universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Genetic variability of geographically widely separated mediterranean populations of *Plakosyllis brevipes* (Polychaeta, Syllidae)

verfasst von

Christine Mayer

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 439
Studienrichtung lt. Studienblatt:	Diplomstudium Zoologie
Betreut von:	a.o. Univ. Prof. Dr. Gerhard Steiner

Table of contents

Abstract

Plakosyllis brevipes belongs to the Syllidae, one of the most diverse families of polychaetes, comprising over 70 genera and 700 species. Syllids are widely distributed and show either a benthic or interstitial mode of life. *P. brevipes* is a representative of the "meiofauna", a group of organisms that inhabit the spaces between sand grains. This challenging habitat requires numerous adaptions, e.g. miniturisation, flexibility, a worm-shaped body, adhesion organs and specialized reproductive modes. *P. brevipes* reproduces by Stolonisation. Due to the small body size, the amount of produced gametes is limited and because of the very restricted pelagic phase of reproductive stages, the dispersal potential of this species is quite low. Even so it is found along the coastline of several countries and occurs for example in the Mediterranean Sea or the Atlantic Ocean.

The aim of this study is to test whether two geographically widely separated populations of *P. brevipes* are conspecific or form cryptic species. To investigate this specimens were collected in the Mediterranean Sea, at the coasts of Croatia and Italy separated by 2500 km of coastline. The mitochondrial CO1 and the nuclear ITS markers were investigated. Maximum likelihood, Bayesian inference, Maximum Parsimony, AMOVA and Coalescence analyses were used to phylogeographically analyse the *P. brevipes* populations. Uncorrected p-distances were used to identify the genetic distance between these two populations and other syllids.

All analyses lead to the assumption that the populations are most likely conspecific, which is supported by low genetic distances between the populations. Genealogical analysis showed two well separated clades, each containing a dominant haplo-lineage and a single haplotype from the other population. Possible causes for this overlap are the different inheritance pattern of genes, gene flow during the last glacial maximum and passive drift. However, there is a high variation among these geographically widely separated populations and the low genetic differentiation within them. Coalescence analyses calculated the point of separation 2.9 myr ago, after the Messinian salinity crisis.

1 Introduction

Syllidae are one of the most diverse families of polychaetes, comprising over 70 genera and 700 species (Aguado and San Martín 2009). They are widely distributed and show either a epibenthic or interstitial mode of life. Interstitial animals inhabit the spaces between sand grains. Their size ranges from 0.5 mm to 3 cm (Swedmark 1964) and is limited by the dimensions of the pore system. This challenging habitat requires numerous adaptions, like miniaturisation, flexibility, a worm-shaped body, adhesion organs and reproductive efficiency to name but a few. Almost all invertebrate groups are represented in this biotope, except Porifera, Ctenophora and Chaetognatha. Polychaetes are also common, like the syllid *Plakosyllis brevipes*, a representative of the sub-family Syllinae.

Syllids exhibit a wide range of spectacular reproductive specialisations. They show two different strategies of reproduction, epigamy and schizogamy. The latter is realised in Syllinae. Due to the small body size, the amount of produced gametes is limited and because of the very restricted pelagic phase of the stolons, which are the reproductive stages, the dispersal potential of this species is low. Even so the animal can be found along the coastline of several countries and occurs for example in the Mediterranean Sea or the Atlantic Ocean.

1.1 Interstitial habitat

The term interstitial fauna by Nicholls (1935) or mesopsammon by Remane (1933) defines all those animals that inhabit the spaces between individual sand grains. This habitat comprises most of the invertebrate groups, like representatives of the turbellarians, crustaceans, molluscs, nematodes and polychaetes. It is affected by grain size and grain shape. Well sorted coarse sediments are packed more lightly and contain up to 45% pore volume (Giere 2009, and references therein). Grain size and shape influence the physical and chemical milieu of the sediment, its sorting and density and the structural and spatial conditions. Other parameters are temperature, salinity, light and oxygen content. All these influences on the interstitial environment pose strict demands on the organisms living in this habitat.

The pore space is a challenging environment, promoting various adaptions. A miniaturised, worm-shaped, flexible body enables the contact with sediment particles. The locomotion, often by body ciliation or wriggling movements, through the pore system can be facilitated with a worm-shaped body. Adhesion organs, static organs and reinforcement of the body are other possible adaptions. The dimensions of the spaces between sand grains determine the size of the organisms, which varies from 0.5 mm to 3 cm. This small body size, which is a consequence of this biotope, limits the amount of produced gametes. Therefore, several different reproductive strategies have evolved to ensure reproduction (Swedmark 1964; Giere 2009, and references therein).

Interstitial polychaetes evolved several times independently from many different macrofaunal forms and are therefore a polyphyletic group. The composition of the polychaete communities is depending on the substrate. According to Martins et al. (2013) coarse sand sediments, mainly in shallower areas, determine for the spatial distribution of syllids abundance and species richness in Portugal. Muddy or sandy mud sediments from sample sites further away from the coast show a decreasing abundance and diversity. *Plakosyllis brevipes*, e.g., was exclusively found in coarser sediments, while *Syllis garciai* (Syllinae) occurs in coarser as well as finer sediments, and is additionally associated with kelp holdfasts. There are habitat specialists and generalists among syllids.

1.2 Classification

Based on the results of Rouse and Fauchald (1997) the traditional Annelida are monophyletic and comprise two clades, Clitellata and Polychaeta, though the monophylum of the latter was not well supported. Westheide (1997) used a new procedure for character evaluation and considered oligochaetes and leeches as derived polychaetes. In the molecular study of McHugh (1997) the Clitellata, Pogonophora and Echiura nested among various polychaetes. The conflict between morphological features and molecular data could have several causes, like the character loss in Clitellata, that would help identify their sister group among the polychaetes.

Palpata and Scolecida are the two major clades of the class Polychaeta Grube, 1850. The placement into the Palpata is based on two synapomorphies for this group, namely the presence of palps and a limited peristomium (Rouse and Fauchald 1997). Palpata comprises the vast majority of polychaetes and contains the orders Aciculata and Canalipalpata, which are divided due to their structurally different palp types, former are ventral "sensory" palps and letter are grooved "feeding" palps (Orrhage 1980).

According to Rouse and Fauchald (1997) Aciculata are devided into two major clades, Eunicida and Phyllodocida, with a few taxa *incertae sedis*. They have their name from a special chaetal type called aciculae. These stout chaetae are not evertable due to their attachment to the musculature and are therefore forming internal "skeletal" rods for the parapodia. Additional apomorphic features are the presence of prostomial antennae, dorsal cirri, ventral cirri, one pair of pygidial cirri, and segmental organs in most segments (Rouse and Pleijel 2001).

Phyllodocida were introduced by Dales (1962). Strong support for this clade is provided by synapomorphies, like the presence of anterior enlarged cirri, the loss of dorsolateral folds, the presence of an axial muscular proboscis (eversible pharynx) and paired pharyngeal retractor muscles, the presence of compound chaetae with a single ligament, and the presence of metanephromixia (Rouse and Fauchald 1997). According to Pleijel and Dahlgren (1998) Aphroditiformia and Nereidiformia are major clades within Phyllodocida. Böggemann (2002) supports an additional clade, the Glyceriformia.

Rouse and Fauchald (1997) placed the families Chrysopetalidae, Hesionidae, Nautiliniellidae, Nereididae, Pilargidae and Syllidae in the Nereidiformia. Their internal relationship is discussed by many authors (e.g. Pleijel and Dahlgren 1998; Dahlgren et al. 2000). They show at least one pair of antennae, at least one pair of tentacular cirri and the palps are short and usually distally blunt, frequently biarticulated. The pharynx is eversible, if armed, with one pair of lateral jaws and sometimes with accessory denticles. The first parapodia are lateral (Fauchald 1977).

Syllidae Grube, 1850 is a highly diverse family that comprises over 70 genera and 700 species (Aguado and San Martín 2009) and has a global distribution. The classification based on

Fauchald (1977) includes four sub-famillies: Syllinae Grube, 1850; Exogoninae Langerhans, 1879; Eusyllinae Malaquin, 1893; and Autolytinae Langerhans, 1879. According to San Martín (2003) only a few characteristics were used in the diagnosis of the sub-families that led to an uncertainty in the classification of the monophyletic character of the sub-families. In 2009, Aguado and San Martín described a new sub-family, Anoplosyllinae, including *Anoplosyllis*, *Syllides*, *Streptosyllis*, *Asterptosyllis* and *Streptospinigera*, as sister group to the rest of the syllids. The pharyngeal armature is the synapomorphic feature for all syllids, except Anoplosyllinae, *Murrindisyllis*, *Anguilosyllis* and *Bollandia,* which explains the division into these two clear groups. Aguado and San Martín's (2009) analysis confirmed the traditional sub-families Autolytinae, Syllinae and Exogoninae as monophyletic groups, whereas "Eusyllinae" is considered to be polyphyletic in accordance with other authors (Nygren 1999; Nygren and Sundberg 2003; Aguado et al. 2007).

1.3 Reproduction in the Syllidae

Syllids evolve remarkable specialisations in reproduction, but the common strategy is epitoky. Epitoky involves modifications in morphological, physiological and behavioural traits. The most obvious changes in syllids affect the sensory, locomotory and nephridial system. There are two different kinds of epitoky, epigamy and schizogamy (Franke 1999).

In epigamy, the whole animal is modified into a sexual individual, which swims to the surface to spawn. This leads to several modifications such as the modification of sense organs, the enlargement of the prostomial eyes (especially in males), the development of notopodia and natatory notochaetae in median segments, and the modification of the nephridial system to store and extrude gametes. The changes in musculature are minimal, and the alimentary canal stays functional. There are records of syllids surviving spawning and returning to a benthic life with the possibility to breed again (Franke 1999).

In schizogamy (or stolonization) only one section of the animal is modified into a sexual unit, the stolon, which breaks away to become pelagic. The stolons have no mouth or pharyngeal structures. They show reorganisation of the musculature, large eyes and swimming notochaetae. Schizogamic stolons live only to mate, while the stock parent is able to

regenerate and reproduce again (Franke 1999, Glasby 2000). There are two types of schizogamy known, scissiparity, where stolons are developed from already existing segments of the stock parent and gemmiparity, where stolon groups are produced from newly added segments (Nygren 1999, and references therein).

The epigamous type occurs in Exogoninae, some Anoplosyllinae and "Eusyllinae", while schizogamy is present in Syllinae and Autolytinae (excepting the genus *Epigamia*; Nygren 2004) (Aguado et al. 2012). Gemmiparity occurs only in the two genera *Myrianida* (Autolytinae) and some species of *Trypanosyllis* (Syllinae) (Nygren and Sundberg 2003; Aguado et al. 2012, and references therein). The mode of reproduction is characteristic for each sub-family.

Syllids also show different modes of brood care, like ventral or dorsal brooding, brooding in ventral egg-sacs or in gelatinous masses, and vivipary (Aguado and San Martín 2009; Aguado et al. 2012). According to Aguado et al. (2012) epigamy is the primitive type, and schizogamy is derived from epigamy and evolved twice. They also assume that gemmiparity in *Myrianida* is derived from scissiparity, and vivparity derived from schizogamy in Syllinae.

1.3.1 Reproduction in the Syllinae

The stolons of the sub-family Syllinae can be differentiated from autolytine stolons. Stolon of the sub-family Autolytinae show a clear sexual dimorphism, they are comprised of three different regions and appear from the midbody. Syllinae stolons can not be distinguished by sex and are developed from the posterior segments of the body (Aguado and San Martín 2009; Aguado et al. 2012).

The stolon types in the Syllinae can be differentiated regarding their anterior appendages (Malaquin 1893; Aguado et al. 2012). Listing them from the simplest to the more complex: acephalic (with no anterior appendages developed; capillary setae are present before separation); acerous or tetraglene (with two pairs of eyes; lacking antennae and palps); dicerous or chaetosyllis (well defined, cleft prostomium; two pairs of eyes and one pair of small unarticulated antennae); tetracerous (well- defined prostomium; two pairs of eyes and one pair of well developed and articulated antennae; one pair of palps); and pentacerous or

ioda (two pairs of eyes, one pair of lateral articulated antennae and an additional central antenna; one pair of palps).

The different stolon types present a useful source of phylogenetic information on the relationships in the Syllinae. According to Aguado et al. (2012) species with acerous stolons form a well separated clade from the remaining Syllinae. The sister group of pentacerous and acephalous stolons are related to the clade of tetracerous and dicerous types.

1.4 Anatomical features and classification of *Plakosyllis brevipes* Hartmann-Schröder, 1956

Plakosyllis brevipes was described by Hartmann-Schröder in 1956 and placed in the sub-family Exogoninae. San Martín (2003) redescribed *P. brevipes* in the sub-family Syllinae. The body is small, up to 4 mm long, dorsally flattened and oval-elongated (Fig. 1A), suitable for an interstitial animal. The prostomium is trapezoidal with two pairs of lens eyes (one dorsal pair, one ventral), three antennae insert at the anterior margin of the prostomium and two spherical, unfused palps, are located ventrally (Fig. 1B). Tubercles on the dorsum are absent. The short and slender pharynx is equipped with one big tooth (Fig. 1C). Characteristic for syllids is the short, ovoid proventricle, a specialized muscular region of the alimentary canal. This synapomorphic feature is barrel-shaped and has 11 to 14 muscle cell rows. The parapodia have 10 to 12 compound chaetae, which are short falcigers (Fig. 1D) and the aciculae are stout and almost straight (Fig. 1E) (Hartmann-Schröder 2006, San Martín et al. 2008). Adhesion organs enable close contact with the substrate (Swedmark 1964).

The current classification of *P. brevipes* is given in Table 1. Aguado and San Martín (2009) described a clade within Syllinae in their morphological study, consisting of the genera *Xenosyllis* Marion and Bobretzky, 1875, *Plakosyllis* Hartmann-Schröder, 1956 and *Eurysyllis* Ehlers, 1864 due to one apomorphy, the ventrally attached palps on the prostomium. *Eurysyllis* and *Plakosyllis* share the shape of antennae and dorsal cirri. They can be distinguished by their palps, aciculae and the presence or absence of tubercles. *Eurysyllis tuberculata* has fused palps, distally expanded aciculae with a distal short tip, and the dorsum is covered with rows of tubercles (San Martín et al. 2008; Hartmann-Schröder 2006).

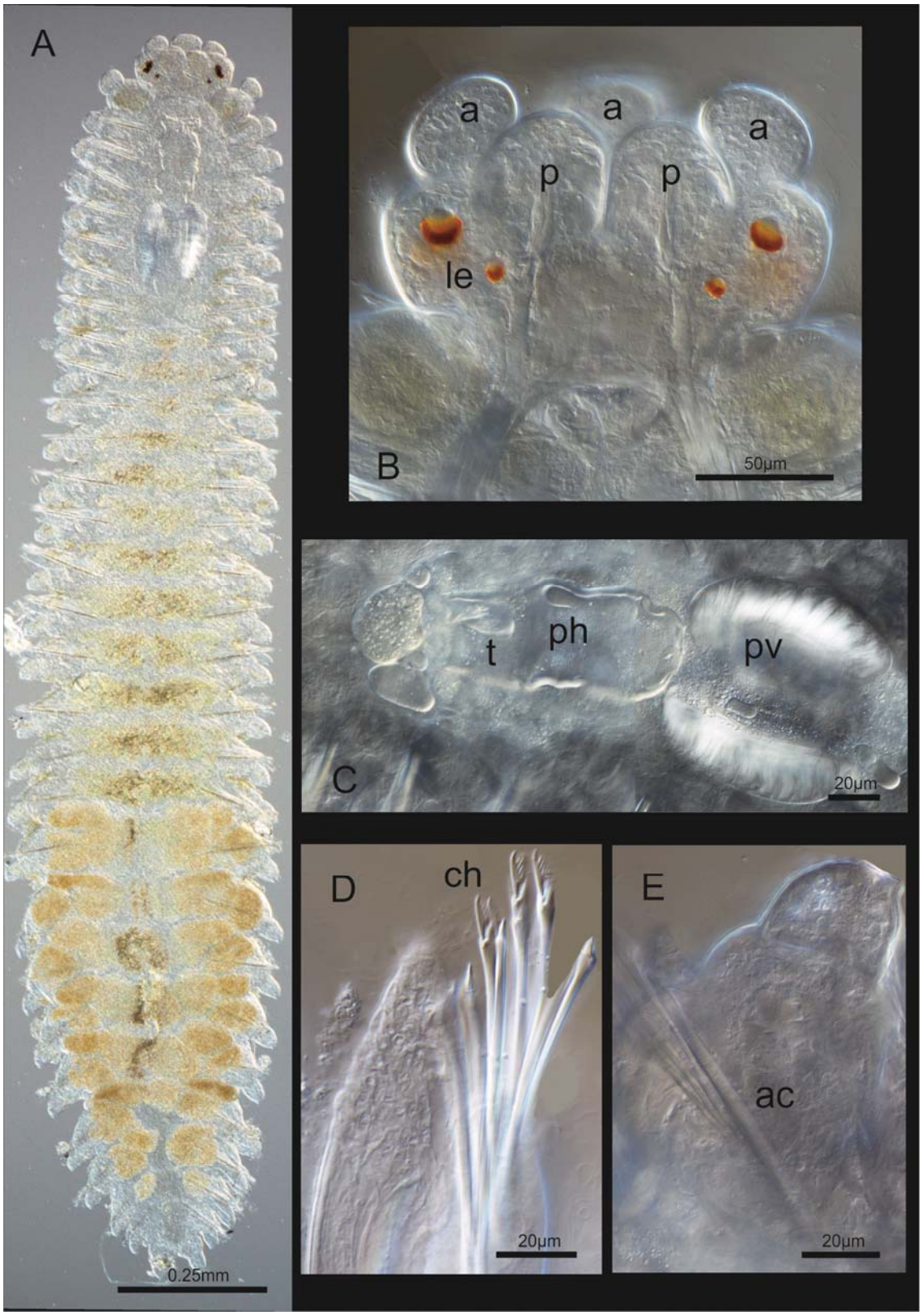

Figure 1 : *Plakosyllis brevipes*. A: Dorsal view; B: Ventral view of the prostomium; two well separated palps, three antennae and two pairs of lens eyes are visible; C: Pharynx, equipped with one tooth, and the muscular proventricle; D: Part of a parapodium with compound chaetae; E: Part of a parapodium with aciculum; [a-antennae; ac-aciculum; ch-chaetae; le-lens eyes; p-palps; ph-pharynx; pv-proventricle; t-tooth]

Table 1 : Classification of *Plakosyllis brevipes* according to Rouse and Fauchald (1997)

Annelida
Polychaeta
Palpata
Aciculata
Phyllodocida
Nereidiformia
Syllidae Grube, 1850
Syllinae Grube, 1850
Plakosyllis Hartmann-Schröder, 1956

1.4.1 Reproduction

Plakosyllis brevipes reproduces by schizogamy, more precisely by scissiparity and develops acerous (tetraglene) stolons. There is no accurate information on reproduction of *P. brevipes*, but there is on Syllinae in general. During reproduction, when the stolons are ready to spawn, small eggs and sperms are released into the sea. This primitive sperm type is called ect-aquasperm (Franke 1999, and references therein). The fertilised eggs soon sink to the bottom and after 24 to 48 hours a trochophore (e.g. *Typosyllis pulchra*, *Typosyllis variegata*) or metatrochophore (e.g. *Typosyllis prolifera*) larvae hatches (Franke 1999, and references therein). The following pelagic phase is very restricted and takes hours up to a few days. According to Franke (1980) there are two reasons. First there already is a pelagic state, the stolon, which ensures a certain amount of dispersal, and second habitats usually have a restricted range and by restricting the pelagic phase of larvae they stay near adequate habitats.

