universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Species delineation and genetic variation of *Hiatella* "arctica" (Bivalvia, Heterodonta) in the Mediterranean Sea

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer.nat.)

Verfasserin	Miriam Oberlechner
Matrikel-Nummer:	9703291
Studienrichtung (lt. Studienblatt):	Biologie
Betreuer:	a.o. Univ. Prof. Dr. Gerhard Steiner
Wien, am	13. August 2008

Ich komme mir wie ein Mädchen vor, das am
Strand spielt und sich damit vergnügt, ein noch
glatteres Kieselsteinchen oder eine noch schönere Muschel
als gewöhnlich zu finden, während das große Meer
der Wahrheit völlig unerforscht vor mir liegt.

frei nach Sir Isaac Newton

Contents

I Abstract

Hiatella is a common bivalve genus with a widespread marine distribution except for tropic waters. On a vertical scale it occurs from the shore down to a depth of 200 m with the main distribution in shallow waters of the rocky subtidal. This organism lives either in rock crevices as bioeroder, nestling in kelp holdfasts, or within sponges. Therefore, the shell morphology is often influenced by the shape of the crevices and the valves are eroded during boring. It is not clear, however, how many species are represented in the Mediterranean Sea due to the high plasticity in shell form and the resulting problems in species delineation. In the North Atlantic including the Mediterranean Sea, two species of Hiatella are described: *Hiatella arctica* (L., 1767) und *Hiatella striata* (Fleuriau De Bellevue, 1802). The diagnostic features – such as shell morphology, reproduction and development, habitat choice and fossil record – are controversial and it is important to figure out whether these characteristics are due to ecophenotypic variation of a single species or represent more species. The present study focuses on a molecular approach to distinguish between morphological and ecological similar species. Specimens were collected from different habitats and geographically distinct parts of the Mediterranean Sea and the Atlantic Ocean. The collecting sites comprised the North Adriatic coast (Croatia), the Ligurian coast (North-West Italy), and the South coasts of Spain and Portugal. Two fragments of the mitochondrial genome were used as genetic markers. They included parts of ribosomal (16S rRNA) and protein coding genes (COI, COIII and ATP6) as well as two non coding regions. The resulting sequence data were used for genealogical analyses (Bayesian Inference and Network analysis) as well as for analyses of molecular variance (AMOVA).

All analyses show conspecifity of bioeroders, nestlers and sponge-dwelling specimens with a maximum genetic divergence of 2.7%. Furthermore, the genetic differentiation among populations is strongly correlated with the geographic distances. The conclusion of the present study therefore is that in the Mediterranean Sea only one species of *Hiatella* occurs. This species, referred to as *Hiatella* “arctica”, has the ability to colonise geographically distinct parts as well as to live in different ecological environments. This capacity probably results from a sum of different factors: the high dispersal abilities, the different modes of habit, the continuous habitat, the sea currents and the anthropogenic influence. There don’t seem to be any physic barriers, with the exception of geographic

distance, that limit gene flow between populations of *Hiatella* “arctica” in the Mediterranean Sea.

II Deutsche Zusammenfassung

Hiatella ist eine marine Gattung innerhalb der Klasse Bivalvia, die mit Ausnahme von tropischen Gewässern weltweit vorkommt. Diese Organismen haben ihre größte Verbreitung im felsigen Sublitoral, wobei sie bis zu Tiefen von 200 Metern auftreten können. Die Lebensweise von *Hiatella* ist vielfältig: sie lebt in Felsspalten bzw. in Bohrlöchern primärer Bioerodierer (bohrende Form), epiphytisch oder epizoisch (nestelnde Form) oder auch im Gewebe von Schwämmen (Schwammbewohner). Die Schalenform ist maßgeblich der ökologischen Nische angepasst und durch die mechanische Abnutzung der Schale aufgrund von Bohrtätigkeiten hoch variabel. Es ist nach dem heutigen Wissensstand unklar, wie viele Arten von *Hiatella* im Mittelmeer vorkommen. Im Nord-Atlantik inklusive Mittelmeer sind zwei Arten beschrieben: *Hiatella arctica* (L., 1767) und *Hiatella striata* (Fleuriau De Bellevue, 1802). Die Bewertungen „artspezifischer" Merkmale – wie Schalenmorphologie, ökologische Nische, Fortpflanzung und Entwicklung – sind umstritten und es gilt herauszufinden, ob es sich um eine ökophänotypisch hoch variable Art, oder aber um zwei bzw. mehrere Arten handelt. Zu diesem Zweck wurde in dieser Diplomarbeit ein molekularer Ansatz gewählt. Es wurden vier Populationen im Mittelmeer bzw. Atlantik (Nordadria - Kroatien, Ligurisches Becken - Italien, Alboran Becken - Spanien, Atlantischer Ozean - Portugal) anhand von molekularen Markern untersucht. Ein Genbereich von 2745 Basenpaaren (bp), bestehend aus Teilen der COI, COIII, ATP6, 16SrRNA sowie zwei „non coding" Regionen, wurde bei 97 Individuen sequenziert. Mit diesem Datensatz wurden eine Bayes'sche Analyse, eine Netzwerkanalyse sowie eine molekulare Varianzanalyse (AMOVA) durchgeführt.

Die Analysen ergaben, dass die maximale genetische Divergenz 2,7% beträgt und somit die Lebensweise - bohrend, nestelnd, Schwamm bewohnend - kein artspezifisches Merkmal ist. Weiters gibt es eine starke Korrelation zwischen genetischer Differenzierung und der geographischen Distanz. Diese molekularen Daten unterstützen die Hypothese, dass es sich im Mittelmeer um eine einzige, phänotypisch variable Art handelt, die in dieser Studie als *Hiatella* „arctica" bezeichnet wird. Die Oberflächenströmungen im Mittelmeer scheinen das Verbreitungspotential von *Hiatella* „arctica" demzufolge nicht negativ zu beeinflussen. Mögliche Erklärungen für die beträchtliche intraspezifische Variation liegen zum einen im hohen Verbreitungspotential der Larven, sowie der juvenilen und adulten Tieren. Andererseits kann durch die Fähigkeit, unterschiedliche Lebensräume zu besiedeln das Habitat als durchgehend betrachtet werden. Somit wird

zusätzlicher Genfluss durch „gene hopping“ ermöglicht. Weiters kann der anthropogene Einfluss als zusätzliche Verbreitungsquelle nicht ausgeschlossen werden. Folglich scheint es im Mittelmeer mit Ausnahme der geographischen Distanz kaum physischen Barrieren zu geben, welche den Genfluss zwischen Populationen von *Hiatella* „arctica“ einschränken.

1. Introduction

1.1. Morphological and systematic assignment

1.1.1. Bivalvia

The bivalves are the second-largest class of molluscs (Giribet and Distel 2003) and they are represented by tens of thousands species (Cope 1996). They are limited to aquatic systems, occurring in both salt and fresh water (Kilias 1982).

Several morphological features characterise the mollusc class of bivalves. The body is laterally compressed. The mouth is located at the anterior and the anus at the posterior end of the body (Giribet and Wheeler 2002). All bivalves lack a radula, a buccal apparatus and salivary glands. The head is reduced. In the adult state, they have two valves, which are connected on the dorsal margin by an elastic ligament. The valves are held together by one or two adductor muscles (Morton 1979). The mantle is laterally elongated enclosing the whole body. The mantle edge consists of three mantle folds: the inner mantle fold is muscular and enables to close the mantle cavity; the middle mantle fold has sensory function and is endowed with sensory cells and sometimes with tentacles or eyes. The inner surface of the outer mantle fold produces the periostracum, whereas the hypostracum is secreted by the entire mantle surface (Kilias 1982). A pair of lateral gills is connected to the ciliated labial palpes. The gills are responsible for respiration and in autobranch bivalves also for providing and transporting food to the mouth (Giribet and Wheeler 2002). The digestive system, especially the gut with the ciliary sorting areas and the style sac, is highly specialised (Purchon 1990). The extensible foot is usually compressed and muscular, enabling to burrow into soft sand or mud (Morton 1979).

1.1.2. Classification of the Bivalvia

The Bivalvia are divided into two main groups: the Protobranchia and the Autobranchia. The first fossil record of the Bivalvia can be traced back to the early Cambrian (Morton 1996, Cope 1996) and belong to the Protobranchia (Cope 1996). Therefore, the Protobranchia, a group of deposit-feeders, are considered the most ancient clade within the Bivalvia (Adamkewicz et al. 1997) and they do not form a monophyletic group (Giribet and Wheeler 2002, Giribet and Distel 2003). The Protobranchia are divided in two subclasses, the Palaeotaxodonta and the Cryptodonta (Newell 1969), which share primarily

plesiomorphic characters (Purchon 1987, Cragg 1996, Healy 1996, Salvini-Plawen and Steiner 1996).

The second clade comprises the monophyletic Autobranchia (Salvini-Plawen and Steiner 1996, Giribet and Wheeler 2002). The Autobranchia includes the following subclasses: Pteriomorphia, Palaeoheterodonta and Heterodonta (Giribet and Wheeler 2002). The Anomalodesmata were believed for a long time to form an additional subclass within the Autobranchia, but is now commonly accepted that this monophyletic group clusters within the Heterodonta *s.s.* (Salvini-Plawen and Steiner 1996, Steiner and Hammer 2000, Giribet and Wheeler 2002, Dreyer et al. 2003, Giribet and Distel 2003, Harper et al. 2006, Taylor et al. 2007).

Heterodonta

The subclass Heterodonta is a highly diverse group, which had the major radiation in the late Mesozoic. The Heterodonta are characterised by the presence of siphons, eulamellibranch gills and hinges with cardinal and lateral teeth (Taylor et al. 2007) and they show a high variation in sperm types (Healy 1996).

Hiatelloidea

The superfamily Hiatelloidea belongs to the subclass Heterodonta and comprises a sole family, the Hiatellidae (Coan et al. 2000).

Several molecular analyses suggest that the Hiatelloidea are a basal lineage of the Heterodonta (Giribet and Wheeler 2002, Dreyer et al. 2003, Giribet and Distel 2003, Harper et al. 2006, Taylor et al. 2007). Therefore, Taylor et al. (2007) suggest an origin of the Hiatelloidea in the late Palaeozoic. It is not clear yet, which family represents the sister group to the Hiatelloidea. In some of these investigations, the Hiatelloidea form a monophyletic clade with the Solenoidea (Dreyer et al. 2003, Giribet and Distel 2003, Taylor et al. 2007). In the combined morphological and molecular analysis performed by Giribet and Wheeler (2002) however, there is no evidence for a sister group relationship between these two taxa.

Hiatellidae

The family Hiatellidae consists of five genera: *Hiatella*, *Saxicavella*, *Panomya*, *Panopea* and *Cyrtodaria*, whereas the genus *Hiatella* is the most common. Based on morphological

characteristics (mantle, ligament and dentition) and habitat, *Saxicavella* is supposed to form the basal group within the family of the Hiatellidae. The genera *Panomya* and *Panopea* are closely related, but the relationship of *Hiatella*, *Cyrtodaria* and the *Panomya*/ *Panopea* clade is not clear (Yonge 1971). However, the fossil record suggests that the genus *Cyrtodaria* derived from *Hiatella* (Gordillo 2001).

1.2. The genus *Hiatella* Bosc, 1801, *ex* Daudin MS

1.2.1. Morphological characteristics

The shell is solid, white, usually wrinkled and almost rectangular (Narchi 1973, Gordillo 2001), although the shape of the shells can be very irregular (among others: Russell-Hunter 1949, Yonge 1951, Micali and Solustri 2004). The size of the shells varies between 6 and 45 mm, whereas larger specimens are generally found in colder waters (Strauch 1968). The valves are often unequal, usually the right one being slightly larger than the left one. However, *Hiatella* with larger left valves are also found (Khalaman 2005). The shell usually gapes, especially at the posterior end (Gordillo 2001). The umbo is set on the anterior part of the shell and two spinous ribs can lead from the umbo to the posterior margin (Forbes and Hanley 1853, Odhner 1914, Russell-Hunter 1949). The elastic ligament is external and opisthodelic (Yonge 1951). The heterodont dentition is reduced with maximally one cardinal tooth in each valve (Forbes and Hanley 1853, Yonge 1971, Gordillo 2001). The adductor muscles are of the isomyarian type (Yonge 1971). The mantle cavity is elongated on the posterior margin and the pallial sinus is distinct (Yonge 1971).The muscular mantle is completely fused (Type C, *sensu* Yonge 1971), except for the opening for the narrow foot and the opening for siphons. The siphons are red (Yonge 1951) and covered with periostracum (Riedl 1983) and variable in their length, diverging only at their ends (Russell-Hunter 1949). The stomach is of type IV (*sensu* Purchon 1958). The eulamellibranch gills are of type C (*sensu* Atkins 1937). The byssus gland is always present, but byssus threads are not always secreted (Russell-Hunter 1949, Yonge 1951).

1.2.2. Geographic range

Hiatella has a widespread marine distribution and is ubiquitous except for tropic waters. The genus is best represented in cool-temperate to polar waters of the Northern and Southern Hemisphere. At present, *Hiatella* is found in the North Atlantic from Greenland

to Mexico and from Europe to Africa (Gabon) as well as in the Mediterranean Sea. In the North Pacific *Hiatella* is found from Alaska to Panama as well as in the Sea of Japan. Furthermore, the genus is ubiquitous in the Arctic Sea. In the Southern Hemisphere, the genus *Hiatella* is widely distributed around South America, Antarctica, Australia and New Zealand (Gordillo 2001).

On the vertical scale, *Hiatella* is widely distributed as well. The genus occurs from the shore down to a depth of 200 m with the main distribution in shallow waters of the rocky intertidal (Forbes and Hanley 1853, Strauch 1968, Yonge 1971).

1.2.3. Habitat

The genus colonises a variety of substrates. It occurs as nestler, borer and sponge dweller. Nestling organisms attach by mean of byssus threads to other organisms, like other bivalves, commonly *Mytilus edulis* (Russell-Hunter 1949), crustaceans (Isaeva et al. 2001), macro-algae (Trudgill 1987) etc. The adults are able to reproduce byssus threads when retrieved from their substrates (Russell-Hunter 1949).

The boring organisms often lack byssus threads (Yonge 1971). They settle in crevices or in boreholes of primary bioeroders, like *Pholas* sp. (Beu 1971) which primarily occurs in soft substrates like sandstone (Trudgill 1987), or *Gastrochaena dubia* which occurs in limestone (Trudgill 1987). *Hiatella* is able to enlarge the boreholes mechanically by means of the valves (Yonge 1971). The mechanical boring activity leads to abrasion of the valves and therefore to high variation in shell forms (Yonge 1971). However, also chemical boring by acid secretion is not excluded (Trudgill 1987), although no acid secreting glands are reported by Russell-Hunter (1949). Yonge (1949, p.176) concluded that “the precise mode of boring is not too obvious” in *Hiatella*.

The sponge dwelling organisms occur mainly as endobiont in the exhalant canals of *Spongia*, *Geodia*, *Sarcotragus*, *Ircinia* (Micali and Solustri 2004), *Cacospongia* (Riedl 1983) etc. Juvenile *Hiatella* are mainly found in the external parts of the sponges and older specimens in deeper parts respectively. Therefore, it is likely that the sponges overgrow the bivalves with time (Micali and Solustri 2004). Within sponge dwellers both specimens with and without byssus threads are found (Micali and Solustri 2004).

Generally, the shape of the shells is influenced by the habitat in which the specimens occur (Russell-Hunter 1949).

1.2.4. Fossil record

The first fossil record of *Hiatella* is reported from the late Jurassic and suggests the presence of the genus for at least 150 million years (Ma). This earliest known record, known as *Hiatella* (*Pseudosaxicava*) *foetica*, are found in temperate waters of the Northern Hemisphere (present-day France and England) in two habitats; as nestlers attached on hard substrate by means of byssus threads (Figure 1,c) and as borers within *Gastrochaenolites* boreholes (Figure 1,d) (Kelly 1980). The shape of the shells of the Jurassic *Hiatella* is quite similar to that of recent specimens (nestler Figure 1,a; borer Figure 1,b). Therefore, both the shell morphology and the ecology of the genus *Hiatella* have not changed much in the last 150 Ma.

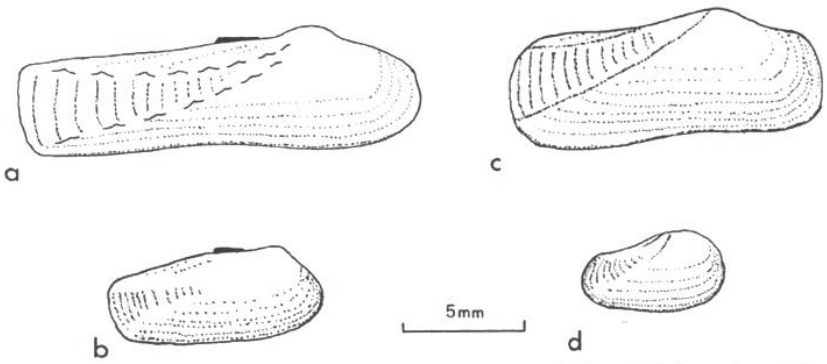

Figure 1: *Hiatella* shells from fossil records (Jurassic) (c, d) and recent specimens (a, b). The upper shells represent nestling, the lower shells boring organisms respectively (Kelly 1980, p. 773, Figure 4).

1.2.5. Nutrition

Hiatella is a filter feeder by means of the siphons ("siphon feeder", Yonge 1971) and its diet consists of planktonic organisms that occur in the water column. The filtration rate rises with increasing water temperature and reaches a maximum between 15° and 17°C. Higher temperatures result in a quick decrease of the feeding rate, a fact that agrees with the distribution in cool-temperature waters (Ali 1970).

1.2.6. Development

Hiatella is, like most bivalves, dioeciously with fertilisation of the eggs in the water column. The size of the eggs varies between 0.05 and 0.08 mm in temperate and in arctic waters respectively (Strauch 1968). The larvae are abundant and planktotrophic (Kilias 1982). They undergo a long pelagic stage, which can last up to six weeks (Lucas 1990).

The late larvae represent one of the largest veligers in the plankton (0.36 mm in *Hiatella arctica* and 0.32 mm in *Hiatella striata* respectively, Lebour 1938). In this stage the velum reaches about the same length of the prodissoconch in *Hiatella arctica* and two-third the length in *Hiatella striata* respectively (Lebour 1938). Before metamorphosis, they develop the foot and undergo a transitional phase between planktonic and bentic life, known as pediveliger stage (Carriker 1990). In this stage, the larvae have a statocyst with one statolith (Odhner 1914). The provinculum is equipped with one tooth in each valve (Booth 1983). The juveniles can migrate up to six months in order to find the appropriate place to settle (Russell-Hunter 1949). It is unknown at which age the specimens gain maturity. In the Clyde Sea they breed all over the year, independently from the habitat or adult shell form, but given that the water temperature does not exceed 12°C (Russell-Hunter 1949). *Hiatella arctica* shows a maximum in planktonic larvae from autumn to spring, and *Hiatella striata* from summer to early autumn respectively (Lebour 1938) in some geographic areas (Plymouth). However, the larvae of both species occur at small numbers throughout the year (Russell-Hunter 1949). It is known that the adults of *Hiatella arctica* can get very old, reaching an age of over 100 years (Sejr et al. 2002).

1.2.7. Taxonomy

The genus *Hiatella* was erected by Bosc (1801, *ex* Daudin MS) for two species, *H. biaperta* (Daudin, 1800) and *H. monoperta* (Daudin, 1800). Both species were synonymised with *Hiatella arctica* (L., 1767) by Lamarck (1818). Linné (1767) described two species of *Hiatella* in the Atlantic Ocean (Norway): *Mya arctica* (=*H. arctica*) and *Mytilus rugosus* (=*H. rugosa*). Winckworth (1932) assigned the type specimen *H. rugosa* (L., 1767) to *H. arctica* (L., 1767). Therefore, *H. rugosa* (L., 1767) is a junior synonym for *H. arctica*, although in literature, *H. rugosa* (L.) is often used to describe another species, *Hiatella striata* (Fleuriau De Bellevue, 1802). However, the differentiation of *H. arctica* and *H. striata* is dubious, and *H. striata* is discussed as a synonym for *H. arctica* (Dodge 1950, Coan et al. 2000). In the past, the high plasticity in shell forms led to a series of names for the same species (see Forbes and Hanley 1853, p.141; p.146), whereas in more recent literature several authors tend to recognise only one, highly variable species (Dodge 1950, Jeffreys 1865, Strauch 1968).

