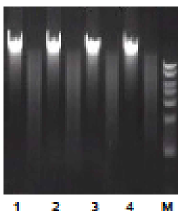

Figure 1. Genomic DNA from a variety of tissue samples was extracted using the Genomic DNA Mini Kit (Tissue). The purified genomic DNA (30-40 kb) was *Eco*RI digested and analyzed by electrophoresis on a 1% agarose gel.
1 = Mouse Liver
2 = Mouse Tail
3 = Fish Muscle
4 = Fruit Fly (Drosophila)
M = Geneaid 1 Kb DNA Ladder

Genomic DNA Mini Kit (Tissue) Protocol

IMPORTANT BEFORE USE

- Add ddH_2O (see the bottle label for volume) to prepare Proteinase K (vortex to dissolve and spin down) and store at 4°C
- Add absolute ethanol (see the bottle label for volume) to the Wash Buffer prior to initial use
- Additional requirements: microcentrifuge tubes, absolute ethanol, (optional) RNase A (10 mg/ml), ddH_2O

Tissue Dissociation	• Cut up to **30 mg of animal tissue (or 0.5 cm of mouse tail)** then transfer it to a 1.5 ml microcentrifuge tube. NOTE: If tissue has a higher number of cells (e.g. spleen or liver), reduce the starting material to 10 mg. • Use the provided **Micropestle** to grind the tissue to a pulp. • Add **200 µl of GT Buffer** to the tube and homogenize the sample tissue by grinding. • Add **20 µl of Proteinase K** to the sample mixture then shake vigorously and incubate at 60°C for 30 minutes. NOTE: During incubation, invert the tube every 5 minutes.
Step 1 Lysis	• Add **200 µl of GBT Buffer** then shake vigorously for 5 seconds. • Incubate at 60°C for at least 20 minutes to ensure the lysate is clear. NOTE: During incubation, invert the tube every 5 minutes. If insoluble material is present following incubation, centrifuge for 2 minutes at 14-16,000 x g then transfer the supernatant to a new 1.5 ml microcentrifuge tube. At this time, preheat the required **Elution Buffer** (200 µl per sample) to 60°C (for Step 4 DNA Elution). **Optional Step: RNA Degradation** (If RNA free gDNA is required, perform this optional step) • Following 60°C incubation, add 4 µl of RNase A (10 mg/ml) to the sample lysate then shake vigorously. • Incubate at room temperature for 5 minutes.
Step 2 DNA Binding	• Add **200 µl of absolute ethanol** to the lysate then immediately shake vigorously for 10 seconds. NOTE: If precipitate appears, break it up as much as possible with a pipette. • Place a **GS Column** in a **2 ml Collection Tube**. • Transfer the mixture (including any precipitate) to the **GS Column** then centrifuge at 14-16,000 x g for 2 minutes. • Discard the **2 ml Collection Tube** then transfer the **GS Column** to a new **2 ml Collection Tube**.
Step 3 Wash	• Add **400 µl of W1 Buffer** to the **GS Column** then centrifuge at 14-16,000 x g for 30 seconds. • Discard the flow-through then place the **GS Column** back in the **2 ml Collection Tube**. • Add **600 µl of Wash Buffer** (make sure ethanol was added) to the **GS Column**. • Centrifuge at 14-16,000 x g for 30 seconds. • Discard the flow-through then place the **GS Column** back in the **2 ml Collection Tube**. • Centrifuge for 3 minutes at 14-16,000 x g to dry the column matrix.
Step 4 DNA Elution	Standard elution volume is 100 µl. If less sample is to be used, reduce the elution volume (30-50 µl) to increase DNA concentration. If higher DNA yield is required, repeat the DNA Elution step to increase DNA recovery and the total elution volume to approx. 200 µl. • Transfer the dried **GS Column** to a clean 1.5 ml microcentrifuge tube. • Add **100 µl of pre-heated Elution Buffer** or TE to the CENTER of the column matrix. • Let stand for at least 5 minutes to ensure the **Elution Buffer** or TE is completely absorbed. • Centrifuge at 14-16,000 x g for 30 seconds to elute the purified DNA.

