University of South Bohemia in České Budějovice

Faculty of Sciences

Characteristics of the mitochondrial genome of the roundworm *Oscheius myriophila* (Rhabditidae)

Bachelor thesis

Hana Pechová

Supervisor: PaedDr. Martina Žurovcová, Ph.D.

Co-supervisor, specialist: RNDr. Vladimír Půža, Ph.D

České Budějovice 2017

Pechová H., 2017: Characteristics of the mitochondrial genome of the roundworm *Oscheius myriophila* (Rhabditidae), Bc. Thesis, in English – 48 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation:

Genus *Oscheius* represents specific group of nematodes which are facultatively entomopathogenic, i.e. they occur in transient stage between necromenic and parasitic life style. Therefore, it represents an interesting model for evolutionary studies; however, the taxonomy of the genus is still not clear. Analysis of the full mitogenome of *Oscheius myriophila* may help to enhance our knowledge on the selected nematode species with respect to the taxonomy, phylogeny as well as its biology in general.

I hereby declare that I have worked on my bachelor thesis independently and used only the sources listed in the bibliography.

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in shortened form resulting from deletion of indicated parts to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

České Budějovice, 15.5.2017

……

Hana Pechová

Acknowledgements:

On this place I would like to thank to my supervisor, PaedDr. Martina Žurovcová, Ph.D. for guiding me in the spheres of evolutionary genetics and bioinformatics, supportive approach in my research and many valuable advices. Also, big thanks belong to my co-supervisor and specialist RNDr. Vladimír Půža, Ph.D, who provided me the samples of *O. myriophila* and helped me to understand its taxonomy. Furthermore, I would like to thank my lab colleagues, namely to Mgr. Lucie Faktorová for her advices. And I immensely thank to my family for their support during the studies.

Content

1. Introduction

1.1 Mitochondria: structure and function

Mitochondria are double membrane organelles found in all eukaryotic cells. It is the place where ATP (adenosine triphosphate) molecules are synthesized. The ATP provides the energy that is necessary for synthesis and catalysis of biomolecules and is thereby essential for all biochemical processes in eukaryotic cells. The energy it contains is acquired by the oxidative metabolism of nutrients. ATP releases that energy when it undergoes hydrolysis.

Mitochondrion has its own genetic information, mitochondrial DNA (mtDNA) that occurs in the mitochondrial cytosol. The genetic role of mtDNA is conserved, typically seen as a circular molecule (Bendich, 1993) but the linear molecular arrangements are known as well (Coleman, 1991). The size of mitochondrial genome is diverse, ranging from less than 6 kilobase pairs (kbp) within *Plazmodium* species (*P.falciparum*, the human blood parasite causing malaria) to large and complex genomes of land plants exceeding 200 kbp. The coding-size between genera significantly differs. For *Arabinopsis thaliana* more than 80% of it sequence is noncoding within mtDNA compared to *Rickettsia prowazekii* and its 24% of noncoding sequence (Gray, 1998).

The mitochondrial DNA encodes a number of RNAs and proteins essential for formation of functional organelle (Tzagoloff, 1986) once the most of the genes for mitochondrial biogenesis resides in the nucleus (Alberts, 2001). Thus mitochondrion contains in large part of protein coding genes (PCGs) that are components of electron transport chain. Taking into account animal mtDNA the electron transport chain consists of respiratory complexes I (7 *nad* genes, ubiquinone oxidoreductase), III (*cob* gene, ubiquinol: cytochrome c oxidoreductase), IV (3 *cox* genes, cytochrome c oxidase) and complex V (2 *atp* genes, ATP synthase). The mitochondrial translation system is composed of small subunit (SSU) and large subunit (LSU) rRNAs, rarely with 5S rRNA and usually of up to 22 genes coding tRNAs (Lang et al., 2004).