Since in the Atlantic and Mediterranean Sea many authors distinguish between *Hiatella arctica* (L., 1767) and *Hiatella striata* (Fleuriau De Bellevue, 1802), in the following sections of this study the diagnostic characters are summarised.

1.3. Differentiation of *Hiatella arctica* and *Hiatella striata*

As discussed above, many investigations regarding the number of species of *Hiatella*, found in the Mediterranean and the Atlantic Sea, were conducted. Species diagnosis mainly included shell characteristics, habitat choice, larval development and reproductive cycle, and fossil record. However, the attempts to define species-specific characters presented a number of difficulties.

1.3.1. Species diagnosis: Shell morphology

For a long time, *Hiatella arctica* was thought to be the nestling form with two spinous ribs that lead from the umbo to the posterior margin (Forbes and Hanley 1853), one cardinal tooth in each valve (Forbes and Hanley 1853) and byssus threads to attach to other organisms. However, Forbes and Hanley (1853) also mentioned that the spines can be lost in adults. Other morphological features assigned to *Hiatella arctica* include: the inequivalve shell; the two valves gape only on the ventral side, where the byssus threads protrude (Bucquoy et al. 1896); the posterior part of the shell being broader than the anterior, pointed, part ; a “lunule-like excavation in front of the beak” (Forbes and Hanley 1853, p.141). However, all these characters are highly variable (Russell-Hunter 1949, Micali and Solustri 2004, Khalaman 2005).

Hiatella striata was described as boring organism without byssus threads (Bucquoy et al. 1896), hinge teeth and spinous ridges (Forbes and Hanley 1853). However, in juvenile *Hiatella striata* both hinge teeth and spinous ridges can occur (Forbes and Hanley 1853). Furthermore, Russell-Hunter (1949) reported that the byssus gland is also present in *Hiatella striata*, even if byssus threads are not always secreted.

1.3.2. Species diagnosis: Habitat choice

The common opinion that *Hiatella arctica* only nestles and that *Hiatella striata* only bores (Forbes and Hanley 1853, Odhner 1914, Lebour 1938) seems to be incorrect (Barrett and Yonge 1958, Russell-Hunter 1949, Strauch 1968). The mode of life is rather determined by the substrate, where the settlement of the larvae takes place (Russell-Hunter 1949). Furthermore, both species are also found within sponges (Micali and Solustri 2004).

1.3.3. Species diagnosis: Reproductive cycle and larval development

Odhner (1914) first described the veliger larvae of the genus *Hiatella* in the Adriatic Sea. He distinguished between the two species using shell characters of the prodissoconch. However, he did not report any seasonality of these two larva-types and he admitted that intermediate forms of the shell characters may exist, too.

In British waters, Lebour (1938) distinguished between the two species, beside size and prodissoconch features, by means of their reproduction cycle occurring in different times of the year: larvae of *Hiatella arctica* are common from July to November, larvae of *Hiatella striata* from winter to early spring. However, Russell-Hunter (1949) reported that both larval types occur in small numbers throughout the year.

1.3.4. Species diagnosis: Fossil record

Based on fossil record, Gordillo (2001) suggests the existence of the two species in the North Atlantic as they originated in different parts of the Northern Hemisphere: *Hiatella striata* in the North Atlantic and *Hiatella arctica* in the North Pacific with the last common ancestor (*Hiatella vera*) in the Eocene (at least 34 Ma ago). It is hypothesised that *Hiatella arctica* invaded the North Atlantic during the Pliocene (5-2 Ma) with the opening of the Bering Strait (Figure 2). However, palaeontologists assign species primarily on the base of shell characters. Since the shape of the shell in *Hiatella* is highly variable and is strongly influenced by the habitat, it is very difficult to classify fossil record with a species.

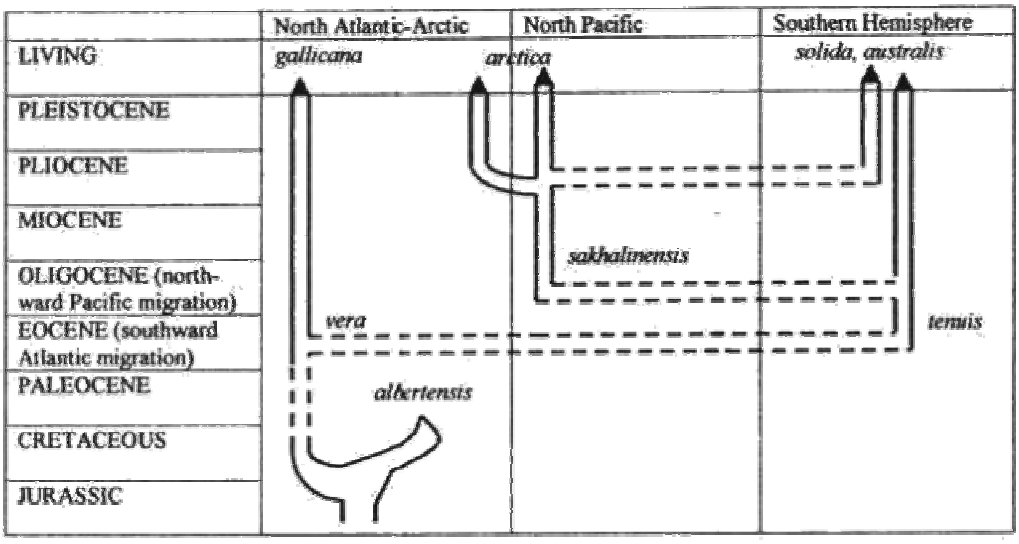

Figure 2: Hypothetical migration of the genus *Hiatella* based on fossil record (Gordillo 2001, p.244, Figure 5).

It is not clear, however, how many species of *Hiatella* are represented in the Atlantic Ocean and the Mediterranean Sea, and whether the characteristics mentioned above – shell morphology, ecology, development and fossil record – are due to ecophenotypic variation of a single species or represent more species.

1.4. Aims of this study

1.4.1. Species delineation

Since the differentiation of the two species *Hiatella arctica* and *Hiatella striata* is dubious and no species-specific characters are unambiguous, the present study focuses on a molecular approach to distinguish between these morphological and ecological similar species.

Assuming that *Hiatella arctica* and *Hiatella striata* are different species, molecular markers can be very useful to distinguish between intra-specific variation and inter-specific divergence. It is commonly accepted that individuals that show a genetic divergence of 5% can be attributed to different species. However, the existence of different species depends more on the gap between intra-specific variation and inter-specific divergence than the 5% threshold, which is valid only for genes, such as the cytochrome oxidase I (COI) or genes with similar substitution rates (Meyer and Paulay 2005). Low intra-specific variation and high inter-specific differences are indicators for species delineation. By contrast, high intra-specific variation and low inter-specific differences support the existence of only one, highly variable species.

1.4.2. Genetic variation

The genetic variation of *Hiatella* is investigated for different habitats and for geographically distinct parts of the Mediterranean Sea and the Atlantic Ocean. This study focuses on how genetically variable the populations are and whether the respective populations are in a panmictic state with almost unrestricted gene flow or rather genetically distinct. Furthermore, the genetic differences between populations are investigated for correlation with the geographic distances. Finally, the specimens are analysed regarding the habitat in which they were collected (nestling, boring, sponge dwelling) in order to examine the genetic variation within and among habitats. These results are used to determine whether the specimens from different habitats are genetically distinct or not.

1.5. Molecular markers

Mitochondrial genes are widely used molecular markers to investigate the relationships at inter- and intra-specific level. However, there is no clear rule concerning which genes are most suitable for analysis, because the substitution rates vary between different genes and across taxa (Hebert et al. 2003). Generally, the transfer RNA genes and the ribosomal RNA genes evolve slowly (Saccone et al. 1999) and therefore, these genes are good molecular markers at higher taxonomic level. In the phylogeny of bivalves, however, the mitochondrial ribosomal gene 16S rRNA is appropriate to investigate phylogenetic relationships at the genus level (Saavedra and Pena 2006, Stepien et al. 2001) as well as at the intra-specific level (Turner et al. 2000).

Protein coding genes, especially the COI, are good molecular markers at the species and population level and are widely used in population genetics (Arnaud et al. 2000; Terranova et al. 2006; Hare and Weinberg 2005). However, only the third codon positions are weakly constrained by selection because of their two-fold or four-fold degeneracy. By contrast, the first and especially the second codon positions are highly conserved as nucleotide changes in these positions often result in a change of the amino acid (Saccone et al. 1999).

Non coding regions in mitochondrial genomes are useful genetic markers at the intra-specific level. They are highly variable because they are not constrained by selection. Therefore, they show higher genetic variation than coding regions, like the COI and the 16S rRNA genes (Aranishi and Okimoto 2005).

2. Material and Methods

2.1. Data collection

2.1.1. Sample collection and conservation

Ninety-six specimens were collected from different habitats and geographically distinct parts of the Mediterranean Sea (North Adriatic Sea, Ligurian Sea, Alboran Sea) and the Atlantic Ocean. Specimens of the Croatian population were sampled in the North Adriatic Sea at: Senj, Pag and Rovinj. *Hiatella* from the Italian population was collected in the Ligurian Sea at Capo Noli and Pontetto Bogliasco. Representatives from the Spanish population were sampled in the Alboran Sea at the sites Calahonda, Malaga and Calaburra and the samples from the Portuguese population were collected from the Atlantic Ocean in the southern part of the country at Ingrina and Rogil (Figure 3).

Figure 3: Map of the collecting sites (earth.google.com).

All *Hiatella* were sampled in the period from April to July 2006 by snorkelling at depths down to 6 meters. Kelp holdfasts and *Mytilus* sp. aggregations were investigated for "nestling" *Hiatella* attached by byssus threads; calcareous boulders were examined for "boring" specimens living in boreholes, and sponges (*Geodia* sp.) were checked for specimens living in the tissue.

The specimens were grouped and coded with regard to the microhabitat they were found in (Table 1).

Table 1: Sampling sites, codes, habitats and numbers of specimens collected.

collecting site	code	habitat			
		"nestling"		"sponge dwelling"	"boring"
		algae	mussels	sponges	boreholes
Croatia - Rovinj	KR_I	-	-	-	10
Croatia - Pag	KR_II	-	-	-	30
Croatia - Rovinj	KR_III	-	-	11	-
Croatia - Senj	KR_IV	-	-	-	19
Italy - Pontetto Bogliasco	IT_II	3	-	-	-
Italy - Capo Noli	IT_III	4	-	-	-
Spain - Malaga	SP_I	1	-	-	-
Spain - Calaburra	SP_II	-	1	-	-
Spain - Calahonda	SP_III	1	-	-	-
Portugal - Rogil	P_I	4	-	-	-
Portugal - Ingrina	P_II	10	2	-	-

Once collected, the adductor muscles of each animal were cut to prevent specimens to close the valves and all samples were preserved in 96% ethanol. The ethanol was changed twice to speed up dehydration and, thus, inhibit DNA degradation.

2.1.2. DNA isolation

A small part of the adductor muscle was placed in a 1.5 ml tube containing 1 ml of sterile deionised H_2O to remove the ethanol. As soon as the tissue sunk to the bottom of the tube, the entire liquid was removed. Genomic DNA extraction was performed using the E.Z.N.A. Tissue DNA Mini Kit (Peqlab) following the instruction manual for eukaryotic cells and tissue. Final elution of genomic DNA was performed once with 100 µl buffer, instead of two times (as suggested in the protocol), in order to get a working solution with higher DNA concentration.

The concentration and the purity of DNA were measured using the spectrophotometer Bio-Photometer 6131 (Eppendorf).

2.1.3. PCR

Based on the gene order of the mitochondrial genome of *Hiatella arctica* (Dreyer and Steiner 2006; Figure 4), two suitable fragments were selected for amplification. The first fragment included a major part of the COI gene (1310 bp), a non coding region (169 bp)

and a small part of the ATP6 gene (12 bp). The second fragment included half of the 16S rRNA gene (rrnL, 676 bp) a non coding region (312 bp) and 199 bp of the COIII gene.

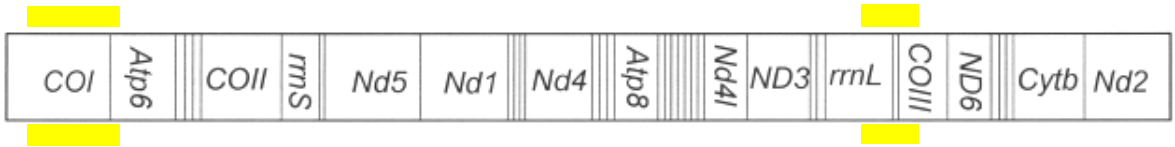

Figure 4: Gene order of the mitochondrial genome of *Hiatella arctica* (Dreyer and Steiner 2006, p. 3 Fig. 1) including the fragments selected for amplification (yellow boxes).

The Polymerase Chain Reaction was performed on a Primus 96 advanced Gradient Thermalcycler (Peqlab) in volumes of 25 µl. PCR solutions contained the ingredients listed in Table 2. The amount of DNA applied for PCR reaction was dependent on the DNA concentration after the extraction.

Table 2: PCR components and their respectively volumes.

PCR component	volume (µl)
$MgCl_2$ (50 mM)	0.8 - 1
dNTP mix (2.5 mM)	2.5
Mango Taq reaction buffer (10x conc.)	2.5
Mango Taq Polymerase (1 unit/µl)	0.75
Primer forward (50 pM)	0.1 – 0.5
Primer reverse (50 pM)	0.1 – 0.5
Deionised H_2O	17.25 – 18.25
DNA	0.2 - 5

All primers used in PCR and sequencing were specifically designed for this study using the OligoAnalyser program (Integrated DNA technologies http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer). All the available sequences of heterodont bivalves were aligned and the most conserved parts were used to design the primer combination for the 16S-COIII fragment (Hi-16for509, Hi-co3rev238) and the COI-ATP6 fragment (Hi-co1for80, hi-atp6rev-46). To increase the efficiency of the amplification reaction of the COI-ATP6 fragment, the first sequences obtained were used to design more specific primers: the CO1forward5_01 primer was used instead of the Hi-co1for80 primer; the Co1-3?neufor primer was used, together with a newly created COII primer (COII369HI rev), to amplify the ATP6 gene. Sequences resulting from the fragment COI-COII were used to design a more specific ATP6 primer (ATP6neu rev).

The primers used to amplify and sequence the genes under study (together with the corresponding protocols) are listed in Table 3 and Table 4.

Table 3: Primers used for PCR and sequencing with the corresponding binding sites and programs. The binding sites are referred to the sequence (gi_1155316) available in the genbank.

primer name	**primer sequence (5' → 3')**	**binding site**	**program**
Hi-16for509 Hi-co3rev238	AACTCGGCAAAACCGCCCTC TCCCTAMACGAAAGCCACGCAC	bp 12951 bp 14198	16S-COIII
Hi-co1for80 hi-atp6rev-46	AAGTWGTAATCATMGAGATATTGG CTGAGCAGGACTATGGTCC	bp 80 bp 1864	COI-ATP6 (old)
Co1-3?neufor COII369HI rev	GTTAACTCTTTAGGGTGGTTAG CAATATCACTGGTGTCCAATAGC	bp 1641 bp 3177	COI-COII
CO1forward5_01 ATP6neu rev	GGGTTTGATTGGCACTTCTTTTAGG GAACCAAACTGAGCAGGACTATG	bp 150 bp 1850	COI-ATP6
Hia16int-for Hi16int-rev	GTACGAAAGGACCAAGATCC GTTTATCGCAGGAGAAGATTTGAC	bp 13529 bp 13711	Sequencing
COIseq294f ATP6neu rev	GGTTGGAGGTTTTGGGAATTGG GAACCAAACTGAGCAGGACTATG	bp 294 bp 1850	Sequencing

Table 4: Amplification and sequencing protocols used in this study.

16S-COIII		
step	**temperature**	**time**
initial denaturation	94°C	2 min
loop 1 - 35		
denaturation	94°C	30 sec
annealing	57.5°C	20 sec
primer extension	72°C	1 min 45 sec
final extension	72°C	7 min
COI-ATP6 (old)		
step	**temperature**	**time**
initial denaturation	94°C	2 min
loop 1 - 35		
denaturation	94°C	25 sec
annealing	50°C +/- 8°C	10 sec
primer extension	72°C	1 min 45 sec
final extension	72°C	7 min
COI-COII		
step	**temperature**	**time**
initial denaturation	94°C	2 min
loop 1 - 35		
denaturation	94°C	30 sec
annealing	52°C +/- 6°C	10 sec
primer extension	72°C	1 min 45 sec
final extension	72°C	7 min
COI-ATP6		
step	**temperature**	**time**
initial denaturation	94°C	2 min
loop 1 - 35		
denaturation	94°C	30 sec
annealing	57°C +/- 3°C	30 sec
primer extension	72°C	2 min
final extension	72°C	7 min
Sequencing		
step	**temperature**	**time**
initial denaturation	--	--
loop 1 - 35		
denaturation	96°C	20 sec
annealing	48°C	10 sec
primer extension	60°C	4 min
final extension	--	--

The results of the PCR reaction were checked with 1% TAE agarose gel electrophoresis stained with 1.3 µl ethidium bromide. A mixture of 3 µl PCR product and 3 µl loading buffer (bromphenol blue) was loaded in the gel chambers. Additionally 4 µl of 1 kb DNA ladder (Peqlab) were loaded to verify the size of the products. The conditions for running the gel were 100 V for 30 minutes.

2.1.4. Purification of PCR products

The purification was performed using the E.Z.N.A. Cycle-Pure-Kit Classic-Line (Peqlab) following the instruction manual. The "washing step" (see the manual) was carried out with 700 µl of SPW buffer instead of 750 µl. The PCR product was eluted in 30 µl of elution buffer in order to increase the concentration.
The results after purification were checked with 1% agarose gel electrophoresis under the same conditions described above.

2.1.5. Sequencing reaction

For each sequence reaction a total volume of 10 µl was necessary. The required reagents were: 2 µl primer (5 pmol/µl), 1 µl BigDye (v.3.1), 2 µl deionised H_2O and 5 µl of PCR product.
The primers used in the sequencing reaction and the corresponding protocol are listed in Table 3 and Table 4. The PCR products were sequenced on an ABI 3130xl capillary sequencer at the University of Vienna.

2.1.6. Alignment and sequence analysis

Sequences were first checked with the NCBI Nucleotid BLAST search (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) to exclude contamination. The fragments were corrected and manually edited with the software Finch TV v.1.4. (http://www.geospiza.com/finchtv). Both DNA strands of each fragment were assembled to generate a consensus sequence by either using the CAP3 Sequence Assembly Program (Huang and Madan 1999) or manually in Bioedit v.7.0.4.1 (Hall 1999).

In addition to the 96 sequences obtained in this study, one more sequence from NCBI genbank (http://www.ncbi.nlm.nih.gov/) (gi_1155316; Dreyer and Steiner 2006) was used for sequence analyses, genealogical analyses and population genetics.

Sequences were analysed with the tRNA-scan-SE v.1.21 (Lowe and Eddy 1997) in order to detect the presence of transfer RNAs. The protein coding genes were analysed with the NCBI Open Reading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf/html).
The multiple sequence alignment was performed using ClustalX v.1.8 (Thompson et al. 1997) and then manually edited in Bioedit v.7.0.4.1 (Hall 1999) and GeneDoc v.2.6.002 **(**Nicholas and Nicholas 1997). The files for the following analyses were converted in a nexus format using ForCon v.1.0 (Raes and Van de Peer 1999).

