



DEPARTMENT OF BIOLOGICAL AND
ENVIRONMENTAL SCIENCES

CRYPTIC SPECIATION IN THE PADDLE WORMS

Nereiphylla rubiginosa & *Pterocirrus macroceros*



Felicia Ulltin

Degree project for Bachelor of Science with a major in Biology

BIO603 - Biologi: Examensarbete – Kandidatexamen, 30 hp

First cycle

Semester/year: Spring 2019

Supervisor: Arne Nygren, Department of Marine Sciences

Examiner: Urban Olsson, Department of Biological and Environmental Sciences

Front page photos: *Pterocirrus macroceros* & *Nereiphylla rubiginosa* taken by me, Felicia Ulltin.

Table of Contents

Abstract	3
Sammanfattning	3
Introduction	4
Definition of species	4
Polychaetes	4
Studied species	5
Molecular markers.....	5
Separation of clades	6
Purpose & hypotheses.....	6
Materials & methods	7
Sampling & collection	7
Amplification & sequencing.....	8
Analyses.....	8
Haplotype networks	8
Phylogenetic inference	8
Results	10
Sequencing success.....	10
<i>Nereiphylla rubiginosa</i>	11
Uncorrected pairwise distances for <i>N. rubiginosa</i>	14
<i>Pterocirrus macroceros</i>	18
Uncorrected pairwise distances for <i>P. macroceros</i>	21
Discussion	23
<i>Nereiphylla rubiginosa</i>	23
<i>Pterocirrus macroceros</i>	24
Conclusions.....	25
Acknowledgements.....	25
References.....	26
Appendix A.....	28
Appendix B.....	29
Appendix C.....	30
Appendix D	32

Abstract

Since the dawn of systematics, species have been separated according to their morphological characters. As the use of molecular methods has advanced, the separation of species has grown even more complex. Individuals that were previously considered to be the same species based on morphology, turn out to be different species when studying their genetic code - so called cryptic species. In this study the two polychaetes *Nereiphylla rubiginosa* and *Pterocirrus macroceros* were analysed separately to detect potential presence of cryptic species. Samples were collected from several locations in the north-eastern Atlantic and the Mediterranean. Haplotype networks were made in the program TCS, where several genetic clusters were detected. Further analysis of the mitochondrial marker COI and the nuclear markers ITS1, ITS2 and 28S were conducted in MrBayes. Congruence between the phylogenetic trees and TCS showed that the taxa *Nereiphylla rubiginosa* consists of four species and *Pterocirrus macroceros* consists of three species.

Keywords: *Phylogenetic inference, Phyllodocidae, North Eastern Atlantic, Mediterranean*

Sammanfattning

Ända sedan Linnés tid var skillnader i morfologi det huvudsakliga sättet att skilja olika arter ifrån varandra. I takt med att molekylära metoder för artavgränsning blivit standard, har den sedan tidigare invecklade indelningen av arter vuxit sig alltmer komplex. Linjer som ansetts morfologiskt identiska och tidigare förts till samma art kan visa sig höra till olika arter när deras DNA analyserats. Dessa identiska linjer kallas kryptiska arter. I den här studien har de två havsborstmaskarna *Nereiphylla rubiginosa* och *Pterocirrus macroceros* undersökts med syftet att se om respektive art består av en eller flera arter. Individerna samlades in från ett antal lokaler i nordöstra Atlanten och Medelhavet. De genetiska markörerna COI, ITS1, ITS2 och 28S amplifierades och sekvenserades från så många individer som möjligt, och analyserades sedan i programmen TCS och MrBayes. Kongruens mellan COI och ITS1, ITS2 och 28S i både haplotypnätverk och fylogenetiska träd visar att *N. rubiginosa* och *P. macroceros* består av fyra respektive tre arter vardera.

Nyckelord: *Fylogenetisk inferens, Phyllodocidae, Nordöstra Atlanten, Medelhavet*

Introduction

Definition of species

The importance of distinguishing different species may not be obvious. But the concept of species is an elemental unit in biology; it is important not just in the field of systematics but also crucial to other areas of biology (De Queiroz 2007). There have been large controversies throughout the years regarding how to define species. In this study the unified species concept is applied; where the only requirement for a species is that it is a separately evolving lineage. The method of inferring species, whether it is a phylogenetic, morphological or ecological approach, is used as lines evidence for the existence of the lineage (De Queiroz 2007).

Cryptic species

Cryptic species are morphologically very similar species that are genetically distinct. Even though they look the same, they may not have interbred for millions of years. It has been suggested that a great part of the cryptic diversity is contained in the ocean (Knowlton 2000, Bickford, Lohman et al. 2007). The nature of the marine environment can make it hard, if not impossible, to observe traits that can help distinguish species from one other. Chemical signalling for example, is a common way of communication in the sea, and is used by different organisms to recognise conspecifics (Knowlton 1993). However, it is a medium that is largely unavailable to humans, and therefore hinders our recognition of different species that might be obvious to the organism itself. Other properties of species such as varying life histories, depth preferences as well as other ecological traits may also be hard to observe. Since most collected material is fixed before it is properly examined, a lot of the characteristics such as colour, pigmentation and as well as extremities of the living specimen can be lost. By photographing the specimen before fixation, some of the hinders to distinguish species from one another can be counteracted.

Polychaetes

Polychaetes are usually the most prominent organisms in benthic communities, both in terms of species richness as well as biomass (Nygren and Pleijel 2018). The multitude of different lifestyles represented within the phylum have made them successfully conquer almost every marine environment there is; where they play key roles as bioturbators, predators and as prey for other organisms among other things (Rouse and Pleijel 2001).

The species in focus of this study are *Nereiphylla rubiginosa* and *Pterocirrus macroceros*. Both are members of the family Phyllodocidae, also known as paddle worms due to their paddle shaped dorsal cirri. Phyllodocids are free-living predators with an eversible pharynx that they use to seize prey (Rouse and Pleijel 2001).

Previous studies on other polychaetes, among them phyllodocids, have revealed major cryptic species complexes. Where species with a previously wide geographic range have been split into several new species with much narrower distributions (Grassle and Grassle 1976, Nygren, Pleijel et al. 2011, Nygren, Parapar et al. 2018). This suggests that other species with large geographic ranges are likely to consist of multiple, potentially cryptic, species. Morphologically very similar species have also been shown to exist in the same geographic locations (Nygren, Eklöf et al. 2010, Nygren, Pleijel et al. 2011).

Studied species

Nereiphylla rubiginosa was originally described from Dinard on the Atlantic coast of France. It is recorded from western Ireland throughout the south coast of England and along the Atlantic coast down into the Mediterranean (Pleijel 1993). It also occurs in the Black sea where it is used as an indicator species of areas that have low levels of eutrophication (Surugiu 2005). *N. rubiginosa* can be found from the intertidal to around 100 m depth; among shell gravel, holdfasts of *Laminaria* and in rock crevices (Pleijel 1993).

The size of *N. rubiginosa* range from 5-100 mm (Pleijel 1993, Hayward and Ryland 2017). The body of the animal is yellowish, two median dark lines are usually present along the dorsal side of the body. Its dorsal cirri are rounded and pointy. Both dorsal and ventral cirri are usually darker yellow to orange and can have unevenly scattered dark brown spots. On the head they have one pair of red eyes (Pleijel 1993).

Pterocirrus macroceros was originally described in 1860 from the Adriatic Sea and is found throughout the Mediterranean extending along the Atlantic west coast of Portugal and France, through the English Channel further to the west coast of Ireland. However, it has also been recorded from places such as the Black Sea, the Gulf of Mexico, and the Sea of Japan and have thus been considered a cosmopolitan species.

P. macroceros can be found in shell gravel and sand, among shell hash and rock crevices, at depths ranging from shallow water down to around 50 m (Pleijel 1993).

P. macroceros has a body length ranging from 30-100 mm (Pleijel 1993, Hayward and Ryland 2017). The colour of the animal is white where some individuals have scattered brown pigmentation. On the head there are two large dark red eyes. The dorsal cirri of the mid-body are lanceolate and pointed. The ventral anterior cirri of segment two is broader and more leaf-like than the rest (Pleijel 1993, Hayward and Ryland 2017).