1.2 Mitochondria: origin and evolution

Mitochondrion and its evolutionary origin has been subject of study for many researches within last 50 years. Based on the different approaches of scientists to the mitochondrial evolution it has been suggested broad variety of possible mitochondrial ancestors. Margulis claimed the mitochondria and plastids evolved from free-living bacteria as result of symbiosis within a eukaryotic host cell (Margulis, 1970).More than one decade later the most often accepted endosymbiotic hypothesis was described, proving that the mitochondrion arose by the engulfment of ancient free-living bacteria by a eubacterial ancestor thereby allowing the free-living proto-mitochondrion species to remain live aerobically (Gray 1982, 1983, 1999 and Cavalier-Smith1987b).Partially opposing character might have the study of Abhishek in 2011 referring that proto-mitochondria might have been related to facultatively anaerobic species as *Rhodobacter* capable of anoxygenic photosynthesis (Abhishek, 2011). However, recent observations have shown that the proto-mitochondrial relatives (by means of the bioenergetic systems) might have lived under anaerobic or micro-oxic conditions, gaining the energy from the nitrogen metabolism (Müller, 2012). Although, according to the most reliable phylogenetic analyses of protein coding genes (PCG) and rRNA genes of mitochondrial genomes, the mitochondrion is most closely related to a single ancestor that is related to the α-class of Proteobacteria (*Alphaproteobacteria*) (Yang et al.1985). It is a group of obligate intracellular parasites as well as *Rickettsia*, *Anaplasma* and *Ehrlichia* species which are considered to be the closest known eubacterial relatives of mitochondria (Gray, 1996). *Midichloria* species seems to be the only one representative of the *Rickettsiales* that have maintained the ancestral features typical of free-living bacteria. Other bacterial species have been considered to be related to the mitochondria like *Rhodospirillum*on the basis of extensive protein analysis (Esser, 2004), and *Caulobacter* on the basis of the sequence similarity of its homologues to the mitochondrial transport protein (Clements, 2009). However, the actual ancestor of mitochondrion remains still elusive (Gray, 2012).

1.3 Mitochondrion as a molecular marker

The mtDNA has proven to be useful tool in determination the phylogenetic relationships. More than thirty years it has been the most popular and mostly used molecular marker (Avise, 1987). In fact, the mitochondrial protein coding genes (PCG) are those that are valuable for studying relationships among various related eukaryotic species. As already mentioned the mtDNA is tightly packed with 13 mRNA PCGs, 2 rRNAs and 22 tRNAs and this arrangement appears to be highly stable among taxonomic classes or phyla. Nonetheless the rate of the evolution on the basis of mitochondrial nucleotide sequences level is very high, much faster than that on nuclear DNAs (Brown, 1979). The most frequent changes in mitochondrial genetic information are simple base substitutions, additions or deletions. According to Gyllensten (1985) and Avise (1987) the mtDNA is inherited strictly on the maternal mode, which was disproved by study of Gyllensten just six years later meaning that there is an evidence of paternal inheritance at lower frequency as in the maternal one (Gyllensten, 1991). Regarding the composition of mitochondrial DNA introns, repetitive DNA, pseudogenes and pseudogenes and sizeable spacer sequences between genes always seem to be absent, which proves that these genes had been probably lost in a common mitochondrial ancestor. Hence, the monophyletic origin enables the mtDNA to be used as one of the best molecular markers (Avise, 1987).

Furthermore, the amplification of mtDNA is not complicated since the molecule as well as the gene content is very small. It has about 16 to 20 kbp genome size length, in comparison to nuclear DNA that has around 3 000 Mbp, moreover mtDNA occurs in eukaryotic cells in multiple copies (around >12 copies) whereas nuclear DNA just in two copies.