The nucleotide alignments were used to perform both “single gene” and “combined gene” analyses. The tested data partition were the following: “proteins” (COI, COIII and ATP6), “16S rRNA” (ribosomal RNA), “NC_tRNA” (non coding regions and the transfer RNA Glycin) and “all data” (“proteins”, “16S rRNA”, “NC_tRNA”). Proteins were translated into amino acids using the NCBI Invertebrate Mitochondrial Code (http://www.bioinformatics.org/JaMBW/2/3/TranslationTables) to detect changes in the amino acid sequence.
Base frequencies, variable sites as well as parsimony informative sites in the alignment were calculated for all data partitions using Mega v.4 (Tamura et al. 2007).

2.2. Genealogical analyses

2.2.1. Bayesian Inference

Bayesian Inference was performed with MrBayes v.3.1.2 (Huelsenbeck and Ronquist 2001). Both single ("proteins", "16S rRNA", "NC_tRNA") and combined data sets were investigated using only unique sequences in the respectively datasets to reduce the calculation time. Gaps were treated as missing in all data partitions and all three codon positions were included in the analyses. The model of evolution was calculated for each analysis using Modeltest v.2.2 (Posada and Crandall 1998) and the AIC criterion was implemented in the analysis. For the combined analysis, three partitions were set, discriminating between "proteins", "16S rRNA" and "NC_tRNA". A method to approximate the posterior probabilities of trees was obtained using the Markov chain Monte Carlo (MCMC; Metropolis et al. 1953). As implemented in this version of MrBayes, 2 x 4 mcmc chains were generated. The number of generations (ngen) varied between 10^6 and 10^7 in order to achieve a standard deviation of split frequencies between the two independent Markov chains below 0.01. The sample frequency was set to 100 in all analyses. The resulting log likelihood (lnL) values of the trees were plotted against the numbers of generations. LnL values increased through the generation and reached a plateau. Trees with lnL values below the plateau were discharged with the burn-in command. In the present study, the burn-in had values between 300 and 2000.

Trees were visualised using Treeview v.1.6.6 (Page R.D.M., 1996) and subsequently edited with Corel Draw v.10 (Corel Comporation 2000).

2.2.2. Network analysis using the statistical parsimony approach

A haplotype is a unique nucleotide sequence (Halliburton 2004). Due to sexual reproduction in a population, the different haplotype lineages link and therefore the relationship between haplotypes result in a network. At intra-specific level, multiple substitutions rarely occur, because the time of divergence of two haplotypes is short (Posada and Crandall 2001).

Network analyses were performed in order to reconstruct genealogical relationships. For this purpose the statistical parsimony approach as implemented in TCS v.1.18 (Clement et

al. 2000) was used. The TCS algorithm estimated the maximum number of mutational steps between haplotypes that could be explained by single substitutions with a statistical confidence of 95%. The resulting parsimony connection limit determined how many steps between haplotypes could maximally occur with a probability greater than 95% to be parsimonious. Haplotypes that differed in more nucleotides than stabilised by the connection limit were not connected to the same network. When the network was established, TCS determined the hypothetical haplotype ancestor, which in this study was used to select the root in the Bayesian topology.

The statistical parsimony network analyses were performed with both single ("proteins", "16S rRNA", "NC_tRNA") and combined data sets. The probability of parsimony in the respective data sets was set to 90 or 95% in order to get one single network. Gaps were treated as missing in all data partitions.

The vector based graphs resulting from TCS were saved as postscript file (.ps) and subsequently converted into pdf format with an online converter (Neevia technology http://convert.neevia.com). The graphs were edited by hand with Corel Draw v.10 (Corel Corporation 2000).

2.3. Genetic variation

In population genetics, allele frequencies and genotype frequencies are used to estimate the genetic variation within a population as well as to indicate, how similar or different two populations are. A basic principle of population genetics is the Hardy-Weinberg equilibrium. It indicates that in a population, where no evolutionary forces and random mating occur, allele and genotype frequencies remain constant from generation to generation (Guo and Thompson 1992).
Evolutionary effects, such as mutation, genetic drift, natural selection and gene flow lead to changes in allelic frequencies and influence the distribution of the allele frequencies in a population. Mutation, natural selection and genetic drift are processes that can cause divergence between populations over time, whereas gene flow retards these divergences. The balance between these processes determines how different populations are (Halliburton 2004).

There are several approaches to estimate whether populations are in a panmictic state or genetically distinct.
In this study, two methods were used to determine whether the four populations (KR, IT, SP, P) and the three habitats (boring, sponge dwelling, nestling) were genetically different.

2.3.1. Distance method

Based on nucleotide differences in the alignment, the p-distances (pairwise differences divided by the length of the sequence) for all pairs of haplotypes were calculated in PAUP*4.0b10 (Swofford 2003). The p-distance matrix was implemented in SPSS v.7.5 (SPSS Inc., 1997). A non parametric rank-sum test (Mann Whitney U Test) was used to test the distribution of the ranks of two hierarchical structures (within versus between populations/habitats) under the null hypothesis of no differentiation.

The results were visualised as box plots or as 3D Bar charts using Sigmaplot v.10 (Systat Software, Inc. 2006).

2.3.2. Analysis of molecular variance (AMOVA)

The analysis of molecular variance (AMOVA) estimated the genetic distances between haplotypes taking into account the allele frequency variance within and between hierarchical structures. The total variance resulted as the sum of their covariance components in the different hierarchical level. The covariance components were tested for significance using a non-parametric permutation approach (10000 permutations) under the null hypothesis of no differentiation. The covariance components were then used to compute the fixation index (F_{st}) (Excoffier et al. 1992).

The analysis of molecular variance was performed using the software package Arlequin v.3.11 (Excoffier et al. 2005) taking into account all nucleotide sites (allowed level of missing data: 100%). The pairwise differences between haplotypes were used to infer the analysis. The fixation indices were calculated for the sum of populations as well as for pairwise population comparisons.

The fixation index (F_{st}) is a measure to determine how different populations are. It can be expressed as the relation of genetic variations within and among populations. Based on the

fixation indices, levels of differentiation between populations can be defined (Wright 1978, Table 5).

Table 5: Differentiation of populations based on the fixation index (F_{st}) (Wright 1978)

fixation index (F_{st})	differentiation of populations
0	none
$0 < F_{st} < 0.05$	low
$0.05 < F_{st} < 0.15$	moderate
$0.15 < F_{st} < 0.25$	strong
> 0.25	very strong

F_{st} values were used as indicators for gene flow between populations. The relationship between gene flow (in terms of number of migrants per generation, Nm) and molecular variance among populations is described as: $F_{st} = 1/(4\ Nm+1)$ (Wright 1951).Therefore, high F_{st} values are indicators for population differentiation and consequently for restricted gene flow.

In this study, a modified equation of Wright's approach was used to estimate gene flow among populations. This alteration consisted in estimating gene flow, or rather the number of migrants between populations per generation, for haploid genomes in dioecious species: $Nm = (1/ F_{st}) -1$ (Terranova et al. 2006).

To test whether genetic differences were correlated to the geographic distances, Slatkin distances ($F_{st}/1- F_{st}$) (Rousset 1997) from pairwise comparisons between sites were plotted as a function of geographical distance along the coast. Geographical distances between sampling sites were estimated with Google Earth (www.earth.google.com), taking into account the average distance between two populations. The level of correlation (r) between two dependent variables was defined following Zöfel (1992, p. 211) (Table 6).

Table 6: Correlation between two dependent variables based on the correlation coefficient r (Zöfel, 1992).

correlation coefficient (r)	degree of correlation
$0 < r < 0.2$	very low
$0.2 < r < 0.5$	low
$0.5 < r < 0.7$	moderate
$0.7 < r < 0.9$	strong
$0.9 < r < 1$	very strong

The significance of the correlation between genetic differences and geographic distances were tested with the Manteltest (Smouse et al. 1986) by means of 10000 permutations.

The results were visualised using Sigmaplot v.10 (Systat Software, Inc. 2006).

3. Results

3.1. Sequence analyses

The size of the amplified COI-ATP6 fragment is 1491 base pairs (bp), and the 16S-COIII fragment is 1254 bp. The length and the position of each component are listed in Table 7.

Table 7: Genes and the corresponding lengths and positions in the alignment. NC_1, NC_2 are non coding regions.

gene	length (bp)	nucleotide position
COI	1310	1 - 1310
NC_1	169	1311 - 1479
ATP6	12	1480 - 1491
16S rRNA	676	1492 - 2167
tRNA Gln	68	2168 - 2235
NC_2	311	2236 - 2546
COIII	199	2547 - 2745

3.1.1. Proteins (COI, ATP6, COIII)

The aligned dataset for the 97 specimens consists of 1521 bp, of which 140 positions are variable (9.2%) and 76 substitutions are parsimony informative. Most substitutions occur at the third codon position (119 variable sites), 13 at the first codon position and the remaining 8 substitutions at the second position. The average base composition is 38.7% thymine, 14.9% cytosine, 23.9% adenine and 22.5% guanine. Individual populations have very similar values.

The amino acid alignment is highly conserved in the three protein coding genes. All 12 changes in the amino acid alignment are parsimony uninformative and occur in different specimens of three populations (Croatia, Italy and Portugal). Nine amino acid changes are detected in the COI gene: three of 12 substitutions at the first codon position and all six changes at the second position lead to amino acid changes in different specimens of three populations. The COIII gene shows three amino acid changes: the single substitution at the first codon position and both changes at the second codon position lead to amino acid changes in specimens of the Croatian population. The ATP6 gene shows no differences in the amino acid alignment.

The open reading frame determined in this study suggests that the COIII gene starts with the codon ATT. The nucleotide C in the genbank sequence therefore was deleted from the alignment in order not to violate the open reading frame (Figure 5, position 33).

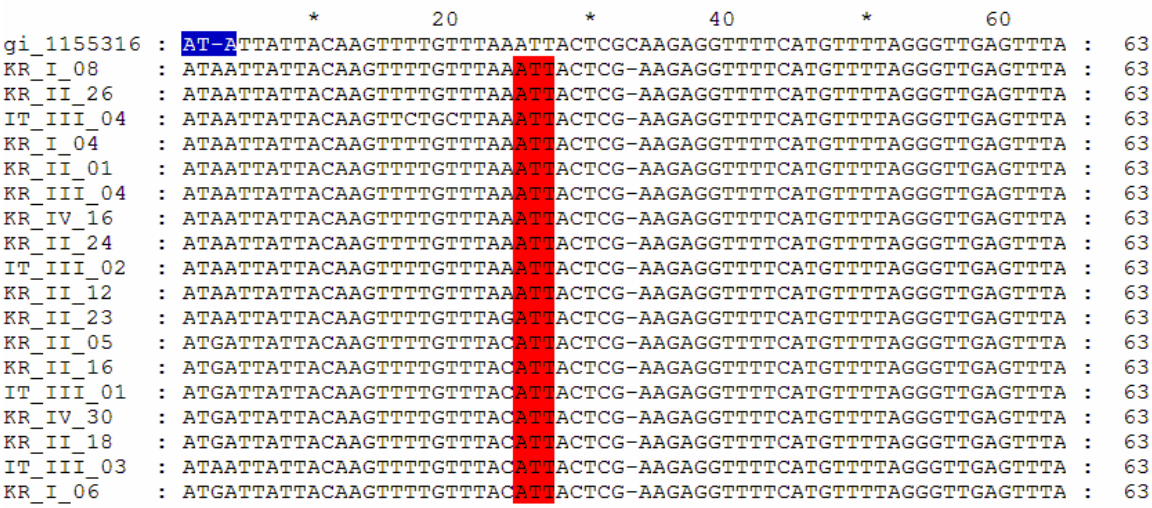

Figure 5: Beginning of the open reading frame of COIII genes in Dreyer and Steiner 2006 (blue) and of this study (red).

3.1.2. 16S rRNA

Of the 676 bp of the 16S rRNA gene, 54 nucleotides are variable (8.0%), 22 of them have parsimony informative character. Base frequencies are similar across populations. The average base frequencies are 32.2% thymine, 15.4% cytosine, 30.7% adenine and 21.7% guanine.

3.1.3. NC_tRNA (NC_1, NC_2, tRNA Gln)

This dataset consists of 548 bp of which 175 are variable (31.9%) and 86 are parsimony informative. The average base composition is 38.6% thymine, 9.7% cytosine, 33% adenine and 18.6% guanine. Again, the base frequencies are homogeneous throughout all populations.

Six individuals, all of them Croatian specimens, show an indel of 67 bp in the NC_2 region (Figure 6).

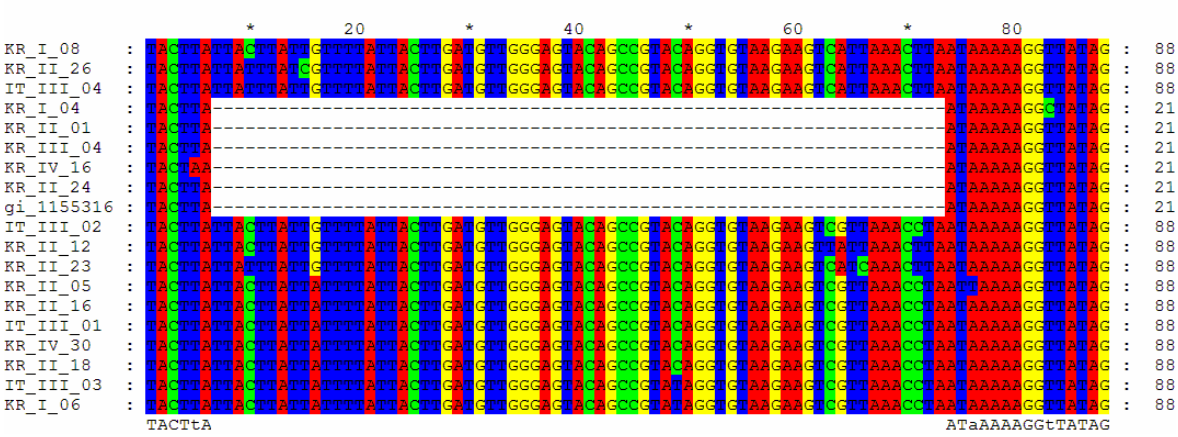

Figure 6: Part of the NC_2 alignment of selected sequences.

3.1.4. All data

The size of the alignment, corresponding to the combination of the previous partitions, is 2745 bp. The total number of variable sites is 369 (13.4%) and 184 are parsimony informative. Base frequencies are similar among populations and the average frequencies across the alignment are 37.1% thymine, 14.0% cytosine, 27.4% adenine and 21.5% guanine.

3.2. Genealogical relationships

All the topologies obtained by statistical parsimony analyses and Bayesian Inference from all data partition are compared and checked for similarities. Based on these results, clades of related haplotypes (haplogroups) are defined. A detailed description of the haplogroups according to origin and habitat follows in the section “Haplogroups”.

3.2.1. Proteins

Bayesian Inference

Bayesian Inference from the protein coding genes was performed with all the three codon positions. The most appropriate model of evolution for this dataset, the GTR+I+G, was implemented in the analysis. The number of generations was set to 10^6, sampling every 100^{th} tree; the burnin was set to 500.

The Bayesian tree shows a clear split of the specimens in the three clades, here referred to as haplogroup A, B and C.
The haplogroup A is supported with a posterior probability of 94%. The phylogenetic signal and the statistical support within haplogroup A are rather too low to unambiguously resolve the relationships within this clade. Note the two clades within haplogroup A separated by a posterior probability of 0.99. As in the network analysis inferred from “proteins” these two groups are not distinct, an additional haplogroups is not justified. The haplogroup B shows a posterior probability of 1 and the relationships within this group are mostly resolved. Haplogroup C is highly supported as well (96%), but the relationships within haplogroup C are again poorly resolved, therefore causing several polytomies at the base of the clade (Figure 7).

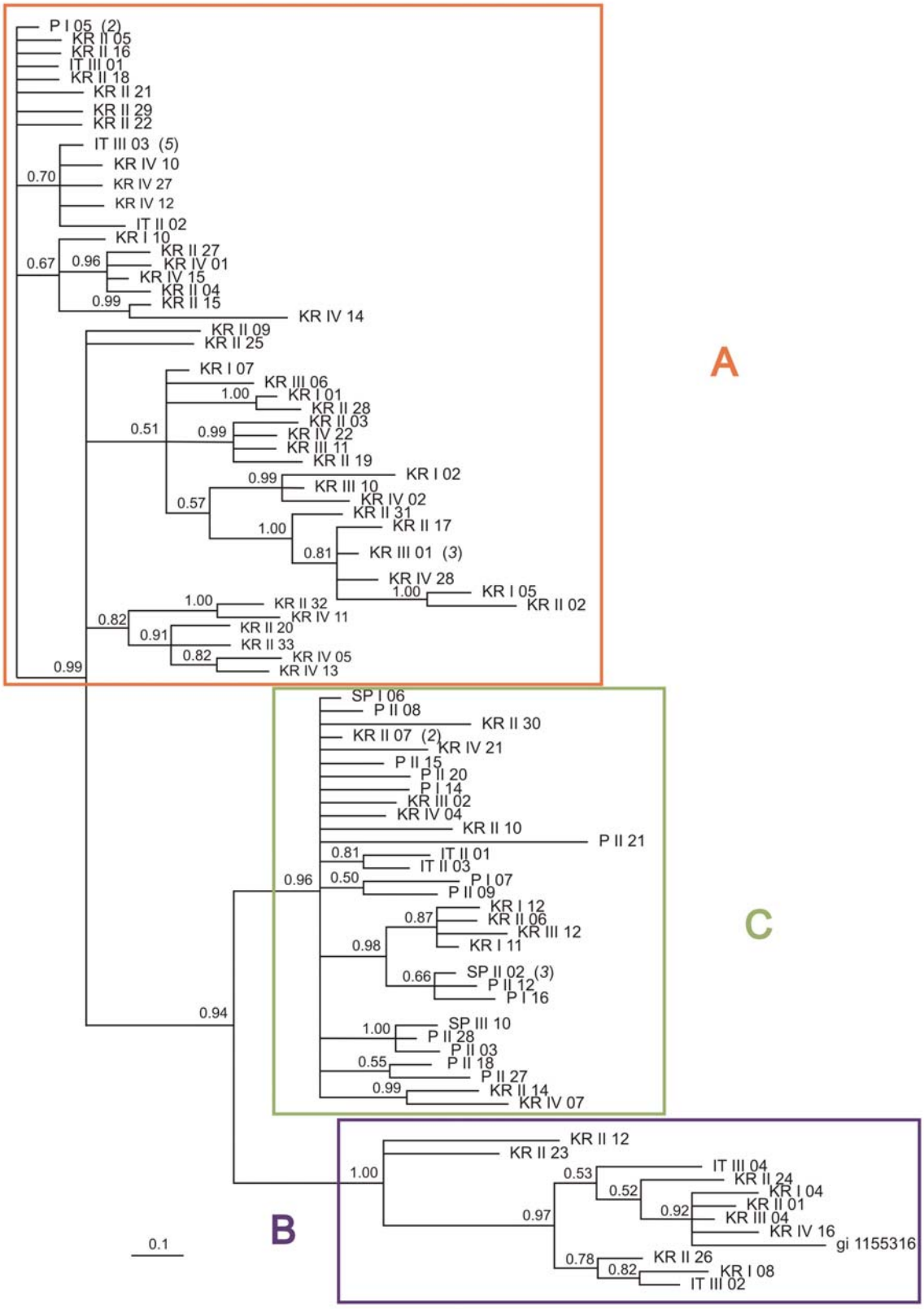

Figure 7: Unrooted Bayesian tree inferred from protein coding genes. The values at the nodes indicate their posterior probabilities. The scale indicates 0.1 substitutions per site.

Network analysis

Of 87 observed haplotypes, 82 occur once. Of five shared haplotypes, two occur twice, two three times, and one five times. Two of these haplotypes occur in Croatia only, whereas Croatia and Portugal, Croatia and Italy as well as Spain and Portugal share one haplotype respectively (Table 8).

Table 8: Shared haplotypes found in the protein coding genes. Names on the left appear in Figure 8.