1.4.2 Distribution and habitat

Plakosyllis brevipes is widely distributed and can be found in the Mediterranean Sea, the north-east and north-west Atlantic Ocean, the Red Sea, the Indian Ocean, New Caledonia and Australia (Western Australia and New South Wales) (San Martín et al. 2008).

This syllid occurs intertidally and at depths up to 60 meters. The favoured habitat is coarse sand, but it is also found on algae (*Halopteris sp.*), sponges, and rhizomes of seagrasses, like

Posidonia oceanica (San Martín 2003). Martins et al. (2013) found *P. brevipes* in depths up to 69 m at the Portuguese Coast.

1.5 Aims of the study

In this study, two populations from different sample sites (Croatia and Italy) are phylogeographically analysed to evaluate their genetic diversity. Even though the dispersal potential is limited, *Plakosyllis brevipes* is distributed along the coastline of several countries and occurs for example in the Mediterranean Sea or the Atlantic Ocean.

These two separated populations are either conspecific due to gene flow or cryptic species because of a reproductive separation event. The work hypotheses expects cryptic species with a distinct differentiation in haplotypes. This is due to the limited dispersal potential, the restricted pelagic phase and the wide geographic distance of 2500 km coastline between the Italian and Croatian populations. Coalescence analyses are used to calculate the point of separation.

2 Material and Method

2.1 Data collection

2.1.1 Population sampling

Adult individuals of *Plakosyllis brevipes* were collected along the Adriatic and Tyrrhenian coastlines in the Mediterranean Sea. I took samples of the specimens in Croatia at the Isle of Pag. The Italian samples were collected along the Italian coast around Naples (Fig. 2). There were no syllids along the east coast of Italy between the cities Pescara and Bari. The coastline from the Ile of Pag to Neaples is about 2500 km long. Coordinates of the sample sites of all specimens are listed in Table 2.

Figure *2* : Map of the sample sites in Croatia (Adriatic Sea) and Italia (Tyrrhenian Sea). Stars indicate cities and the Ile of Pag and red dots show the sample sites.

Table 2 : Coordinates of sample sites in Croatia and Italy and sampling depths.

Country	Location	Coordinates	Depth
Croatia	Ile of Pag	44°26'30'' N, 15°2'40'' E	1-4
Italy	Meta (Strada Statale della Penisola sorrentina)	40°38'5'' N, 14°24'19'' E	3-4
	Marina del Cantone	40°35'00'' N, 14°21'27'' E	4-8
	Laura	40°26'09'' N, 14°58'3'' E	3
	Viale mare cristallo	40°17'59'' N, 14°17'59'' E	3-5
	Marina diaequa	40°39'34'' N, 14°24'54'' E	3-4
	Agnone	40°12'54'' N, 14°59'41'' E	2-4
	Lago di Patria	40°55'40'' N, 14°00'58'' E	5-10

The sand was collected in 1 to 10 meters depth by snorkeling (Tab. 2). After 24 hours I removed the top layer of the sand, put it in a flask and treated it with seawater-isotonic magnesiumchloride-hexahydrat for 25 minutes to anaesthetise the animals. The animals were detached from the sand grains by shaking and sieved through nets of 250 µm and 125 µm mesh size. The net was rinsed with seawater to collect the animals in a petri dish. Then I identified and collected the polychaetes under a binocular and fixed them in RNAlater. I changed the RNAlater once before preserving the animals in it. At first they were stored in liquid nitrogen, later at -80°C in a freezer. I preserved some specimens in 96% ethanol for further identification. 34 specimens from Croatia and 25 from Italy were used for further analyses.

2.1.2 DNA extraction, PCR reaction and purification

Before DNA extraction I rinsed the specimens with distilled water. The extraction was performed with peqGOLD Tissue DNA Mini Kit (PEQLAB, Germany), following the manufactor's protocol. Due to their small body size, the entire animal was used. 110 µl of pre-heated elution Buffer (70°C) were used to get a higher amount of DNA. The isolated DNA was stored at -20°C.

Polymerase chain reaction (PCR) amplification of the nuclear internal transcribed spacer (ITS 1, 5.8S and ITS 2) and the mitochondrial cytochrome c oxidase subunit 1 (CO1) was accomplished with the primers listed in Table 3. Amplification of ITS yielded fragments of approximately 980 bp, CO1 gained about 670 bp. I used different PCR mixtures for both fragments. The solution used for the ITS region contained 0,5 µl of each primer (100 pM), 1

to 2 µl DNA template, 2.5 µl dNTPs (2.5 mM) and Mango Taq reaction buffer (6X), 1.3 µl magnesium chloride (20 mM), 0.25 µl of Mango Taq DNA Polymerase (5 unit/µl) (Bioline, Germany) and filled up with ddH2O to a total volume of 30 µl. The amplification reaction mixture used for the CO1 gene included 0.8 µl of each primer (10 pM), 1 to 2 µl DNA template, 2.5 µl of Dream buffer (10X) and dNTPs (2.5 mM), 0.1 µl of Dream Taq Polymerase (5 unit/µl) (Thermo Fisher Scientific, USA) and filled up with ddH20 to a total of 25 µl reaction volume. The amount of DNA for both reactions was determined by the intensity of the product in the agarose gel after gel electrophoresis. PCR reactions were performed in Eppendorf and Biometra Mastercyclers.

I used different temperature profiles to amplify the gene fragments depending on the gene and population (Tab. 4). Some CO1 PCR products were re-amplified to gain a higher amount of the wanted product. I used the same reaction mixture, but taking only a pipette tip of the former reaction solutions as template. The annealing temperature for re-amplification was 48°C for 25 cycles.

PCR products were visualized on 1% TAE agarose gels. 5 µl of PCR solution, 2.5 µl Loading dye (PEQLAB, Germany) and 2.5 µl Sybr Green (1:500X) were mixed and loaded in the gel chambers. Sybr Green was diluted in TAE buffer (1X) before usage. Additionally 3 µl of FastRuler Low Range DNA Ladder (Thermo Fisher Scientific, USA) between 50 and 1500 base pairs were utilised to verify the fragments length. Gels were run for 30 minutes at 100 volt and visualized with a transilluminator (UV-Documentation System; PEQLAB, Germany).

Amplified products were purified using peqGOLD Cycle-Pure Kit (PEQLAB, Germany), following the manufactor's protocol. I used 40 µl of elution buffer to dissolve the product, which was finally stored at -20°C.

Table 3 : Primers used in PCR and sequencing

Gene	Primer name	Sequence 5' to 3'	Reference
CO1 - mitochondrial	LCO 1490	GGTCAACAAATCATAAAGATATTGG	Folmer et al. (1994)
	HCO 2198	TAAACTTCAGGGTGACCAAAAAATCA	Folmer et al. (1994)
ITS - nuclear	ITS18SFPOLY	GAGGAAGTAAAAGTCGTAACA	Nygren et al. (2009)
	ITS28SRPOLY	ATGCTTAAATTCAGCGGGT	Nygren et al. (2009)

Table 4 : Temperature profiles for PCR and sequencing

Nuclear ITS			
Step		**Temperature**	**Time**
Initial denaturation		96°C	4 min
Loop 1-45			
	Denaturation	94°C	30 sec
	Annealing	47-48°C	30 sec
	Primer extention	72°C	1 min
Final extention		72°C	8 min

Mitochondrial CO1			
Step		**Temperature**	**Time**
Initial denaturation		95°C	3 min
Loop 1-30			
	Denaturation	95°C	30 sec
	Annealing	47-48°C	30 sec
	Primer extention	72°C	1 min
Final extention		72°C	5 min

Sequencing			
Step		**Temperature**	**Time**
Initial denaturation		--	--
Loop 1-25			
	Denaturation	96°C	20 sec
	Annealing	48°C	10 sec
	Primer extention	60°C	4 min
Final extention		--	--

2.1.3 DNA Sequencing and Alignment

The amplified products were sequenced in both directions. Each reaction mixture included 1 µl Big Dye (v. 3.1), 0.7 to 1 µl primer (10 pM), 1 to 6 µl of purified PCR product and ddH_2O for a total volume of 10 µl. I used the same primers as for PCR amplification. The amount of DNA for sequencing was determined by the intensity of the purified product in the agarose gel after gel electrophoresis. All products were sequenced according to the Sanger method using an ABI 3130xl capillary sequencer at the University of Vienna.

The NCBI Nucleotide BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) validated the obtained sequences. Editing was done with the software Finch TV v. 1.4.0 and Genedoc v.

2.7.000 (Nicholas, Nicholas and Deerfield 1997) to generate a consensus sequence originating from the forward and reverse strands. ClustalX v. 2.0 (Larkin et al. 2007) performed the multiple sequence alignment, which I corrected manually in Genedoc. The file was converted into an interleave of Nexus format with ForCon v. 1.0 (Raes and Van de Peer 1999), which created a sequential of Nexus format.

I used the nucleotide alignments to perform "single gene" as well as "combined gene" analysis. The proteins were translated into amino acids using the Invertebrate Mitochondrial Code (http://www.bioinformatics.org/JaMBW/2/3/TranslationTables) to detect changes in the amino acid sequence of the CO1 gene.

2.2 Phylogenetic analysis

Neighbor-Joining (NJ) trees from single gene and combined gene datasets were calculated with PAUP* 4.0b10 (Swofford 2003) using p-distances (the uncorrected percentage of nucleotide differences). The calculated p-distance matrix made it possible to distinguish the haplotypes of a dataset. Haplotypes are unique sequences which were used in parsimony analyses. The software SPSS Statistics for Windows v. 21.0 (IBM Corp.) was used to visualize the distribution of p-distances within and among the populations from Italy and Croatia. The nonparametric Mann-Whitney-U Test was used to test if there are differences in the amount of p-distances within and among these two populations. Hebert et al. (2004) proposed a standard screening threshold of sequence difference of at least 10X average intra-specific distance. A distance value higher than this calculated threshold would indicate specific distinctness, a lower value taxonomic identity.

Parsimony bootstrap analyses were conducted using PAUP* 4.0b10 (Swofford 2003) with the heuristic search option, 5000 bootstrap replicates and 10 random addition sequences. The starting trees were computed using stepwise addition. Parsimony analyses were made with the heuristic search option, implementing stepwise addition using between 5000 and 50000 replicates. The rearrangements were calculated via tree-bisection-reconnection (TBR). Maximum likelihood analyses were made using the parsimony consensus tree and the best-fit model for the analysed dataset. The branch-swapping algorithm was TBR. The models of

evolution best-fitting were determined with Modeltest 3.7 (Posada and Crandall 1998) for each dataset following the AKAIKE Information Criterion (AIC) or the Hirarchical Likelihood Ratio Tests (hLRTs) (Tab. 5).

Bayesain inference analyses were conducted with Mr. Bayes v. 3.1.2 (Ronquist and Huelsenbeck 2003). Combined dataset analyses were done by calculating three different combined analyses. The first analysis combined the coding CO1 with noncoding ITS, the second noncoding with each codon position of the CO1 dataset, and the third used one model for the entire combined dataset. The partitions and corresponding models are shown in Table 5. Gaps were treated as missing. The 2 X 4 Markov-chain Monte Carlo ran for 10.000.000 generations. The sample frequency was set to 100 in all analyses and the burnin values changed depending on the dataset (400 in the CO1, 700 in the ITS and between 300 and 800 in the combined dataset). The burnin was determined manually by the curve of the log likelihoods.

Trees were visualized using FigTree v. 1.4.0 (http://tree.bio.ed.ac.uk). and edited with CorelDRAW Graphics Suite X5 v. 15.2.0.661 (http://www.corel.com).

Table 5 : Data partitions and corresponding models

data partition	selected model
ITS	SYM+I+G
CO1 (all codon positions)	TIM+I
CO1 (1st codon position)	TrN+I
CO1 (2nd codon position)	F81+I
CO1 (3rd codon position)	GTR
Complete data set	HKY+I+G

2.3 Population genetics

Analysis of the molecular variance (AMOVA) was performed with Arlequin v. 3.11 (Excoffier 2005) to test the degree of differentiation within and among populations. The fixation index (F_{ST}) estimates the amount of genetic variation that is due to population differences. Wright (1978) described the relationship between gene flow and molecular variance among populations for haploid genomes in diocious species as: $(Nm)_F = 0.5 X ((1/F_{ST}) -1)$. Nm is the

number of migrants per generation. The higher the F_{ST} value is, the lower is the frequency of migrants per generation.

The statistical parsimony software TCS v.1.21 (Clement et al. 2000) estimates networks of DNA sequences. Therefore the sequences were collapsed into haplotypes and the haplotype frequency is measured. The outgroup probabilities, which correlate with a relative haplotype age, were estimated using these frequencies. Haplotypes are unique sequences, which can be united to haplotype lineages. The linkage between these lineages is due to sexual reproduction in a population and forms a network. The maximum number of mutational steps between haplotypes could be explained by single substitutions with a statistical confidence of 95%. The resulting parsimony connection limit specifies how many steps between haplotypes can emerge with a probability higher than 95% to still be parsimonious. Haplotypes are not connected to a network, if there are more differences in the nucleotide sequence than allowed by the connection limit. The parsimony probability was set between 90 and 91% to assure a connected network. Gaps were treated as 5th state and not as missing data, because flanking and tailing nucleotides were cut of in the alignments and only observed indels are left.

2.4 Saturation analysis

The saturation of ITS and CO1 datasets was tested using PAUP* 4.0b10 (Swofford 2003) to ensure coalescence analysis. Therefore patristic distances from maximum parsimony trees and absolute distances were calculated. Patristic distances are calculated from the sum of lengths of the branches that link two terminal nodes in a tree and therefore indicate evolutionary steps between two taxa or extant gene sequences. Visualization of the saturation plots and the calculation of the linear regression were performed with SPSS Statistics for Windows v. 21.0 (IBM Corp.).

Another saturation test was performed with Dambe v. 5.3.52 (Xia 2013), which calculates the ratio of transitions and transversions in a dataset and plots them along model calculated distances.

2.5 Coalescence analysis

The estimation of coalescence was executed with Beast v. 1.7.5 (Drummond, Suchard and Rambaut 2012). ITS and CO1 datasets were analysed with the best fitting model of evolution and empirical base frequencies. Different substitution rates especially for polychaetes were 0.3354% per million years for ITS and 0.37% for CO1 (Nygren et al. 2009) and for general invertebrate CO1 0.65% (Crandall et al. 2012). The 1^{st} and 2^{nd} codon positions were also analysed together with the aforementioned substitution rates. Additionally 1^{st}+2^{nd} and 3^{rd} codon positions were analysed using substitution rates based on Nygren et al. (2009), 0.001 for the first and 0.01 for the latter. Bivalve substitution rates were also used on the CO1 dataset, one for each codon position (Marko 2002).