1.4 Entomopathogenic nematodes

Entomopathogenic nematodes (EPNs) are a group of nematodes that live symbiotically with pathogenic bacteria in order to kill the insect host and to live from the tissues of the insect's cadaver and from the proliferating bacterial cells. It is a sort of mutualistic partnership between the entomopathogenic bacteria of *Xenorhabdus* and *Photorhabdus* species (Gram-negative Enterobacteriaceae) and their nematode hosts belonging to Rhabditida group (Steinernematidae and Heterorhabditidae). The capability to infect the insect's body by bacteria is a property of

just the modified third larval stage- infective juvenile (IJ) that is the only free-living stage of entomopathogenic nematodes. The IJ is stress-tolerant, non-feeding stage. As depicted on the Figure 1, IJs carry the pathogenic bacteria, actively seek for a possible insect host and penetrate into the host's body. Once the IJs enter the hemocoel they release the bacteria into the nutrient-rich hemolymph thereby causing death of the insect after the bacterial attack. The nutrients from the insect's carcass are used by the bacteria to proliferate and for reproduction of nematodes. The new generation of IJs reassociates with the pathogenic bacteria as all the nutrients from the insects' body were used and the juveniles leave the body and emerges into the soil in order to search for a new host.

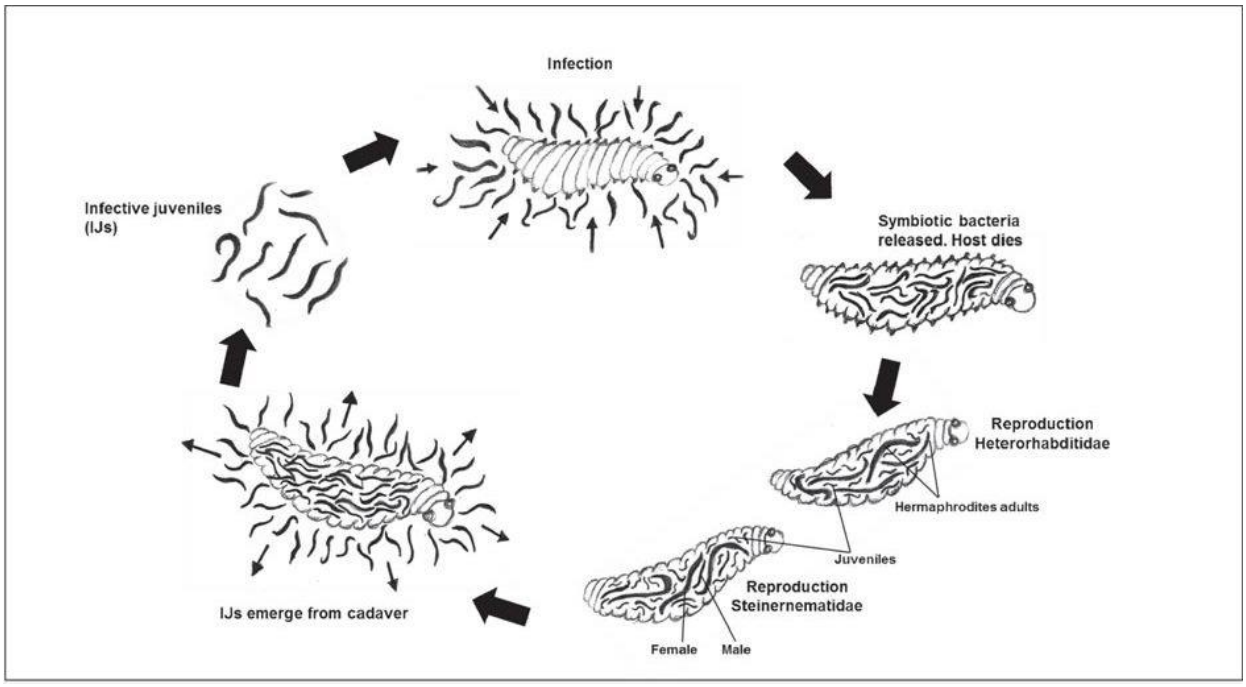

Figure 1: Life cycles of entomopathogenic nematodes (available from http://www.grainsa.co.za/how-compatible-is-entomopathogenic-nematodes-with-other-pest-management-strategies)

Different species of EPNs show variations in their behavior, host selection, effectivity and virulence of infection and tolerance to different environmental conditions. Moreover, many recent genetic methods have allowed studying the physiological, structural and evolutionary aspects of symbiosis that could not be achieved earlier. This criterion makes the EPNs and their

symbiotic partnership an ideal model for reconstruction and full understanding of the evolution of symbiosis (Flores-Lara et al., 2007).