Proteins	
KR III 01	KR III 03, KR III 09
KR II 07	KR IV 09
P I 05	KR IV 30
SP II 02	P II 10, P II29
IT III 03	KR I 06, KR III 07, KR III 08, KR IV 24

The genealogical relationships are shown in Figure 8. All haplotypes are connected to a single network with a statistical parsimony probability of 95% equivalent to a connection limit of 18 steps between haplotypes. The haplotype ‘P I 05, KR IV 30’ results as the most probable ancestor with 11.11%.

The haplogroups A and C include three and two shared haplotypes respectively, whereas in haplogroup B all haplotypes are unique. Haplogroup A is separated from haplogroups B and C by eight and nine unobserved haplotypes, respectively. The haplogroups B and C are not directly connected.

Within the haplogroup A, up to seven unobserved haplotypes are detected. These relatively high numbers of differences cause that these specimens are not closely related in the Bayesian tree. Within the haplogroup B, the two specimens, KR II 12 and KR II 23, are set apart from the remaining specimens of the group. The same pattern is supported in the Bayesian Inference where these specimens branch off first and result as sister group to all other haplotypes of the group. Since network analyses inferred from the “16S rRNA” dataset clearly show that these two specimens are not separated from the rest of the group, an additional haplogroup is not justified. Within haplogroup C, a single specimen from Portugal (P II 21) is quite different compared to other specimens in this group connected by a long branch of unobserved haplotypes (Figure 8). In the Bayesian Inference, this specimen is also characterised by a long branch.

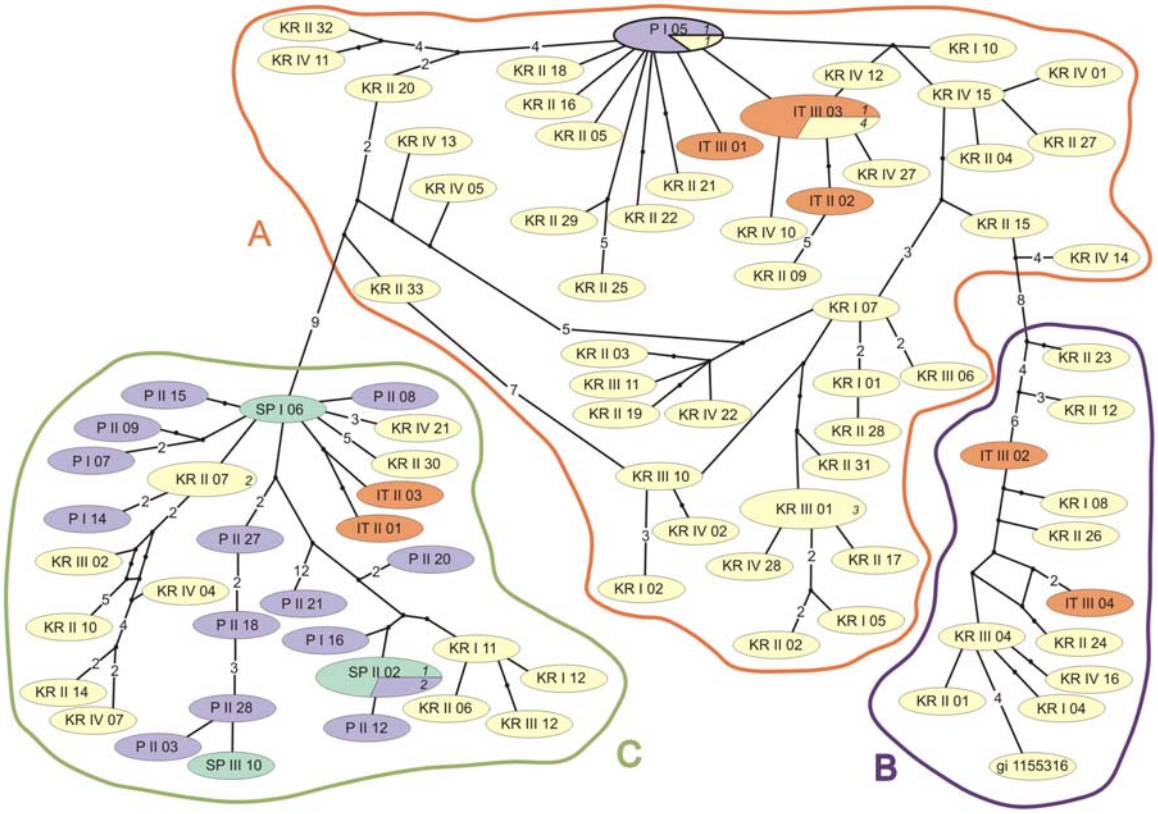

Figure 8: TCS network inferred from protein coding genes, including all haplotypes found in the Mediterranean and Atlantic Ocean. Parsimony probability: 95%; Outgroup probability: 11.11%. Annotation for this and following TCS networks: Haplotypes of different populations are visualised by different colours. The coloured lines represent the haplogroups A, B and C respectively, and are not related to the colours used to visualise the populations. The haplotype with the highest outgroup probability is thicker framed. The size of the ovals and the included italic numbers are correlated to the haplotypes frequencies. When the same haplotype is shared by different populations, the relative frequencies of each population are visualised. Nodes and numbers between haplotypes represent unobserved haplotypes; but the length of the lines is not correlated to the genetic distance.

3.2.2. 16S rRNA

Bayesian Inference

Bayesian Inference from the 16S rRNA gene was performed including all positions and treating gaps as missing character. The most appropriate model of evolution selected for this dataset was the HKY+I. The number of generations was set to 10^6, sampling every 100^{th} tree; the burnin was set to 300.

In contrast to the "protein" dataset, the Bayesian tree obtained from the "16S rRNA" data does not support the haplogroups defined before. The low phylogenetic signal in this gene causes many polytomies. The haplogroup B is largely resolved and, except for the genbank sequence, well supported (Figure 9).

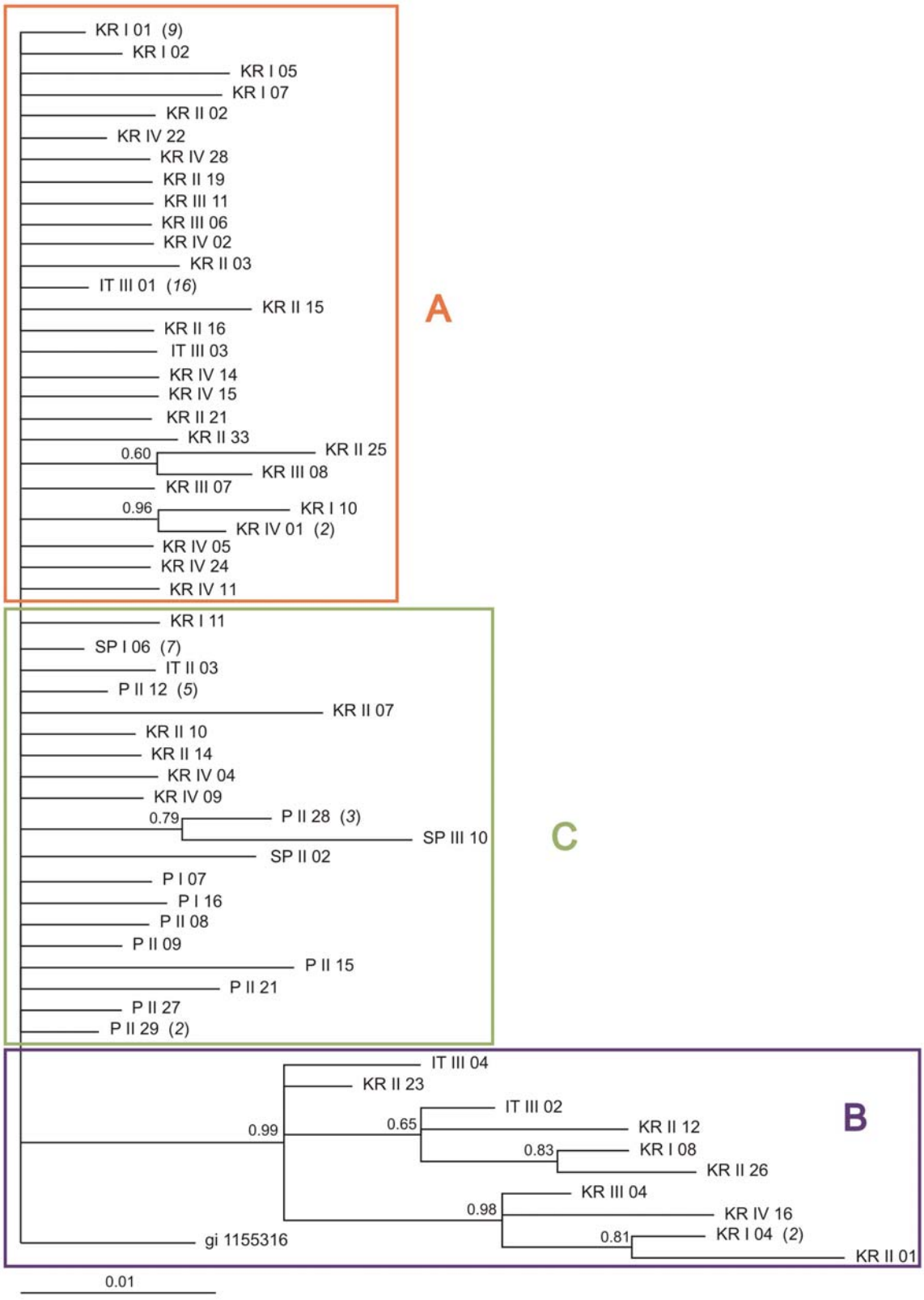

Figure 9: Unrooted Bayesian tree inferred from the 16S rRNA gene. The values at the nodes indicate their posterior probabilities. The scale indicates 0.01 substitutions per site.

Network analysis

The sequence variability of this partition revealed 59 different haplotypes, 51 of them occurring once. Of eight shared haplotypes, four occur rather frequently (5, 7, 9 and 16 times respectively), while three haplotypes occur twice and one haplotype three times. One haplotype is present in all four populations, one haplotype occurs in three populations (except Spain), one haplotype occurs in two populations (Croatia and Portugal) and five haplotypes are limited to one population, i.e. three in Croatia and two in Portugal (Table 9).

Table 9: Shared haplotypes found in the 16S rRNA genes. The names on the left appear in Figure 10.

16S rRNA	
KR I 04	KR II 24
KR IV 01	KR II 09
KR I 01	KR II 17, KR II 28, KR II 31, KR III 01, KR III 03, KR III 09, KR III 10, KR IV 13
P II29	P II 10
P II 28	P II 03, P II 18
IT III 01	KR I 06, KR II 04, KR II 05, KR II 18, KR II 20, KR II 22, KR II 27, KR II 29, KR II 32, KR IV 10, KR IV 12, KR IV 27, KR IV 30, IT II 02, P I 05
IT II 01	KR II 30, KR III 02, KR IV 21, SP I 06, P I 14, P II 20
KR I 12	KR II 06, KR III 12, KR IV 07, P II 12

The network inferred from the “16S rRNA” dataset shows complete connection of all haplotypes at the 95% level equivalent to a maximum of 11 connecting steps. The hypothetical ancestor differs from the “proteins” presenting the haplotype-cluster KR I 01, which in this dataset occurs nine times, as outgroup (13.58% probability).

The network shows low genetic variation within and among haplogroups. Figure 10 shows that haplogroup B and C are not connected and both branch out from the haplogroup A as observed before. In this partition, however the TCS analysis shows no unobserved haplotypes between the haplogroup A and the other two ones. Since the two most frequent haplotypes both occur within haplogroup A and show no unobserved haplotypes between them, this network analysis could suggest the existence of a fourth haplogroup. However, the two clades of haplotypes within haplogroup A are not identical with those observed in the network inferred from the “proteins”. The Portuguese specimen (P II 21) that resulted as single outlier in the network analysis inferred from “proteins”, in this analysis is clearly

included in the haplogroup C, as well as the two specimens, KR II 12 and KR II 23, from the haplogroup B (Figure 10).

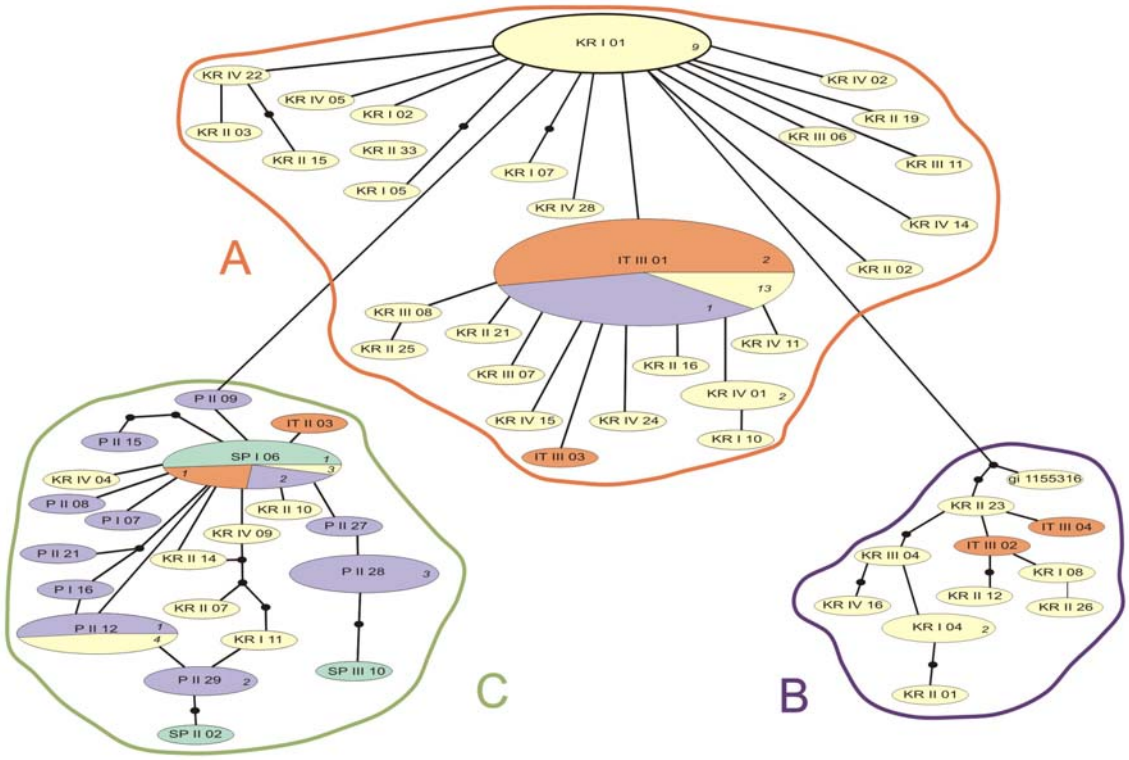

Figure 10: TCS network from 16S rRNA genes, including all haplotypes found in the Mediterranean and Atlantic Ocean. Parsimony probability: 95%; Outgroup probability: 13.58%.

3.2.3. NC_tRNA

Bayesian Inference

Bayesian Inference from the NC_tRNA data partition was performed including all positions of the alignment and treating the gaps as missing. The most appropriate model of evolution for this dataset is GTR+I+G. The number of generations was set to 2 x 10^6, sampling every 100^{th} tree; the burnin was set to 1540.

The resulting topology supports the existence of the three haplogroups described in the "protein" partition. However, the statistical support for the haplogroup C and the separation of the B-C clade from the haplogroup A are lower. The resolution within the haplogroup A is lower compared to the "proteins" resulting in several polytomies at the base of the tree. On the other hand, the haplogroup C is better resolved in this dataset. All nodes that occur in the haplogroup B are well supported. It is interesting to note that inside

this haplogroup the cluster sustained by a posterior probability support of 1 includes only specimens with the indel in NC_2 (Figure 11).

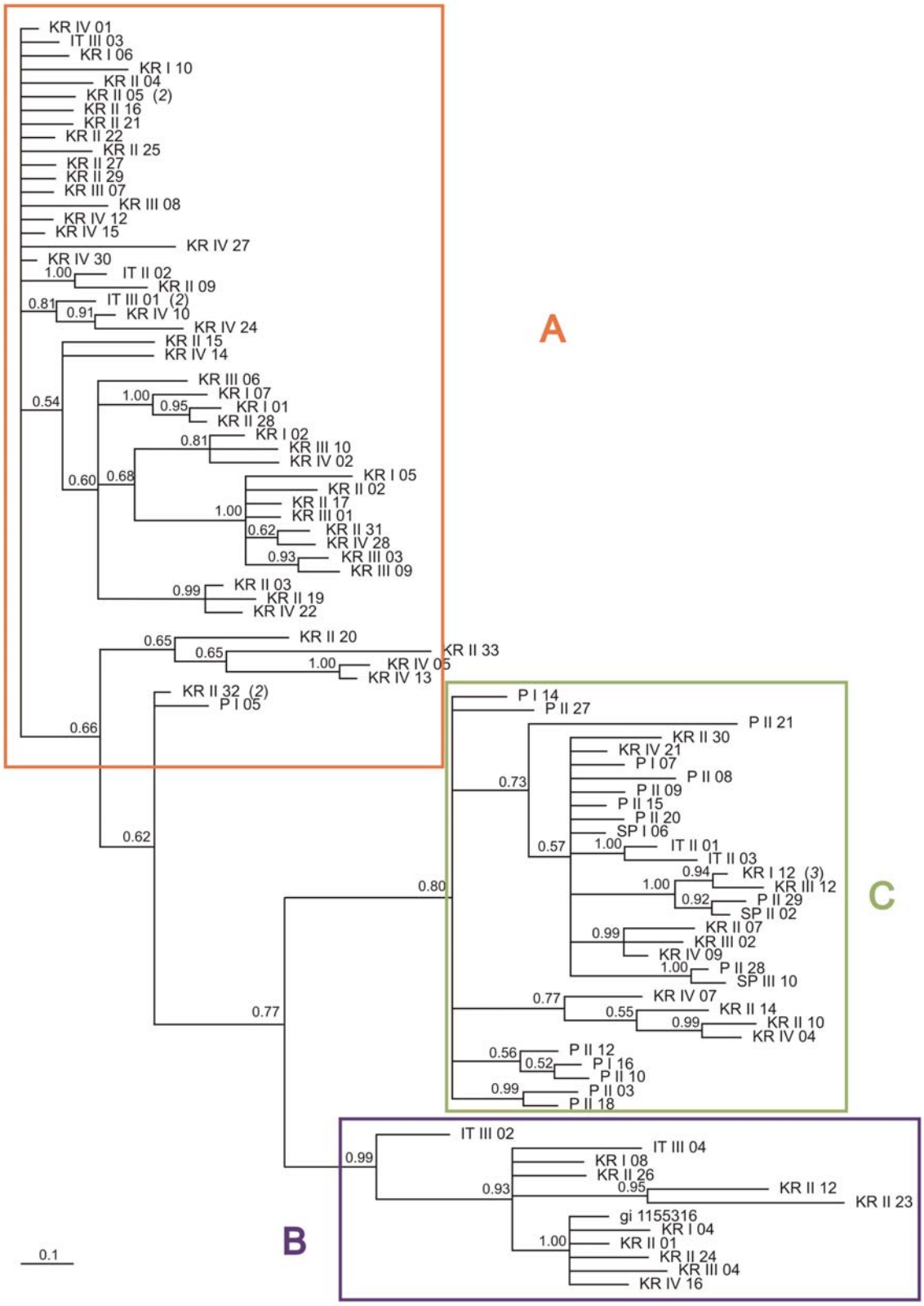

Figure 11: Unrooted Bayesian tree inferred from NC_tRNA dataset. The values at the nodes indicate their posterior probabilities. The scale indicates 0.1 substitutions per site.

Network analysis

The dataset reveals 92 different haplotypes, 88 of them occurring only once. Of four shared haplotypes, three occur twice and one three times. One of these haplotypes is shared by Croatia and Italy, whereas the remaining three occur in Croatia only (Table 10).

Table 10: Shared haplotypes found in the NC_tRNA genes. The names on the left appear in Figure 12.