Molecular markers

The molecular markers used in this study are the cytochrome c oxidase unit 1 (COI) gene found in the mitochondrial genome and the internal transcribed spacer regions (ITS1 & ITS2) as well as the 28S gene, situated in the nuclear genome. The COI gene, also known as the DNA barcoding gene, is used in phylogenetic studies because it has an evolutionary rate that allows for observation of recent speciation events (Hebert, Ratnasingham et al. 2003).

The mitochondrial genome is inherited from the maternal side, which means that the COI gene is only inherited from the mother. To counteract this maternal bias another gene is used that represent the paternal line as well. Ribosomal DNA includes the internal transcribed spacer regions, ITS1 and ITS2 that are situated in between the genes encoding the small ribosomal subunit (18S) and the large ribosomal subunit (28S) the nuclear genome (Hillis and Dixon 1991). There are multiple copies of ribosomal DNA in each cell, these ribosomal regions are homogenised through concerted evolution which means that the different copies evolve dependent on each other. This makes the copies of the ribosomal region very similar, both within the same individual as well as between members of the same species (Hillis and Dixon 1991).

Separation of clades

A haplotype is a variant of a certain gene. In the program TCS, haplotypes are congregated into clusters based on how similar they are. Every haplotype is represented by a coloured circle, identical haplotypes will cluster together. For every difference between two haplotypes these two haplotypes will end up further apart, with a line representing every difference. Over a certain number of differences the other haplotype will end up in a separate, unconnected cluster (Clement, Posada et al. 2000).

Inference of phylogenetic trees will be done by using Bayesian statistics in the program MrBayes. The program searches for the set of trees in the tree space that best fits the data and the selected model of evolution. In the end of the run the most likely trees are summarised into a consensus tree, which is the output used to infer phylogenies (Salemi and Vandamme 2003, Hall 2011).

Uncorrected pairwise distances are used to compare the number of sites where nucleotides differ between two sequences. This method does not account for different evolutionary rates between sites, substitution rate biases or multiple substitutions at the same site (Salemi and Vandamme 2003, Hall 2011).

In this study, lineages are considered as separate species if there is congruence between the COI and ITS trees as well as in the TCS plots. The lineages should also be clearly distinct from each other in the sense that there is no evident genetic exchange between the taxa (Nygren 2014). Uncorrected pairwise distances as well as morphological analyses based on live specimen from the different clusters are used as further support for the different clades.

Purpose & hypotheses

Cryptic speciation is a widespread phenomenon that has implications for our understanding of biological diversity (Brasier, Wiklund et al. 2016). To be able to make correct statements of species and their qualities it is vital that this largely understudied diversity is identified and described. The aim of this study was to take the first steps in this process; to find out whether each of the two taxa *Nereiphylla rubiginosa* and *Pterocirrus macroceros*, were made up of one or several species respectively.

Materials & methods

Sampling & collection

Data were collected during the period 2001-2018 from several different locations throughout the North Eastern Atlantic and the Mediterranean (see fig. 1).

I participated in the collection and preparation of the animals from Roscoff in 2018, where the following procedures were applied.

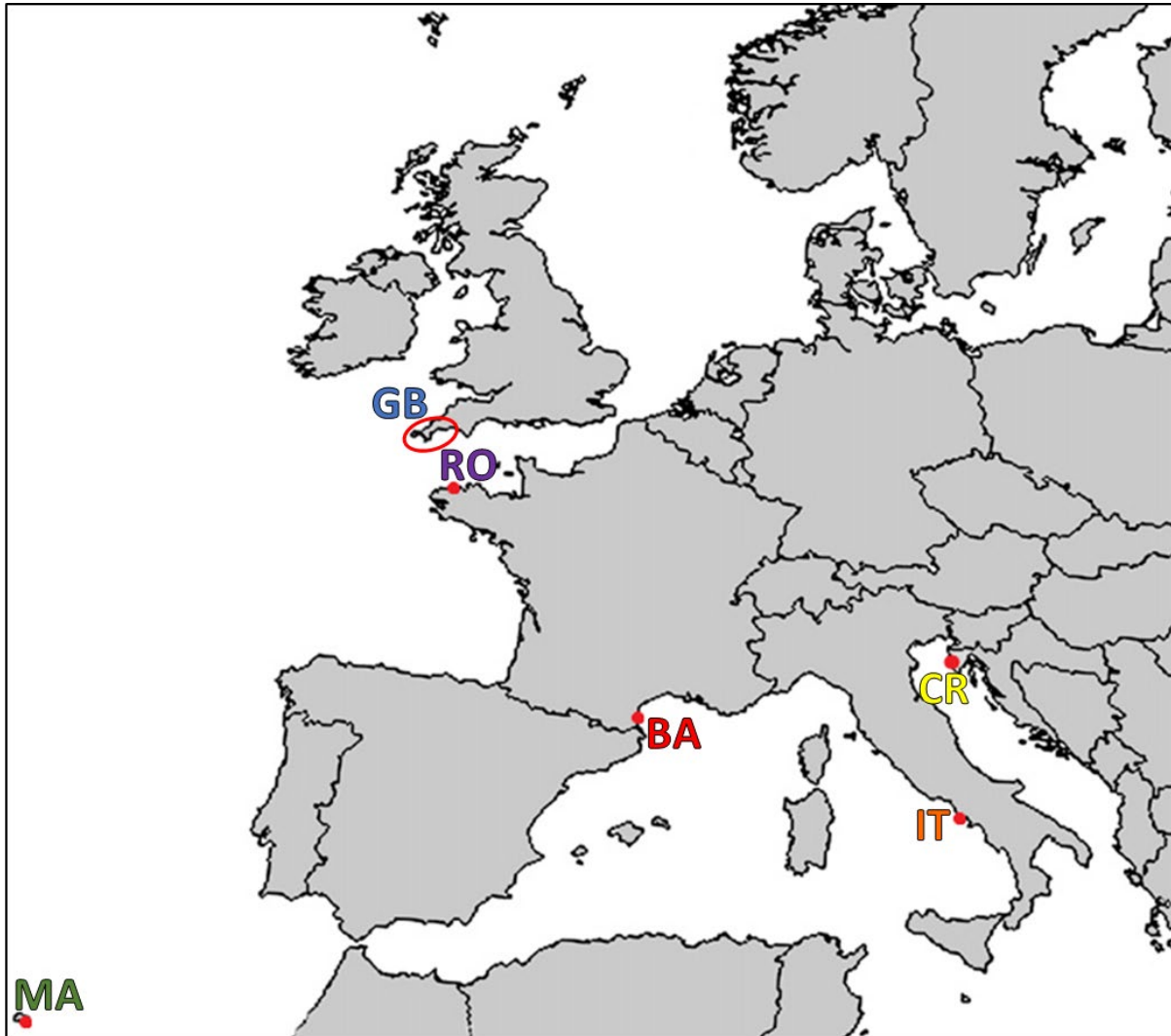


Figure 1. The collection sites in North Eastern Atlantic and the Mediterranean. Abbreviations stand for: GB = Great Britain (blue); RO = Roscoff, France (purple); MA = Madeira, Portugal (green); BA = Banyuls, France (red); IT = Italy (orange); CR = Croatia (yellow).

Samples were collected with a round trawl from several locations around Roscoff. The acquired material was then left without water circulation (for oxygen to be depleted) so the polychaetes would crawl up and be collected without being damaged. Following collection, the animals were photographed and then anaesthetized with $MgCl_2$ before preservation in 95% ethanol.

For the DNA extraction, 1-3 parapodia or the posterior segments of the worm were removed and put in 50ul QuickExtract DNA Extraction (Epicentre). The extraction fluids were then put on a heating block of 65 degrees Celsius for 45 minutes followed by 2 minutes on 95 degrees Celsius.

Amplification & sequencing

The primers used to amplify the mitochondrial DNA-barcoding gene, and the nuclear marker from the ITS1, ITS2 and/or 28S region can be found in appendix A.