A strict molecular clock model, a UPGMA starting tree and the Yule Process of speciation were implemented. Clock priors were set to normal. The best fitting model of evolution for the ITS dataset is TrN93+GI and TrN93+G for CO1. Each analysis was run four times for $1X10^7$ million steps, the burnin was set to 250 for each replicate. The log-file obtained from Beast v. 1.7.5 was analysed in Tracer v. 1.5.0 to determine the burnin value. LogCombiner v. 1.7.5 combined the logged parameter values and trees from replicated runs. Trees were calculated with the software TreeAnnotator v. 1.7.5 and visualized in FigTree v. 1.4.0 (http://tree.bio.ed.ac.uk).

3 Results

3.1 General sequence analysis

3.1.1 Combined dataset

The alignment of the combined dataset comprises the nuclear CO1 gene, 658 base pairs (bp), and the mitochondrial ITS gene, 965 base pairs. The length of the dataset is 1623 base pairs from 43 individuals. There are 42 unique sequences. The combined dataset comprises 142 parsimony informative characters and 176 variable sites. The combined Croatian population shows less indels compared with the single gene ITS. Information on usable loci, polymorphic sites, transitions, transversions, substitutions and indles of both populations in each dataset is listed in Table 6. The average nucleotide composition of the combined dataset is 26.80% adenine, 27.28% thymine, 22.43% guanine and 23.51% cytosine. The base frequencies for all datasets are listed in Table 6.

3.1.2 ITS dataset

The ITS alignment consists of 965 characters, of which 53 are variable and 33 are parsimony informative. The Croatian population represents 31 sequences and 28 haplotypes. The Italian population comprises 25 individuals and 16 unique sequences. The average base frequencies are 24.49% adenine, 22.71% thymine, 27.94% guanine and 24.87% cytosine. The ITS alignment has 22 observed indels in the Croatian and 17 in the Italian population (Tab. 6). Haplotypes and the specimens with identical sequences are listed in Table 7.

3.1.3 CO1 dataset

The CO1 alignment consists of 658 characters, 131 sites are variable and 108 of them are parsimony informative. There are fewer haplotypes and no indels than in the ITS dataset. The dataset comprises 17 individuals and 11 haplotypes in the Italian population, and 27 specimens and 15 unique sequences at the Croatian site. The base frequencies are similar across populations. The average base frequencies are 21.56% cytosine, 33.86% thymine,

30.12% adenine and 14.47% guanine. All information on usable loci, polymorphic sites, transitions, transversions, substitutions and exact nucleotide composition is presented in Table 6. The different haplotypes used in this dataset are shown in Table 7.

Table 6 : Comparison of all three datasets between Croatian and Italian populations

	CO1 dataset (44 sequences)		ITS dataset (56 sequences)		Combined dataset (42 sequences)	
	Croatia	Italy	Croatia	Italy	Croatia	Italy
Nr. sequences	27	17	31	25	26	17
Nr. haplotypes	15	8	28	16	26	16
Nr. base pairs	658	658	965	965	1623	1623
Nr. usable loci	658	658	964	959	1621	1617
Nr. polymorphic sites	117	9	60	51	157	53
Nr. observed transitions	101	3	30	25	128	21
Nr. observed transversions	27	6	12	10	34	16
Nr. substitutions	128	9	42	35	162	37
Nr. observed indels	0	0	22	17	8	17
Nucleotide composition						
Adenine	30.10%	30.13%	24.85%	24.12%	27.00%	26.59%
Thymine	34.31%	33.41%	22.86%	22.56%	27.53%	27.02%
Guanine	14.36%	14.57%	27.63%	28.25%	22.22%	22.63%
Cytosine	21.23%	21.89%	24.66%	25.07%	23.25%	23.76%

Table 7 : Haplotypes of the ITS and CO1 datasets and the specimens with identical sequences

ITS dataset		**CO1 dataset**	
Croatian haplotypes	Italian haplotypes	Croatian haplotypes	Italian haplotypes
K1	I1	K1 (K2, K3, K30)	I2 (I6)
K2	I2 (I21)	K4	I8
K3	I3 (I5, I7)	K5 (K6, K10, K11)	I16
K4 (K28, K29)	I4	K7	I18 (I7, I9, I17, I22, I24)
K5	I8	K8	I20 (I4, I21, I27)
K6	I9 (I6, I19, I23, I24, I26, I29)	K9	I23
K7	I10	K12	I26
K8	I16	K14 (K15)	I30
K9	I17	K16	
K10	I18	K17	
K11	I20	K18	
K12	I22	K19 (K23)	
K13	I25	K20 (K21, K24, K29)	
K14	I27	K26 (K27)	
K15	I28	K28	
K16	I30		
K17			
K18 (K32)			
K19			
K21			
K22			
K23			
K24			
K26			
K27			
K30			
K31			
K33			

3.2 Phylogenetic relationships

3.2.1 ITS dataset

Maximum likelihood (ML) and Bayesian inference (BI) analyses show two well separated clades, a green and a blue one (Fig. 3). The green clade contains almost exclusively all Croatian specimens, except I18, the only Italian haplotype. The blue clade contains almost exclusively all Italian haplotypes, except K12, the only Croatian sequence.

ML and BI analyses show almost the same topology for the blue clade. The only difference is I8, which has a sister relationship to the rest of the Italian haplotypes in the ML analysis, or is part of this clade in the BI analysis. K12 and I20 form a single clade as sister group to the remaining blue haplotypes, highly supported by posterior probability values (pp=1.0). Due to differences in the formation of haplotypes (Tab 7.) (which specimens have unique sequences and therefore form a haplotype), the haplotype clusters of CO1 and combined data analysis cannot be found in the ITS topology. There are similarities, but an exact allocation was not possible.

The BI and ML analyses indicate that there are three Italian haplotype lineages, I18, clustering with Croatian samples, I20 and the remaining Italian haplotypes.

The BI tree of the "Croatian" green clade is not as resolved as the ML tree, due to the lower phylogenetic signal (Fig. 4). K13 has in all analyses a sister relationship to the remaining green haplotypes. In the BI analysis I18 is the sister specimen to the big "Croatian" clade (except K13), while I18 clusters in ML with K7, K8 and K21. The grouping of K4, K27, K31, K1 and K10 with moderate support, and K5 and K16 with high support values (pp=0.94) is indicated in both analyses. The Croatian ITS haplotype clusters did not appear in CO1 single gene or combined gene analyses.

The Croatian haplotype lineages of the ITS dataset are hardly identifiable.

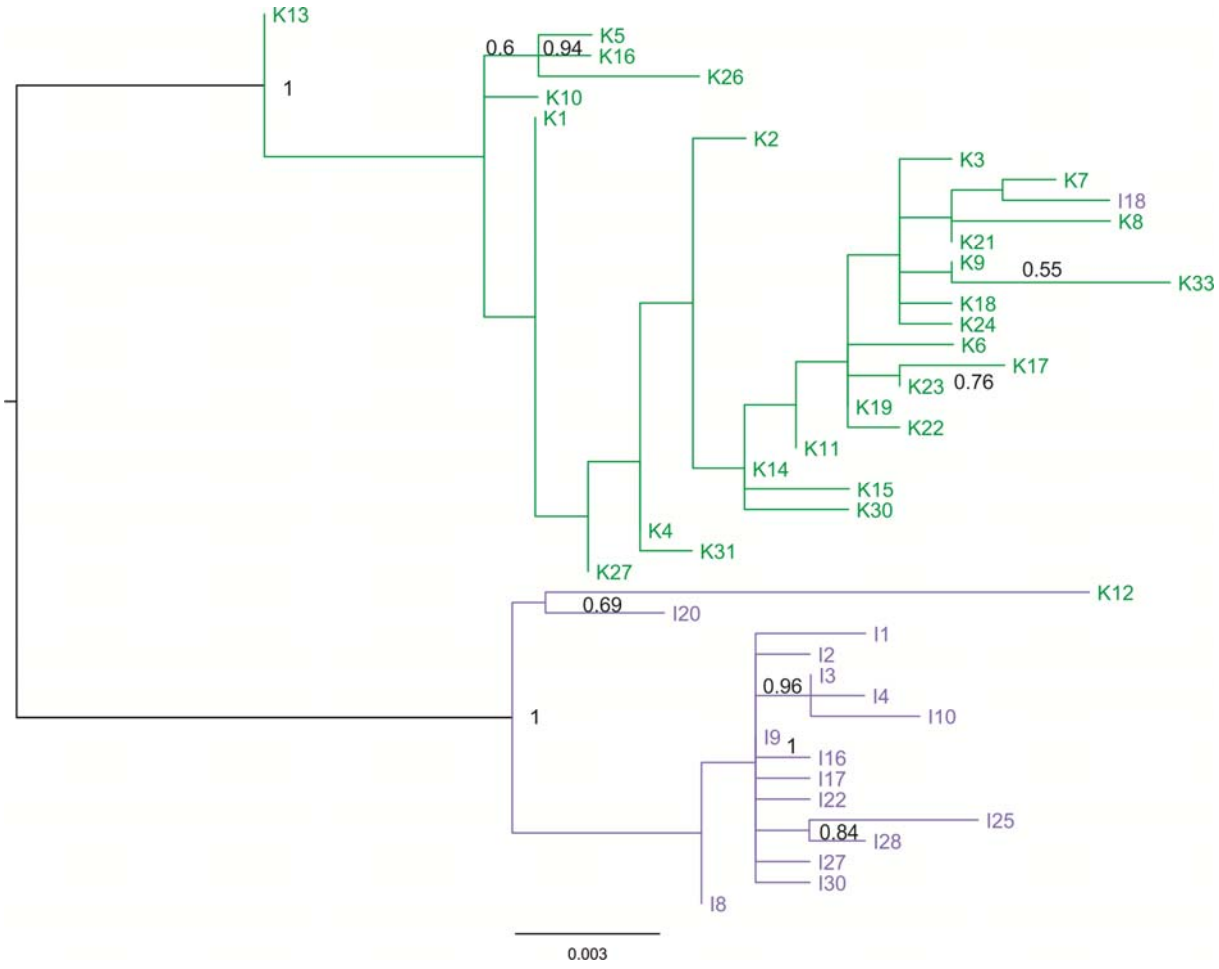

Figure *3* : Maximum likelihood tree of the ITS haplotype dataset with posterior probability values. Haplotypes of different populations are visualised in different colors, blue for Italian and green for Croatian haplotypes. The tree is deliberately rooted on the longest branch. number of trees=9750; tree length=68; CI=0.5294; RC=0.4728; -lgL=1941.21231; model=SYM+I+G; The scale bar indicates 3 substitutions per 1000 sites. [CI-consitency index; RC-rescaled consitency index]

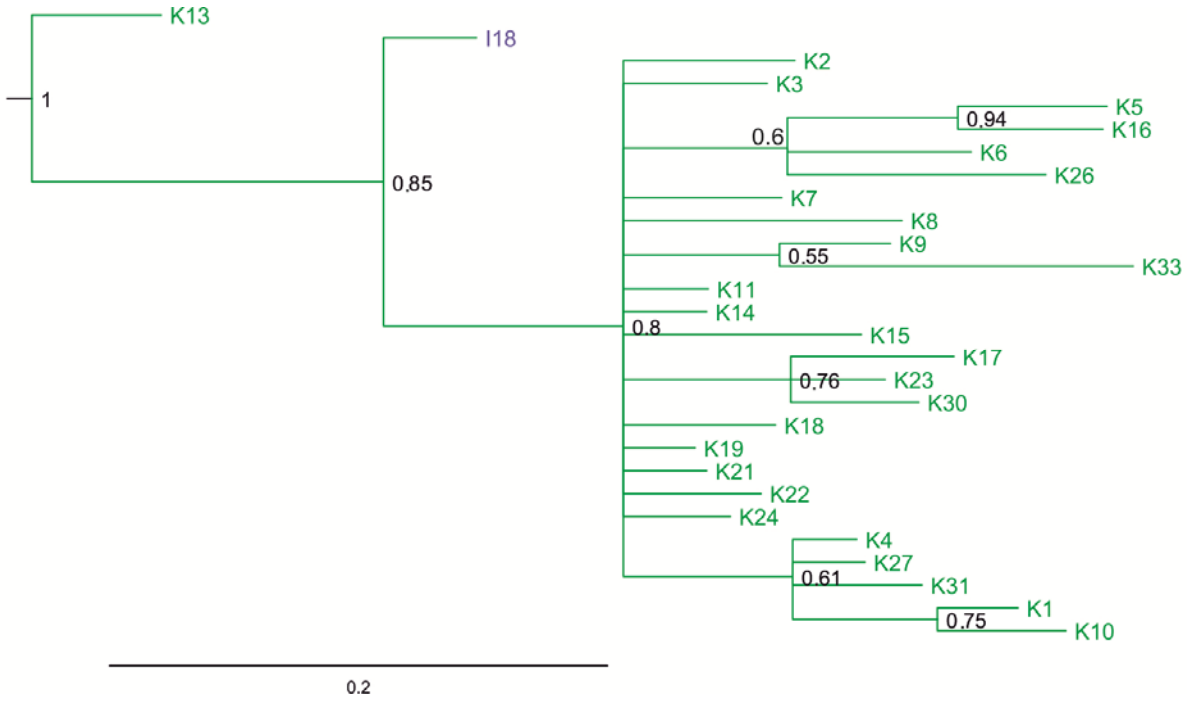

Figure *4* : "Croatian" sub-tree of the Bayesian inference tree of the ITS haplotype dataset. Values of nodes are posterior probabilities. Due to similarities in ML and BI analyses regarding the blue clade only the "Croatian" green clade is pictured in this figure. model=SYM+I+G; The scale bar indicates 2 substitutions per 10 sites.

3.2.1.1 Network analysis

Statistical Parsimony analysis of the ITS dataset connects the Italian and Croatian populations. There are two well distinguishable networks, linked via K13 (Fig. 5). The hypothetical ancestor is I26 with an outgroup probability of 23% in the Italian part of the network. The green part of the network contains all Croatian haplotypes and one Italian specimen (I18). The blue part contains all unique Italian sequences.

The Italian part of the ITS network, except I18, is clearly defined. Most of the blue haplotypes are closely positioned to the hypothetical ancestor, differing only in one base. I10, I4 and I25 are the only Italian haplotypes, which are not directly linked with the assumed ancestor. The specimens I3, I4, I10, I25 and I28 are clustering together, like in the ML analysis. The Croatian and Italian haplotype lineages are connected via K13, as explained from the ML and BI analyses. The connection of K12 with the Croatian part of the network is in contrast to ML and BI analyses, where K12 always clusters with the Italian specimens. K12 is connected with

K10, but differing in 15 bases. Compared to the Italian part, the Croatian part of the network is more differentiated. Due to the distinct network character of the Croatian population it is not possible to define haplotype lineages.

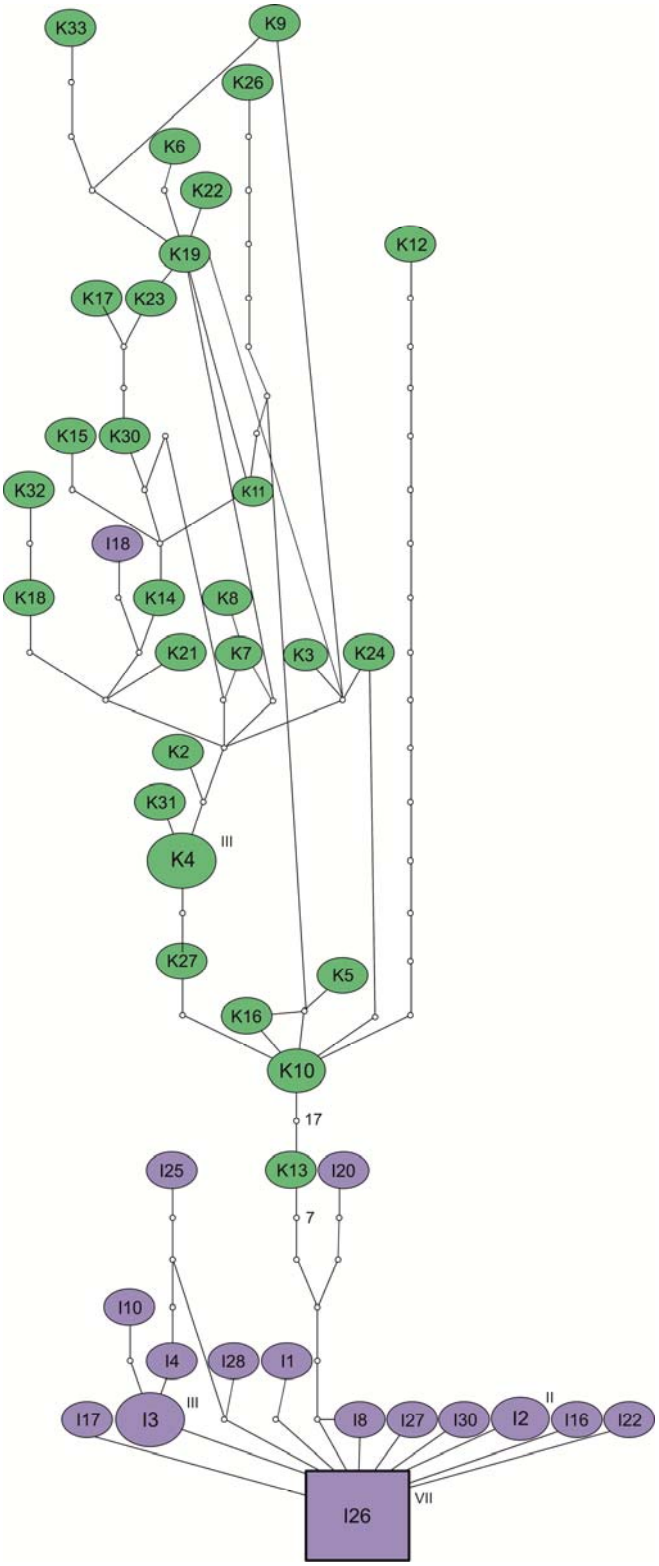

Figure 5 : TCS network from ITS data. Parsimony probability: 91%; Outgroup probability: 23%; Haplotypes of different populations are visualised in different colors; blue for Italian and green for Croatian haplotypes. Roman numerals on the right side of the spheres and the rectangle indicate haplotype frequencies. Nodes and arabic numbers between haplotypes represent unobserved haplotypes; The length of the lines is not representing genetic distance.