NC_tRNA	
KR II 05	KR III 11
IT III 01	KR II 18
KR I 12	KR I 11, KR II 06
KR II 32	KR IV 11

The network analysis of the NC_tRNA dataset required a decrease of the parsimony probability to 90% (connection limit of 15 steps) in order to get one single network. In contrast to both previews networks, in this dataset the haplotype KR IV 01 has the highest outgroup probability (8.75%).

In Figure 12, the genealogical relationships are shown. The haplogroup A is separated from haplogroup B by six unobserved haplotypes. The haplogroups B and C are separated by nine unobserved haplotypes.
In contrast to the other data partitions, the haplogroup A is not directly connected to the sub-network C, but via haplogroup B. Similar to the network inferred from “proteins”, the Portuguese specimen (P II 21) in haplogroup C is separated from others by 10 unobserved haplotypes. In contrast to the networks “proteins” and “16S rRNA”, the Italian specimen IT III 02 cannot clearly be assigned to the haplogroup B.
Eight unobserved haplotypes separate IT III 02 from the remaining haplogroup B and this amount is similar between IT III 02 and P II 10 (Haplogroup C) as well as between IT III 02 and KR II 32 (Haplogroup A) with ten and nine unobserved haplotypes respectively. In the Bayesian analysis inferred from this dataset, IT III 02 branches off first and results as sister to all other specimens of haplogroup B.

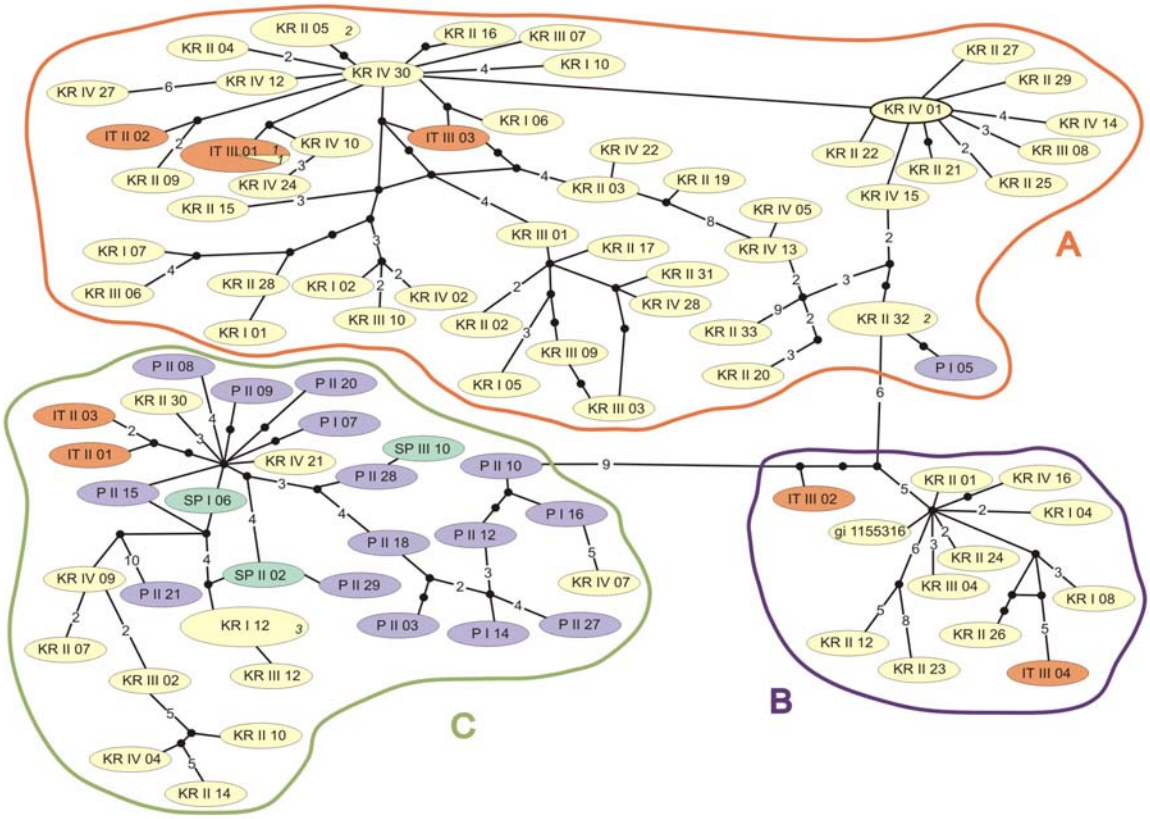

Figure 12: Network of the partition NC_tRNA, including all haplotypes found in the Mediterranean and Atlantic Ocean. Parsimony probability: 90%; Outgroup probability: 8.75%.

3.2.4. All data

Bayesian Inference

Bayesian Inference from "all data" was performed on an alignment resulting by the concatenation of the three single partitions "proteins", "16S rRNA" and "NC_tRNA" and their corresponding models of evolution. In this analysis the number of generations was set to 10^7, sampling every 100^{th} tree; the burnin was set to 2000.

The Bayesian Inference of this combined dataset clearly supports the split in three haplogroups A, B and C. The posterior probability support of the nodes corresponding to the haplogroups B and C is comparable to that of the "protein" partition. The separation of the B-C clade from haplogroup A is supported as well even if the posterior probability is lower than in the "protein" dataset. The haplogroups A and C show many polytomies indicating a network character of these genealogies (Figure 13).

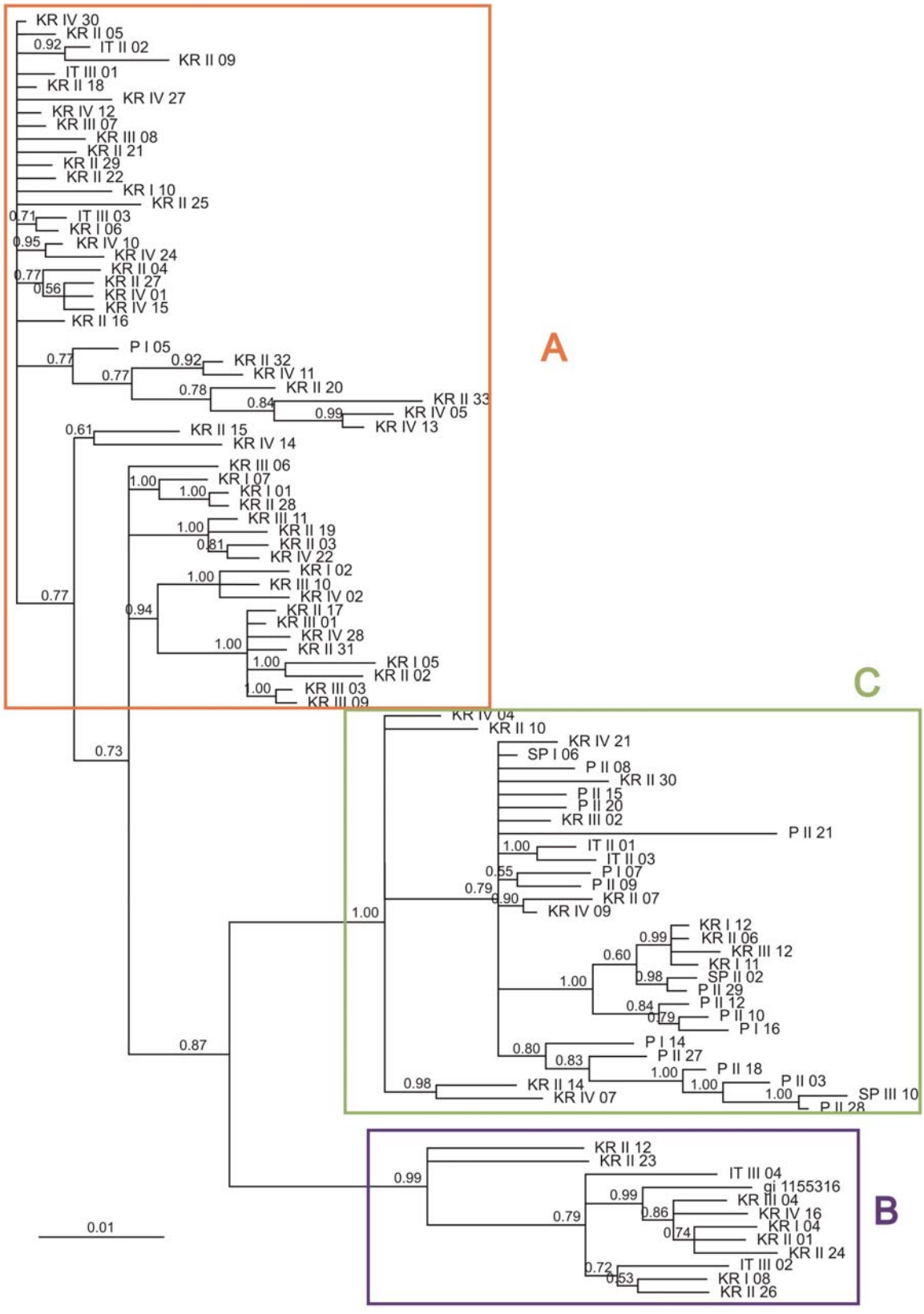

Figure 13: Unrooted Bayesian tree inferred from all data. The values at the nodes indicate their posterior probabilities. The scale indicates 0.01 substitutions per site.

Network analysis

According to the "NC_tRNA" network, the statistical parsimony analysis inferred from "all data" was performed with a parsimony probability of 90% (connection limit: 33 steps). The haplotype KR IV 30 results as the most probable outgroup (3.7%).

In the combined dataset all haplotypes are occur only once. The statistical parsimony network shows a clear slit of the taxa in the three haplogroups, A, B and C. The haplogroup A is separated from haplogroup B and C by 23 and 28 unobserved haplotypes respectively. Similar to the networks of the partitions "proteins" and "16S rRNA", the haplogroups B and C are not directly linked but connect to haplogroup A. There is no evidence for the existence of a fourth haplogroup as suggested in the network analysis inferred from "16S rRNA". The TCS analysis of the combined data set shows that the two putative groups are not separated. Similar to the "protein" and "NC_tRNA" networks, the Portuguese specimen (P II 21) in haplogroup C is separated from the others by 21 unobserved haplotypes. Although this specimen is genetically quite different from the other members of haplogroup C, no additional haplogroup is created for this single sequence. Within the haplogroups A and B, up to 18 unobserved haplotypes are detected. Although within single haplogroups considerable numbers of unobserved haplotypes occur, no additional haplogroups are defined because they are not supported by single data partitions (Figure 14).

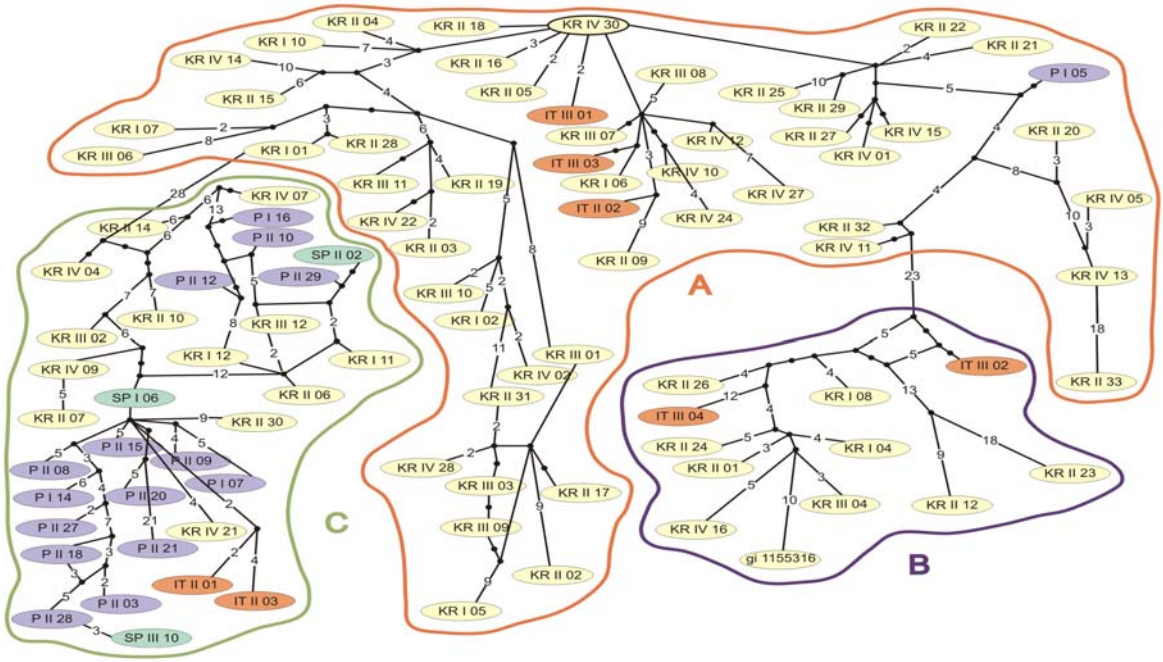

Figure 14: Statistical parsimony network performed with all data and all haplotypes. Parsimony probability: 90%; Outgroup probability: 3.7%.

3.2.5. Haplogroups

Both Bayesian Inferences and TCS show that the populations from Croatia and Italy are a mix of all haplogroups, whereas specimens from Spain belong to haplogroup C only and specimens from Portugal belong to the haplogroups A and C. The maximum genetic distance is similar both within haplogroups (1.5%-1.7%) and among haplogroups (2.6%-2.7%).

Haplogroup A includes 52 specimens from three populations and therefore 53.6% of all samples. In all TCS networks the haplogroup A includes the haplotype outgroup. The group is mostly represented in the Croatian population with 48 individuals from both habitats, boreholes and sponges. The three Italian specimens and a single specimen from Portugal belong to haplogroup A and come from the nestling habitat. None of the Spanish samples belongs to this group.

Haplogroup B is the smallest haplogroup with 12 specimens (12.4%) from Croatia and Italy. The 10 Croatian specimens are borers and sponge dwellers, the two Italian individuals are nestlers. This haplogroup includes the Croatian samples with the deletion in the NC_2 region (Figure 6). This feature is not restricted to a collecting site or habitat.

Haplogroup C includes 33 individuals from all populations and all habitats. It includes two Italian samples, all Spanish samples and 15 specimens from Portugal, all of them belonging to the nestling habitat. The 13 Croatian samples are borers and sponge dwellers. In contrast to the other groups, the haplogroup C is represented mostly in the Western Mediterranean Sea and Atlantic Ocean (Figure 15).

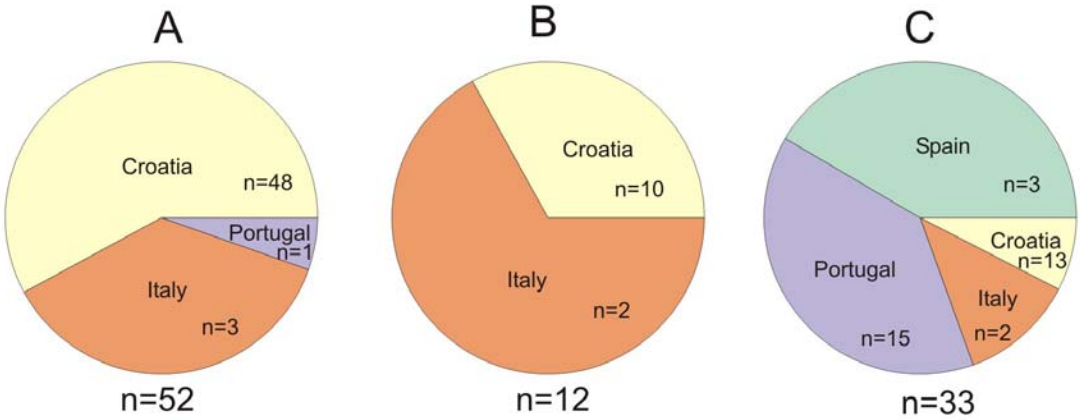

Figure 15: Relative frequencies of each population within the haplogroups A, B and C respectively. The size of the circles is not correlated with the size of the haplogroups.

3.3. Genetic variation

3.3.1. Populations

Among all 71 Croatian specimens, 48 belong to the haplogroup A (67.6%). Haplogroup B and C are detected similar frequent, exhibiting 10 samples (14.1%) in the former group and 13 specimens (18.3%) in the latter respectively.

Although only seven specimens were collected from the Ligurian Sea, they are homogenously distributed among all three haplogroups. Haplogroup A includes three samples and both B and C two individuals respectively.

The three samples from the Mediterranean coast of Spain belong to the same haplogroup (C). Of 16 specimens collected from Portugal, all except one belong to the haplogroup C (Figure 16).

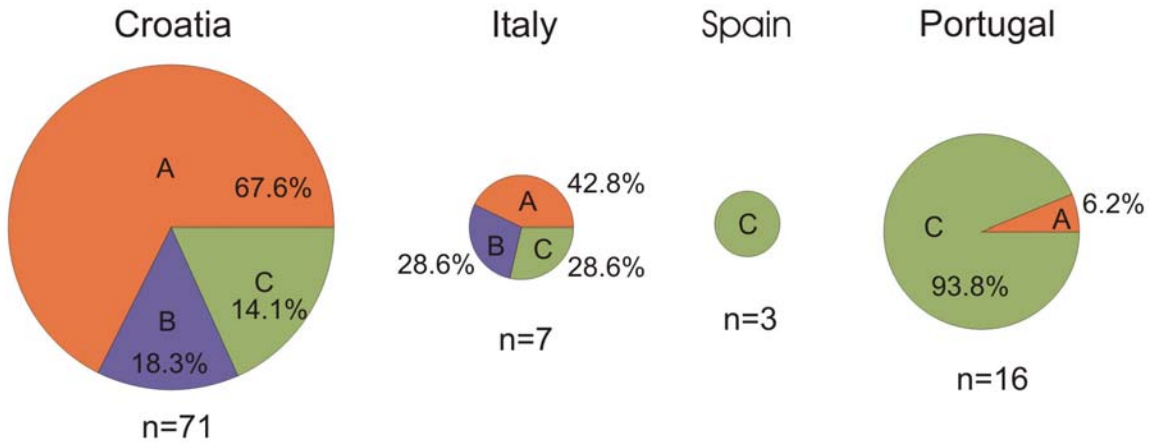

Figure 16: Haplogroup frequencies within populations. The size of the circles is standardised to the relative numbers of specimens per sampling area.

Analyses to estimate the genetic variation within and among populations were performed with all data partitions. As the results are very similar in the different data sets, only the "all" data partition is described in detail. The results of the partitions "proteins", "16S rRNA" and "NC_tRNA" are shown in Table 11.

Croatian specimens show high genetic variation within the population. The average and the median pairwise genetic distance is 1.4% and reaches a maximum of 2.6% between KR I 05 (Haplogroup A, borer) and KR II 30 (Haplogroup C, borer).

The Italian specimens display high genetic variation according to the uniform distribution of haplogroups. The average genetic distance of 1.8% and the median value of 2.0% are noticeable higher than in the Croatian population. The maximum genetic distances of 2.4%

results between IT II 03 (Haplogroup A, nestler) and IT III 04 (Haplogroup B, nestler), but 50% of the distances lie in the interval between 1.4 and 2.2%. The low genetic distances between specimens belonging to the same haplogroup lead to a left-skewed shape of the distance distribution.

The three Spanish specimens belong to haplogroup C. The average and median genetic distances are 0.8 and 0.7% respectively and the maximum of 0.9% occurs between SP II 02 and SP III 10 (both nestlers).

The Portuguese specimens belong to haplogroup C, with one exception. The average genetic distance is 0.9%, slightly higher than the median of 0.8%. The maximum genetic distance of 1.9% occurs between P I 05 (Haplogroup A, nestler) and P II 28 (Haplogroup C, nestler), but 50% of the distances lie in the interval between 0.7 and 1%. The specimen from haplogroup A and P II 21 contribute mostly to the right-skewed shape of the distance distribution (Figure 17).