Troubleshooting

Problem	Possible Reasons/Solution
Clogged Column	**Too much tissue was used** • If using more than 30 mg of tissue, separate into multiple tubes. **Sample tissue was not lysed completely** • Add additional Proteinase K and extend the incubation time in the Lysis Step. • Following the Lysis Step, centrifuge for 2 minutes at 14-16,000 x g to remove sample debris. Transfer the supernatant to a new microcentrifuge tube and proceed with the DNA Binding Step. **Precipitate was formed at DNA Binding step** • Reduce the sample material. • Following ethanol addition, break up any precipitate as much as possible prior to loading GS Column.
Low Yield	**Sample tissue was not lysed completely** • Add additional Proteinase K and extend the incubation time in the Lysis Step. **Column was clogged at DNA Binding step** • Following the Lysis Step, remove the insoluble debris by centrifugation. • Prior to loading the column, break up the precipitate in the ethanol-added lysate. **Incorrect DNA Elution Step** • Ensure that the Elution Buffer or TE is added to the center of the GS Column matrix and is absorbed completely. **Incomplete DNA elution** • Elute twice to increase the DNA recovery.
Eluted DNA does not perform well in downstream applications	**Residual ethanol contamination** • Following the Wash Step, dry the GS Column by centrifuge at 14-16,000 x g or incubate at 60°C for 5 minutes. **RNA/Protein contamination** • Perform optional RNA Degradation step/reduce the sample amount. **Genomic DNA was degraded** • Use fresh samples or freeze fresh samples in liquid nitrogen immediately and store at -80°C.

Changes applied to the manufacturer protocol are:

-Tissue dissociation step: addition of the GT buffer to the tubes before adding the tissue; addition of the tissue to the tubes with GT buffer and cut with the help of scissor, instead of using the provided micro-pestles, to minimize the loss of tissue.

-Tissue dissociation step: incubation time and temperature modified to 56 °C overnight or 60 °C during 4 hours incubation, to ensure the complete dissolution of the tissue by proteinase K.

-Step 4, DNA elution: 75 µl of Elution Buffer were added to the centre of the column matrix, instead of 100 µl, and it was incubated during 10 min. This step was done twice, with a final volume of 150 µl of eluted DNA.

-No RNase A was added.

Appendix 4 – Details of haplotype network reconstruction for nuclear markers: visualization of alleles position in the network.