3.2.2 CO1 dataset

ML and BI analyses show two well separated clades. The green clade is consisting of almost all Croatian haplotypes and the blue clade contains all Italian haplotypes, except K12, the only Croatian specimen (Fig. 6).

ML and BI analyses of all positions in the CO1 dataset show an identical topology for the Italian population with high posterior probability values. I26 and I20 are clustering together (pp=1) and form a sister clade to the remaining haplotypes. K12 clusters with the Italian haplotypes (pp=1). The haplotype clusters of the Croatian clade are highly supported and identical in ML and BI analyses. The only difference is the topology of the clusters in both analyses. In ML, K19, K26 and K20 form a clade to the remaining specimens. The group K8, K7 and K1 is the sister clade to K14, K4, K5, K9 and K28 with the group K16, K18 and K17.

There are four Croatian and one Italian haplotype lineage in the CO1 dataset. The Croatian haplo-lineages are K12, which is clustering with the blue clade, K16, K17 and K18, the group K1, K7 and K8, and the remaining haplotypes.

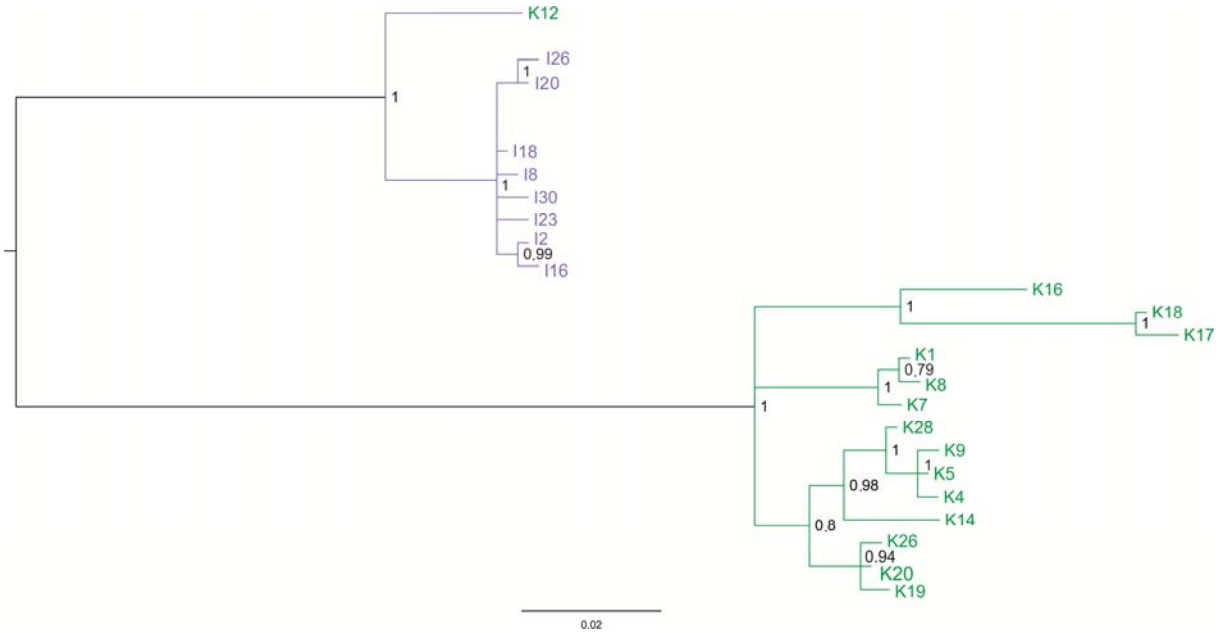

Figure *6* : Bayesian inference tree of the CO1 dataset. Values at nodes are posterior probabilities. The tree is deliberately rooted on the longest branch. Haplotypes of different populations are visualised in different colors, blue for Italian and green for Croatian haplotypes. model=TIM+I; The scale bar indicates two substitutions per 100 sites.

3.2.2.1 Network analysis

Statistical parsimony analysis of the CO1 dataset shows a well defined Italian and a split up Croatian network. The separated Croatian haplotypes are reconnected to a whole network as expected from the ML and BI analyses (Fig. 7).

All Italian haplotypes form a single network (blue network in Fig. 7), which is separated from the Croatian sequences. I18 is the hypothetical ancestor of this network. The blue haplotypes are close to the assumed ancestor. In the unconnected green "Croatian" network three haplotype clusters are recognised. The clusters are K1 with K7 and K8, the group K16, K17 and K18, and the remaining haplotypes. K12 is positioned close to I18 as in the BI and ML topologies.

The haplotype lineages for both populations formed in ML and BI analyses are also shown in the network.

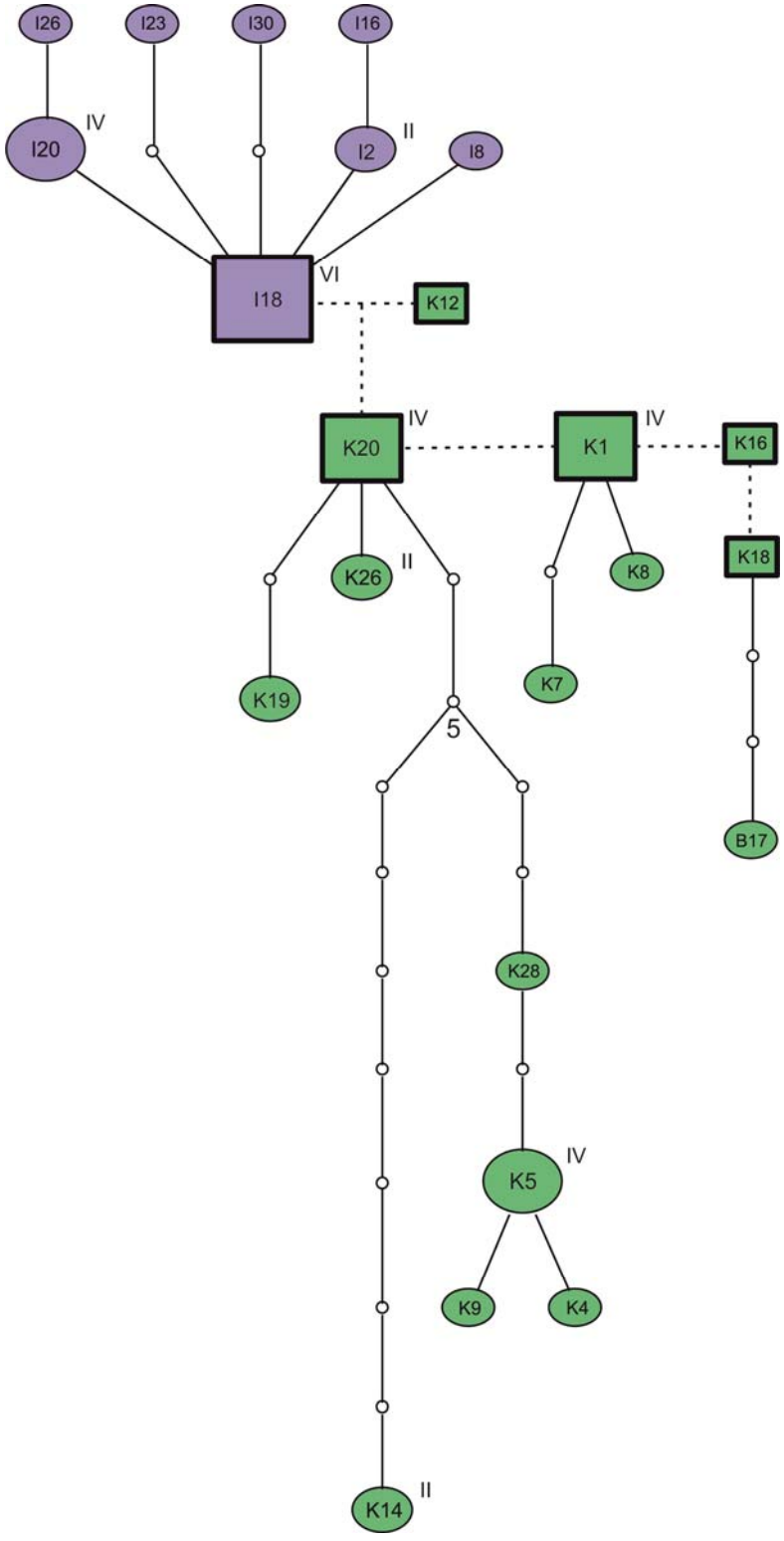

Figure *7* : TCS network of CO1 data. Parsimony probability: 90%; Haplotypes of different populations are visualised in different colors; blue for Italian and green for Croatian haplotypes. Roman numerals on the right side of the spheres and rectangles indicate haplotype frequencies. Nodes and arabic numbers between haplotypes represent unobserved haplotypes; The length of the lines is not correlated to the genetic distance. The dotted lines indicate the hypothetical connection between haplogroups according to ML and BI trees.

3.2.3 Combined dataset

BI and ML analyses of the combined dataset show two well separated clades (Fig. 8; Fig. 9). The green clade consists of almost all Croatian haplotypes, except K12, which is clustering with the blue clade. The blue clade contains all Italian haplotypes and K12.

Three different combined BI analyses were calculated. The first combined the coding CO1 with noncoding ITS, the second noncoding with each codon position of the CO1 dataset, and the third used one model for the entire combined dataset. The blue clade showed the same topologies in all BI analyses. I18 has a sister relationship to K12, and also to the group of remaining haplotypes. Within this group the better resolved ML tree showed a different topology (Fig. 9). The group I21, I26 and I27 has a sister relationship to the group of remaining haplotypes, I4 and I7, and I2, I6 and I16 cluster together. In BI analyses I4, I21, I26 and I27, and I2, I6 and I16 form haplotype groups. Compared with ITS single gene analysis I18 is not clustering with the Croatian haplotypes. The different combined BI analyses showed different topologies for the clusters in the green clade, but with similar support values. The haplotype clusters A, B, C, and D are the same in BI and ML analyses, only the topology differs and the ML tree is better resolved. Posterior probability values of internal nodes are high in all combined analyses.

The topologies of ML and BI analyses indicate three Italian haplotype lineages, I18, I20 and the group of remaining haplotypes. The Croatian population showed four lineages, the clusters A, D, K12 and the groups B and C together.

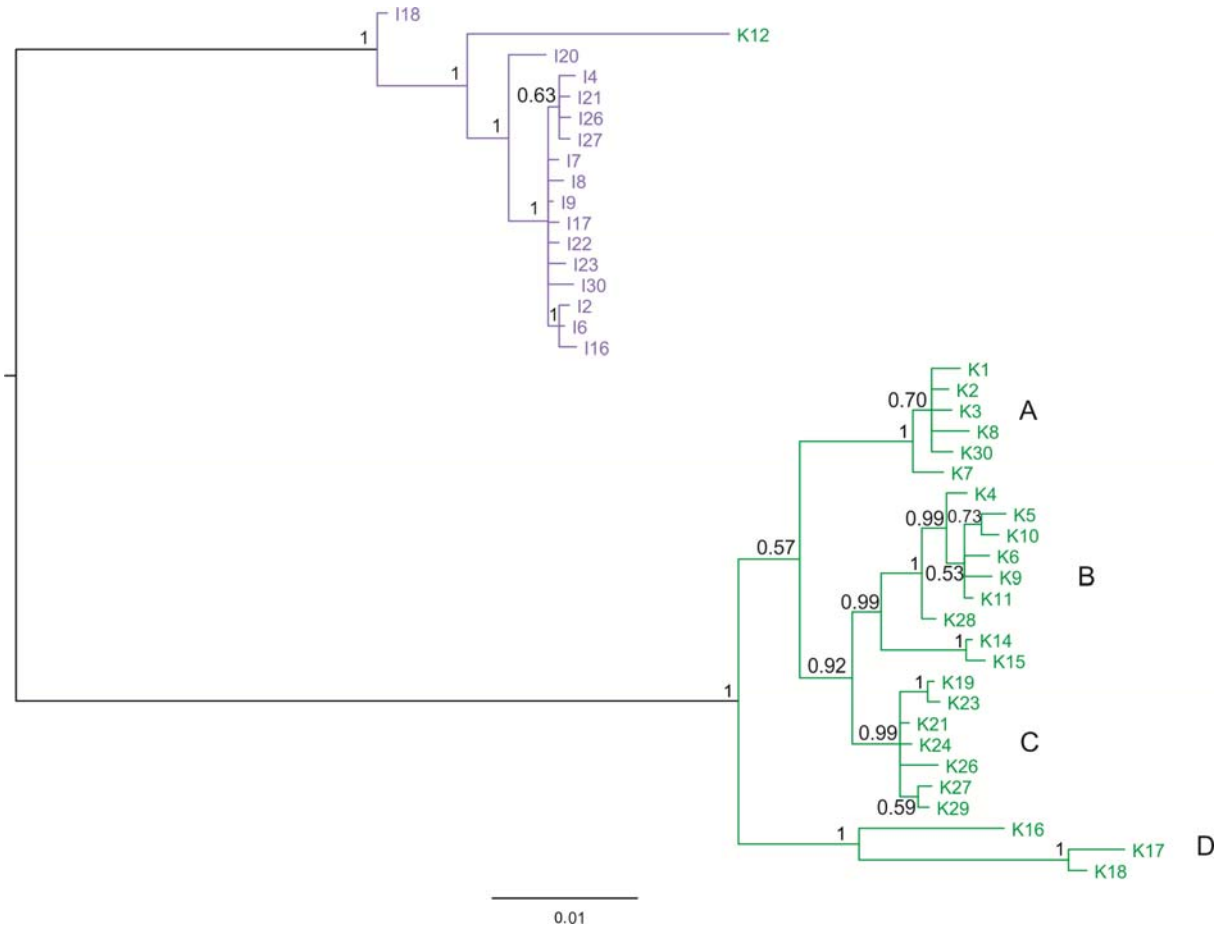

Figure *8* : Bayesian inference tree of the combined ITS and CO1 datasets using 2 partitions (ITS; CO1 all positions). Values at nodes are posterior probabilities. Haplotype clades are labelled with letters. Haplotypes of different populations are visualised in different colors, blue for Italian and green for Croatian haplotypes. The tree is deliberately rooted on the longest branch. model=SYM+I+G (ITS); TIM+I (CO1 all codon positions); The scale bar indicates one substitution per 100 sites.

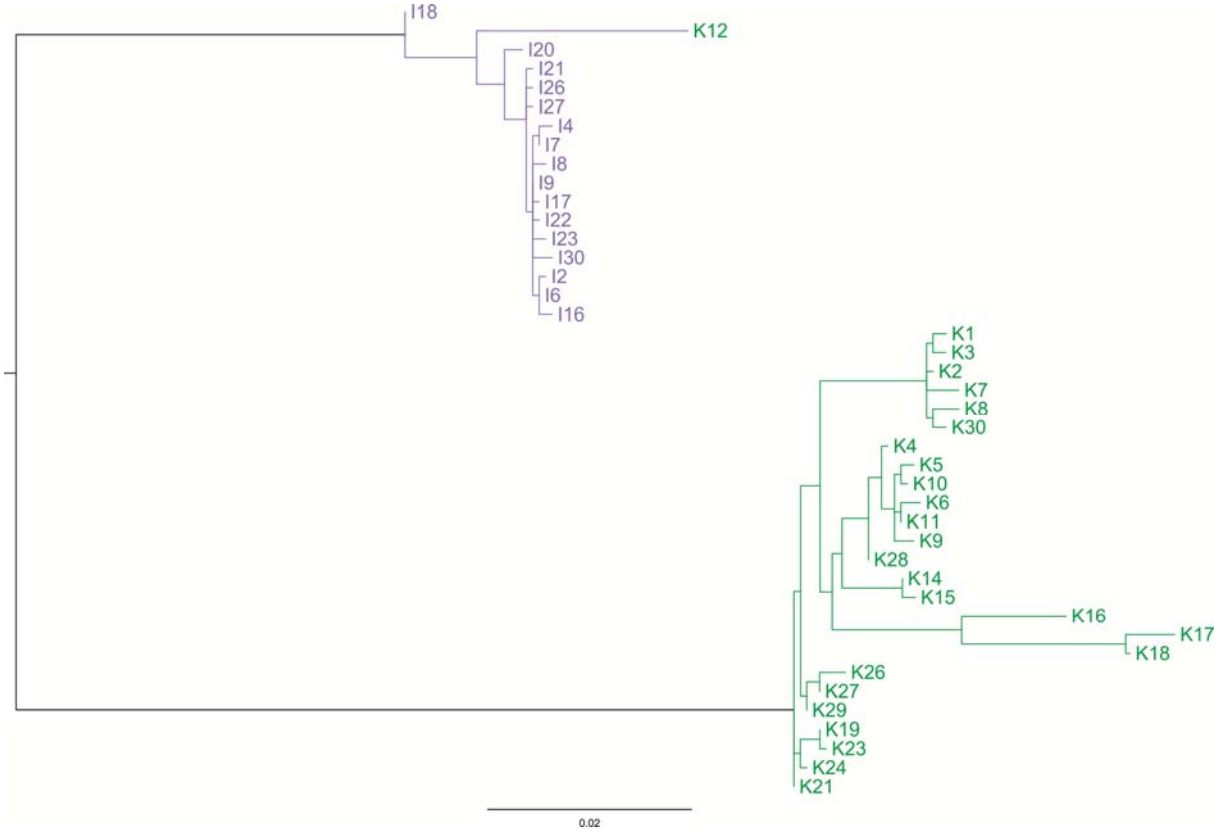

Figure 9 : ML tree of the combined ITS and CO1 datasets. Haplotypes of different populations are visualised in different colors, blue for Italian and green for Croatian haplotypes. The tree is deliberately rooted on the longest branch. number of trees=593; tree length=246; CI=0.6423; RC=0.6084; -lgL=3939.91827; model=HKY+I+G; The scale bar indicates two substitutions per 100 sites. [CI-consitency index; RC-rescaled consitency index]

3.2.3.1 Network analysis

TCS analysis of the combined dataset produced two separated networks and three single haplotypes (Fig. 10). The green network contains almost all Croatian haplotypes, but four sequences are separated. The blue network consists of almost all Italian specimens, only one haplotype is separated. The detached sequences are reconnected to a whole network as expected from the ML and BI analyses.