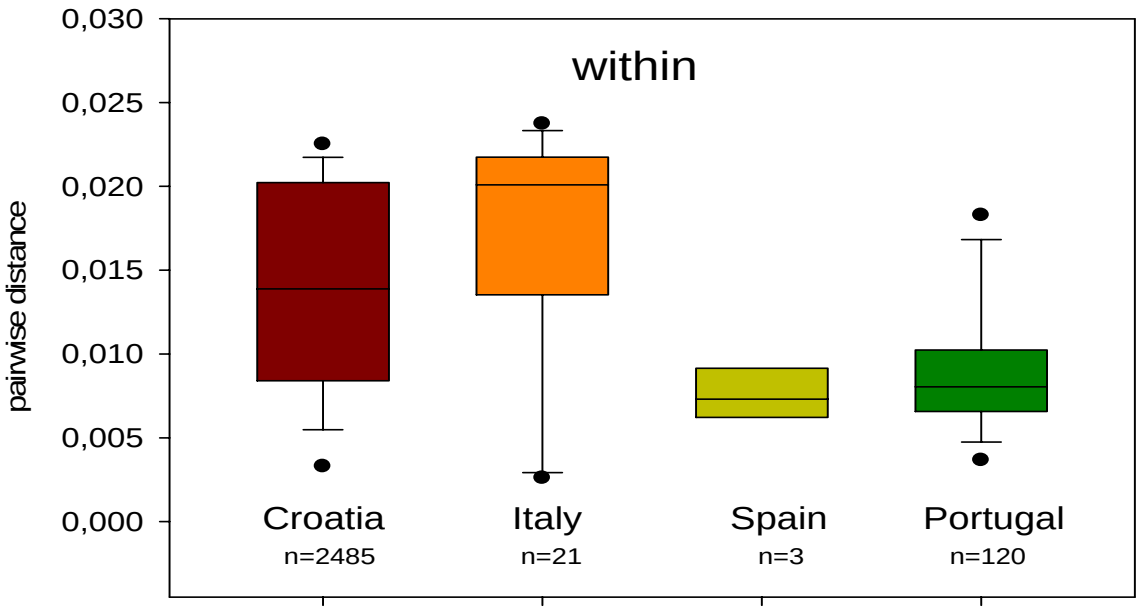

Figure 17: Pairwise genetic distances (p-distances) within each population. n indicates the number of pairwise comparisons.

Figure 18 shows the distribution of the genetic distances within and among populations. At intra-population level, the genetic differences follow a bimodal distribution with two peaks in the intervals 0.6-1% and 1.9-2.3% respectively. In among population comparisons the genetic distances are clearly higher and most of them occur between 1.8 and 2.3%. The maximum genetic distance of 2.7% among populations occurs between KR II 12 (Haplogroup B, borer) and SP III 10 (Haplogroup C, nestler). However, also at inter-

population level several comparisons with low genetic differences occur, leading to a left-skewed shape of the distance distribution.

The AMOVA reveals a strong differentiation of these two genetic structures (F_{st}=0.25, p=0.000). The Mann Whitney U-Test shows that the p-distances among populations are significantly higher than within populations (p=0.000).

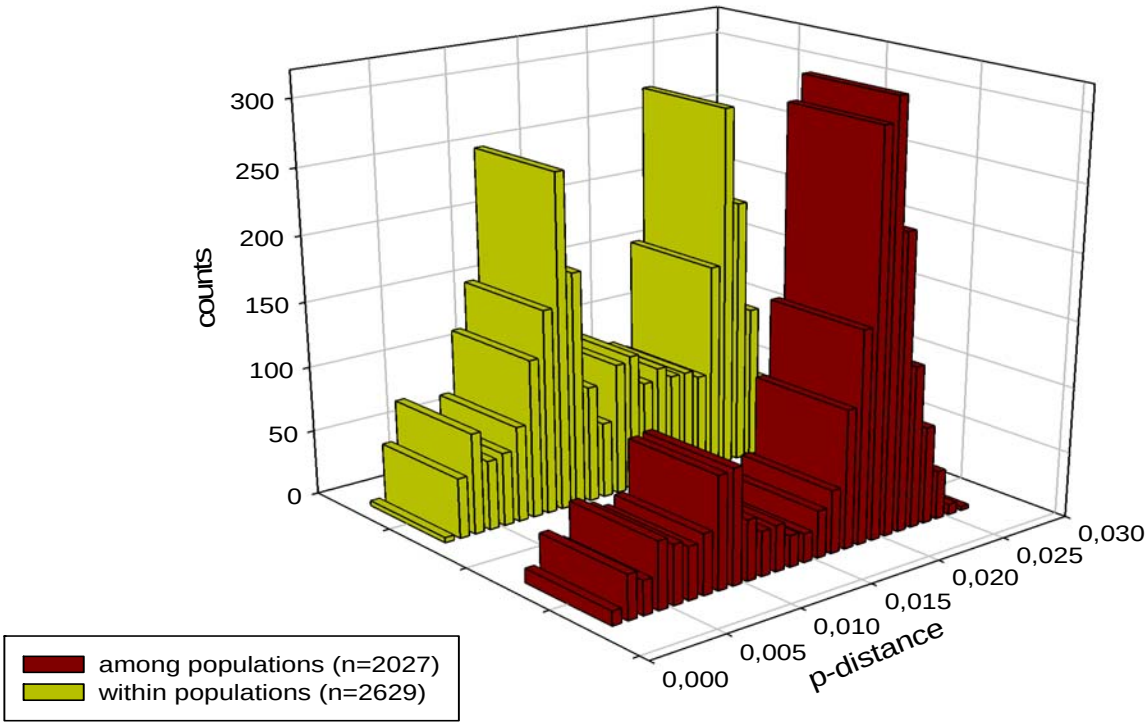

Figure 18: Distribution of p-distances within and among populations. n indicates the number of pairwise comparisons.

In pairwise population comparisons, the pattern is highly different between pairs of populations (Table 11).

Between Croatia and Italy, the analysis of molecular variance reveals low differentiation between the two populations (F_{st}=0.003, p=0.29). However, this result disagrees with the Mann Whitney U-Test. The p-distances between the two populations are significantly higher than within (p=0.000).

The populations from Spain and Portugal show no differentiation (U-Test: p=0.30). This result is supported by AMOVA. The genetic variation within the two populations is remarkable higher than between them. Therefore, no differentiation between the Spanish and the Portuguese population is detectable (F_{st}=0, p=0.73).

The Spanish population shows a very strong differentiation from the Croatian population (AMOVA: F_{st}=0.34 p=0.003). The distribution of pairwise distances from within population comparisons is significantly different from between population comparisons as well (U-Test: p=0.000).

AMOVA shows a strong differentiation between the Spanish and the Italian population (F_{st}=0.21), although the differentiation is not supported by the AMOVA significance test (p=0.14). According to this result, the Mann Whitney U-Test shows that the genetic distances within the two populations are not significantly different from between population comparisons (p=0.57).

The Portuguese specimens belong mostly to the haplogroup C. The low frequency of other haplogroups in this population sets it apart from Croatia and Italy. The Portuguese population shows a very strong differentiation from both, Croatia (F_{st}=0.32) and Italy (F_{st}=0.24) with high statistical supports of p=0.000 and p=0.003, respectively. The Mann Whitney U-Test shows a similar result. The genetic distances from within population comparisons are highly significant different from between population comparisons (p=0.000).

All data partitions (“all data”, “proteins”, “16S rRNA” and “NC_tRNA”) agree in a strong differentiation between Croatia and Spain, between Croatia and Portugal as well as between Italy and Portugal. Although AMOVA shows a strong differentiation between Italy and Spain in all data partitions, the differentiation between these two populations is supported only by the “16S rRNA” data set. According to the combined dataset, the single data partitions show no differentiation between Croatia and Italy in the analysis of molecular variance, but significant differences in the distribution of p-distances between these two populations (Table 11).

Table 11: Comparison of the significance tests (U-Test, AMOVA: non-parametric permutation) both under the null hypothesis of no differentiation for all pairs of populations and for all data partitions. av.F_{st}=average weighted variation among pairs of populations. * indicates significant differences on the 5% level, ** significant differences on the 1% level.

Population comparisons	**AMOVA**			**Mann Whitney U-Test**	
all data	av. F_{st}	p	differentiation	p	differentiation
Croatia – Italy	0.003	0.29	no	0.000	yes**
Croatia – Spain	0.34	0.003	very strong**	0.000	yes**
Croatia – Portugal	0.32	0.000	very strong**	0.000	yes**
Italy – Spain	0.21	0.14	strong	0.57	no
Italy – Portugal	0.24	0.003	strong**	0.000	yes**
Spain - Portugal	0.000	0.73	no	0.30	no
proteins					
Croatia – Italy	0.005	0.34	low	0.000	yes**
Croatia – Spain	0.32	0.004	very strong**	0.000	yes**
Croatia – Portugal	0.32	0.000	very strong**	0.000	yes**
Italy – Spain	0.18	0.19	strong	0.30	no
Italy – Portugal	0.25	0.005	strong**	0.000	yes**
Spain - Portugal	0.000	0.99	no	0.10	no
16S rRNA					
Croatia – Italy	0.000	0.68	no	0.000	yes**
Croatia – Spain	0.32	0.005	very strong**	0.000	yes**
Croatia – Portugal	0.24	0.000	strong**	0.000	yes**
Italy – Spain	0.28	0.05	very strong*	0.08	no
Italy – Portugal	0.25	0.000	very strong**	0.000	yes**
Spain - Portugal	0.04	0.28	low	0.23	no
NC_tRNA					
Croatia – Italy	0.006	0.33	no	0.000	yes**
Croatia – Spain	0.36	0.001	very strong**	0.000	yes**
Croatia – Portugal	0.33	0.000	very strong**	0.000	yes**
Italy – Spain	0.21	0.13	strong	0.37	no
Italy – Portugal	0.22	0.002	strong**	0.000	yes**
Spain - Portugal	0.000	0.47	no	0.34	no

Based on the F_{st} values obtained from the "all data" partition, the numbers of migrants per generation (Nm) are calculated for all pairs of populations. The lack of differentiation between Croatia and Italy results in a high number of migrants (Nm=332) like between Spain and Portugal, where the F_{st} value of 0 results in an infinite high number of migrants.

By contrast, only two migrants occur between Croatia and the Western Mediterranean (Spain, Portugal). Similar results arise in comparisons between Italy and the Western Mediterranean (Spain, Portugal) with four and three migrants per generation respectively.

To determine whether the genetic differences between pairs of populations are correlated with geographic distances, Slatkin distances (F_{st} /(1- F_{st}) are compared to the geographic distances in kilometres coastline (Table 12, Figure 19).

Table 12: Gene flow between populations in terms of number of migrants per generation [$Nm=(1/F_{st})-1$]. Slatkin distances result from the term $F_{st}/(1-F_{st})$.

Population comparisons	av. F_{st}	Nm	Slatkin distance	Geographic distance [km]
Croatia – Italy	0.003	332	0.003	1800
Croatia – Spain	0.34	2	0.515	3000
Croatia – Portugal	0.32	2	0.471	3600
Italy – Spain	0.21	4	0.266	1300
Italy – Portugal	0.24	3	0.316	1900
Spain - Portugal	0.000	inf.	0.000	600

Figure 19 shows the comparison of the pairwise Slatkin distances with the pairwise geographic distances along the coast. The correlation coefficient of 0.74 indicates a strong correlation between genetic differences and geographic distances, although the correlation is not supported by the Manteltest (p=0.08).

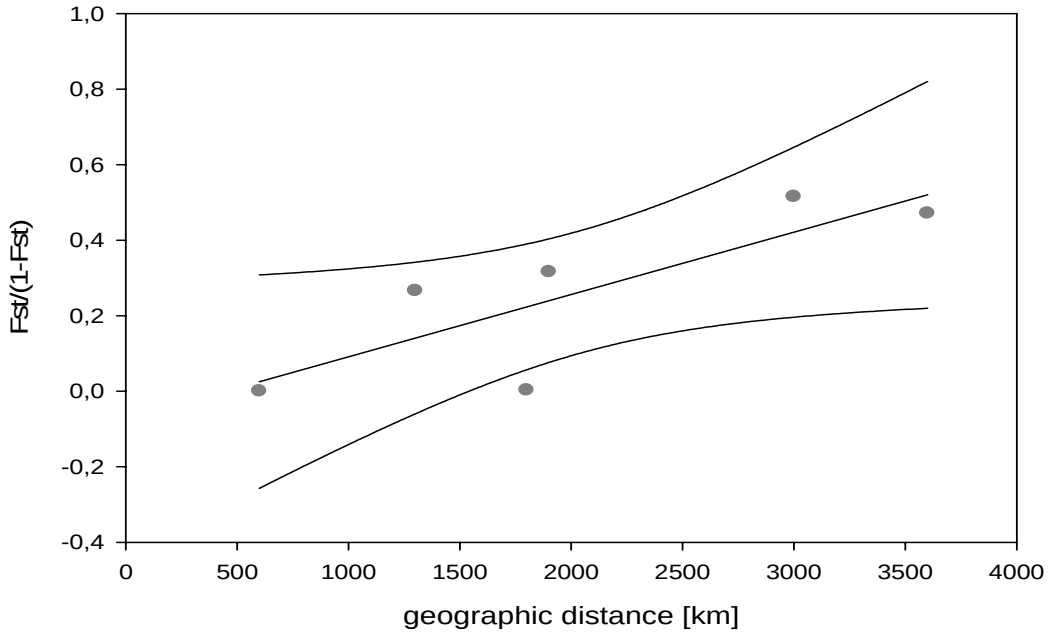

Figure 19: Linear regression line and 95% confidential intervals inferred from Slatkin distances plotted against geographic distance along the coast. Correlation coefficient r= 0.74. Significance: p=0.08.

3.3.2. Habitats

All 60 specimens obtained from boreholes come from the North Adriatic Sea (Croatia). Among these, 40 belong to the haplogroup A (66.6%). Haplogroup B and C are detected similar frequent, showing nine (15%) boring specimens in the former group and 11 (18.3%) in the latter respectively.

The 11 sponge dwelling specimens are obtained from *Geodia* sp. and all of them were sampled in Croatia. Most of these specimens belong to the haplogroup A (72.7%), whereas haplogroup B and C are represented by one (9.1%) or two (18.2%) specimens only.

The nestling specimens were obtained from kelp holdfasts or between *Mytilus* aggregations. All Italian, Spanish and Portuguese specimens were obtained from this habitat, whereas no Croatian specimen belongs to this habitat. Of all 26 nestling specimens, 20 belong to the haplogroup C (76.9%). The haplogroups A and B are represented by four (15.4%) and two (7.7%) specimens respectively (Figure 20).

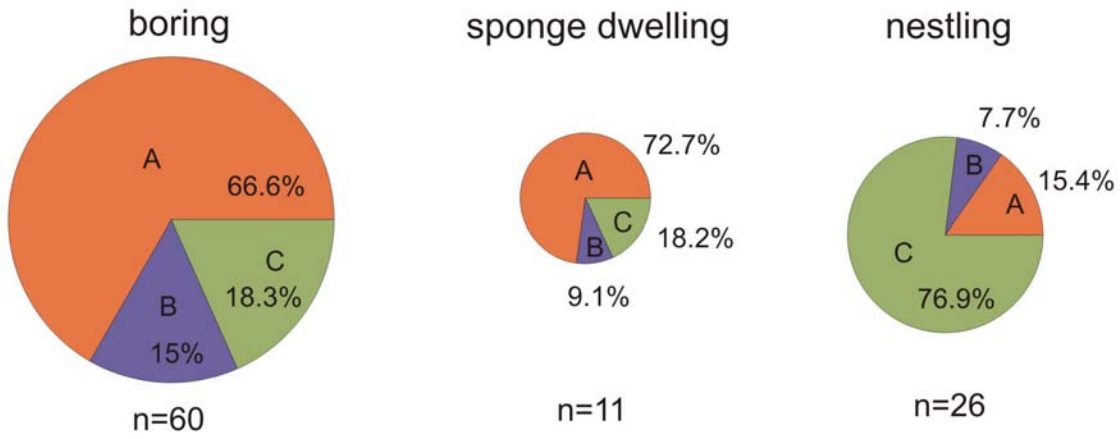

Figure 20: Haplogroup frequencies in the three habitats. The size of the circles is standardised to the relative numbers of specimens per habitat.

The haplogroup distribution in the three habitats is similar for the boring and the sponge dwelling specimens, whereas the nestling specimens mainly belong to the haplogroup C.

Analyses to estimate the genetic variation within and among habitats were performed with all data partitions. As the results are very similar in the different data sets, only the combined data partition is described in detail. The results of the single partitions “proteins”, “16S rRNA” and “NC_tRNA” are shown in Table 13.

The genetic variation is high in all the three habitats. The boring specimens show an average and median pairwise distance of 1.4% and reaches a maximum of 2.6% between

KR I 05 (Haplogroup A) and KR II 30 (Haplogroup C). The specimens obtained from the sponge *Geodia* sp. show an average genetic distance of 1.3%. The right skewed shape of the distance distribution leads to a lower median value (1%). The maximum p-distance of 2.3% occurs between KR III 09 (Haplogroup A) and KR III 12 (Haplogroup C). The nestling specimens display a similar distribution as the sponge dwelling samples. The average and the median pairwise distances are 1.2% and 0.9% respectively. The minimum p-distance of 0.1% occurs between SP II 02 and P II 29 (both haplogroup C), the maximum of 2.6%between SP III 10 (Haplogroup C) and IT III 04 (Haplogroup B) respectively (Figure 21).

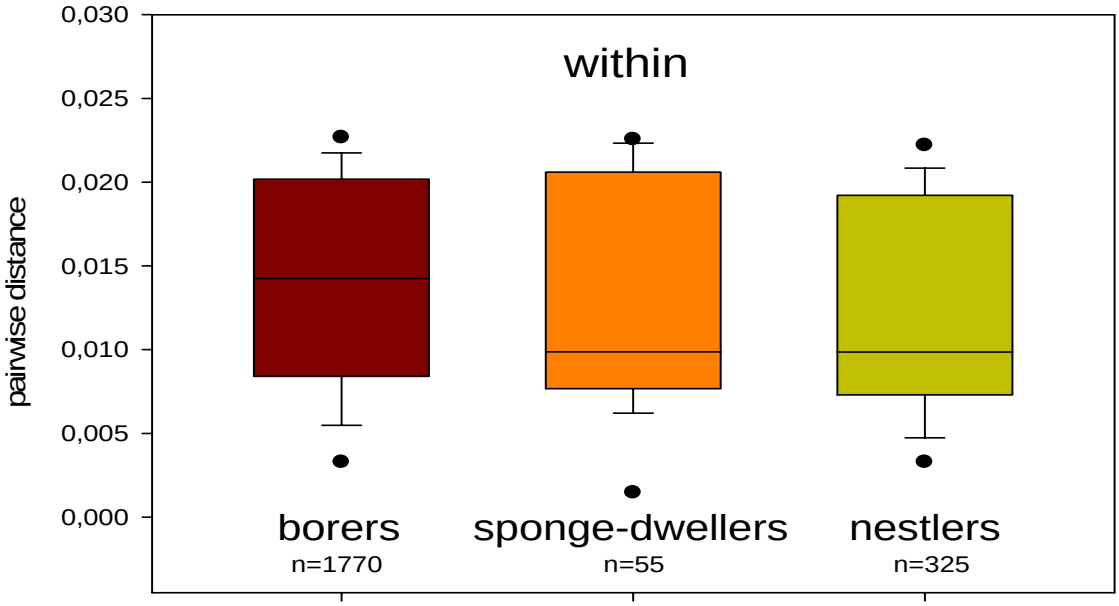

Figure 21: Pairwise genetic distances within the three habitats. n indicates the number of pairwise comparisons.

Within all habitats, the genetic distances span from 0.04% to 2.6%. The similar mean and median values (both 1.4%) result from a bimodal distance distribution. Among habitats, both the mean value of 1.6% and the median value of 1.9% are higher compared to within habitat comparisons. The p-distances range from 0.1 to 2.7%. Both, the bimodal distribution of the p-distances as well as the interval of the two peaks are similar to the within habitat comparisons (Figure 22).

However, the peak at 2% is significantly higher in the comparisons between habitats (U-Test: p=0.000). The AMOVA reveals a strong differentiation between within and among habitat comparisons (F_{st}=0.17, p=0.000).