The DNA templates were diluted to increase the probability of successful amplification. The extracts were diluted in MilliQ-water with a dilution factor of 1:5 (one-part DNA and four parts distilled water). Prior to PCR, the following products were mixed for each individual sample: 1 µl of DNA template, 16 µl of VWR Red Taq DNA polymerase Master Kit, 0,7 µl of forward primer and 0,7 µl reverse primer. The PCR was done in a Veriti® 96-Well Thermal Cycler (Applied Biosystems), the protocols used can be found in appendix B. To make sure that the amplification was successful; the PCR products were run through a 1% agarose gel with 3 µl Gel red (Biotium).

Before sequencing, the DNA was purified by addition of 2 µl ExoSAP-IT™ PCR Product Cleanup to 5 µl of the PCR product from each individual sample. Lastly, the samples were put in the thermal cycler again. The protocol for this can also be found in appendix B under 'Purification'. The sequences were then sent to the lab Eurofins Europe for sequencing.

Analyses

All available DNA sequences (see appendix C) were prepared in Geneious version 5.4.6 (Biomatters, LTD, Auckland, NZ). The sequences were trimmed, and low-quality reads pruned. The following outgroups were chosen from within Phyllodocidae: *Notophyllum foliosum*, *Paranaitis spp.*, *Mystides sp* and *P. macroceros* was used as additional outgroup for *N. rubiginosa* and vice versa.

The sequences of good enough quality were then aligned with the outgroups using MAFFT v.7.017 (Kato, Misawa et al. 2002) with the following settings: algorithm = E-INS-i, scoring matrix = 200PAM / k = 2, gap open penalty = 1.53, to align COI mDNA, ITS1 rDNA, ITS2 rDNA and 28S rDNA according to Nygren, Parapar et al. (2018).

Haplotype networks

The alignments were converted into NEXUS-format and run in the program TCS (Clement, Posada et al. 2000) with settings as follows: connection limit 95%, gaps treated as uninformative. Separate haplotype networks for COI, ITS1, ITS2 and 28S were generated and then visualised with the help of PowerPoint for Windows.

Phylogenetic inference

In the analyses the data was partitioned into different sets, where the parts of the data with similar rates of substitutions were put in the same matrix. The COI sequences were sorted into one matrix with the first and second position of each codon in the sequence, and into another matrix with only the third position in every codon. The separate parts of the internal transcribed spacer region including ITS1, ITS2 and 28S were left as they were, as they are not part of any coding sequence.

Appropriate models of evolution were selected for each of the different data partitions, using the Akaike information criterion, in the program jModelTest version 2 (Darriba, Taboada et al. 2012). The models used can be found in appendix D.

The alignments were then concatenated in Mesquite version 2.75 (Maddison and Maddison 2009) resulting in one COI matrix and one ITS matrix for each species. Matrices were run in MrBayes version 3.2 (Ronquist, Teslenko et al. 2012). The analyses were run for 2 million generations in two independent analyses, every 1000th generation was sampled. Out of 2 million generations 25% were discarded as burn-in. Four parallel chains were run, out of which three were hot and one was cold. For parameters statefreq, revmat, shape and pinvar the partitions were unlinked. Rateprior was set to be variable for the partition rate multiplier.

Uncorrected pairwise distances were calculated using PAUP* 4.0b10 (L. Swofford 2002), then the distances between the different clades were compiled into matrices in Microsoft Excel for Windows.

Results

Sequencing success

Out of all the available sequence data, the individuals with good quality DNA reads for COI were chosen as well as the ones that also had corresponding ITS1, ITS2 and/or 28S sequences.

In *N. rubiginosa* the amount of good quality sequences for COI equalled 47 individuals, out of these were 20 individuals that had ITS1, ITS2 and/or 28S reads of good enough quality to be used in the analyses. Representatives for each cluster in COI were found in the ITS1, ITS2 and/or 28S sequences except for cluster α and β .

There were 58 individuals with good quality COI sequences in *P. macroceros* out of which 22 had ITS1, ITS2 and/or 28S sequences. All clusters in COI had representation in ITS1, ITS2 and/or 28S.

Nereiphylla rubiginosa

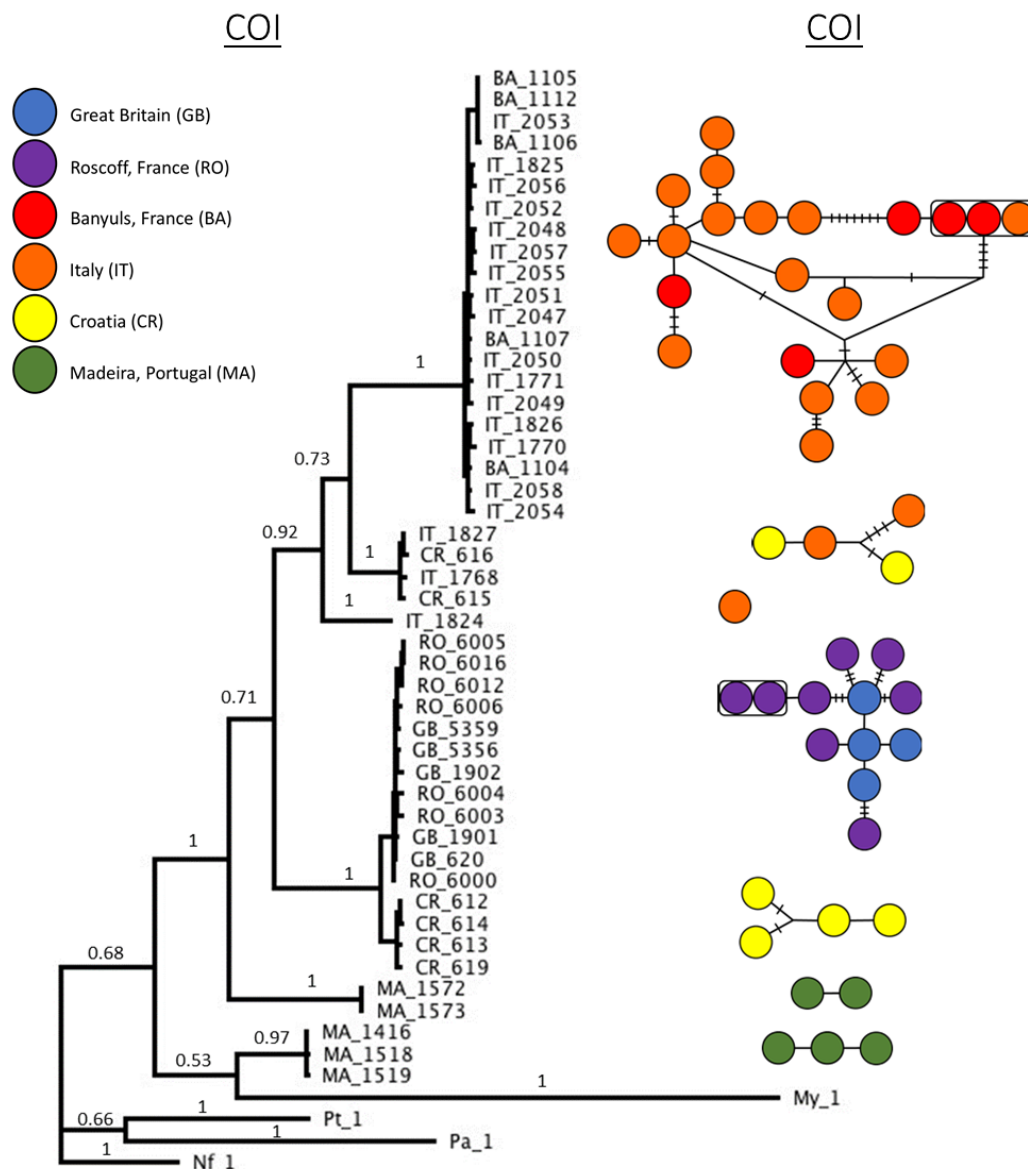


Figure 2 showing COI tree and TCS plot for *N. rubiginosa*. My_1, Pt_1, Pa_1 & Nf_1 are outgroups. The length of the branches depicts amount of differences, the numbers represent the probability of each branch.

There were seven groups in the COI (see fig. 2) tree with seven corresponding clusters in the TCS plot for *N. rubiginosa*. The first group/cluster was made up of individuals from Italy and Banyuls in France. The second group/cluster consisted of individuals from Italy and Croatia. The third group consisted of a single individual from Italy. The fourth group/cluster was made up of individuals from Great Britain and Roscoff, France. In the group that followed were four specimens from Croatia. The last two of the clusters were both from Madeira.