As EPNs kill insects to obtain nutrition fundamental for their survivor, they have been found to be advantageous pest-control agents in agriculture. Moreover, the pathogenicity to vertebrates, humans and environment has been tested and it was proved, that there is either no risk to warm-blooded animals nor to plants (Boemare, 1996). EPNs have been even commercialized on some continents and are used in big-scale agricultures as well as by smaller growers. However, the efficacy of EPNs as biological control agents prevent them to be used as the only pest control thereby avoid the worldwide usage of chemical control. Nematodes are poorly tolerant to many factors and the most critical ones are moisture, temperature, pathogenicity for the targeted insect, and foraging strategy (Kung et al., 1991; Campbell et al., 2003; Grewal et al., 2005b). Many researchers have already worked on enhancement of the efficacy of EPNs against arthropods by enhancing their tolerance to multiple critical factors using strategies like artificial selection and genetic improvement via mutagenesis (Hiltpold, 2015 and Gaugler et al 1997). Although improvements of this kind have usually led to instabilities of newly generated genotypes (Glazer, 2014). The probable more effective approach is to use modern genetic and molecular tools yet not have been fully applied (Glazer, 2015).

1.4.1 Rhabditida

Order Rhabditida contains more than 20 known families so far, from whose Heterorhabditidae and Steinernematidae are highly adapted for entomopathogenicity and are the best described representatives of EPNs being considered as “true entomopathogens”. Steinernematidae (as the family’s name indicates) have been first studied by Steiner in 1923, followed by Weiser (1955). Few years later Poinar (1967) focused more onto the symbiotic bacterium *Xenorhabdus* and the location of the bacteria in stages of infective juvenile. Heterorhabditidae were first described by Poinar (1976) who first observed the symbiotic bacteria of these species that were able to produce bioluminescence (Poinar et al., 1980a) later described as bacterial genus *Photorhabdus* (Boemare et al., 1993). Beyond the morphological and ecological similarities, it has been proved that both groups (Heterorhabditidae and Steinernematidae) evolved independently from two common ancestors (Poinar 1993). The

phylogenetic study of these two Rhabditis families has shown that the Heterorhabditidae could have arisen from free-living bacterivorous antecedent (most closely related to free-living nematodes *Pellioiditis*) The entomopathogenicity of Steinernematidae is obviously caused by their bacterial symbionts of *Xenorhabdus* species, moreover the Steinernematidae are life-dependent upon their bacteria by means of gaining the nutrients and secondary metabolites as production of bacterial metabolism (Goodrich-Blair, 2007).

As a representative of the family Heterorhabditidae *Heterorhabditis bacteriophora* can be pointed out. It is the mostly used species in biological control of insects thanks to the symbiotic *Photorhabdus luminiscens* bacteria. It is an ideal model of living animal to study the parasitism, symbiosis, vector-born disease and heterogonic sex determination. It is typical for the alternating of sexual and hermaphroditic generations and unique for the presence of a dorsal "hook" on the tip of the head that allows entering the hosts' body more effectively (Bedding and Molyneux, 1982).

As the identification of new nematode species proceeds fast, another family of Rhabditidae from Rhabditoidea suborder has been described to cooperate with pathogenic bacteria in their manner of life cycle. One of the most important species is *Caenorhabditis elegans* as it has been used as a molecular phylogenetic marker within the phylum Nematoda. Next to the *Drosophila* species, *C. elegans* has been used as a model system for microevolutionary studies (Delattre and Félix, 2001). In those it has been shown, that *C. elegans* is even more powerful genetic model because of its hermaphroditic reproduction, fast production of homozygotes and its short life. These features make *C. elegans* an appropriate organism for analysis of phenotypic variations within phyla. The genes of *C. elegans* seem to undergo very fast molecular evolution in respect to other animal species and this evolutionary rate difference has placed the origin of *C. elegans* as well as all nematodes in early metazoan molecular phylogenies. So it has been disproved the close relation of nematodes to arthropods and vertebrates and the nematodes have been moved before the arthropod-vertebrate branches in phylogeny (Aguinaldo et al., 1997).