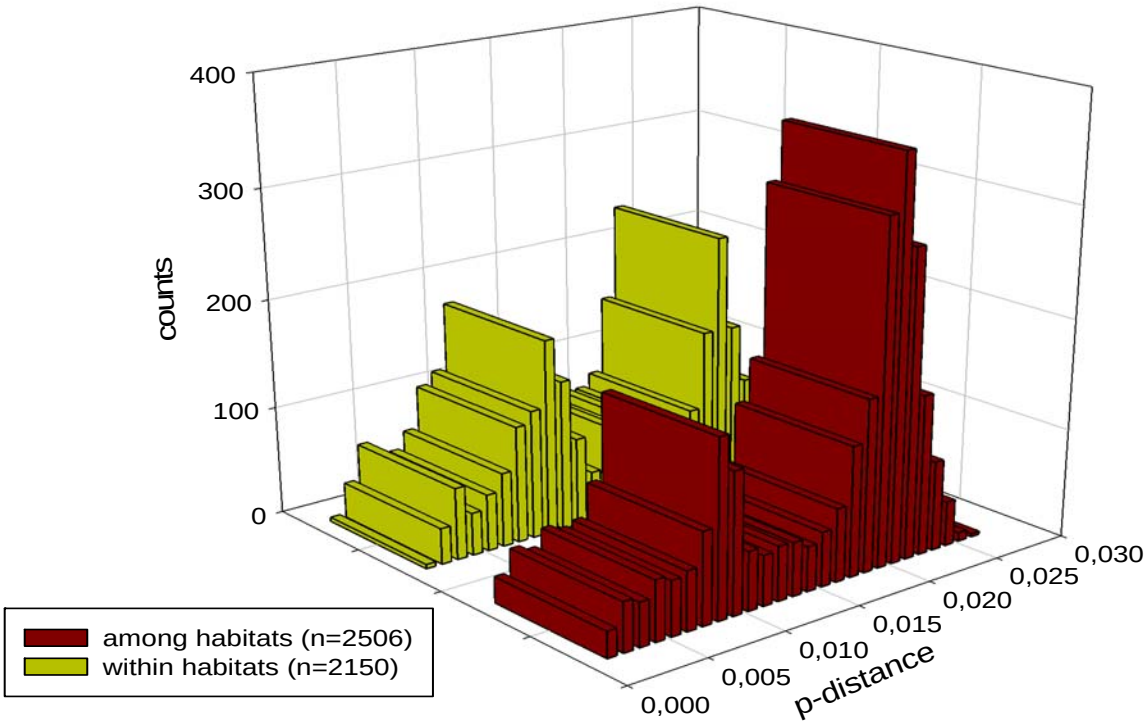

Figure 22: Genetic variation within and among habitats. n indicates the number of pairwise comparisons.

The pairwise habitat comparisons reveal no genetic differentiation between specimens living in boreholes and sponge-dwelling specimens, whereas specimens obtained from the nestling habitat are significantly different from specimens obtained from both, the boring and sponge-dwelling habitats. These results are supported in all data partitions ("all data", "proteins", "16S rRNA", "NC_tRNA") (Table 13).

Table 13: Comparison of the significance tests (U-Test, AMOVA: non-parametric permutation) both under the null hypothesis of no differentiation for all pairs of habitats and for all data partitions. av.F_{st}=average weighted variation among pairs of populations. * indicates significant differences on the 5% level, ** significant differences on the 1% level.

Habitat comparisons	**AMOVA**			**Mann Whitney U-Test**	
all data	av. F_{st}	p	differentiation	p	differentiation
boring – sponge dwelling	0.000	0.76	no	0.72	no
boring – nestling	0.18	0.000	strong**	0.000	yes**
sponge dwelling – nestling	0.23	0.000	strong**	0.000	yes**
proteins					
boring – sponge dwelling	0.003	0.56	low	0.72	no
boring – nestling	0.16	0.000	strong**	0.000	yes**
sponge dwelling – nestling	0.25	0.000	strong**	0.000	yes**
16S rRNA					
boring – sponge dwelling	0.002	0.57	low	0.64	no
boring – nestling	0.15	0.000	strong**	0.000	yes**
sponge dwelling – nestling	0.22	0.000	strong**	0.000	yes**
NC_tRNA					
boring – sponge dwelling	0.000	0.68	no	0.57	no
boring – nestling	0.18	0.000	strong**	0.000	yes**
sponge dwelling – nestling	0.20	0.000	strong**	0.000	yes**

Based on the F_{st} values obtained from the "all data" partition, the numbers of migrants per generation (Nm) are calculated for all pairs of habitats. The lack of differentiation between borers and sponge dwellers results in an infinite number of migrants. By contrast, only five migrants occur between borers and nestlers, three migrants between sponge dwellers and nestlers respectively (Table 14).

Table 14: Gene flow between habitats in terms of number of migrants per generation (Nm). $Nm=(1/F_{st})-1$.

Habitat comparisons	**av. F_{st}**	**Nm**
boring – sponge dwelling	0.000	inf.
boring – nestling	0.18	5
sponge dwelling – nestling	0.23	3

Based on the facts that all boring and sponge dwelling specimens are obtained from one population only (Croatia) and the nestling specimens are retrieved from the three other populations, no comparisons between genetic and geographic distances are possible.

4. Discussion

In this study, mitochondrial DNA sequence data are used to investigate whether *Hiatella* is one highly variable species or whether it is divided in two or more species in the Mediterranean Sea. For this purpose, specimens from geographically distinct parts of the Mediterranean Sea and Atlantic Ocean are collected and analysed. In addition, the populations are tested for genetic differentiation and whether this result correlates with the geographic distance. Finally, the samples are examined under the aspect of the habitat they occurred in, in order to test genetic differentiation among habitats.

4.1. Comparison of the different datasets

The protein coding genes show that in 9.2% positions substitutions occurred. As expected, the genetic variation in protein coding genes can be attributed almost exclusively to changes at the third codon position, whereas changes at the first and especially at the second position rarely occur. The 16S rRNA is less variable than the “proteins” with 8%. By contrast, the “NC_tRNA” dataset clearly shows with 31.9% the highest substitution rate. These results agree with the common opinion that non coding regions show higher genetic variation than coding regions (Aranishi and Okimoto 2005) and that protein coding genes show higher substitution rates than ribosomal genes (Saccone 1999).

This study reveals that the protein coding gene COIII starts 24 bp after the position hypothesised by Dreyer and Steiner (2006). The open reading frame determined in this study suggests that the COIII gene starts with the codon ATT that in mitochondrial genomes can be used as alternative initiation codon (http://www.bioinformatics.org/JaMBW/2/3/TranslationTables). The nucleotide C in the genbank sequence was probably counted wrongly (Figure 5) (Hermann Dreyer personal communication).

In this study, mitochondrial DNA sequence data seem appropriate to resolve the relationships at the population level. In the analysis of molecular variance (AMOVA), all data sets, single and combined, show similar results. By contrast, the genealogical analyses show a lower resolution in the 16S rRNA data set, which probably depends on the genetic variability of this molecular marker. In the present study therefore, especially the protein coding genes and non coding regions shape up as suitable genes at population level.

4.2. Species delineation

All analyses show that *Hiatella* is genetically variable, both under the aspect of populations as well as under the aspect of habitats. The maximum genetic difference among specimens of 2.7% occurs between Croatia (KR II 12, borer) and Spain (SP III 10, nestler) and is noticeably below the 5% threshold commonly used to distinguish between two species (Meyer and Paulay 2005). Both the network analyses and the Bayesian Inference suggest the occurrence of three distinct haplogroup lineages, referred to as A, B and C. However, the haplogroups are restricted neither to a population nor to a habitat, but all haplogroups are a mix of at least two populations and all habitats. This fact let suggest a remarkable level of gene flow between both, populations and habitats.

The genetic distances within and among populations (Figure 18) show a similar distribution. Therefore, low genetic differences occur not only within populations but also between specimens from geographically distinct parts of the Mediterranean. Furthermore, high genetic differences occur between both within and among populations. Due to the high overlap of the distance distributions, no genetic gap can be detected. Similar to the population comparisons, the comparisons of habitats (Figure 22) show that low genetic differences occur between specimens from the same habitat as well as between specimens from different habitats. In addition, high genetic differences occur not only between specimens from different habitats, but also between specimens from the same habitat. The high overlap in the distance distributions of within and among habitat comparisons therefore shows that a genetic gap does not occur. The lacking genetic gap regarding both, populations as well as habitats, suggest that all specimens belong to the same species of *Hiatella*.

The conclusion of the present study therefore is that in the Mediterranean Sea occurs only one species with a large geographic range and different modes of habit. Based on the results obtained in this study however, it is not possible to conclude whether these specimens belong to *Hiatella arctica* (L., 1767) or to *Hiatella striata* (Fleuriau De Bellevue, 1802). Since the first description of *Hiatella* returns to Linnean's *Mya arctica* (1767, p. 1113), in the following sections *Hiatella* "arctica" is used to refer to this Mediterranean species.

4.3. Genetic differentiation of populations

This investigation shows a very strong genetic differentiation among populations (*sensu* Wright, 1978). In pairwise comparisons between populations, this study clearly supports a panmictic state between Spain and Portugal. The two populations are genetically not distinct and therefore these specimens can be attributed to the same West-Mediterranean population.

The results regarding the genetic structure between Croatia and Italy are contradictive. Both populations show a high intra-population variation as visible from the distribution of haplotypes in the genealogical analyses (Figures 8 - 14). Therefore, no genetic structure is supported by the analysis of molecular variance (Table 11). The amount of genetic variation can be attributed almost exclusively to within population comparisons and consequently the genetic differentiation between the two populations is low (*sensu* Wright, 1978). By contrast, the Mann Whitney U-Test supports a genetic differentiation of these two populations (Table 11). Even if the distribution of the haplogroups is similar in these two populations, the significant differentiation in the Mann Whitney U-Test could result from the high genetic differences between specimens from these populations. Especially the Croatian and Italian specimens belonging to the haplogroups B and C show remarkable genetic differences as it results from the large numbers of intermediate haplotypes in the network analyses. Therefore, the distribution of the genetic distances between the two populations is significantly higher than within the populations. However, in order to resolve the genetic structure of these two populations additional samples from the Italian population are necessary.

The Croatian population shows a very strong genetic differentiation from both the Spanish and the Portuguese population. The AMOVA test of significance as well as the Mann Whitney U-Test support these results (Table 11).

Similar to the Croatian population, the Italian population shows a strong differentiation from the Portuguese population in both significance tests. Due to small samples sizes, the strong differentiation between Italy and Spain is significant neither in the AMOVA test nor in the Mann Whitney U-Test (Table 11). Therefore, additional specimens are necessary to find out the genetic structure of these two populations.

4.4. Correlation between genetic differentiation and geographic distance

Isolation by distance results from the spatially limited gene flow between populations. It was first described by Wright (1943) and is defined as the increase of genetic differentiation with increasing geographic distance. Comparing the populations under study, *Hiatella* "arctica" shows a strong correlation (*sensu* Zöfel, 1992) between genetic and geographic distances (Figure 19). Except for the comparison between Croatia and Italy, the genetic differentiation continuously increases with the increase of the geographic distance. Therefore, the strong genetic differentiation between the North- and the West-Mediterranean can be attributed to the spatially limited gene flow. However, the strong differentiation is not supported by the Manteltest ($p=0.08$) since the genetic differentiation between the Croatian and Italian population is too low as expected for the geographic distance. It is also worth mentioning that the geographic distances used in this study are only approximations and it cannot be excluded that in reality they could differ somewhat from the values used. As mentioned above, the genetic structure of the Croatian and Italian population is not completely resolved and more samples from the Italian population could give clearance whether the genetic differentiation significantly correlates with the geographic distance.

4.5. Genetic differentiation of habitats

This investigation reveals a strong genetic structure between within and among habitats (*sensu* Wright, 1978). It results from the limited gene flow between the nestling organisms and, the boring and the sponge dwelling specimens respectively (Table 13). The specimens from the boring and sponge dwelling habitats cannot be distinguished (Table 13). Since all these specimens belong to the Croatian population, it is not astonishing that unrestricted gene flow between these two habitats occurs. By contrast, the nestling specimens belong to the Italian, Spanish and Portuguese population. For that reason, it is likely that the strong genetic structure between the nestling and, the boring and the sponge dwelling habitat respectively, is the result of the spatially limited gene flow between these collecting sites. Therefore, it is necessary to sample all the three habitats in each population. However, since the habitat choice is the result of chance where the settlement of the larvae takes places (Russell-Hunter 1949) it is assumable that a better sampling could reveal no genetic differentiation of habitats.

4.6. Gene flow

A high genetic variation in a species can indicate high genetic exchange between populations (Halliburton 2004). The fact that *Hiatella* "arctica" colonises a large geographic area and different habitats let suggest that gene flow must occur between populations as well as between habitats. Several mechanisms can contribute to gene flow, and it is questionable whether it is possible to define a chief cause. The high intra-specific variation of *Hiatella* "arctica" and the high level of gene flow probably result from a sum of different factors: the dispersal abilities, the mode of habit, the habitat distribution, the sea currents and the anthropogenic influence.

4.6.1. Dispersal abilities

Hiatella "arctica" shows a variety of dispersal abilities. As mentioned in the introductive part of this study, the genus *Hiatella* is dioecious. The gametes, sperms and eggs, are released into the water, where the fertilisation takes place. The planktotrophic larvae undergo a long pelagic stage that can last up to six weeks (Lucas 1990). In this time, the distribution of the larvae occurs mainly by passive transport through surface currents (Scheltema 1992). After settlement, the young specimens undergo a motile phase in search for suitable substrates (Russell-Hunter 1949). These migrations give *Hiatella* "arctica" an additional possibility for dispersal by "byssus drifting". Furthermore, the possibility to live as facultative epibiont on vagil fauna (Isaeva et al. 2001) is a supplementary way of dispersal. Finally, the nestling adults may disperse by rafting on debris or kelp. Helmuth et al. (1994) showed that bivalves could disperse up to 2000 km by kelp rafting. This method of dispersal may play a role also in the dispersal of *Hiatella* "arctica", especially at large geographic scale. Gordillo (2001) suggests that nestling *Hiatella* could have invaded the Atlantic Ocean from the North Pacific in the Pliocene and Pleistocene by rafting on *Zostera* and *Laminaria* respectively.

4.6.2. Mode of habit

Hiatella "arctica" has a great variability in the mode of habits. The organisms occur as nestler, sponge dweller as well as borer. As nestling organism, *Hiatella* "arctica" attaches to different organisms (e.g. *Mytilus*, algae). The boring and the sponge dwelling behaviour provide a chance to colonise substrates unsuitable for nestling organisms. This study

agrees with Russell-Hunter (1949) that the mode of life depends exclusively on the substrate where the settlement takes place. This ecological plasticity gives *Hiatella* "arctica" the opportunity to live in different habitats and to expand by using different habitats.

4.6.3. Habitat distribution in the Mediterranean Sea

The habitat distribution can influence the dispersal capability of a species. A continuous habitat can contribute to gene flow whereas barriers like habitat fragmentation can counteract gene flow (Arnaud et al. 2000).

Hiatella is known to occur within sponges of the genera *Spongia*, *Ircinia* (Micali and Solustri 2004), *Geodia* (this study) and *Cacospongia* (Riedl 1983). They show a wide distribution in the Mediterranean Sea (Hofrichter 2004). Although the sponge dwelling habitat was investigated for Croatia only, it is assumable that it occurs in all populations under study.

The boring behaviour of *Hiatella* "arctica" depends on the occurrence of suitable substrates for primary bioeroders. Especially limestone is a good substrate for chemically boring organisms like *Lithophaga* and *Gastrochaena*. The Mediterranean coast is very heterogenic from a geomorphologic point of view (Cavazza et al. 2004), but carbonates are widely distributed on the Mediterranean coast. The Messinian salinity crisis at the end of the Miocene (6-5 Ma) led to a mass mortality of organisms in the Mediterranean and enabled the deposition of high amounts of calcareous shells in a short period of time and to the formation of carbonates. Therefore, this boring habitat is common in the Mediterranean (Hofrichter 2002), although calcareous rocks were not found in the collecting sites of the Ligurian coast (Stefano Schiapparelli personal communication). In addition, the Portuguese coast, especially the coast of Algarve, is mainly characterised by carbonates (Forth et al. 1999). In conclusion, the boring habitat is common in the Mediterranean as well as in the South of Portugal.

As nestling organism, *Hiatella* "arctica" can be found throughout the Mediterranean Sea. It lives by means of byssus threads epiphytically (macro-algae, sea grass) as well as epizoically (*Mytilus*, barnacles, crustaceans etc.). Since these habitats are common in the Mediterranean, it is assumable that the nestling habit is possible throughout the Mediterranean Sea.

To summarise, the habitat where *Hiatella* "arctica" occurs can be supposed to be continuous in the Mediterranean Sea. The habitat distribution seems not to be a barrier to

gene flow and therefore it can be guessed that gene flow partly can be attributed to “gene hopping”, i.e. via intermediate sites.

4.6.4. Sea currents

Sea currents play a central role in the dispersal of marine organisms (Scheltema 1992). The currents in the Mediterranean vary in both, temperature and salinity. The Mediterranean Sea is connected to the Atlantic Ocean via the Strait of Gibraltar. Through this 14.5 km wide and 300 meters deep strait, cold water with low salt concentration flows continuously as surface current into the Mediterranean Sea. The main surface sea currents follow a cyclone. This means, that the surface water entering the Strait of Gibraltar flows mainly on the African coast eastwards. On the Strait of Sicily, the surface current flows partly along the West Italian coast northwards and along the Ligurian coast to the southeast coast of Spain (Hofrichter 2002). An other part of the surface water flows eastwards into the East-Mediterranean basin. The water of the East Mediterranean generally is warmer and, due to evaporation, saltier than the Atlantic water. The Adriatic is a mix of both, surface water from the Atlantic Ocean (Ott 1996) and intermediate Levantine water (Robinson et al. 2001). Finally, water exchange between the Adriatic and the Tyrrhenian occurs through the Strait of Messina (Leier 2001).

In conclusion, it seems that the surface currents in the Mediterranean Sea do not limit the dispersal of larvae. Furthermore, *Hiatella* “arctica” is euryoecious and therefore tolerant in both, water salinity and water temperature. Therefore, it seems that no physic barriers inhibit gene flow in *Hiatella* “arctica” in the Mediterranean Sea.

4.6.5. Anthropogenic influence

The anthropogenic influence in the dispersal of organisms cannot be neglected. The transport of larvae in ballast water or the transport of juveniles and adults attached by means of their byssus on the hulls of ships could be an additional possibility of dispersal for *Hiatella* “arctica”. Finally, the creation of docks could be an ulterior opportunity to colonise areas otherwise unsuitable as habitats.

However, on global scale human impact is probably only a small contribution to the distribution of *Hiatella* because fossil record shows an almost worldwide distribution of the genus *Hiatella* since at least the Pleistocene (2 Ma).

4.7. Summary and conclusions

This molecular study shows that in the Mediterranean Sea only *Hiatella* "arctica" occurs. The maximum genetic difference of 2.7% and the missing genetic gap between intra-specific variation and inter-specific divergence show that all specimens under study are conspecific. *Hiatella* "arctica" has the ability to colonise geographically distinct parts as well as to live in different ecological environments. This capacity probably results from a sum of different factors: the dispersal abilities, the mode of habit, the habitat distribution, the sea currents and the anthropogenic influence.

The genetic differentiation between populations is strongly correlated with the geographic distance. The genetic differentiation between habitats is supposed to be an artefact resulting from the geographic distances between the collecting sites. There seems not to be physic barriers, like surface currents, temperature and salinity, that limit gene flow between the populations. Furthermore, the habitat distribution seems not to be a barrier to gene flow and therefore it can be guessed that gene flow partly can be attributed to "gene hopping", i.e. via intermediate sites.