The Italian network is well defined and almost all haplotypes are closely positioned to their hypothetical ancestor I9. I20, I21, I2 and I16 are not directly linked with the assumed ancestor. I4 and I7 are clustering together like in the ML analysis. The only detached haplotype from the network is I18. The Croatian haplotypes result in an interlinking,

stretched network. Partially the haplotypes are distant from the assumed ancestor K19. The highest base difference of 52 is between the K19 and K15 or K8. K12, K16 and K17 with K18 are detached from the other haplotypes and reconnected as expected from ML and BI trees.

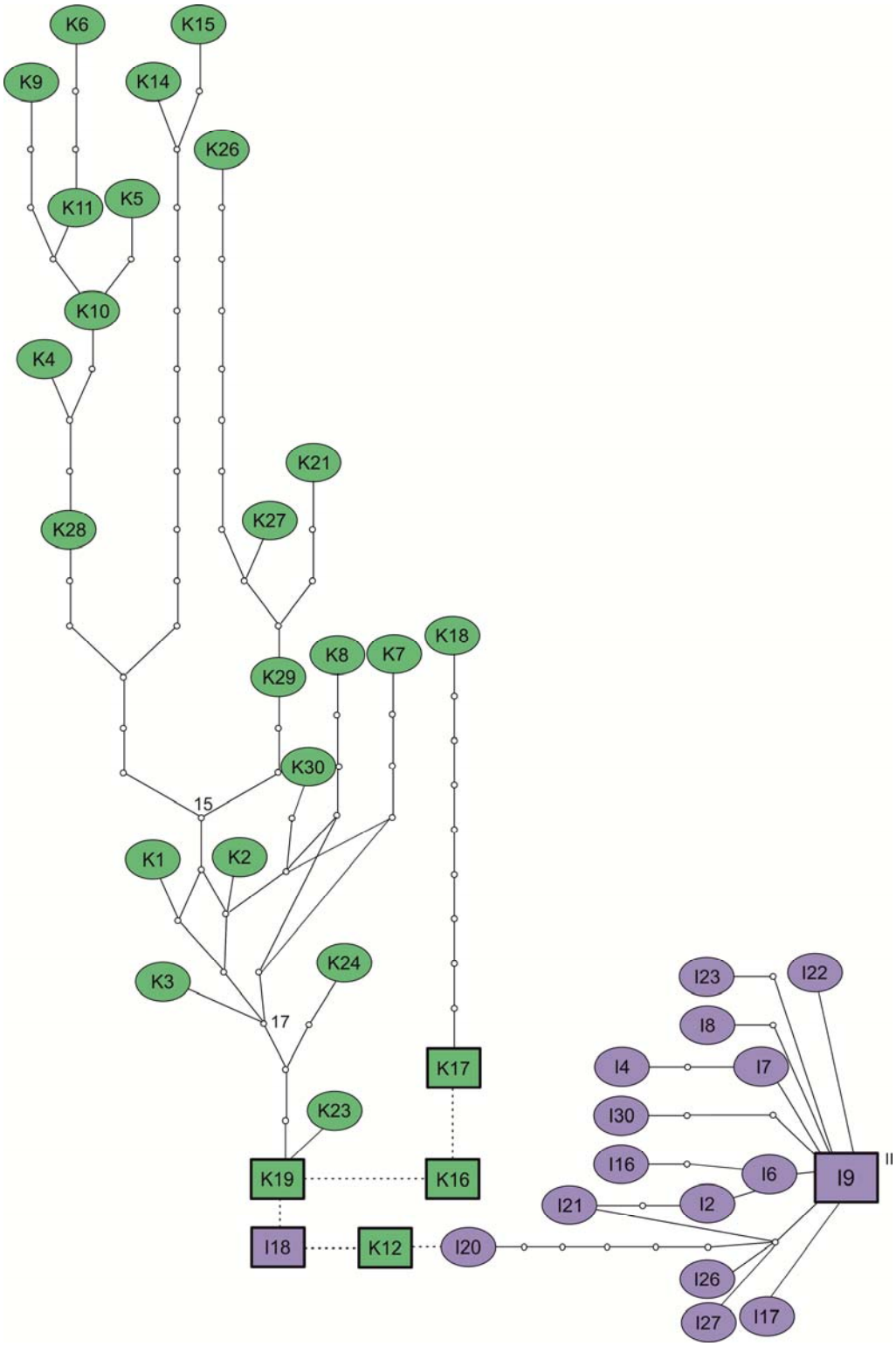

Figure *10* : TCS network of the combined dataset. Parsimony probability: 90%; Haplotypes of different populations are visualised in different colors, blue for Italian and green for Croatian haplotypes. Roman numerals on the right side of the spheres and rectangles indicate haplotype frequencies. Nodes and arabic numbers between haplotypes represent unobserved haplotypes. The length of the lines is not correlated to the genetic distance. The dotted lines indicate the hypothetical connection between haplogroups according to ML and BI trees.

3.2.4 Comparison between *Plakosyllis brevipes* and other syllids

The NJ tree of the CO1 dataset showed that the *Plakosyllis brevipes* populations form a single clade. The branch lengths between the Italian and Croatian populations are shorter than the branches between the other species, which shows a high genetic distance between *P. brevipes* and the other syllids (Fig. 11). The genetic distance among these populations is 2.1-6.7% and the distance to other syllids is between 25% and 35%. Genetic distances between members of the families Exogoninae (*Sphaerosyllis sp*.; *Prosphaerosyllis isabellae*) and Syllinae (*Eurysyllis tuberculata*; *Parahaplosyllis brevicirra*) are 28.5% and 29.4%.

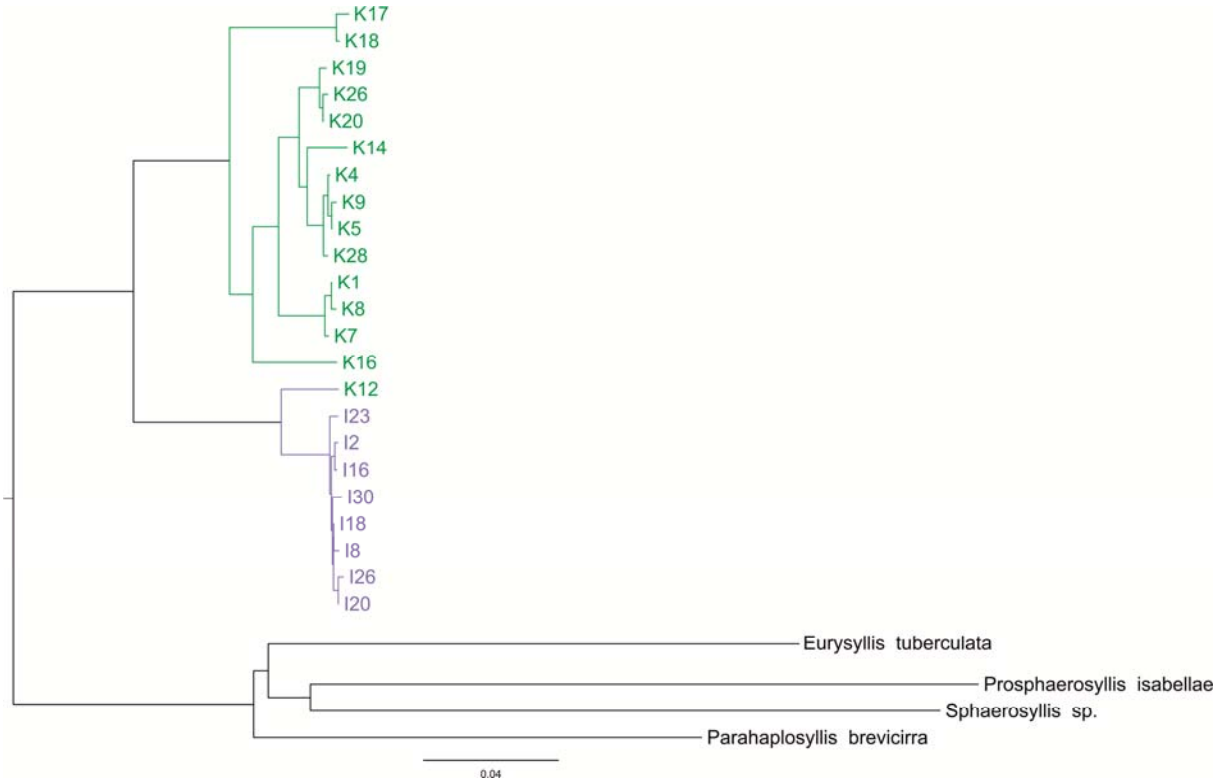

Figure 11 : NJ tree of uncorrected p-distances of the CO1 dataset. Haplotypes of different populations are visualised in different colors, blue for Italian and green for Croatian haplotypes. The tree is deliberately rooted on the longest branch The scale bar indicates four substitutions per 100 sites. *Eurysyllis tuberculata* (EF123748.1); *Prosphaerosyllis isabellae* (JF903764.1); *Sphaerosyllis sp.* (EF123767.1); *Parahaplosyllis brevicirra* (JF903784.1)

3.3 Statistical analysis and genetic variation

The distributions of genetic distances of all three datasets are shown in Figure 12. The p-distance distributions among and within populations are bimodal and overlapping. A so-called barcoding gap is absent between the two maxima.

The ITS dataset shows the lowest pairwise distances among (0.3-2.5%) and within (0-2%) populations, and CO1 the highest (among: 3-12.9%; within: 0-12.8%). The combined dataset shows intermediate distances. There are 0-6.1% of sequence differences within and 2.1-6.7% among these geographically separated populations. The Mann-Whitney-U Test of uncorrected p-distances among and within populations in all three datasets confirmed, that the distances within populations are significantly lower than among them ($p<0.001$).

The matrices of uncorrected p-distances of the ITS and CO1 datasets are shown in Table 8 and Table 9. The average genetic distance within populations is 0.48% for the ITS dataset, 1.34% for combined sets and 2.57% for CO1. The calculated standard thresholds according to Hebert et al. (2004) are 4.8% for ITS, 25.7% for CO1 and 13.4% for the combined dataset.

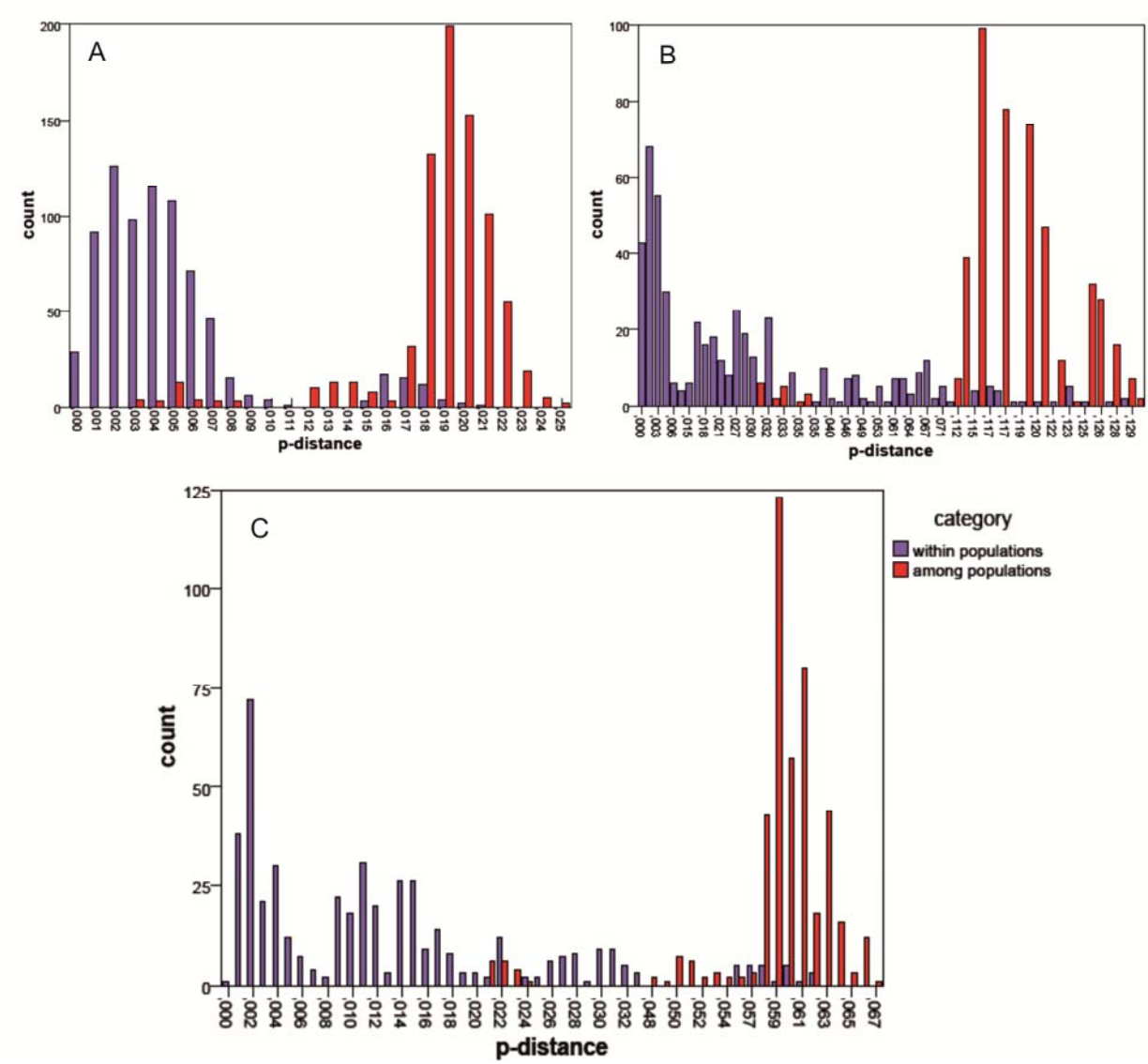

Figure *12* : Genetic distances within and among populations; the uncorrected p-distances of all datasets are given. A: ITS single gene dataset; B: CO1 single gene dataset; C: combined dataset