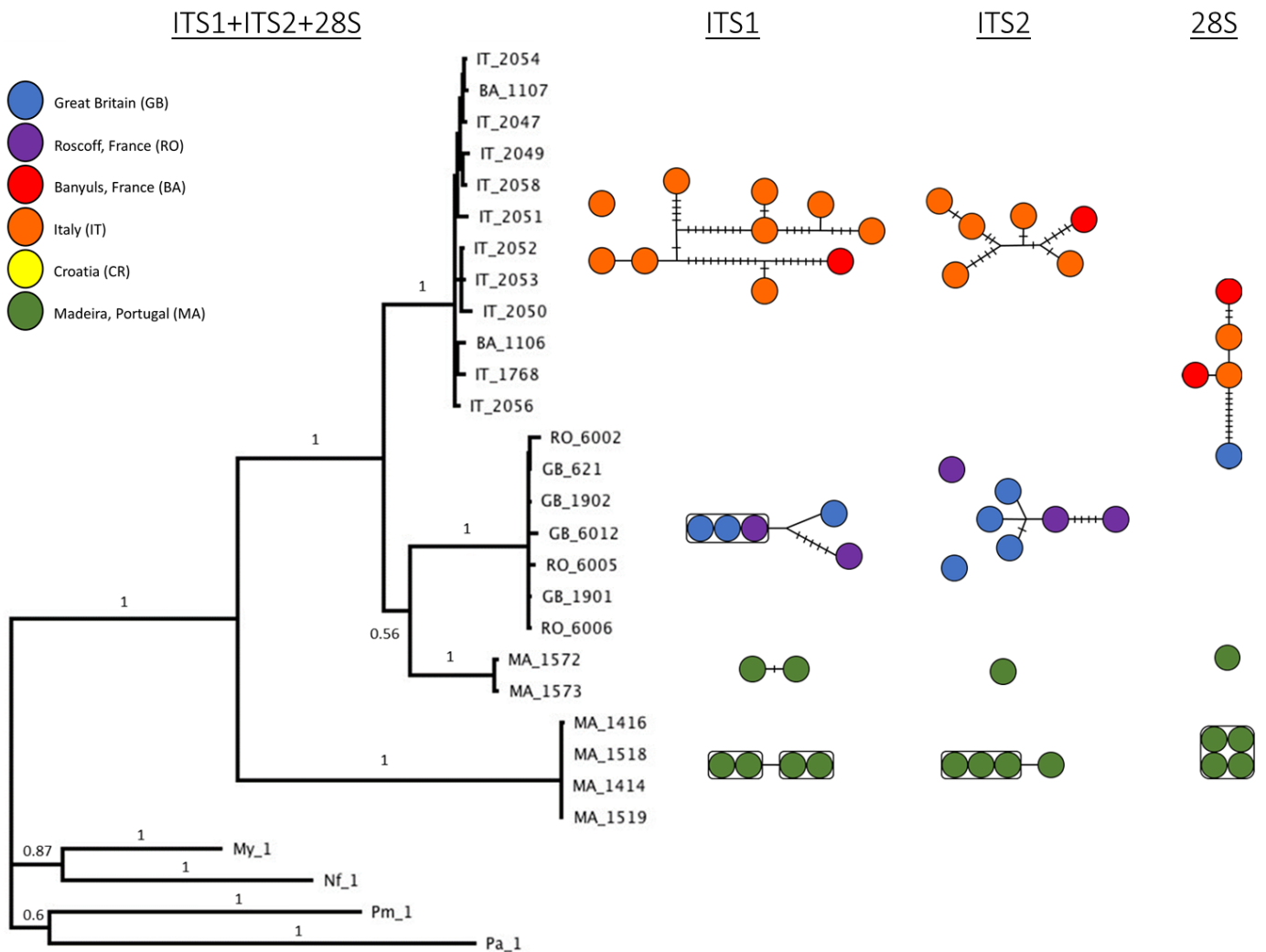


Figure 3 showing ITS1+ITS2+28S tree and TCS plot for *N. rubiginosa*. *My_1*, *Pt_1*, *Pa_1* & *Nf_1* are outgroups. The length of the branches depicts amount of differences, the numbers represent the probability of each branch.

There were four groups in the ITS tree (see fig. 3). The TCS plots for the ITS1, ITS2 and 28S largely corresponded to the phylogenetic tree from MrBayes. They included four clusters with one haplotype in the first cluster in ITS1 that appeared as separate and two haplotypes in the second cluster in ITS2 that were separated from the larger cluster nearby. The reason was that they had passed the limit (with one unit) for the maximum number of differences within one cluster in TCS. However, they were included in the same groups in the tree, therefore the singles were counted as parts of the larger surrounding clusters.

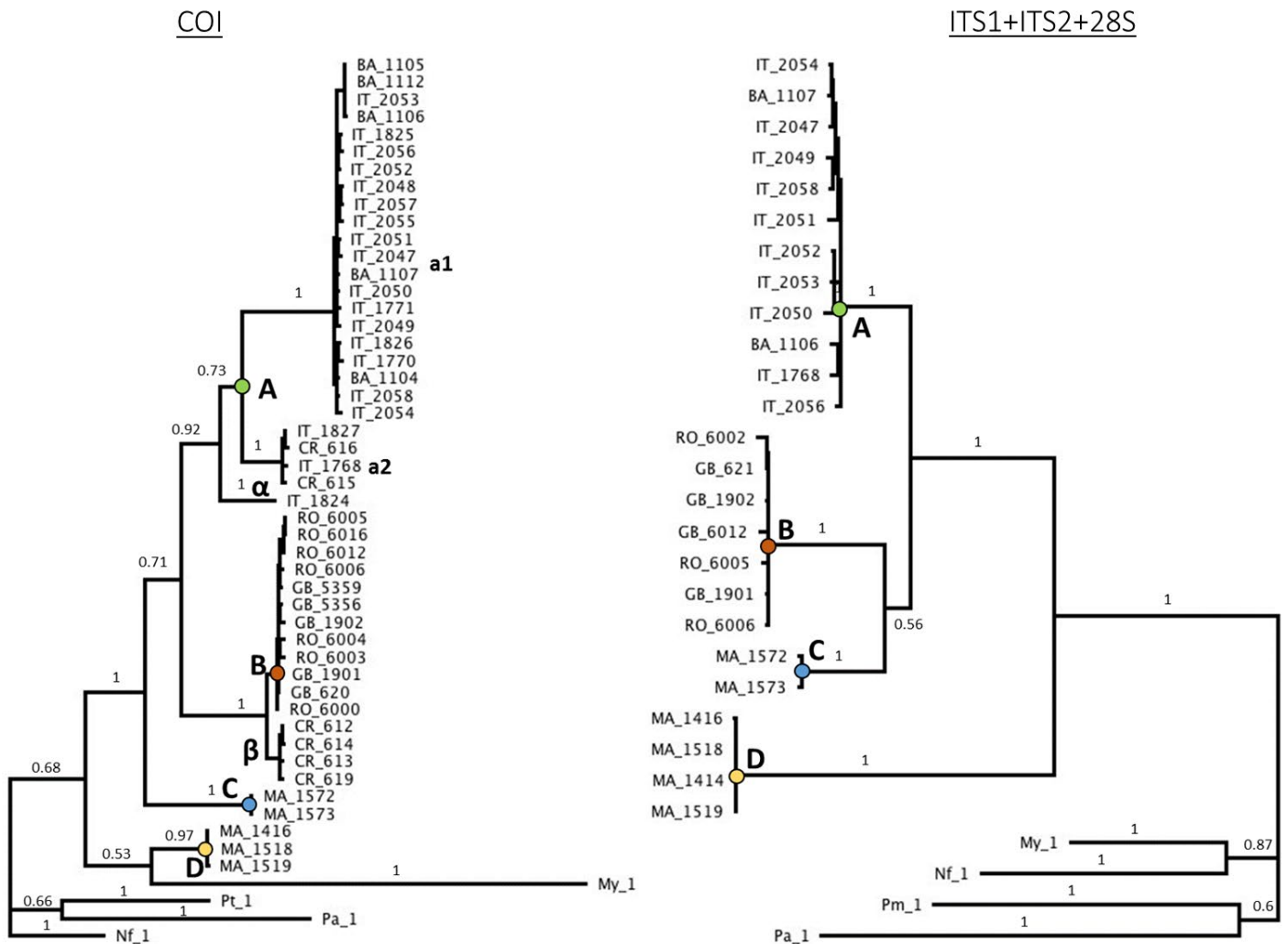


Figure 4 showing congruence in phylogenetic trees for COI and ITS1+ITS2+28S in *N. rubiginosa*. My_1, Pt_1, Pa_1 & Nf_1 are outgroups. The length of the branches depicts amount of differences, the numbers represent the probability of each branch.