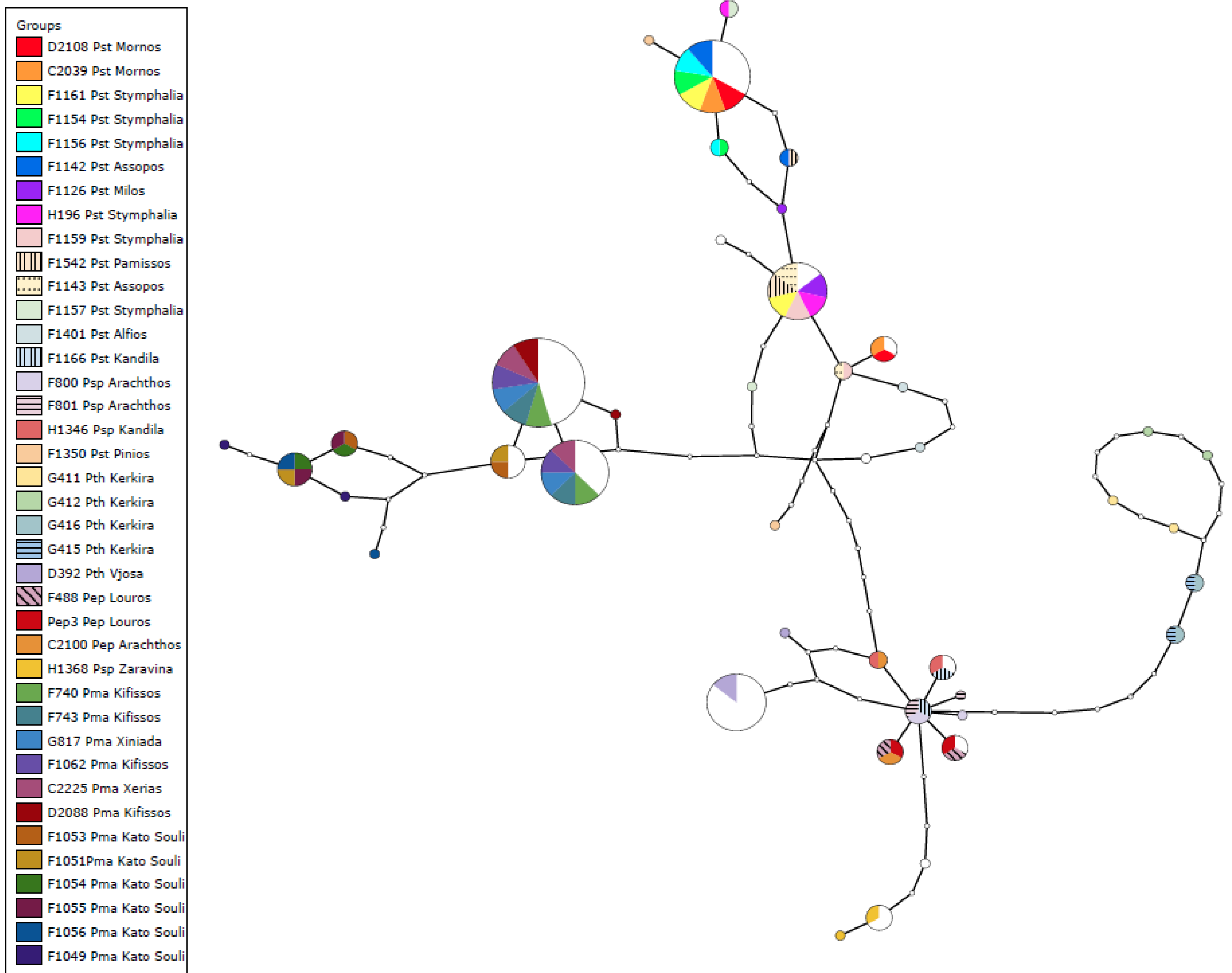

Figure 1 - Detail of haplotype network for the long sequence of S7, representing the position of the alleles on individuals with heterozygote positions. Modified with tcsBU (TCS Beautifier).