CO1	I2	I26	I6	I22	I21	I17	I7	I4	I20	I27	I16	I9	I8	I18	I30	I23	I24	K12	K16	K29	K21	K30	K2	K1	K23	K7	K3	K28	K14	K26	K15	K19	K9	K11	K10	K6	K5	K4	K8	K18
I2																																								
I26	0.0046																																							
I6	-	0.0046																																						
I22	0.0015	0.0030	0.0015																																					
I21	0.0030	0.0015	0.0030	0.0015																																				
I17	0.0015	0.0030	0.0015	-	0.0015																																			
I7	0.0015	0.0030	0.0015	-	0.0015	-																																		
I4	0.0030	0.0015	0.0030	0.0015	-	0.0015	0.0015																																	
I20	0.0030	0.0015	0.0030	0.0015	-	0.0015	0.0015	-																															
I27	0.0030	0.0015	0.0030	0.0015	-	0.0015	0.0015	-	-																														
I16	0.0015	0.0061	0.0015	0.0030	0.0046	0.0030	0.0030	0.0046	0.0046	0.0046																														
I9	0.0015	0.0030	0.0015	+	0.0015	-	-	0.0015	0.0015	0.0030	0.0015																													
I8	0.0030	0.0046	0.0030	0.0015	0.0030	0.0015	0.0015	0.0030	0.0030	0.0046	0.0015	0.0015																												
I18	0.0015	0.0030	0.0015	+	0.0015	-	-	0.0015	0.0015	0.0030	-	0.0015																												
I30	0.0046	0.0061	0.0046	0.0030	0.0046	0.0030	0.0030	0.0046	0.0046	0.0061	0.0030	0.0046	0.0030																											
I23	0.0046	0.0061	0.0046	0.0030	0.0046	0.0030	0.0030	0.0046	0.0046	0.0061	0.0030	0.0046	0.0030	0.0061																										
I24	0.0015	0.0030	0.0015	+	0.0015	-	-	0.0015	0.0015	0.0030	-	0.0015	-	0.0030	0.0030																									
K12	0.0334	0.0350	0.0334	0.0319	0.0334	0.0319	0.0319	0.0334	0.0334	0.0350	0.0319	0.0334	0.0319	0.0350	0.0350	0.0319																								
K16	0.1261	0.1277	0.1261	0.1246	0.1261	0.1246	0.1246	0.1261	0.1261	0.1246	0.1246	0.1261	0.1246	0.1277	0.1277	0.1246	0.1246	0.1277																						
K29	0.1155	0.1170	0.1155	0.1140	0.1155	0.1140	0.1140	0.1155	0.1155	0.1170	0.1140	0.1155	0.1140	0.1170	0.1170	0.1140	0.1140	0.1155	0.0471																					
K21	0.1155	0.1170	0.1155	0.1140	0.1155	0.1140	0.1140	0.1155	0.1155	0.1170	0.1140	0.1155	0.1140	0.1170	0.1170	0.1140	0.1140	0.1155	0.0471	-																				
K30	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0532	0.0274	0.0274																			
K2	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0532	0.0274	0.0274	-																		
K1	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0532	0.0274	0.0274	-	-																	
K23	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0456	0.0030	0.0304	0.0030	0.0030	0.0274																
K7	0.1140	0.1155	0.1140	0.1125	0.1140	0.1125	0.1125	0.1140	0.1140	0.1155	0.1125	0.1140	0.1125	0.1155	0.1155	0.1125	0.1125	0.1170	0.0532	0.0304	0.0304	0.0274	-	-	0.0274															
K3	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0532	0.0274	0.0274	-	-	-	0.0274	0.0030													
K28	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0532	0.0274	0.0274	0.0137	0.0319	0.0319	0.0167	0.0289	0.0319												
K14	0.1261	0.1277	0.1261	0.1246	0.1261	0.1246	0.1246	0.1261	0.1261	0.1246	0.1246	0.1261	0.1246	0.1277	0.1277	0.1246	0.1246	0.1292	0.0471	0.0213	0.0213	0.0380	0.0380	0.0380	0.0228	0.0380	0.0380	0.0274	0.0137	0.0213									
K26	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0486	0.0015	0.0015	0.0289	0.0289	0.0289	0.0046	0.0319	0.0289	0.0152	0.0152	0.0228									
K20	0.1155	0.1170	0.1155	0.1140	0.1155	0.1140	0.1140	0.1155	0.1155	0.1170	0.1140	0.1155	0.1140	0.1170	0.1170	0.1140	0.1140	0.1155	0.0471	-	-	0.0274	0.0274	0.0274	-	0.0304	0.0274	0.0137	0.0213	0.0213	0.0015								
K27	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0486	0.0015	0.0015	0.0289	0.0289	0.0289	0.0046	0.0319	0.0289	0.0152	0.0228	-	0.0015								
K24	0.1155	0.1170	0.1155	0.1140	0.1155	0.1140	0.1140	0.1155	0.1155	0.1170	0.1140	0.1155	0.1140	0.1170	0.1170	0.1140	0.1140	0.1155	0.0471	-	-	0.0274	0.0274	0.0274	-	0.0304	0.0274	0.0137	0.0213	0.0015	-	0.0015							
K15	0.1261	0.1277	0.1261	0.1246	0.1261	0.1246	0.1246	0.1261	0.1261	0.1246	0.1246	0.1261	0.1246	0.1277	0.1277	0.1246	0.1246	0.1292	0.0471	0.0213	0.0213	0.0380	0.0380	0.0380	0.0228	0.0380	0.0380	0.0274	0.0167	0.0228	0.0213								
K19	0.1170	0.1185	0.1170	0.1155	0.1170	0.1155	0.1155	0.1170	0.1170	0.1185	0.1155	0.1170	0.1155	0.1185	0.1185	0.1155	0.1155	0.1170	0.0456	0.0030	0.0030	0.0274	0.0274	0.0274	-	0.0304	0.0274	0.0167	0.0228	0.0046	0.0030	0.0228							
K9	0.1216	0.1231	0.1216	0.1201	0.1216	0.1201	0.1201	0.1216	0.1216	0.1231	0.1201	0.1216	0.1201	0.1231	0.1231	0.1201	0.1201	0.1246	0.0471	0.0162	0.0162	0.0334	0.0334	0.0334	0.0213	0.0334	0.0334	0.0046	0.0213	0.0198	0.0162	0.0198	0.0213						
K11	0.1201	0.1216	0.1201	0.1185	0.1201	0.1185	0.1185	0.1201	0.1201	0.1216	0.1185	0.1201	0.1185	0.1216	0.1216	0.1185	0.1185	0.1231	0.0456	0.0167	0.0167	0.0319	0.0319	0.0319	0.0196	0.0289	0.0319	0.0030	0.0198	0.0162	0.0167	0.0182	0.0198	0.0015					
K10	0.1201	0.1216	0.1201	0.1185	0.1201	0.1185	0.1185	0.1201	0.1201	0.1216	0.1185	0.1201	0.1185	0.1216	0.1216	0.1185	0.1185	0.1231	0.0456	0.0167	0.0167	0.0319	0.0319	0.0319	0.0196	0.0289	0.0319	0.0030	0.0198	0.0162	0.0167	0.0182	0.0198	0.0015	-				
K6	0.1201	0.1216	0.1201	0.1185	0.1201	0.1185	0.1185	0.1201	0.1201	0.1216	0.1185	0.1201	0.1185	0.1216	0.1216	0.1185	0.1185	0.1231	0.0456	0.0167	0.0167	0.0319	0.0319	0.0319	0.0196	0.0289	0.0319	0.0030	0.0198	0.0162	0.0167	0.0182	0.0198	0.0015	-	-			
K5	0.1201	0.1216	0.1201	0.1185	0.1201	0.1185	0.1185	0.1201	0.1201	0.1216	0.1185	0.1201	0.1185	0.1216	0.1216	0.1185	0.1185	0.1231	0.0456	0.0167	0.0167	0.0319	0.0319	0.0319	0.0196	0.0289	0.0319	0.0030	0.0198	0.0162	0.0167	0.0182	0.0198	0.0015	-	-	-		
K4	0.1201	0.1216	0.1201	0.1185	0.1201	0.1185	0.1185	0.1201	0.1201	0.1216	0.1185	0.1201	0.1185	0.1216	0.1216	0.1185	0.1185	0.1231	0.0441	0.0152	0.0152	0.0304	0.0304	0.0304	0.0162	0.0274	0.0304	0.0046	0.0162	0.0167	0.0152	0.0167	0.0182	0.0030	0.0015	0.0015	0.0015		
K8	0.1185	0.1201	0.1185	0.1170	0.1185	0.1170	0.1170	0.1185	0.1185	0.1201	0.1170	0.1185	0.1170	0.1201	0.1201	0.1170	0.1170	0.1185	0.0547	0.0289	0.0289	0.0015	0.0015	0.0015	0.0289	0.0046	0.0015	0.0304	0.0289	0.0304	0.0289	0.0395	0.0389	0.0350	0.0334	0.0334	0.0334	0.0319	
K18	0.1261	0.1277	0.1261	0.1246	0.1261	0.1246	0.1246	0.1261	0.1261	0.1246	0.1246	0.1261	0.1246	0.1277	0.1277	0.1246	0.1246	0.1277	0.0608	0.0669	0.0669	0.0608	0.0608	0.0608	0.0623	0.0669	0.0608	0.0623	0.0638	0.0608	0.0638	0.0623	0.0638	0.0608	0.0623	0.0623	0.0623	0.0608	0.0684
K17	0.1292	0.1307	0.1292	0.1277	0.1292	0.1277	0.1277	0.1292	0.1292	0.1277	0.1277	0.1292	0.1277	0.1307	0.1307	0.1277	0.1277	0.1216	0.0517	0.0653	0.0653	0.0714	0.0714	0.0714	0.0653	0.0714	0.0669	0.0653	0.0653	0.0669	0.0653	0.0653	0.0669	0.0684	0.0669	0.0669	0.0653	0.0729	0.0046

Table 8 : Uncorrected p-distances (uncorrected percentage of nucleotide differences) of the complete CO1 dataset. 1 indicates 100% difference between sequences. "-" indicate zero distances

Table 9 : Uncorrected p-distances (uncorrected percentage of nucleotide differences) of the complete ITS dataset. 1 indicates 100% difference between sequences. "-" indicate zero distances

The amount of genetic variation that is due to population differences is calculated by F_{ST}. The average and specific values are given in Table 10. The average fixation index for the combined dataset is 0.81. Therefore, the calculated number of average migrants per generation is 0.11.

Table 10 : Average F_{ST} and specific F_{ST} are given for each dataset and population

dataset	Average F_{ST}	Croatia F_{ST}	Italy F_{ST}
Combined	0.81343	0.80958	0.81932
CO1	0.81156	0.80647	0.81964
ITS	0.81931	0.81836	0.82049

Table 11 shows the percentage of variation among and within the populations in reference to each dataset. 81% of variation is explained by the comparison among Italian and Croatian populations and only less than 19% can be explained by comparison within populations. The genetic differentiation of these populations is high, which is also visible in the well separated green and blue clades in ML and BI analyses.

Table 11 : The differences between the sums of squares (SSD), the variance among (Va) and within (Vb) populations and the total percentage of variation are given

Dataset		SSD	Va/Vb	% variation
Combined				
	Among populations	939.283	45.185	81.34
	Within populations	424.903	10.364	18.66
	Total	1364.186	55.549	
CO1				
	Among populations	671.131	31.813	81.16
	Within populations	310.255	7.387	18.84
	Total	981.386	39.201	
ITS				
	Among populations	378.03	13.55	81.93
	Within populations	161.363	2.988	18.07
	Total	539.393	16.538	

3.4 Saturation analysis

The saturation analysis compares patristic and absolute distances of each dataset. Patristic distances show the amount of genetic change and absolute distances the absolute number of nucleotide differences between these two populations. The plots show a linear trend in both datasets, with a slight saturation in the ITS set (Fig. 13). Patristic distances in the CO1 dataset are higher (0-89) than in ITS (0-27) and have an interrupted trend. The amount of observed transitions and transversions is higher in the Croatian population. The difference between transitions and transversions is higher in CO1 than in the ITS dataset (Tab. 6). Each set shows a higher amount of transitions than transversions.

CO1 (all codon positions) and ITS saturation plots of TrN93 distances show a linear trend in transitions and transversions with high regression values (Fig. 14). All analyses have a higher rate of transitions than transversions due to a higher amount of transitions in the datasets. CO1 (1^{st} and 2^{nd} codon position) shows a low R^2-value in the transversion rate.

Due to the fact that both datasets are not or only low saturated coalescence analyses are possible (Wilke et al. 2009).

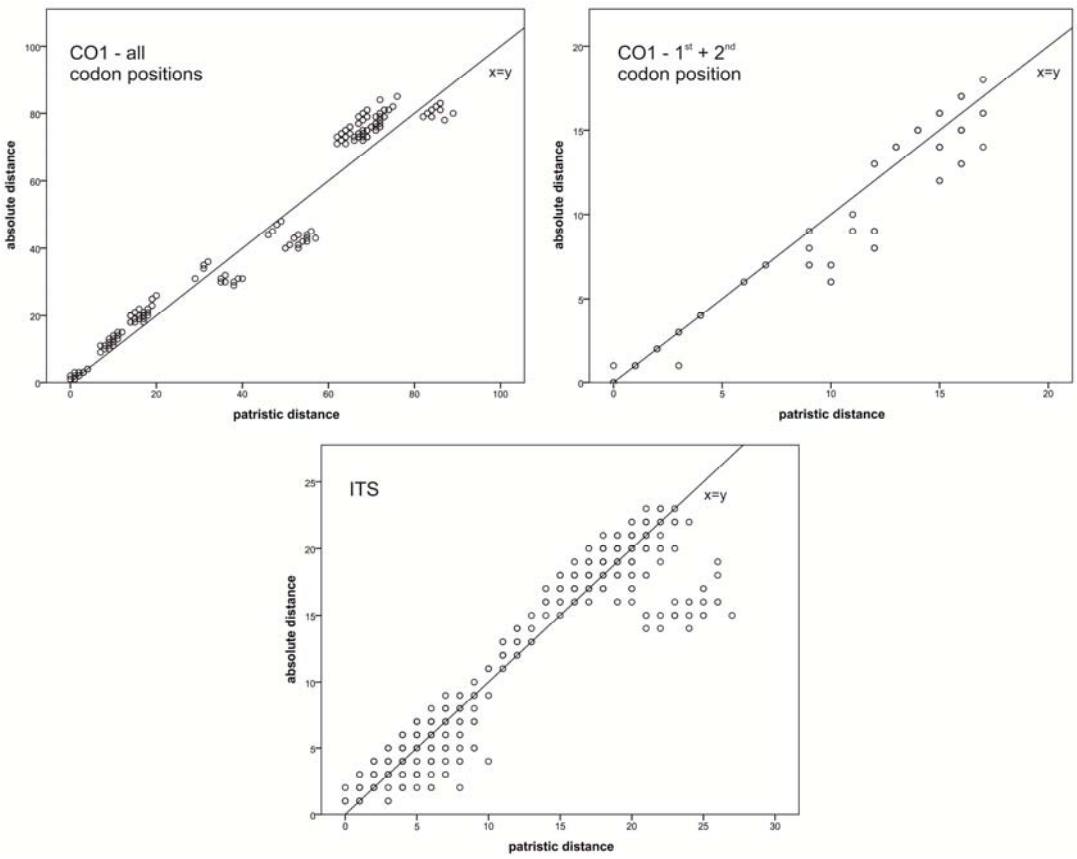

Figure 13 : Saturation plots calculated from patristic and absolute distances of the CO1 and ITS datasets, with a reference line (x=y).

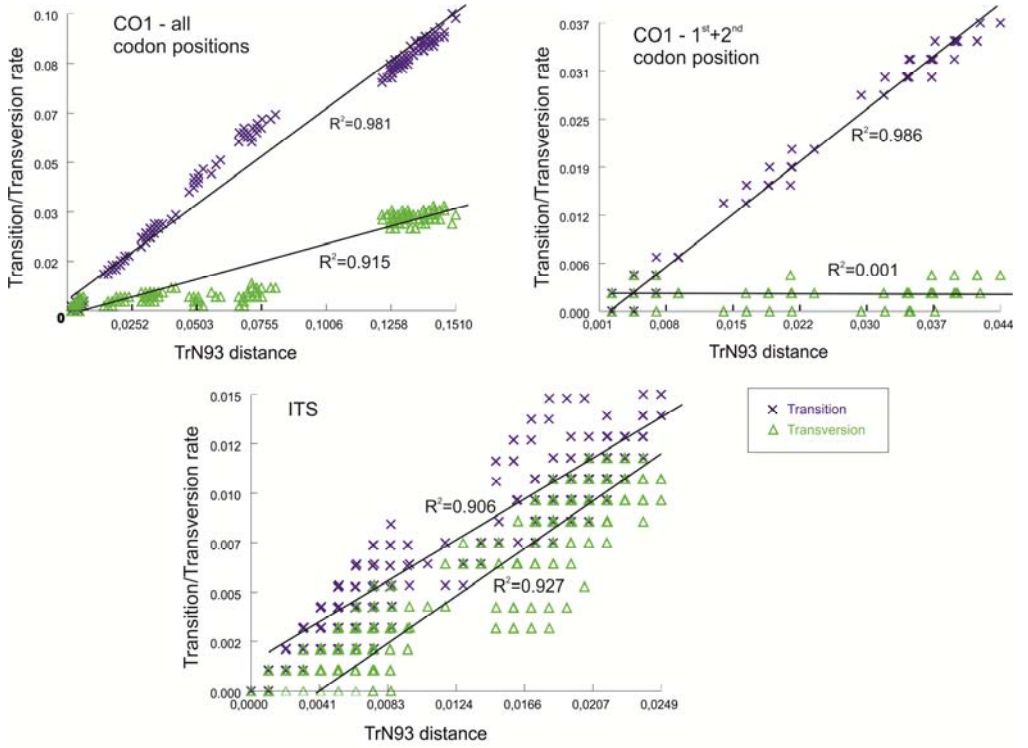

Figure 14 : Saturation plots of transition and transversion rates along TrN93 distances from CO1 and ITS datasets, with linear regression lines and R^2-values.

3.5 Coalescence analysis

Coalescence analysis determined the point of separation for Croatian and Italian haplotype lineages in the ITS dataset about 2.85 myr ago (Tab. 12; Fig. 15). The separation of K12 and the Italian haplotypes happened 2.06 myr ago. I18 clusters with K7 like in ML analysis and they separated 0.31 myr ago. Coalescence topologies are similar to genealogical analyses (Fig. 3; Fig. 4).