Species of genus *Oscheius* are commonly found in soil. In opposite to genus *Caenorhabditis*, they have been proved to share common ancestor with *Heterorhabditis* and some vertebrate parasites of Strongylida genus (Blaxter et al., 1998). And besides

Heterorhabditis, the entomopathogenicity has been established in some species of genus *Oscheius*.

Genus *Oscheius* consists of two subgroups: Dolichura and Insectivora (Sudhaus, 1976). The main difference between these subgroups is significantly smaller size and characteristic morphology of the Dolichura group species (Flemming et al., 2000). As depicted in Figure 2, phylogenetic scheme by Blaxter (1998) mentions 6 species belonging to the Dolichura group- *O. guentheri*, *O. dolichuroides*, *O. dolichura*, *O. tipulae* and other two Dolichura group species, from these *O. dolichura* is among the first described nematode species that reproduce in androdioceous manner (Schneider, 1866). On the other hand, according to the same scheme Insectivora group comprised of two species- *O. insectivora* and *O. myriophila*.

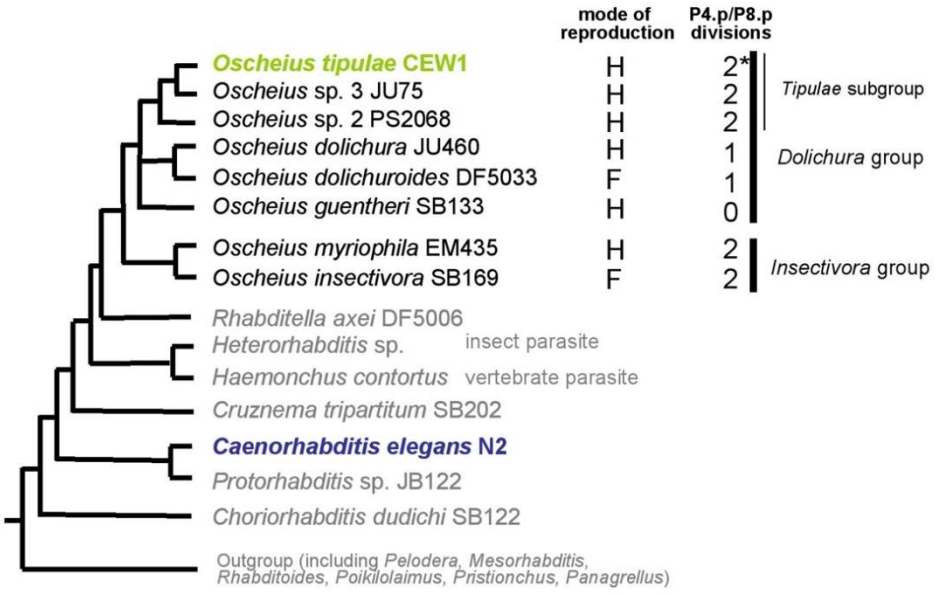

Figure 2: Phylogenetic relationships of some nematodes, showing the position of *Oscheius tipulae.* After Blaxter et al. (1998), Félix et al. (2001), The phylogenetic relationships of *Caenorhabditis* and other *rhabditids* (available at http://www.wormbook.org/chapters/www_genomesOtipulae/genomesOtipulae.html#sec1).

In recent years the characterization of *Oscheius* genus has progressed even further and according to Tabassum (2016) there are already 42 described valid species within the *Oscheius* genus, among them 28 insectivorous species and 14 species belong to the Dolichura group (see Table 1).

From those newly described species from the Dolichura group reported to be associated with insects, some of them have been indicated as potential insects' pathogens in a common experiment where laboratory insect larvae *G. mellonella* were used as a model host (Zhang, 2008; Ye, 2010, Abebe, 2010, Torrini et al., 2016 and Tabassum, 2016).