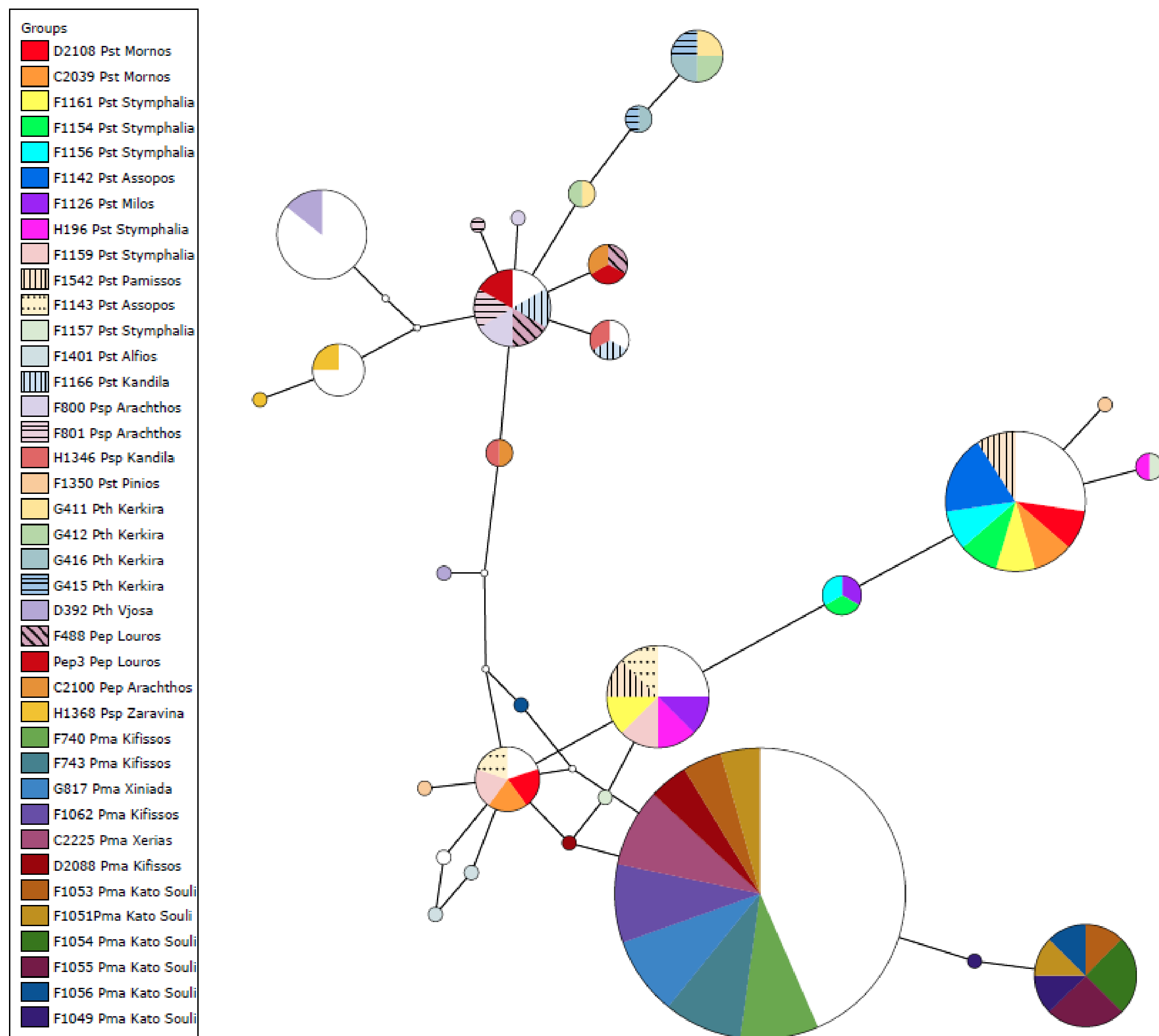

Figure 2 - Detail of haplotype network for the short sequence of S7 representing the position of the alleles on individuals with heterozygote positions. Modified with tcsBU (TCS Beautifier).