Comparison of the COI tree and ITS1, ITS2 and 28S tree (see fig. 4) showed that there were four main groups that were congruent between them – A, B, C and D. All of which had support in the TCS plots in figure 2 and 3. There were two subdivisions in group A - a1 and a2, in both the COI tree as well as the TCS plot for COI. These subdivisions did not appear in the ITS-tree, where the only representative from cluster a2, namely IT_1768, was situated within cluster A in the ITS tree.

Two other clusters were found in the COI-tree, α and β . The fact that no ITS or 28S sequences were available for α and β , explains their lack of representation in the ITS1, ITS2 and 28S tree and haplotypes networks.

Table 1 Uncorrected pairwise distances for COI in *N. rubiginosa* shown in percent.

		COI					
		a1	a2				
			0-12,95				
COI		A	α	B	β	C	D
A		0-12,95					
α		12,22-12,92	0				
B		13,6-16,52	14,29-14,46	0-1,55			
β		14,46-15,66	14,11-14,46	3,01-4,48	0,17-0,69		
C		13,94-16,28	16,52-16,82	15,49-16,75	15,81-16,07	0	
D		14,63-17,56	16,01-16,18	14,97-16,35	14,6-15,15	15,49-16,16	0,17-0,34

Table 2 Uncorrected pairwise distances for ITS1 in *N. rubiginosa* shown in percent.

ITS1	A	B	C	D
A	0,16-3,85			
B	22,22-25,27	0-1,2		
C	16,7-18	19,2-21	0,38	
D	33,33-35,18	36,67-38	36,91-37,2	0,21

Table 3 Uncorrected pairwise distances for ITS2 in *N. rubiginosa* shown in percent.

ITS2	A	B	C	D
A	0,44-3,14			
B	20,6-29,97	0,41-5,97		
C	21,7-27,23	29,23-32	0	
D	32,23-34,95	34,88-36,3	37-37,3	0-0,23

Table 4 Uncorrected pairwise distances for 28S in *N. rubiginosa* shown in percent.

28S	A	B	C	D
A	0,13-0,68			
B	1,37-1,78	0		
C	1,63-1,9	1,9	0	
D	8-8,3	8,85-8,86	8,69-8,7	0

Uncorrected pairwise distances for *N. rubiginosa*

Uncorrected genetic distances between the main groups A, B, C and D in the COI sequences varied between 12-18%. The variation within groups ranged from around 0-1,6% in B, C and D, while group A stood out with a genetic variation ranging from 0-13% in COI (see table 1). The genetic differences in ITS1 and ITS2 between A, B, C and D ranged from almost 17% to 37%, the variation went from 0 to almost 4% within groups (see table 2 and 3). In 28S the distances were approximately 1,4-9% between groups and varied between 0-0,7% within the groups (see table 4).



Figure 5. Two of the individuals in cluster A, subgroup a1. Photos taken by Fredrik Pleijel (above) and Arne Nygren (below).



Figure 6. Two of the brown individuals from group A, subgroup a2. Photos taken by Fredrik Pleijel (above) and Arne Nygren (below).



Figure 7. Two of the individuals from Cluster B from GB and BA. Photos taken by Arne Nygren.



Figure 8. The two individuals in cluster C from Madeira. Photos taken by Fredrik Pleijel.



Figure 9. Two of the individuals in cluster D from Madeira. Photos taken by Arne Nygren.

Pterocirrus macroceros

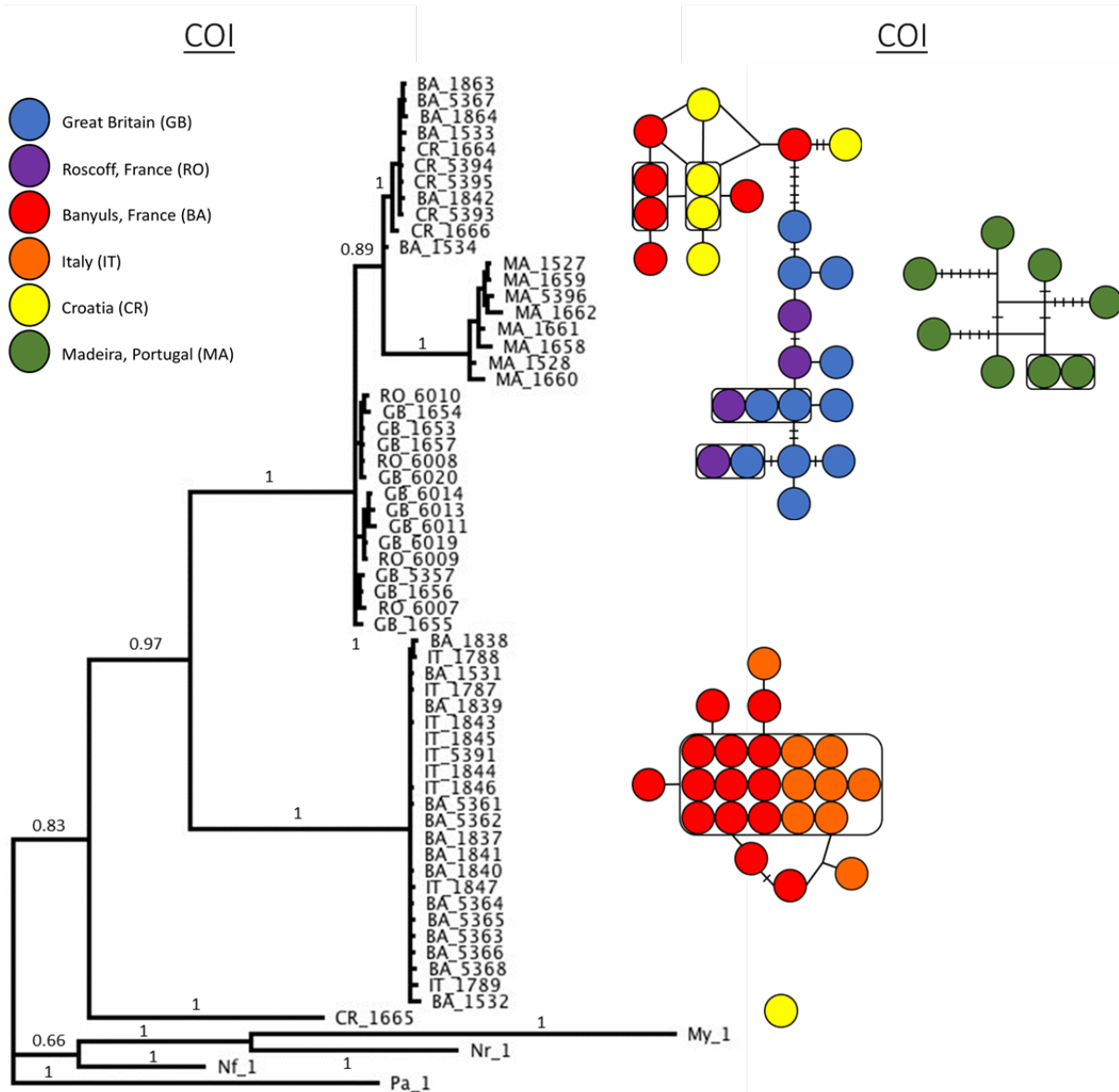


Figure 10 Showing COI tree and TCS plot for *P. macroceros*. My_1, Nr_1, Pa_1 & Nf_1 are outgroups. The length of the branches depicts amount of differences, the numbers represent the probability of each branch.