CO1 coalescence analyses show higher separation dates for the Italian and Croatian populations, 15.34 myr for invertebrate substitution rates and 27 myr for polychaete rates. The formed haplotype clusters are the same as in genealogical analyses (Fig. 6). The lowest separation date is 2.89 myr calculated with invertebrate substitution rates, where only 1^{st} and 2^{nd} codon position were used. The calculated divergence dates and corresponding substitution rates are shown in Table 12.

Table 12 : Summery of substitution rates for each partition with calculated divergence date. References are given. [cp-codon position]

Dataset	Partition	Model	Substitution rate per myr (%)	Divergence date (myr)	Reference
ITS	all positions	TrN93+GI	0.3354%	2.85 (+1.25/-1.13)	Nygren et al. 2009
CO1	all positions	TrN93+G	0.37%	27 (+8.39/-7.65)	Nygren et al. 2009
	all positions		0.65%	15.34 (+4.72/-4.45)	Crandall et al. 2012
	1^{st}+2^{nd} cp		0.37%	5.06 (+2.41/-2.12)	Nygren et al. 2009
	1^{st}+2^{nd} cp		0.65%	2.89 (+1.39/-1.21)	Crandall et al. 2012
	1^{st}+2^{nd} and 3^{rd} cp		0.1% for 1^{st}+2^{nd} cp 1% for 3^{rd} cp	23.14 (+5.36/-5.07)	based on Nygren et al. 2009
	1^{st}, 2^{nd} and 3^{rd} cp		0.41-0.46% for 1^{st} cp 0.17-0.19% for 2^{nd} cp 6.11-6.84% for 3^{rd} cp	4.54 (+1.17/-1.24)	Marko 2002

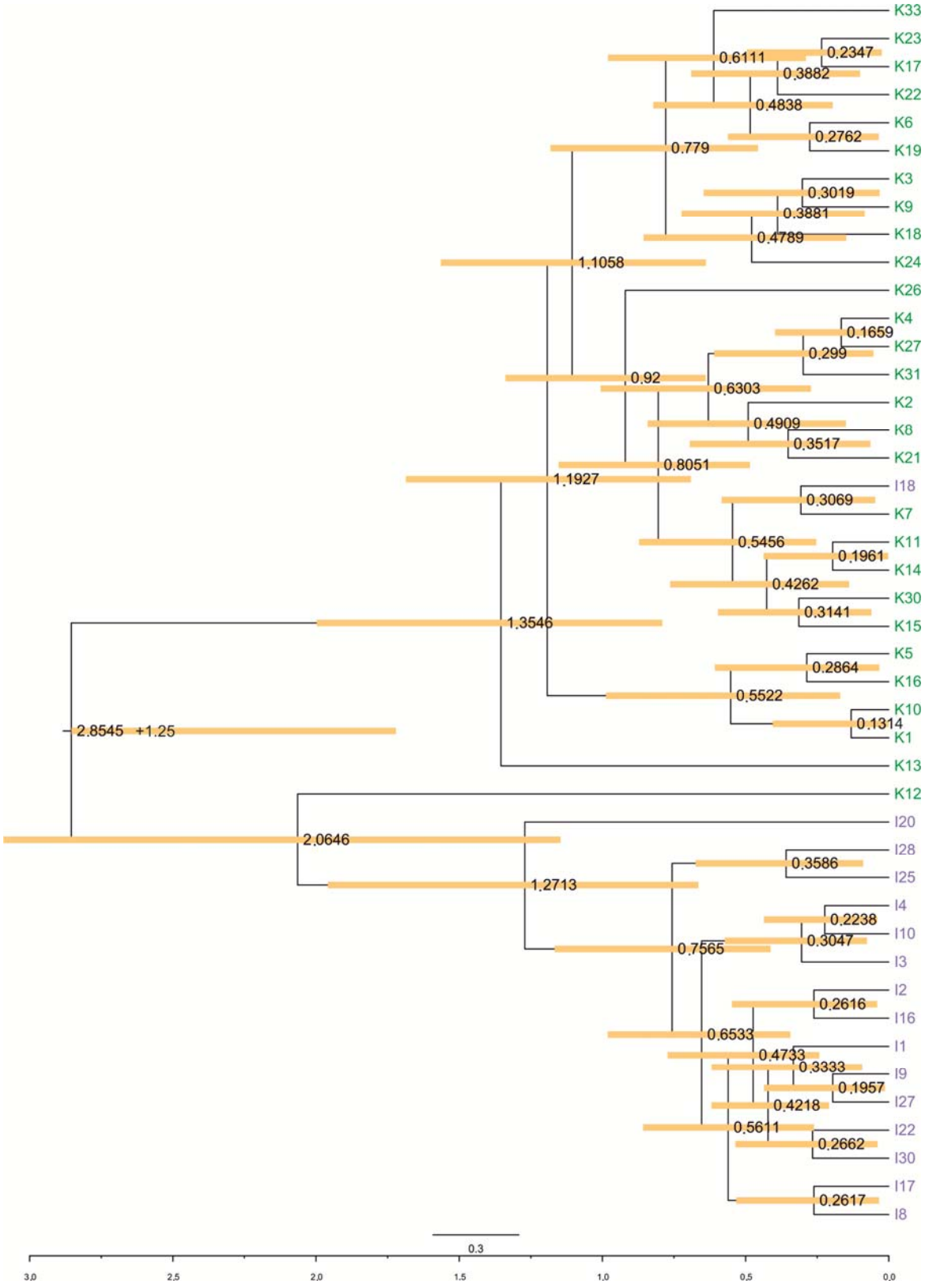

Figure 15 : Coalescence tree of the ITS dataset. The Croatian haplotypes are shown in green, the Italian ones in blue. Reverse scale axis indicating million years. Node bars indicate 95% HPD. 1.25 myr from the error bar of the basal node is omitted.

4 Discussion

Mitochondrial and nuclear markers were used to investigate whether the widely separated populations of *Plakosyllis brevipes* are conspecific, or the genetic differentiation of these populations leads to the hypothesis of cryptic species.

4.1 Comparison of all three datasets

All analyses show that the geographically separated populations of *Plakosyllis brevipes* are genetically distinct. The F_{ST} values are above 0.80 for all three datasets and the migration rate is 0.11, which indicates a high differentiation. Furthermore, the variation among populations is higher, about 81%, than within (19%). The ML and BI analyses show two well separated clades, which indicate two genetically different lineages. However the clades do not strictly correspond to locations. High posterior probability values support this topology. However, the separation of the Italian and Croatian populations is not complete, since there are Croatian individuals clustering with the Italian ones and vice versa. Furthermore a low genetic distance among these Mediterranean populations indicates less distinctness. Coalescence analyses calculated the point of separation of both populations about 2.9 myr ago. It is possible, that not enough time has passed to genetically separate these populations completely.

4.2 Coalescence analysis

A faunistically important event in the Mediterranean Sea was the Messinian salinity crisis, caused by isolation from the Atlantic and progressive evaporation, erosion and deposition in the Mediterranean basin. Massive extinction was the consequence. The crisis lasted between 5.96±0.02 and 5.33 myr (Krijgsman 1999, and references therein). The Zanclean flood occurred as the Strait of Gibraltar, was formed and water from the Atlantic Ocean filled the Mediterranean basin. After a few months to two years 90 per cent of the water was already transferred into the basin (Garcia-Castellanos et al. 2009). Coalescence analyses of the ITS dataset and CO1 (1^{st}+2^{nd} codon position) show the lowest separation dates of about 2.9 myr. This date fits the migration of *Plakosyllis brevipes* into the Mediterranean Sea after the

Zanclean flood with following separation into Croatian and Italian populations. Gene flow during the last glacial maximum was possible, but is not proven by coalescence analyses. A possible explanation is the low sample size and that the haplotypes that were created during the LGM, were not collected. Another cause for absent haplotypes could be extinction.

Molecular clock analyses for *Plakosyllis brevipes* are difficult due to the lack of fossil calibration points. Therefore the use of already calculated substitution rates is necessary. It is possible, that the used substitution rates are not appropriate for *P. brevipes*. Higher rates could lead to a more recent point of separation.

4.3 Cryptic species or conspecific populations

4.3.1 General comparison

Cryptic species cannot be distinguished by morphological features, but by other characteristics, like chemical and auditory signals and the nucleotide sequence. The distribution of genetic distances is a good indicator whether cryptic species are present or not. All datasets show overlapping distributions and maxima in the genetic distances among and within populations. A barcoding gap between these two distributions is absent. It occurs if all intra-population distances are lower than the inter-population distances. This should enable the identification of cryptic species (Hebert et al. 2003; Meyer and Paulay 2005). Due to the overlapping distribution of genetic distances and the absence of a barcoding gap, cryptic species can be excluded in this study.

According to Hebert et al. (2003) the CO1 fragment from the mitochondrial genome is sufficient in most Metazoa to identify to the species level. ITS is also a useful marker to distinguish between closely related species, because it evolves more rapidly than the coding regions (Hills and Dixon 1991). Other studies confirmed the usefulness of ITS (Westheide and Hass-Cordes 2001; Westheide and Schmidt 2003; Nygren et al. 2009). The datasets of sequences obtained with these markers showed genetic distances with different lengths. Distances in the ITS dataset are the lowest and vary between 0 and 2.5%, the CO1 distances lie between 0 and 13%. Hebert et al. (2004) proposed a standard screening threshold of

sequence difference of at least 10 X average intra-specific distance. Genetic distances among the Italian and Croatian populations are lower than the calculated threshold of 25%, which is allowed for the CO1 dataset. Compared with Nygren et al. (2009), who assumed a distance of 14.6±0.35% for their cryptic species (*Paranaitis wahlbergi*, Phyllodocidae) in the CO1 dataset, the distances of *Plakosyllis brevipes* among populations are lower. The ITS1 and ITS2 regions also showed higher values (22.2±0.4% and 16.04±0.33%) compared with *P. brevipes* (2.5%). The p-distance value is also lower than the threshold of 4.8%. Given the fact, that Nygren et al. (2009) used K2P-corrected distances (Kimura-2-Parameter), and corrected distances are usually higher than uncorrected, the genetic distances among populations of *P. brevipes* calculated in this study are still lower and therefore do not indicate different species.

Another indicator for differentiation between populations is the F_{ST} value. It varies between 0.81 and 0.82 and indicates high differentiation between Italian and Croatian populations of *P. brevipes*. Usually such a high value would also indicate separate species, but its susceptibility to a low sample size and a high number of haplotypes occurring of low frequencies complicates the interpretation (Meirmans and Hedrick 2011). The high variation among and the low variation within Italian and Croatian populations also support the hypotheses of different species.

Genealogical analysis of the ITS and CO1 datasets also show two well separated clades. In the ITS dataset these clades contain almost exclusively all Italian or all Croatian haplotypes, except one specimen, which is always clustering with the other population. The CO1 dataset showed that only the Croatian haplotype have a sister relationship to the Italian specimens, which form a monophyletic group. The allocated haplo-lineages also differ in these datasets. The differences between CO1 and ITS topologies are not surprising considering their different inheritance pattern. Mitochondria are attributed to maternal inheritance, while nuclear genes are recombining and biparentally transmitted (Avise 2000). Both populations show a dominant haplo-lineage, a single haplotype from the other lineage and different topologies in genealogical analysis. This condition can be explained by the different inheritance pattern of nuclear and mitochondrial genes. The fact that the haplotypes are found at both locations indicates conspecific populations as opposed to cryptic species.

Genealogical analysis of *Plakosyllis brevipes* and other syllids showed, that the populations of *P. brevipes* form clearly separated clades. The genetic distances between the populations are lower than to the other syllids. The distances between these syllids are higher than among populations of *P. brevipes*. If the populations belong to separate species, the genetic distance should be as high as between other syllids. This indicates, that the geographically separated populations are conspecific.

Since the genetic distances among these geographically separated populations are lower than among other syllids, and also lower than the calculated threshold as well as the values found in the literature, cryptic species are unlikely. Furthermore the distribution of genetic distances are overlapping and a barcoding gap is absent. The fact that there is a dominant haplo-lineage and single haplotypes from the other lineage at each location, also indicates conspecific populations. In favour of cryptic species is the high variation among and the low variation within these populations and the high F_{ST} value.

4.3.2 Comparison with other interstitial organisms

Ritschl (2013) assumed that the geographically separated Mediterranean populations of the interstitial syllid *Sphaerosyllis sp.* are conspecific, due to a low genetic distance (5% in the ITS dataset) and a variation of 78.31% among populations. However, the mitochondrial 16S dataset showed a higher sequence divergence (17%) and a high variation among populations (92.8%). The F_{ST} value varied between 0.78 and 0.93. However, genealogical analyses showed two well separated clades in nuclear and mitochondrial datasets.

Two geographically separated populations of the interstitial gastropod *Philinoglossa praelongata* from similar locations in the Mediterranean were examined by Trpisovsky (2013). The high genetic distance between the populations (CO1: 10.4-14.1%; 16S: 5-6.9%), the presence of a barcoding gap, the very high F_{ST} value (0.91-0.96) and the high variation among populations (91-96%) enabled the differentiation of different species. The genealogical analyses showed two highly supported, well separated clades.

Trpisovsky's (2013) results contrast Werth's (2007) on the relatedness of separated populations of interstitial gastropods. They both occur at the same locations in the

Mediterranean Sea, but the populations show a different structure. *Pontohedyle milaschewitschii* showed genetically mixed clades with maximum sequence distances of 4.83% and a low genetic variability among the widely separated populations. Werth (2007) assumed, that the *P. milaschewitschii* populations are conspecific.

The interstitial species of Ritschl (2013), Trpisovsky (2013) and Werth (2007) are from the Mediterranean Sea. They all occur in the same habitat at the same locations and have a low dispersal potential like *Plakosyllis brevipes*, but show different population structures.

In comparison with these interstitial species, the results of this study showed *that P. brevipes* has a lower genetic distance, fixation index and variation among populations than the presumed cryptic species *Philinoglossa praelongata* (Trpisovsky 2013), but the values were not as low as in Werth (2007). Ritschl (2013) showed similar values. I assume that the geographically separated populations of *P. brevipes* are conspecific.

4.3.3 Dispersal potential and geographic distance

According to Ritschl (2013), Trpisovsky (2013) and Werth (2007), who examined interstitial species from the same locations and habitats, the dispersal potential of these species is low due to the very restricted pelagic larval phase or brood care. However, they have the same distribution in the Mediterranean Sea. Due to the small body size, the amount of produced gametes is limited (Franke 1999).

Sphaerosyllis sp. reproduces by epigamy and exhibits brood care where the larvae are carried ventrally. In contrary *Plakosyllis brevipes* reproduces by stolonisation and therefore ensures a certain amount of dispersal due to the pelagic stolons (Franke 1999). In the first case dispersal is depending on the migration of the mother animal, in the second case on the swimming stolon and water currents. Numerous authors suggested that geographic distribution and the type of larvae are linked. Bhoud (1998) argued, that the area covered by larvae is larger than the adults, but they tend to spread less, than might be expected through water currents. Since the stolons of Syllinae are pelagic for only hours to a few days, and the fertilized eggs soon sink to the bottom of the ocean, the time span for distribution is limited.

On the other hand a short pelagic stage is advantageous if the habitat is restricted in range. So the offspring stays near suitable habitats, were the survivability is higher (Franke 1999).

According to Wright (1943) the genetic differentiation should increase with increasing geographic distance. The populations of *P. brevipes* are separated by a coastline of 2500 km. This geographic distance and the low dispersal potential could lead to a clear separation of Croatian and Italian populations. The high variation among and the low variation within these geographically separated populations supports this assumption. However, the low genetic distance between them and their overlapping distributions do not support Wright (1943).

4.4 How does gene flow happen?

Geographically disjunct populations tend to adapt to different local environments, while gene flow operates against it. Gene flow can maintain due to passive drift or gene hopping e.g., while geographic barriers are able to stop it.

4.4.1 Passive drift

Sand is a dynamic living environment, causing several adaptations to the inhabitants. Strong currents can easily effect the sand layers and move particles. Even weak near-bottom currents, which become enhanced as they pass over bottom ripples, are recorded to transport sediment (Giere 2009, and references therein). Adult *Plakosyllis brevipes* live in the sand where their adhesion organs enable close contact with the substrate (Swedmark 1964). Stolons are the pelagic stage and fertilized eggs and larvae tend to inhabit the upper sand layer (Franke 1999). All four stages are therefore easily transported in suspension by water currents. Due to passive drift along the continental coastlines, dispersal is possible.

4.4.2 Sea level fluctuations and the influence on gene flow

A good indicator for gene flow is the number of migrants per generation (Wright 1978), which is 0.11 in both populations. The variation among populations is high (81.48%). Considering the low number of migrants and the very high variation among Croatian and

Italian populations, the present gene flow is very low. An indication for genetic exchange is the single Croatian haplotype clustering with the Italian population in both datasets.

During reproduction it is possible for genes to spread in a population without long-scale migration of their carriers. Due to this natural process, long distance migration of individuals would not be necessary to maintain gene flow between distant populations. This process is called gene hopping and requires habitat connectivity. Since sand is a dynamic sediment and *Plakosyllis brevipes* is distributed across the Mediterranean Sea, a continuous habitat cannot be excluded.