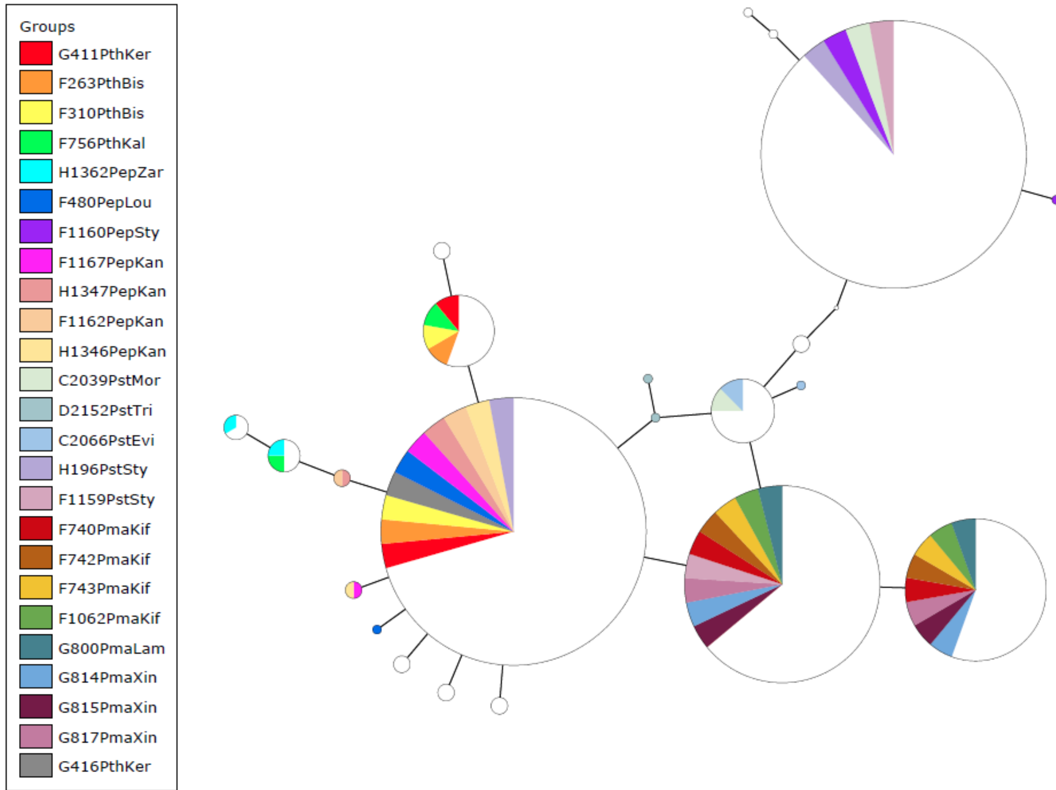

Figure 3 - Detail of haplotype network for RAG representing the position of the alleles on individuals with heterozygote positions. Modified with tcsBU (TCS Beautifier).

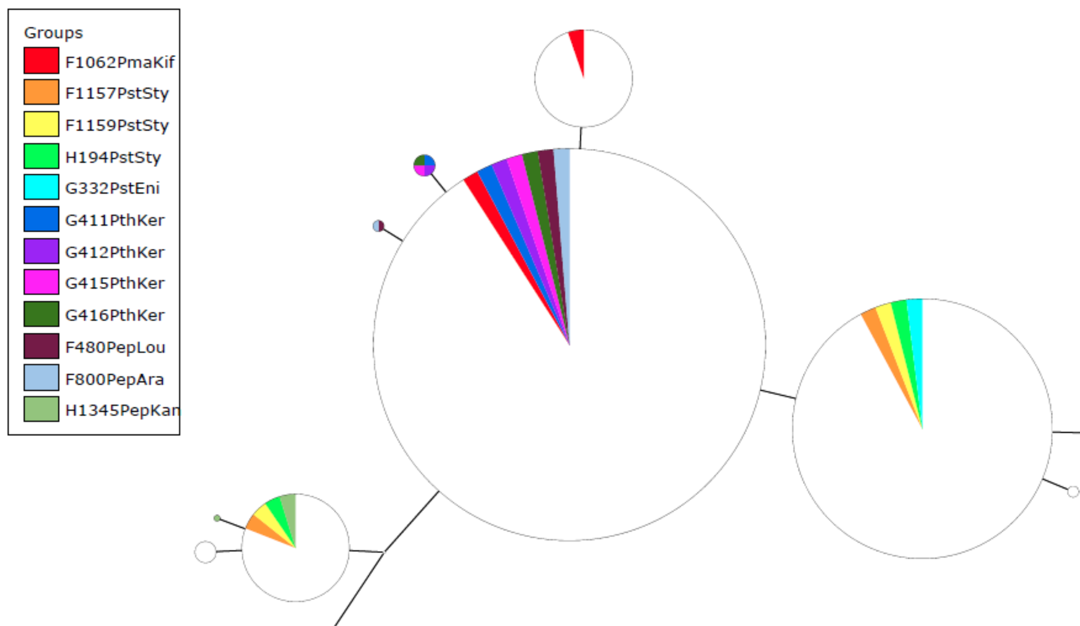

Figure 4 - Detail of Rhodopsin haplotype network, representing the position of the alleles on individuals with heterozygote positions. Modified with tcsBU (TCS Beautifier).