In the phylogenetic tree were five groups, while there were four clusters in the TCS plot (see fig. 10). The first group consisted of individuals from Croatia and Banyuls in France. In the TCS clusters it is linked to individuals from Great Britain and Roscoff, France. The second group in the tree consists of individuals from Madeira. The third group consists of worms from Italy and Banyuls in France. Lastly, there was a single specimen from Croatia that was separated from the rest.

COI

ITS1+ITS2+28S

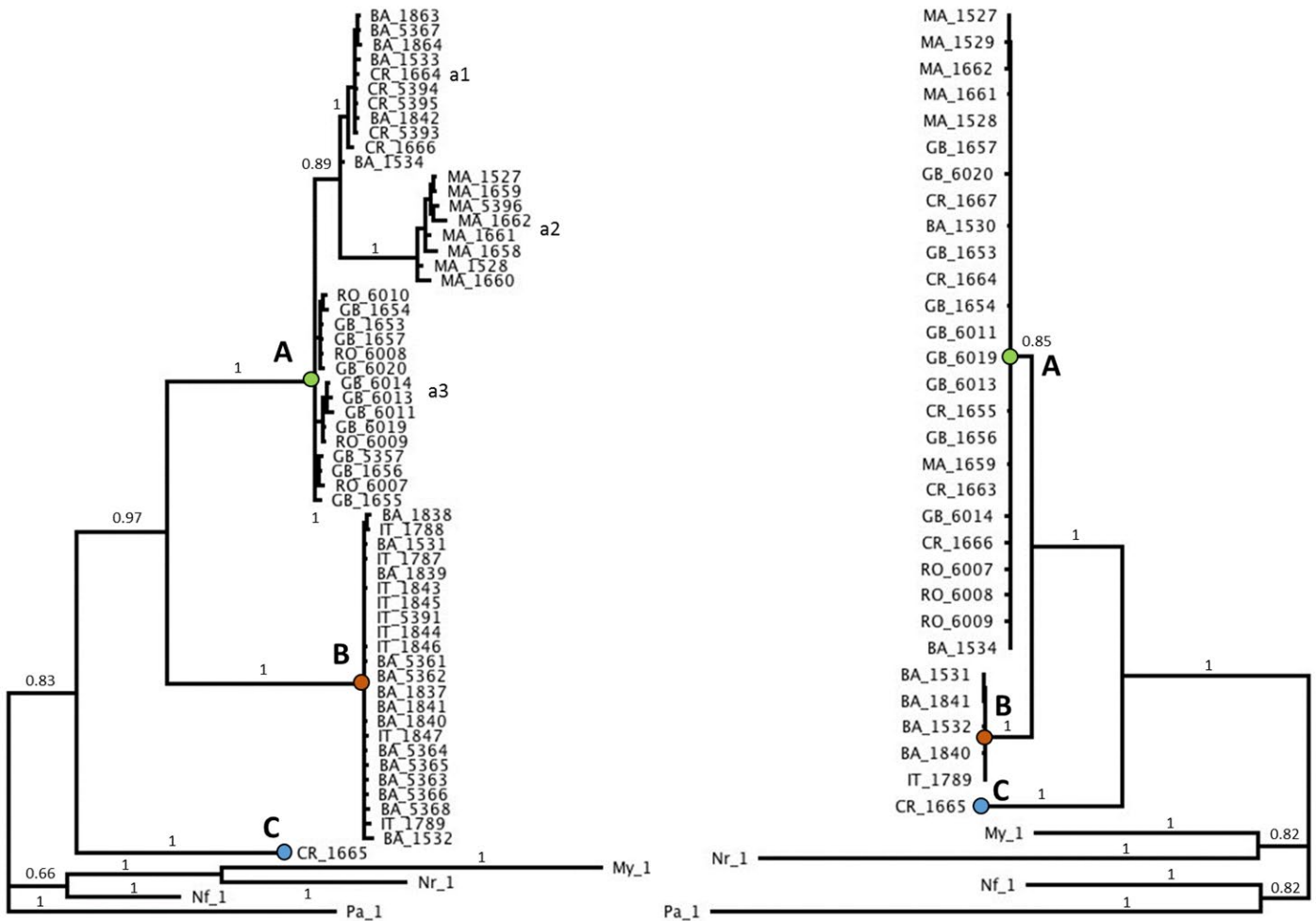


Figure 12. Showing phylogenetic trees for COI and ITS1+ITS2+28S for *P. macroceros*. My_1, Nr_1, Nf_1 and Pa_1 are outgroups. The length of the branches depicts amount of differences, the numbers represent the probability of each branch.

Group A was made up of three subgroups – a1, a2 and a3 (see fig. 12). The first subgroup, a1, was made up of individuals from Banyuls and Croatia. Subgroup a2 consisted of worms from Madeira and in the third subgroup, a3, were worms from the Great Britain and Roscoff in France. Group B consisted of individuals from Banyuls, France and Italy. The last group consisted of the single individual from Croatia.

Table 5 Uncorrected pairwise distances for COI in *P. macroceros* shown in percent.

COI	A	B	C
A	0-8,83		
B	15,14-18,71	0-1,38	
C	18,02-20,61	18,11-20,47	0

Table 6 Uncorrected pairwise distances for ITS1 in *P. macroceros* shown in percent.

ITS1	A	B
A	0-2,15	
B	11,49-13,61	0,13-0,532

Table 7 Uncorrected pairwise distances for ITS2 in *P. macroceros* shown in percent.

ITS2	A	B	C
A	0-3,17		
B	8,83-10,03	0-0,38	
C	27,99-29,06	28,44-28,63	0

Table 8 Uncorrected pairwise distances for 28S in *P. macroceros* shown in percent.

28S	A	B	C
A	0		
B	0,54	0	
C	2,16	2,16	0

Uncorrected pairwise distances for P. macroceros

The genetic distances in COI between the main groups A, B, C and D, were in the range from 15-21%. The variation in A went from 0-9%, and between 0-1,4% within B and C. In ITS1 it went from around 12 to 14% between groups and within the same cluster the range was 0 to 2,2%. In ITS2 the distances between groups varied between 9-29% and within it went from 0 to around 3,2%. Finally, the 28S had no variation within groups, and ranged from 0,5-2,2% between groups.



Figure 13. Individual from group A. Showing a1, a2 and a3. Photos taken by Fredrik Pleijel, Arne Nygren & Felicia Ulltin.



Figure 14. Individual from group B of Pterocirrus macroceros. Photo taken by Arne Nygren.



Figure 15. Individual from group C. Photo taken by Fredrik Pleijel.

Discussion

Nereiphylla rubiginosa

Congruence between the COI and ITS trees from MrBayes as well as the haplotype networks from TCS supports four species within *N. rubiginosa* (see fig. 2, 3 and 4). The genetic distances between A, B, C and D are in the range 12-18% in COI, between 17-37% in ITS1 and ITS2 and ranging between 1,3-8% in 28S. This reinforces the claim that these are separate lineages. The potential presence of even more species in group α and β , is not possible to determine as there is no available ITS data to confirm it.

The first lineage A is solely made up of individuals from the Mediterranean, group a1 were collected in Banyuls, France and Italy and group a2 were found in Italy and Croatia. Individuals from group a1 is a lime-yellow, with faint lines along the dorsal side of the body (see fig. 5). In contrast, worms from group a2 are ranging in colour from dark brown with lighter spots on their dorsal cirri to overall spotted light-brown, with brown lines along the dorsum (see fig. 6). Interestingly, the species overlap in their geographical distributions, despite this they are differing genetically. According to the uncorrected pairwise distances, the COI sequence differs with as much as 13% between a1 and a2.

There are two possible explanations for the differences found in the COI, that was not found in the ITS1, ITS2 and 28S sequences. One explanation could be that the populations have been separated and during separation the COI and ITS sequences diverged, and a2 acquired its brown colouration. Then the lineages were reunited, and the COI sequences remained different, but the ITS sequences were homogenised through concerted evolution (Hillis and Dixon 1991).