Table 1: Described species of *Oscheius* of Insectivora and Dolichura group (Tabassum at al., 2016); species with asterisk indicate potential insect pathogens

Insectivora group		**Dolichura group**
*Oscheius amsactae**	*O. nadarajani*	*O. bengalensis*
O. adrassyi	*O. necromenus**	*O. dolichura*
*O. carolinensis**	*O. niazii**	*O. dolichuroides*
O. caulleryi	*O. rupaekramae*	*O. dux*
O. ciceri	*O. rugoaensis*	*O. guentheri*
*O. chongmingensis**	*O. shamimi*	*O. janetti*
O. colombianus	*O. siddiqii**	*O. latus*
O. esperancensis	*O. wohlgemuthi*	*O. onirici**
O. gingeri	*O. citri**	*O. pheropsophi*
O. hussainii	*O. cobbi**	*O. pseudodolichura*
O. insectivorus	*O. esculentus**	*O. sechellensis*
O. luviani	*O. cynodonti**	*O. tereticorpus*
O. maqbooli	*O. sacchari**	*O. tipulae*
O. myriophila	*O. punctata**	*O. zarinae*

1.4.2 Characteristics of the roundworm *Oscheius myriophila*

O. myriophila has been first described by Poinar in 1986 as an associate of *Oxidis gracilis*, a garden millipede. It is a hermaphroditic species that can be distinguished from other Rhabditida by the typical structure of the spicules and bursa and by the characteristic size of stoma, metacorpus, tail and rectum. Very typical for *O. myriophila* as well as for related nematode *O. necromena* is the manner of their life cycle described by Sudhaus et al. (1989). The infective juveniles penetrate into the body of the developing millipede (or they enter the millipede's digestive system with the food); they enter the gut and other body cavities and develop to the fourth-stage individual in the hosts' hemolymph. In this stage they are very active and can even enter the Malpighian tubules but they cannot ever reach the adult-stage as long as the millipede lives. The fully developed adult capable of hermaphroditic reproduction

occurs only when the host dies and the bacteria spread among the host body. Therefore, the bacteria that enter the dead hosts' body are essential for *O. myriophila* as a nutritional source (Poinar, 1986). The life cycle is completed as the conditions inside the hosts' body equals the environmental conditions in the soil and new generation of infective juveniles is produced. This kind of parasitism was named as necromeny ("waiting for cadaver", Sudhaus et al. 1989) and according to Dillman (2012), it might have evolved as an intermediate evolutionary stage between entomopathogeny and parasitism.

However, the ecology of *O. myriophila* is still not clear. It is possible, that nematodes can be both, necromenic and parasitic depending on the bacteria and host they meet. Consequently, *O. myriophila* might be called facultative entomopathogenic nematode. That means, when the nematode encounters some pathogenic bacterium and carries the bacterium inside the host's body on its cuticle, they kill the host. On the contrary, when the nematode encounters and carries only a harmless bacterium into the insect's body, then they would be called necromenic organisms. Nematodes described by Poinar and Sudhaus (1986 and 1989) might have actually been isolated without their pathogenic bacteria or with just non-pathogenic bacteria. Hence, the specification the *O. myriophila* as necromenic species (only) might have been precocious.

So far genome of *O. myriophila* hasn't been analysed completely. There are only few nucleotide sequences of small and large subunits of ribosomal RNA genes available in GenBank which makes this nematode species benign for genome analysis *de novo*. Moreover, that *O. myriophila* can be classified as potential EPN (Sudhaus et al., 1989 and Dillman, 2012) it is of high interest to specify this kind of potential pathogenicity to the insects. Also, comparing *O. myriophila* to other *Oscheius* species (that in contrast use the real entomopathogenic symbionts to kill the insect host) may enable to bring closer view on the phylogeny of these peculiar roundworms.