The analysed populations of *P. brevipes* were from locations in the northern Adriatic Sea and the Tyrrhenian Sea. The examination of the sand between Pescara and Bari at the east coast of Italy showed, that *P. brevipes* is absent in the local interstitial fauna. This fact indicates an interrupted habitat. A reason for the absence of syllids could be the size of the sand grains in this area. The sand sediment at the east coast of Italy appeared to be finer as the sand found at the west coast. Martins et al. (2013) confirmed the favoured habitat of *P. brevipes* as course sand and its absence in fine sand and muddy sediments. Since the present habitat is interrupted, gene flow via gene hopping can be excluded.

At least a temporarily continuous habitat between the northern Adriatic and the Tyrrhenian Sea could establish gene flow via gene hopping. In the last glacial maximum (LGM), which occurred between 26 and 19 kyr ago, the sea level sank to its lowest position of about -120 meters (Maselli 2011,and references therein). During this period the habitat of interstitial species shifted. The animals had to migrate form the north to south, where the interstitial populations of *P. brevipes* could have come in close contact in the conducted areal. It is possible that during the LGM a continuous habitat was present between both sides of the Adriatic Sea. Therefore gene flow between populations from Croatia and Italy was feasible. The fact, that the clades in genealogical analyses do not completely correspond to locations supports gene flow during the LGM.

The period of the post LGM was characterized by the sea level rising in successive pulses until it reached the modern position 5.5 kyr ago (Maselli 2011, and references therein). During the rise of the sea level the habitats were shifted again and the interstitial species had to migrate

northwards. As a consequence gene flow between Croatian and Italian populations of *P. brevipes* stopped.

The high variation among populations of the geographically widely separated populations of *P. brevipes* and the very low migration rate indicate the absence of gene flow in the present. However, gene flow during the LGM can not be excluded.

5 Conclusion

This investigation shows that two geographically separated populations of *Plakosyllis brevipes* from sample sites in the Adriatic and Tyrrhenian Sea are likely to be conspecific. The spreading potential of stolons, fertilized eggs and larvae is low, but water currents could facilitate their distribution in the Mediterranean Sea, though it is unlikely, since a continuous habitat is not verifiable. Gene flow during the LGM, where a temporary continuous habitat was possible, could explain why a Croatian haplotype clustered with the Italian specimens in all genealogical analyses. It also explains the short genetic distance between populations from these two sample sites. However, AMOVA calculated a high differentiation and variation among populations, supported by well-separated clades in the genealogical analyses.

Coalescence analysis calculated the point of separation for both populations at about 2.9 myr ago. Therefore the separation of haplotype lineages dominant in the Croatian and the Italian populations, respectively, occurred after the Messinian salinity crisis, which ended 5.33 myr ago.

6 References

Aguado, M.T., Nygren, A. and Siddall, M. E. (2007). Phylogeny of Syllidae (Polychaeta) based on combined molecular analysis of nuclear and mitochondrial genes. Cladistics 23: 552–564.

Aguado, T.M. and San Martín, G. (2009). Phylogeny of Syllidae (Polychaeta) based on morphological data. Zoologica Scripta 38: 379-402.

Aguado, T.M., San Martín, G. and Sidall, M.E. (2012). Systematics and evolution of syllids (Annelida, Syllidae). Cladistics 28: 234-250.

Avise, J.C. (2000). Phylogeography: the history and formation of species. Third Edition, Cambridge: Cambridge University Press, 441 pp.

Bhoud, M. (1998). The spreading potential of polychaete larvae does not predict adult distributions; consequences for conditions of recruitment. Hydrobiologia 375/376: 35-47.

Böggemann, M. (2002). Revision of the Glyceridae (Annelida: Polychaeta). Abhandlungen der Senckenbergischen Naturforschenden Gesellschaft 555: 1-249.

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

Crandall, E.D., Sbrocco, E.J., DeBoer, T.S., Barber, P.H. and Carpenter, K.E. (2012). Expansion Dating: Calibrating Molecular Clocks in Marine Species from Expansions onto the Sunda Shelf Following the Last Glacial Maximum. Molecular Biology and Evolution 29 (2): 707-719.

Dahlgren, T.G., Lundberg, J., Pleijel, F. and Sundberg, P. (2000). Morphological and molecular evidence of the phylogeny of Nereidiform polychaetes (Annelida). Journal of Zoological Systematics and Evolutionary Research 38: 249-253.

Dales, R.P. (1962). The polychaete stomodeum and the inter-relationships of the families of Polychaeta. Proceedings of the Zoological Society of London 139: 389-428.

Drummond, A.J., Suchard, M.A., Xie, D. and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology And Evolution 29 (8): 1969-1973.

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

Fauchald, K. (1977). The Polychaete Worms: definitions and keys to the orders, Families and genera. Natural History Museum of Los Angeles County 28: 188 pp.

Folmer, O., Black, M., Hoeh, W., Lutz, R. and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from divers metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294-299.

Franke, H.D. (1999). Reproduction of the Syllidae (Annelida: Polychaeta). Hydrobiologia 402: 39-55.

Franke, H.D. (1980). Zur Determination der zeitlichen Verteilung von Fortpflanzungsprozessen in Laborkulturen des Polychaeten *Typosyllis prolifera*. Helgoländer Meeresuntersuchungen 34: 61–84.

Garcia-Castellanos, D., Estrada, F., Jiménez-Munt, I., Gorini, C., Fernàndez, M., Vergés, J. and De Vicente, R. (2009). Catastrophic flood of the Mediterranean after the Messinian salinity crisis. Nature 462: 778-782.

Giere, O. (2009). Meiobenthology: The Microscopic Motile Fauna of Aquatic Sediments. Second Edition. Springer Verlag, pp 527.

Glasby, C.J. (2000). Family Syllidae. In: Beesley, P.L., Ross, G.J.B., Glasby, C.J. (Eds.), Polychaetes and Allies: The Southern Synthesis. Fauna of Australia. Vol. 4: Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula. CSIRO Publishing, Melbourne, Vic., pp 161–167.

Hartmann-Schröder, G. (2006). Polychaeta. Second Edition. Gustav Fischer Verlag pp 648.

Hebert, P.D., Cywinska, A., Ball, S.L. and de Waard, J.R. (2003). Biological identifications through DNA barcodes. Proceedings of the Biological Society 270: 313-321.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S. and Francis, C.M. (2004). Identification of Birds through DNA Barcodes. PLoS Biology 2 (10): e312.

Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. Quaterly Review of Biology 66: 411–429.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J. and Wilsonk, D.S. (1999).Chronology, causes and progression of the Messinian salinity crisis. Nature 400: 652-655.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.

Malaquin, A. (1893). Recherches sur les Syllidiens. Morphologie, anatomie, reproduction, développement. Memoires de la Societe Royale des Sciences, de I 'Agriculture et des Arts de Lille 1 8: 1-477.

Marko, P.B. (2002). Fossil Calibration of Molecular Clocks and the Divergence Times of Geminate Species Pairs Separated by the Isthmus of Panama. Molecular Biology and Evolution 19 (11):2005-2021.

Martins, R., Magalhaes, L. and Peter, A. (2013). Diversity, distribution and ecology of the family Syllidae (Annelida) in Portuguese coast (Western Iberian Peninsula). Helgoland Marine Research.

Maselli, V. (2011). The Stratigraphic Record of the Quaternary Sea Level Fluctuations and the Impact of the Post-Glacial Sea Level Rise (Termination I) in the Adriatic Basin (Mediterranean Sea). Doktorarbeit, Universität Bologna.

McHugh, D. (1997). Molecular evidence that echiurans and pogonophorans are derived annelids. PNAS USA 94: 8006-8009.

Meirmans, P.G. and Hedrick, P.W. (2011). Assessing population structure: F_{ST} and related measures. Molecular Ecology Resources 11: 5-18.

Meyer C.P. and Paulay G. (2005). DNA Barcoding: Error rates based on comprehensive sampling. *PLoS Biology*, 3(12): 1-10.

Nicholls, A.G. (1935). Copepods from the interstitial fauna of a sandy beach. Journal of Marine Biology. Association of the United Kingdom 20: 379-405.

Nicholas, K.B., Nicholas H.B. Jr., and Deerfield, D.W. II. (1997). GeneDoc: Analysis and Visualization of Genetic Variation, EMBnet.news 4: 14.

Nygren, A. (1999). Phylogeny and reproduction in Syllidae (Polychaeta). Zoological Journal of the Linnean Society 126: 365-386.

Nygren, A., (2004). Revision of Autolytinae (Syllidae: Polychaeta). Zootaxa 680: 1–314.

Nygren, A. and Sundberg, P. (2003). Phylogeny and evolution of reproductive modes in Autolytinae (Syllidae, Annelida). Molecular Phylogenetics and Evolution 29: 235-249.

Nygren, A., Eklöf, J. and Pleijel, F. (2009). Arctic-boreal sibling species of *Paranaitis* (Polychaeta, Phyllodocidae). Marine Biology Research 5: 315-327.

Orrhage, L. (1980). On the Structure and Homologues of the Anterior End of the Polychaete Families Sabellidae and Serpulidae. Zoomorphology 96: 113-168.

Rouse, G.W. and Pleijel F. (2001). Polychaetes. Oxford University Press, New York: pp 354.

Pleijel, F. and Dahlgren, T. (1998). Position and Delineation of Chrysopetalidae and Hesionidae (Annelida, Polychaeta, Phyllodocida). Cladistics 14: 129-150.

Posada, D. and Crandall, K.A. (1998). MODELTEST: testing the model of DNA substitution. Bioinformatics 14 (9): 817-8.

Raes, J. and Van de Peer, Y. (1999). ForCon : a software tool for the conversion of sequence alignments. EMBnet.news 6(1). Distributed for free at: http://vulcan.rug.ac.be/~jerae/ForCon/index.html.

Remane, A. (1933). Verteilung und Organisation der benthonischen Mikrofauna der Kieler Bucht. Wissenschaftliche Meeresuntersuchungen Abteilung Kiel, NF 21: 161-221.

Ritschl, C. (2013). Genetic variability of geographically widely separated mediterranean populations of *Sphaerosyllis sp*. pp 51, Diplomarbeit, Universität Wien.

Ronquist, F. and Huelsenbeck, J.P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Rouse, G.W. and Fauchald, K. (1997). Cladistics and polychaetes. Zoologica Scripta 26 (2): 139-204.

San Martín, G. (2003). Familia Syllidae. In: M. A. Ramos (Ed) Annelida, Polychaeta II. Fauna Ibérica, Vol. 21 (pp 554). Madrid: Museo Nacional de Ciencias Naturales, CSIC.

San Martin, G., Hutchings, P. and Aguado, T.M. (2008). Syllinae (Polychaeta: Syllidae) from Australia. Part 1. Genera *Brachiosyllis*, *Eurysyllis*, *Karroonsyllis*, *Parasphaerosyllis*, *Plakosyllis*, *Thopalosyllis*, *Tetrapalpia* n.gen, and *Xenosyllis*. Record Australian Museum 60: 119-160.

Swedmark, B. (1964). The interstitial fauna of marine Sand. Biological Reviews 39: 1–42.

Swofford, D.L. (2003). PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Trpisovsky, D. (2013). Genetic variability in the Mediterranean mesopsammal mollusc *Philinoglossa praelongata* (Gastropoda, Euopisthobranchia). pp 79, Diplomarbeit, Universität Wien.

Werth, S.M. (2007). Genetic variability in the Mediterranean interstitial gastropod *Pontohedyle milaschewitschii* (Kowalevsky, 1901) (Opisthobranchia, Acochlidia). pp 59, Diplomarbeit, Universität Wien.

Westheide, W., 1997. The direction of evolution within the Polychaeta. Journal of natural History 31: 1–15.

Westheide, W. and Hass-Cordes, E. (2001). Molecular taxonomy: description of a cryptic Petitia species (Polychaeta: Syllidae) from the island of MaheÂ (Seychelles, Indian Ocean) using RAPD markers and ITS2 sequences. Journal of Zoological Systematics and Evolutionary Research 39: 103-111.

Westheide, W. and Schmidt, H. (2003). Cosmopolitan versus cryptic meiofaunal polychaete species: an approach to a molecular taxonomy. Helgoland Marine Research 57: 1-6.

Wilke, T., Schultheiß, R. and Albrecht, C. (2009). As Time Goes by: A Simple Fool's Guide to Molecular Clock Approaches in Invertebrates. American Malacological Bulletin 27(1/2): 25-45.

Wright S. (1943). Isolation by distance. Genetics 28: 114-138.

Wright, S. (1978). Variability within and among natural populations. In: Evolution and genetics of populations. Vol. 4 (pp 490), University of Chicago Press.

Xia, X. (2013). DAMBE 5: A comprehensive software package for data Analysis in molecular biology and evolution. Molecular Biology and Evolution 30 (7): 1720-1728.

Deutsche Zusammenfassung

Plakosyllis brevipes gehört zu den Syllidae, einer der vielfältigsten Familien innerhalb der Polychaeta. Sie umfasst 70 Gattungen und 700 Arten. Syllidae sind weit verbreitet und zeigen entweder eine benthische oder interstitielle Lebensweise. *P. brevipes* ist ein Vertreter der sogenannten "Meiofauna", einer Gruppe von Organismen deren Lebensraum sich in dem Lückensystem zwischen den Sandkörner befindet. Dieses anspruchsvolle Habitat erfordert zahlreiche Anpassungen, wie zum Beispiel Miniaturisierung, Flexibilität, Adhäsionsorgane, einen wurmförmigen Körper und spezielle Fortpflanzungsstrategien. *P. brevipes* pflanzt sich durch Stolonisierung fort. Aufgrund der geringen Körpergröße ist die Anzahl der Gameten begrenzt und die verkürzte pelagische Phase der Fortpflanzungsstadien limitiert das Ausbreitungspotential der Art. Trotz des geringen Verbreitungspotentials ist diese Art weltweit zu finden, wie zum Beispiel im Mittelmeer oder im Atlantischen Ozean.

Das Ziel dieser molekularen Studie ist es zu testen ob zwei geographisch weit getrennten Populationen von *P. brevipes* konspezifisch sind oder es sich dabei um kryptische Arten handelt, die eine hohe Differenzierung aufweisen. Die Versuchstiere wurden im Mittelmeer entlang der Küste von Kroatien und Italien gesammelt. Die Sammelpunkte sind durch eine Küstenlinie von 2500 km getrennt. Mit Hilfe des mitochondriellen CO1 und des nuklearen ITS Markers wurden Sequenzen erstellt. Maximum Likelihood, Bayesian Inference, Maximum Parsimony, AMOVA und Koaleszenz Analysis wurden durchgeführt. Unkorrigierte p-Distanzen wurden verwendet um die genetische Distanz zwischen den Populationen und anderen Sylliden festzustellen.

Alle Analysen führten dazu, dass es sich sehr wahrscheinlich um konspezifische Populationen handelt. Diese Annahme wird durch die geringe genetische Distanz zwischen den Populationen gestützt. Genealogische Analysen ergaben zwei deutlich getrennte Kladen, welche jeweils eine dominante Haplolinie und einen Vertreter der anderen Population beinhalten. Als mögliche Gründe kommen das unterschiedliche Vererbungsmuster der Gene, Genfluss während des letzten glazialen Maximums und passiver Drift in Frage. Jedoch zeigte sich auch eine hohe Variation zwischen den geographisch weit getrennten Populationen und

eine niedrige genetische Differenzierung innerhalb der Population. Die Koaleszenz Analyse berechnete einen möglichen Separierungspunkt der Haplolinien vor 2,9 Millionen Jahren nach der messinischen Salinitätskrise.

Acknowledgments

I want to thank my family for encouraging and motivating me during my work.

Thanks to Dr. Gerhard Steiner for the support, suggestions and patience during this work.

A special thank to Dr. Leyli Samadi Kaleh Masihi for introducing me to the lab work, Mag. Dr. Johannes Rath for sequencing my samples, Dr. Eva Ringler and Mag. Christian Baranyi for helping ideas and suggestions, and Mag. Gerald Timelthaler for solving my biggest problem. I also want to thank Mag. Josef Harl for helping me with Beast.

Thanks to my working group Daniela Trpisovsky and Christoph Ritschl, without you I probably would have starved in Italy and the last years would have been half as funny. Our discussions always helped.

I appreciate the help of Mag. Christine Mayer and Verena Reischl for lingual corrections and mental support.

A special thank to Florian, who always backed me up and was there for me.

Thanks to my friends, who stood by my side and listened to all my problems.

Curriculum vitae

Christine Mayer

Personal data	Place of birth	Vienna - Austria
	Marital status	unmarried
	Citizenship	Austria
Education	Since 2011	University of Vienna, Studies of Zoology
		Specialisation: Phylogeny and reproduction
	2004 - 2011	University of Vienna, Studies of biology
	2003 - 2004	University of Vienna, Studies of medicine
	1995 - 2003	Bundesgymnasium/ Realgymnasium Zirkusgasse, 1020 Vienna
	1991 - 1995	Volksschule Oberhausen/ Probstdorf, Lower Austria
Work experience	2012 - 2014	University of Vienna - Department of Int. Zoology Tutorial: Bauplans of animals 1
	2007 - 2008	University of Vienna - Department of Anthropology Tutorial: Dissection course for anthropologist
Computer literacy	Microsoft Windows, Microsoft Office (Word, Excel, Power Point, Outlook), SPSS, phylogenetic and phylogeographic software	
Languages	German, English	