Hiatella "arctica" is a typical "generalist". The organisms are euryoecious (Ali 1970, Gordillo 2001), habitat generalists (Russell-Hunter 1949) and show high dispersal abilities (Gordillo 2001). This ecological plasticity enables this species to live under various conditions and to expand over a large geographic scale.

4.8. Future prospective

Due to the heterogeneity of the samples sizes, several questions arise. It is not resolved, whether the Croatian and the Italian populations are in a panmictic state or genetically different. Furthermore, it is necessary to figure out whether the West-Mediterranean is genetically less variable compared to the North-Mediterranean and whether the haplogroups A and B, which are mainly restricted to the North-Mediterranean, could be an ecological adaptation to the colder North-Mediterranean waters. Finally, it could be very useful to include specimens from other parts of the North Atlantic Ocean to verify the ancestral haplotype.

5. References

Adamkewicz S.L., Harasewych M.G., Blake J., Saudek D. and Bult C.J., 1997. A Molecular Phylogeny of the Bivalve Mollusks. *Molecular Biology and Evolution*, 14: 619-629.

Ali R.M., 1970. The influence of suspension density and temperature on the filtration rate of *Hiatella arctica*. *Marine Biology*, 6: 291-302.

Aranishi F. and Okimoto T., 2005. Sequence polymorphism in a novel noncoding region of Pacific oyster mitochondrial DNA. *Journal of Applied Genetics*, 46: 201-206.

Arnaud S., Monteforte M., Galtier N., Bonhomme F. and Blanc F., 2000. Population structure and genetic variability of pearl oyster Pinctada mazatlanica along Pacific coasts from Mexico to Panama. *Conservation Genetics*, 1: 299-307.

Atkins D., 1937. On the ciliary mechanisms and interrelationships of lamellibranches. III. Types of lamellibranch gills and their food currents**.** *Quarterly Journal of Microscopical Science (N.S.)*, 79: 375-421.

Barrett J. and Yonge C.M., 1958. In: *Collins pocket guide to the sea shore*, Collins, London.

Beu A.G., 1971. New light on the variation and taxonomy of the Bivalve *Hiatella*. *New Zealand Journal of Geology and Geophysics*, 14: 64-66.

Booth J.D., 1983. Studies on twelfe common bivalve larvae, with notes on bivalve spawning seasons in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 17: 231-265.

Bosc L.A.G., 1801. *Histoire naturelle des coquilles*, Vol. 3, Paris.

Bucquoy E., Dautzenberg P. and Dollfus G., 1896. *Les mollusques marins du Roussillon*, Vol. 2, Paris.

Carriker M.R., 1990. Functional significance of the pediveliger in bivalve development. In: Morton B. (ed.) *The Bivalvia – Proceedings of a Memorial Symposium in Honour of Sir C.M. Yonge*, Hong Kong University Press, pp: 267-282.

Cavazza W., Roure F., Spakman W. and Stampfli G.M., 2004. In: *The TRANSMED Atlas. The Mediterranean region from crust to mantle: Geological and geophysical framework*, Springer, Berlin.

Clement M., Posada D. and Crandall K.A., 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, 9: 1657-1659.

Cope J.C.W., 1996. The early evolution of the Bivalvia. In: Taylor J. (ed.) *Origin and evolutionary radiation of the Mollusca*, Oxford University Press, USA, pp: 361-370.

Coan E.V., Valentich-Scott P., Bernard F.R. and Sadeghian P.S., 2000. *Bivalve Seashells of Western North America; Marine Bivalve Mollusks from Arctic Alaska to Baja California*. Santa Barbara Museum of Natural History, Santa Barbara, California, pp: 483-487.

Corel Corporation 2000: Corel Draw version 10.

Cragg S.M., 1996. The phylogenetic significance of some anatomical features of bivalve veliger larvae. In: Taylor J. (ed.) *Origin and evolutionary radiation of the Mollusca*, Oxford University Press, USA, pp. 371-380.

Daudin F.M., 1800. *Receuil de Mémoires et de notes sur les espèces inédites ou peu connues de Mollusques, de Vers, et de Zoophytes*, Paris.

Dodge H., 1950. *Hiatella* Daudin versus *Saxicava* Bellevue. *The Nautilus*, 63(4): 138-139; 64(1): 29-33.

Dreyer H., Steiner G. and Harper E.M., 2003. Molecular phylogeny of Anomalodesmata (Mollusca: Bivalvia) inferred from 18S rRNA sequences. *Zoological Journal of the Linnean Society*, 139: 229-246.

Dreyer H. and Steiner G., 2006. The complete sequences and gene organization of the mitochondrial genomes of the heterodont bivalves *Acanthocardia tuberculata* and *Hiatella arctica* – and the first record of a putative Atpase subunit 8 gene in marine bivalves. *Frontiers in Zoology*, 3 (13): 1-14.

Excoffier L., Smouse P.E. and Quattro J.M., 1992. Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131: 479-491.

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.

Finch TV v.1.4: Geospiza Inc. (http://www.geospiza.com/finchtv/)

Fleuriau De Bellevue C., 1802. *Bulletin des Sciences, par la Societe Philomatique*, 62: 105-109.

Forbes E. and Hanley S., 1853. In: J. Vaan Voorst Paternoster Row (ed.), *A History of British Mollusca and their shells*, Vol. 4, London, pp: 136-149.

Forth R.A., Butcher D. and Senior R., 1999. Hazard mapping of karst along the coast of the Algarve, Portugal. *Engineering Geology*, 52(1): 67-74.

Giribet G. and Wheeler W., 2002. On bivalve phylogeny: a high-level analysis of the Bivalvia (Mollusca) based on combined morphology and DNA sequence data. *Invertebrate Biology*, 121: 271-324.

Giribet G. and Distel D.L., 2003. In: Lydeard C. and Lindberg D. (eds.) *Molecular Systematics and Phylogeography of Molluscs*, Smithsonian Institute Press, pp: 45-90.

Google Earth: www.earth.google.com.

Gordillo S., 2001. Puzzling distribution of the fossil and living genus *Hiatella* (Bivalvia). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 165: 231-249.

Guo W.S. and Thompson E.A., 1992. Performing the Exact Test of Hardy-Weinberg Proportion for Multiple Alleles. *Biometrics*, 48: 361-372.

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

Halliburton R., 2004. In: *Introduction to population genetics*, Prentice-Hall, United States Ed, pp: 4-5; pp: 51.

Hare M.P. and Weinberg J.R., 2005. Phylogeography of surfclams, *Spisula solidissima*, in the western North Atlantic based on mitochondrial and nuclear DNA sequences. *Marine Biology*, 146: 707-716.

Harper E.M., Dreyer H. and Steiner G., 2006. Reconstructing the Anomalodesmata (Mollusca: Bivalvia): morphology and molecules. *Zoological Journal of the Linnean Society*, 148: 395-420.

Healy J.M., 1996. Molluscan sperm ultrastructure: correlation with taxonomic units within the Gastropoda, Cephalopoda and Bivalvia. In: Taylor J. (ed.) *Origin and evolutionary radiation of the Mollusca*, Oxford University Press, USA, pp: 99-113.

Hebert P.D.N., Cywinska A., Ball S.L. and deWaard J.R., 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B*, 270: 313-321.

Helmuth B., Veit R.R. and Holberton R., 1994. Long distance dispersal of subantarctic brooding bivalve (*Gaimardia trapesina*) by kelp rafting. *Marine Biology*, 120: 421-426.

Hofrichter R., 2002. In: Hofrichter R. (ed.) *Das Mittelmeer: Fauna, Flora, Ökologie*. Band:I *Allgemeiner Teil*. Spektrum Akademischer Verlag, Heidelberg, Berlin, pp: 88; pp: 268.

Hofrichter R. 2004. *Das Mittelmeer: Fauna, Flora, Ökologie*. Band:II/1 *Bestimmungsführer* (ed. Hofrichter). Spektrum Akademischer Verlag, Heidelberg; Berlin.

Huang X. and Madan A., 1999. CAP3: A DNA sequence assembly program. *Genome Research*, 9: 868-877. (http://pbil.univ-lyon1.fr/cap3.php)

Huelsenbeck J.P. and Ronquist F., 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics*, 17: 754-755.

Integrated DNA technologies (IDT). Innovation and precision in Nucleic Acid Synthesis (http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer)

Isaeva V., Kulikova V. and Kasyanov V., 2001. Bivalve molluscs, *Mytilus trossulus* and *Hiatella arctica*, as facultative epibionts of the crab, *Hemigrapsus sanguineus*, infested by the rhizocephalan, *Sacculina polygenea*. *Journal of Marine Biology Association of UK*, 81: 891-892.

Jeffreys J.G., 1865. In: J. Vaan Voorst Paternoster Row (ed.), *British Conchology – The Mollusca*, Vol. 3, London, pp: 77-85.

Kelly S.R.A., 1980. *Hiatella* – A Jurassic bivalve squatter? *Palaeontology*, 23: 769-781.

Khalaman V.V., 2005. Testing the hypothesis of tolerance strategies in *Hiatella arctica* L. (Mollusca: Bivalvia). *Helgoland Marine Research*, 59: 187-195.

Kilias R., 1982. *Lehrbuch der speziellen Zoologie*, Band 1, 3. Teil. Gustav Fischer Verlag, Stuttgart, pp: 152-155.

Lamarck J-B., 1818. In: *Histoire naturelle des animaux sans vertebres, Vol. 6.*

Lebour M.V., 1938. Notes on the breeding of some lamellibranches from Plymouth and their larvae. *Journal of Marine Biology Association of UK*, 23: 119-144.

Leier M., 2001. In: *Weltatlas der Ozeane*, Frederking & Thaler.

Linné C., 1767. In: *Systema naturae, Tom. I Pars II, Editio Duodecima Reformata*, Holmiae.

Lowe T.M. and Eddy S.R., 1997. "tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence" *Nucleic Acids Research*, 25: 955-964. (http://lowelab.ucsc.edu/tRNAscan-SE).

Lucas A., 1990. Feeding and digestion in bivalve larvae. In: Morton B. (ed.) *The Bivalvia – Proceedings of a Memorial Symposium in Honour of Sir C.M. Yonge*, Hong Kong University Press, pp: 177-193.

Metropolis N., Rosenbluth A.E., Rosenbluth M.N., Teller A.H. and Teller E., 1953. Perspective on “Equation of state calculations by fast computing machines”. *Journal of Chemical Physics*, 21:1087-1092.

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

Micali P. and Solustri C., 2004. Osservazioni su *Hiatella rugosa* (Linné, 1767) (Bivalvia Hiatellidae) endobionte di poriferi e differenze conchigliari con *Hiatella arctica* (Linné, 1767). *Bollettino Malacologico*, 40: 49-55.

Morton B., 1996. The evolutionary history of the Bivalvia. In: Taylor J. (ed.) *Origin and evolutionary radiation of the Mollusca*, Oxford University Press, USA; pp. 337-359.

Morton J.E. 1979. In: *Molluscs*, Hutchinson & Co., pp: 26-27.

Narchi W., 1973. On the functional morphology of *Hiatella solida* (Hiatellidae:Bivalvia). *Marine Biology*, 19: 332-337.

NCBI: The Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi)

NCBI: Homepage (http://www.ncbi.nlm.nih.gov/)

NCBI: Open Reading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf/html)

NCBI: The Genetic Codes (http://www.bioinformatics.org/JaMBW/2/3/TranslationTables)

Neevia technology. Document converter express. (http://convert.neevia.com).

Newell N.D., 1969. Classification of Bivalvia. In: Moore R.C. (ed.) *Treatise on invertebrate palaeontology: Mollusca 6: Mollusca Part N*, Geological Society of America, pp: 205-224.

Nicholas K.B. and Nicholas H.B.Jr., 1997. GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the author (www.cris.com/ketchup/genedoc.shtml).

Odhner N. Hj., 1914. Notizen über die Fauna der Adria bei Rovigno. Beiträge zur Kenntnis der marinen Molluskenfauna von Rovigno in Istrien. *Zoologischer Anzeiger*, 19 (4): 156-170.

Ott, J., 1996. *Meereskunde*. Zweite Auflage. Ulmer Verlag. Stuttgart, pp: 69.

Page R.D.M., 1996. Treeview: An application to display phylogenetic trees on personal computers. *Computer Applications in the Bioscience*, 12: 357-358.

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

Posada D. and Crandall K.A., 2001. Intraspecific gene genealogies: Trees grafting into networks. *Trends in Ecology and Evolution*, 16: 37-45.

Purchon R.D., 1958. The stomach in the Eulamellibranchia; stomach type IV. *Proceedings of the Zoological Society of London*, 131: 487-525.

Purchon R.D., 1987a. The stomach in the Bivalvia. *Philosophical Transactions of the Royal Society of London, Series B*, 316: 183-276.

Purchon R.D., 1990. Stomach structure, classification and evolution of the Bivalvia. In: Morton B. (ed.) *The Bivalvia – Proceedings of a Memorial Symposium in Honour of Sir C.M. Yonge*, Hong Kong University Press, pp: 73-82.

Raes J. and Van de Peer Y., 1999. ForCon: a software tool for the conversion of sequence alignments. *Embnet News*, 6:10-12.

Riedl R., 1983. *Fauna und Flora des Mittelmeers*. Verlag Paul Parey, Hamburg & Berlin, pp: 374.

Robinson A.R., Leslie W.G., Theocharis A. and Lascaratos A., 2001. Mediterranean Sea Circulations. *Academic Press*, 1:19.

Rousset F., 1997. Genetic differentiation and estimation of gene flow from F-Statistics under Isolation by Distance. *Genetics*, 145: 1219-1228.

Russell-Hunter W., 1949. The structure and behaviour of *Hiatella gallicana* (Lamarck) and *Hiatella arctica* (L.), with special reference to the boring habitat. *Proceedings of the Royal Society of Edinburgh*, 3: 271-289.

Saccone C., De Giorgi C., Gissi C., Pesole G. and Reyes A., 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as model system. *Gene*, 238: 195-209.

Saavedra C. and Pena J.B., 2006. Phylogenetics of American scallops (Bivalvia: Pectinidae) based on partial 16S and 12S ribosomal RNA gene sequences. *Marine Biology*, 150: 111-119.

Salvini-Plawen L. and Steiner G., 1996. Synapomorphies and plesiomorphies in higher classification of Mollusca. In: Taylor J. (ed.) *Origin and evolutionary radiation of the Mollusca*, Oxford University Press, USA, pp: 29-51.

Scheltema R.S., 1992. Passive dispersal of planktonic larvae and the biogeography of tropical sublittoral invertebrate species. In: *Marine Eutrophication and Population Dynamics*, 25th European Marine Biology Symposium, Olsen & Olsen, University of Ferrara, pp:195-202.

Sejr M.K., Sand M.K., Jensen K.T., Petersen J.K., Christensen P.B. and Rysgaard S., 2002. Growth and production of *Hiatella arctica* (Bivalvia) in a high-Arctic fjord (Young Sound, Northeast Greenland). *Marine Ecology Progress Series*, 244: 163-169.

Sigmaplot v.10 (Systat Software, Inc. 2006).

Smouse P.E., Long J.C. and Sokal R.R., 1986. Multiple regression and correlation extensions of the Mantel Test of matrix correspondence. *Systematic Zoology*, 35(4): 627-632.

SPSS v.7.5 (SPSS Inc., 1997)

Steiner G. and Hammer S., 2000. Molecular phylogeny of Bivalvia (Mollusca) inferred from, 18S rRNA sequences with particular reference to the Pteriomorphia. In: Harper E.M., Taylor J.D., Crame J.A. (eds.) *The evolutionary biology of the Bivalvia.* Geological Society of London Special Publication, 77: 11-29.

Stepien C.A., Morton B., Dabrowska K.A., Guarnera R.A., Radja T. and Radja B., 2001. Genetic diversity and evolutionary relationships of the troglodytic 'living fossil' *Congeria kusceri* (Bivalvia: Dreissenidae). *Marine Ecology*, 10: 1873-1879.

Strauch F., 1968. Determination of Cenozoic Sea Temperatures using *Hiatella arctica* (Linné). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 5: 213-233.

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

Tamura K., Dudley J., Nei M. and Kumar S., 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24: 1596-1599.

Taylor J.D., Williams S.T., Glover E.A. and Dyal P., 2007. A molecular phylogeny of heterodont bivalves (Mollusca: Bivalvia: Heterodonta): new analyses of 18S and 28S rRNA genes. *Zoologica Scripta*, 36(6): 587-606.

Terranova M.S., Lo Brutto S., Arculeo M. and Mitton J.B., 2006. Population structure of *Brachidontes pharaonis* (P. Fisher, 1908) (Bivalvia, Mytilidae) in the Mediterranean Sea, and evolution of a noval mtDNA polymorphism. *Marine Biology*, 150: 89-101.

Thompson J.D., Gibson T.J., Plewniak F., Jeanmougin F. and Higgins D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 24:4876-4882.

Trudgill S.T. and Crabtree R.W., 1987. Bioerosion of intertidal limestone, CO. Clare, Eire – 2: *Hiatella arctica*. *Marine Geology*, 74:99-109.

Turner T.F., Trexler J.C., Harris J.L. and Haynes J.L., 2000. Nested Cladistic Analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. *Genetics*, 154: 777-785.

Winckworth R., 1932. The British Marine Mollusca. *Journal of Conchology*, 19:211-252.

Wright S., 1943. Isolation by distance. *Genetics*, 28: 114-138.

Wright S., 1951. The genetic structure of populations. *Annales of Eugenics*, 15: 323-354.

Wright S., 1978. In: *Evolution and the genetics of populations, Vol. 4: Variability within and among natural populations*, University of Chicago Press.

Yonge C.M., 1949. On the structure and adaptations of the Tellinacea, deposit-feeding Eulamellibranchia. *Philosophical Transactions of the Royal Society of London, Series B*, 234: 29-76.

Yonge C.M., 1951. Observations on *Sphenia binghami*. *Journal of the Marine Biological Association of the United Kingdom*, 30 (2): 387-391.

Yonge C.M., 1971. On functional morphology and adaptive radiation in the bivalve superfamily Saxicavacea (*Hiatella* (=*Saxicava*), *Saxicavella*, *Panomya*, *Panope*, *Cyrtodaria*). *Malacologia*, 11:1-44.

Zöfel P., 1992. *Statistik in der Praxis*, Gustav Fischer Verlag, Stuttgart, Jena, p: 211.

Acknowledgements

I want to thank my family for the attention, advice and emotional support during my life and especially in this work. Without your allowance, my education would not have been possible. I thank my niece Anna for her vitality and her curiosity in life. I am grateful for taking part in your life.

Thanks to Dr. Gerhard Steiner for the support, suggestions and patience during this work. You aided me joining the pieces of a *Hiatella* - puzzle together to an overall picture.

I am grateful for the scientific assist and cordiality of Dr. Carmen Salas, Dr. Serge Gofas and Dr. Jose Rueda (University of Malaga) and Dr. Stefano Schiapparelli (University of Genoa) during my collecting trips in Spain and Italy.

Thanks to my working group (EF - Molecular Phylogeny – University of Vienna) and the members of the central lab, especially to: Dr. Emiliano Dell' Ampio for lively discussions during many coffee breaks; Dr. Hermann Dreyer for inducting me into lab techniques and for sparking interest in molecular research; Dr. Johannes Rath for sequencing my samples.

I appreciate the help of both, Dr. Judith Notdurfter for lingual corrections and Dr. Daniela Oberlechner for formatting my diploma thesis.

Special thanks to Christian for the support and patience during this time. Thanks for standing by my side.

Thanks to Michi Werth. You were, are and will always be an important person in my life. We have been a good team during our diploma work and I am grateful for the time we spent together. Without you, the last years would have been only half as funny.

Thanks to my friends Ingrid, Gloria, Marika, Valli, Christine, Lisi and Tschicki.

Curriculum vitae

Miriam Oberlechner
Geboren am 04.11.1978 in Bozen, Italien

1984 – 1992	Besuch der Volks – und Mittelschule in Bozen
1992 – 1997	Besuch des deutschsprachigen Wissenschaftlichen Lyzeum „Raimund von Klebelsberg“ in Bozen
Seit 1997	Studium der Biologie, Studienzweig Zoologie an der Universität Wien, Österreich