2. Specific aims

- Sequencing and annotation of the mitochondrial genome of the nematode *Oscheius myriophila*
- Comparison of mtDNA of *O. myriophila* with mtDNAs of related Rhabditis (*Caenorhabditis elegans*, *Caenorhabditis briggsae*, *Oscheius chongmingensis* and *Heterorhabditis bacteriophora*) and determination of their phylogenetic relationships using more distinct Ascaridida species (*Ascaris suum*)

3. Material and Methods

3.1 Materials

Samples of *O. myriophila*, indicated as JU1386 originated from the collection of Marie Ann-Félix laboratory. The genotype might belong to wild isolate since it was isolated from the surface of a millipede feeding on rotting banana flowers placed on sample of soil nearby, all collected on the island of La Réunion by Valérie Robert and Loïc Sablé in early January, 2008.

3.2 DNA isolation

For the extraction of the mtDNA from the target species, two types of kits were utilized (in both protocols approximately 200µL of nematodes were used):

3.2.1 Isolation with GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific)

Isolation of DNA was accomplished according to the manual of the Thermo Fisher Scientific Company. A protocol for genomic DNA purification from mammalian tissues was used in order to disrupt the strong cuticle that encloses the body of nematode. To increase the yield of DNA, the elution step was repeated with additional 200 µL of Elution Buffer. The purified DNA from the first elution was used immediately for PCR reaction whereas the second extraction (diluted one) was stored in freezer at -20°C for later usage.

3.2.2 DNA extraction with Extraction Buffer (Cox and Herbert, 2001; Frati et al. 2001)

To the sample of nematodes 20 µL of Extraction Buffer (consisted of 558 µL ddH_2O, 100µL 100 µL 10x PCR Buffer with $MgCl_2$ (TopBio s.r.o.), 10 µL 1% Tween 20 and 5µL proteinase K (100 mg/mL)) was added, vortexed and spinned for few seconds at maximum speed and let to freeze at -20°C for 20 min. After that the sample was incubated at 65°C for 60 min followed by inactivation of proteinase K by increasing the temperature of incubation to 95°C for 8 min. The obtained extracted sample was immediately used for PCR reaction.

3.3 PCR amplification

The mtDNA segments were amplified using polymerase chain reaction (PCR) that was accomplished using the thermocycler TProfessional TRIO (Biometra GmbH). Solution for PCR reaction was always prepared on ice in 0.2 mL PCR-tube. Based on the concentration of template mtDNA 1-2 µL (more template mtDNA when diluted after reamplification) was added.

3.3.1 Reaction mixture

Different sets of chemicals were used in order to enhance the initialization of PCR reaction where the other reaction mixtures failed. It often occurred so at "problematic" AT-rich regions for whose more sensitive ExTaq (TaKaRa Bio, Inc.) and PrimeStar GXL Taq (TaKaRa Bio, Inc.) reaction mixtures worked well. For more "cooperative" regions of mtDNA UNIS Taq (Top-Bio s. r. o.) reaction mixture was used.

Concentrations and amounts of reaction mixtures are summarized in Table 2.

Table 2: Reaction mixtures used for PCR reactions

	UNIS [µL]	**PrimeStar GXL [µL]**	**ExTaq [µL]**
ddH_20	7.25	6.25	7.75
10x Buffer (10x Unis/ 10x PrimeStar GXL/ 10x TaKaRa)	1.25	2.5	1.25
dNTPs (2.5 mM)	1	1	1
Primers (forward/reverse) (5 µM)	0.75/0.75	0.75/0.75	0.75/0.75
Taq polymerase (UNIS Taq (5 U/µL)/ PrimeStar GXL Taq (1.25 U/ µL)/ ExTaq (5 U/ µL)	0.1	0.25	0.05
Template DNA	1-2	1-2	1-2
Total amount	12-13	12.5-13.5	12.5-13.5

3.3.2 PCR primers

Amplification of the genome was done using primers designed in previous studies of *Oscheius chongmingensis* (Jarošová et al., 2015) and *Oscheius guentheri* (Žurovcová, unpublished). For segments where these primers could not anneal properly amplification was

done using specific primers that were designed by the "primer walking" method. List of successful primers is presented in Table 3.