The second explanation, which is the more plausible, is that group a1 and a2 have been separated. During separation the two lineages have diverged in COI, while a2 has acquired brown colouration. Although, both have remained similar in the ITS1, ITS2 and 28S sequences. Because of the large genetic distances between a1 and a2, it is possible that a larger genetic distance in the nuclear genome could be detected if other faster evolving genetic markers than ITS1, ITS2 and 28S were used. Essentially, the yellow group a1 and brown group a2 could turn out to be two different species. But due to the absence of other molecular markers that support a divergence into two species, group a1 and a2 will be considered the same species A.

Species B is made up of individuals from the Great Britain and Roscoff in France. The individuals in this area are quite robust. The body and the dorsal cirri are darker yellow with two distinct dark lines along the dorsum (see fig. 7). Intraspecific variation is between 0-1,55% in COI, 0-1,2% in ITS1, 0,4-6% in ITS2 and lastly there is no variation in 28S. The genetic distance to cluster β is very small. This implies that the distribution of species B stretches to Croatia as well, though this cannot be concluded as there is no data in ITS1, ITS2 or 28S for β .

The two species found in Madeira seem to be similar in appearance. Lineage C is a darker yellow colour with brownish spots on the dorsal cirri (see fig. 8). The individuals from species D are yellow as well (see fig. 9). None of species C or D have darker lines on the dorsum. The genetic distances between the two species in COI is around 16%, in

ITS1 and ITS2 it is around 37% and in 28S almost 9%. This suggests that the lineages have not interbred in a long time, despite being sympatric.

Pterocirrus macroceros

According to the trees from MrBayes (see fig. 12) as well as the haplotype networks in TCS (see fig. 10 and 11), three species exist within *P. macroceros*. The uncorrected pairwise distances between species ranges from 15-21% in COI, 12-14% in ITS1, 9-29% in ITS2 which further supports their separation.

Species A is made up of three clusters (a1, a2 and a3) in the COI tree (see fig. 13). The individuals from a1 are from Croatia and Banyuls, a2 from Madeira and the worms in the last cluster are from the Atlantic coast of France and Great Britain. The genetic distance within the species range from 0 to almost 9% in COI, these differences are most likely there because of the geographic spread of the clusters. One example where the distance is large is the Madeira branch which appears as a separate network in TCS for COI but is included in the same cluster in ITS1, ITS2 and 28S.

Species B is an exclusively Mediterranean species with individuals from Italy and Banyuls in France. The genetic distances within this cluster ranges from 0 to almost 1,4% in COI, ITS1, ITS2 and 28S. There is a slight difference in the morphology of this species compared to the others in that it has little dot-shaped “eyebrows” above each eye (see fig. 14 and 16). Apart from that no obvious morphological differences are found. The reason that some individuals appear slightly darker in colour is probably due to the worms being in poor condition when the photographs were taken.

Species C is made up of one single individual from Croatia (see fig. 15), that comes out as a separate cluster in COI, ITS2 and 28S, no ITS1 sequence was available. Despite sympatric occurrence with species A, it is largely different in the sequence data ranging from 18-21% in the mitochondrial sequence and from 28-29% in the ITS2 sequence, it also showed a difference of about 2% in 28S.



Figure 16. Close-up of “eyebrows” in species B, in *P. macroceros*. Photo taken by Arne Nygren.

If further sampling efforts were undertaken in the different locations, it is possible that more species would be found. It is also likely that individuals from intermediate locations would be gradually different as the geographic distances would increase, due to lower levels of connection between the lineages.

When dealing with molecular markers in phylogenetic inference there is always the risk that the markers used do not represent the true divergence of the species. In combination with conservative criteria for species delimitation there is a risk that the number of lineages present in a dataset are underestimated. By using more types of molecular markers, one could counteract some of these effects.

The colouration and general morphology are used as complementary information to the other analyses in this study. To be able to say for certain if these characters are valid for identification, it would be wise to study a larger set of individuals from the respective groups. But from this smaller subset of individuals it is likely that there are in fact some morphological differences in live specimen. To consider if a species is cryptic or not is in the end a matter of where to draw the line.

Conclusions

This study of two commonly encountered polychaetes has revealed that their distribution ranges are not as previously expected. Neither of the two worms consist of only one species. *Nereiphylla rubiginosa* consists of one lineage unique to the Mediterranean, another lineage occurring around the Atlantic coast of France and southern coast of England and two sympatric lineages in Madeira. *Pterocirrus macroceros* consists of three species, two species exclusively occurring in the Mediterranean and only one species ranging throughout the sampled areas.

It is my hope that these species will be described and acknowledged so that more correct statements can be made about species and species diversity.

Acknowledgements

I would like to thank my supervisor Arne Nygren for all the time and effort he has put into guiding and helping me throughout this project. I would also like to thank Arne and Fredrik Pleijel for letting me use their pretty pictures of most of the worms used in this study. I would also like to thank the people who have helped me out in the lab at Tjärnö.

Last, but not least I would like to send out a thought to all the worms that died in this project. I hope you enjoy worm-heaven, reincarnation or whatever you are up to.

References:

- Bickford, D., et al. (2007). "Cryptic species as a window on diversity and conservation." Trends in Ecology & Evolution **22**(3): 148-155.
- Brasier, M. J., et al. (2016). "DNA barcoding uncovers cryptic diversity in 50% of deep-sea Antarctic polychaetes." **3**(11): 160432.
- Carr, C. M., et al. (2011). "A tri-oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian polychaetes." **6**(7): e22232.
- Clement, M., et al. (2000). "TCS: a computer program to estimate gene genealogies." **9**(10): 1657-1659.
- Darriba, D., et al. (2012). "jModelTest 2: more models, new heuristics and parallel computing." Nature Methods **9**: 772.
- De Queiroz, K. J. S. b. (2007). "Species concepts and species delimitation." **56**(6): 879-886.
- Grassle, J. and J. F. J. S. Grassle (1976). "Sibling species in the marine pollution indicator *Capitella* (Polychaeta)." **192**(4239): 567-569.
- Hall, B. G. (2011). Phylogenetic trees made easy : a how-to manual. Sunderland, Mass., Sunderland, Mass. : Sinauer Associates.
- Hayward, P. J. and J. S. Ryland (2017). Handbook of the marine fauna of North-West Europe, Oxford University Press.
- Hebert, P. D., et al. (2003). "Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species." Proceedings of the Royal Society of London. Series B: Biological Sciences **270**(suppl_1): S96-S99.
- Hillis, D. M. and M. T. J. T. Q. r. o. b. Dixon (1991). "Ribosomal DNA: molecular evolution and phylogenetic inference." **66**(4): 411-453.
- Katoh, K., et al. (2002). "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform." **30**(14): 3059-3066.
- Knowlton, N. (1993). "Sibling Species in the Sea." **24**(1): 189-216.

Knowlton, N. (2000). Molecular genetic analyses of species boundaries in the sea. Marine genetics, Springer: 73-90.

L. Swofford, D. (2002). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10.

Maddison, W. and D. Maddison (2009). MESQUITE: a modular system for evolutionary analysis.

Nygren, A. (2014). "Cryptic polychaete diversity: a review." **43**(2): 172-183.

Nygren, A., et al. (2010). "Cryptic species of Notophyllum (Polychaeta: Phyllodocidae) in Scandinavian waters." **10**(3): 193-204.

Nygren, A., et al. (2018). "A mega-cryptic species complex hidden among one of the most common annelids in the North East Atlantic." **13**(6): e0198356.

Nygren, A. and F. Pleijel (2018). Nationalnyckeln till Sveriges flora och fauna: Annelida: Polychaeta: Aciculata. Ringmaskar: Havsborstmaskar, Artdatabanken, Sveriges lantbruksuniversitet.

Nygren, A., et al. (2011). "From one to ten in a single stroke—resolving the European Eumidasanguinea (Phyllodocidae, Annelida) species complex." **58**(1): 132-141.

Pleijel, F. (1993). Polychaeta Phyllodocidae, Berg Publishers.

Ronquist, F., et al. (2012). "MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space." Systematic Biology **61**(3): 539-542.

Rouse, G. and F. Pleijel (2001). Polychaetes, Oxford university press.

Salemi, M. and A.-M. Vandamme (2003). The phylogenetic handbook : a practical approach to DNA and protein phylogeny. New York, New York : Cambridge University Press.

Surugiu, V. (2005). "The use of polychaetes as indicators of eutrophication and organic enrichment of coastal waters: a study case—Romanian Black Sea Coast." Analele științifice ale Universității "Al. I. Cuza" Iași, Biologie animală **51**: 55-62.

Appendix A

Primers used for extraction of the COI-region and the ITS1, ITS2 and 28S-regions.

REGION	PRIMER	SEQUENCE
COI	LCO1490	GGTCAACAAATCATAAAGATATTGG
COI	HCO2198	TAAACTTCAGGGTGACCAAAAAATCA
COI	dgLCO1490	GGTCAACAAATCATAAAGAYATYGG
COI	dgHCO2198	TAAACTTCAGGGTGACCAARAAYCA
ITS	ITS18SF	GAGGAAGTAAAAGTCGTAACA
ITS	ITS58SR	GTTCAATGTGTCCTGCAATTC
28S	ITS58SF	GAATTGCAGGACACATTGAAC
28S	28SD2R	TCCGTGTTCAAGACGG

Appendix B

PCR protocols, modified from (Carr, Hardy et al. 2011).

COI		
STEP	TEMP	TIME
Initial denaturation	94°C	1 min
5 cycles	94°C	40 sec
	45°C	40 sec
	72°C	1 min
35 cycles	94°C	40 sec
	52°C	40 sec
	72°C	1 min
Final extension	72°C	5 min
Hold	4°C	∞

ITS		
STEP	TEMP	TIME
Initial denaturation	94°C	1 min
35 cycles	94°C	40 sec
	62°C	40 sec
	72°C	1 min
Final extension	72°C	5 min
Hold	4°C	∞

Purification		
STEP	TEMP	TIME
1 cycle	37°C	15 min
35 cycles	80 °C	30 min
Hold	4°C	∞

Appendix C

Tables of specimens used in this study.

ID	SPECIES	LOCATION	YEAR
GB620	<i>Nereiphylla rubiginosa</i>	Bryher, Scilly Islands	2006
CR612	<i>Nereiphylla rubiginosa</i>	Istra, Croatia	2008
CR613	<i>Nereiphylla rubiginosa</i>	Istra, Croatia	2008
CR614	<i>Nereiphylla rubiginosa</i>	Istra, Croatia	2008
CR615	<i>Nereiphylla rubiginosa</i>	Istra, Croatia	2008
CR616	<i>Nereiphylla rubiginosa</i>	Istra, Croatia	2008
CR619	<i>Nereiphylla rubiginosa</i>	Istra, Croatia	2008
BA1104	<i>Nereiphylla rubiginosa</i>	Banyuls, France	2009
BA1105	<i>Nereiphylla rubiginosa</i>	Banyuls, France	2009
BA1106	<i>Nereiphylla rubiginosa</i>	Banyuls, France	2009
BA1107	<i>Nereiphylla rubiginosa</i>	Banyuls, France	2009
BA1112	<i>Nereiphylla rubiginosa</i>	Banyuls, France	2009
MA1416	<i>Nereiphylla rubiginosa</i>	Funchal, Madeira	2009
MA1518	<i>Nereiphylla rubiginosa</i>	Funchal, Madeira	2009
MA1519	<i>Nereiphylla rubiginosa</i>	Funchal, Madeira	2009
MA1572	<i>Nereiphylla rubiginosa</i>	Funchal, Madeira	2009
MA1573	<i>Nereiphylla rubiginosa</i>	Funchal, Madeira	2009
IT1768	<i>Nereiphylla rubiginosa</i>	Ischia, Italy	2010
IT1770	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT1771	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT1824	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT1825	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT1826	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT1827	<i>Nereiphylla rubiginosa</i>	Ischia, Italy	2010
IT2047	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2048	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2049	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2050	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2051	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2052	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2053	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2054	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2055	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2056	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2057	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
IT2058	<i>Nereiphylla rubiginosa</i>	Naples, Italy	2010
GB1901	<i>Nereiphylla rubiginosa</i>	Plymouth, Great Britain	2011
GB1902	<i>Nereiphylla rubiginosa</i>	Plymouth, Great Britain	2011
GB5356	<i>Nereiphylla rubiginosa</i>	Cornwall, Great Britain	2015
GB5359	<i>Nereiphylla rubiginosa</i>	Cornwall, Great Britain	2015
GB6012	<i>Nereiphylla rubiginosa</i>	Plymouth, Great Britain	2017
GB6016	<i>Nereiphylla rubiginosa</i>	Plymouth, Great Britain	2017
RO6000	<i>Nereiphylla rubiginosa</i>	Roscoff, France	2018
RO6003	<i>Nereiphylla rubiginosa</i>	Roscoff, France	2018
RO6004	<i>Nereiphylla rubiginosa</i>	Roscoff, France	2018
RO6005	<i>Nereiphylla rubiginosa</i>	Roscoff, France	2018
RO6006	<i>Nereiphylla rubiginosa</i>	Roscoff, France	2018

ID	SPECIES	LOCATION	YEAR
BA1837	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA1838	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA1839	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5361	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5362	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5363	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5364	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5365	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5366	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5367	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
BA5368	<i>Pterocirrus macroceros</i>	Banyuls, France	2001
GB1653	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2006
GB1654	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2006
GB1655	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2006
GB1656	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2006
GB1657	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2006
CR1664	<i>Pterocirrus macroceros</i>	Istra, Croatia	2008
CR1665	<i>Pterocirrus macroceros</i>	Istra, Croatia	2008
CR1666	<i>Pterocirrus macroceros</i>	Istra, Croatia	2008
CR5393	<i>Pterocirrus macroceros</i>	Istra, Croatia	2008
CR5394	<i>Pterocirrus macroceros</i>	Istra, Croatia	2008
CR5395	<i>Pterocirrus macroceros</i>	Istra, Croatia	2008
BA1531	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1532	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1533	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1534	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1840	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1841	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1842	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1863	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
BA1864	<i>Pterocirrus macroceros</i>	Banyuls, France	2009
MA1661	<i>Pterocirrus macroceros</i>	Canical, Madeira	2009
MA1662	<i>Pterocirrus macroceros</i>	E Porto Moniz, Madeira	2009
MA1527	<i>Pterocirrus macroceros</i>	Funchal, Madeira	2009
MA1528	<i>Pterocirrus macroceros</i>	Funchal, Madeira	2009
MA1658	<i>Pterocirrus macroceros</i>	Funchal, Madeira	2009
MA1659	<i>Pterocirrus macroceros</i>	Funchal, Madeira	2009
MA1660	<i>Pterocirrus macroceros</i>	Funchal, Madeira	2009
MA5396	<i>Pterocirrus macroceros</i>	Funchal, Madeira	2009
IT1843	<i>Pterocirrus macroceros</i>	Capri, Italy	2010
IT1844	<i>Pterocirrus macroceros</i>	Capri, Italy	2010
IT1787	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
IT1788	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
IT1789	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
IT1845	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
IT1846	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
IT1847	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
IT5391	<i>Pterocirrus macroceros</i>	Ischia, Italy	2010
GB5357	<i>Pterocirrus macroceros</i>	Cornwall, Great Britain	2015
GB6009	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2017
GB6011	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2017
GB6013	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2017
GB6014	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2017
GB6019	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2017
GB6020	<i>Pterocirrus macroceros</i>	Plymouth, Great Britain	2017
RO6007	<i>Pterocirrus macroceros</i>	Roscoff, France	2018
RO6008	<i>Pterocirrus macroceros</i>	Roscoff, France	2018
RO6010	<i>Pterocirrus macroceros</i>	Roscoff, France	2018

Appendix D

Evolutionary models selected in jModelTest .

Nereiphylla rubiginosa

PARTITION	EVOLUTIONARY MODEL
1 st & 2 nd position COI	GTR+I+G
3 rd position COI	HKY + G
ITS1	GTR+I+G
ITS2	GTR+G
28S	GTR+I+G

Pterocirrus macroceros

PARTITION	EVOLUTIONARY MODEL
1 st & 2 nd position COI	GTR+I+G
3 rd position COI	GTR+G
ITS1	GTR+I+G
ITS2	HKY+G
28S	GTR+I+G