Table 3: Universal and specific primers used for PCR

Primer	Sequence	Direction
Nem_ND6for (2)	TTCTGGTGCTTTACGTAAAG	F
Nem_12S-960f (1)	TCTAGGTTTGATTTACTGARTAGG	F
Nem_ND1_2327f (1)	CGKGCTCCTTTCGATTTTTC	F
Nem_ND2-3872f (1)	CTTCGACWGARTCTTTTAATTG	F
Nem_CytB-5325f (1)	TGCWATTTTRCGTGCTATTCC	F
Omyri-6297for (3)	AGGTTTTCTATGTCTGATGG	F
Omyri-6816for (3)	CAGAAATATTAGTTTTGGTTTG	F
Nem_COI-8703f (1)	ATTGGTTTAATTGGTTGTGTAGT	F
Nem_COI-8766f (1)	TCACGTGCWTATTTTWCDGC	F
Nem_16S-10767f (1)	TGAAATCTYRATCAAATGTTTT	F
Omyri-12316f (3)	GCGGTTCTATATTCAGGGG	F
Nem_ND5_12706f (1)	CCAGRGGTGCTGTAAGAAAA	F
Nem_Asn-1672r (1)	ACCTTTTGATTAACAGTCAAC	R
Nem_ND1-2536r (1)	ADATCATAACGATAACGDGGA	R
Nem_Ile-4296r (1)	GGAACTYTACCTTATCAAGA	R
Nem_CytB-5497r (1)	GGRTCTTCTACTRTACACTGACC	R
Nem_COIII-6383r (1)	ACWACATCAACAAAATGTCAA	R
Nem_COI-8091r (1)	CAATTACCAAAWCCACCAAT	R
Omyri-8956r (3)	CAACACCTGTTAAACCACCA	R
EafRev_8966 (2)	TTCATCACAATCTAACTCCAG	R
Nem_COII-10244r (1)	AGCACCACAAATTTCTGAAC	R
Nem_16S-10844r (1)	ATGTCCTCACGCTAAGACTG	R
ND5_12250r (1)	GGCTCTTATAGCTTTKGGTA	R
Nem_AT2-r 14507 (1)	ACTCCTATTACCCTGACAAG	R
Nem_Pro-33r (1)	CTCAGYAACTAATCTTARTRACC	R

Note: (1) designed for *O. chongmingensis*; by Jarošová et al., 2015

(2) designed for *O. guentheri*; by Žurovcová; unpublished

(3) specific primers designed for *O. myriophila*; by Pechová; unpublished

3.3.3 PCR profiles for amplification of individual fragments

The amplification was done using standard PCR profile that was based on following steps: pre-denaturation at 94°C for 1 min followed by 35 repetitions of 3 consecutive steps (denaturation at 94°C, 30 s, annealing at 51-59°C, 45 s, elongation at 68 or 72°C, 1 min -2 min 30 s) and the process was completed by post-elongation step at 72°C for 2 min and by further cooling of the reaction product at 16°C.

Settings that resulted in successful amplifications are listed in Table 4.

Table 4: PCR profiles used for successful amplification of individual fragments

Fragment	Primers	Annealing		Elongation		Size	Polymerase
		t [s]	T [°C]	t [s]	T [°C]	[kbp]	
500-1672	Nem_ND6for	45	51	90	72	0.8	UNIS
	Nem_Asn-1672r						
960-2536	Nem_12S-960f	45	55	90	72	1	UNIS
	Nem_ND1-2536r						
2327-4296	Nem_ND1-2327f	45	53	120	72	1.2	UNIS
	Nem_Ile-4296r						
2872-6383	Nem_ND2-3872f	45	55	90	72	1.2	UNIS
	Nem_COIII-6383r						
3872-5497	Nem_ND2-3872f	45	55	120	68	0.5	TaKaRa
	Nem_CytB-5497r						
3872-6383	Nem_ND2-3872f	45	59	90	68	0.6	TaKaRa
	Nem_COIII-6383r						
5325-6383	Nem_CytB-5325f	45	53	60	72	1	UNIS
	Nem_COIII_6383r						
6297-8091	Omyri_6297 for	45	53	90	68	0.6	UNIS
	Nem_COI-8091r						
6816-8091	Omyri_6816 for	45	53	90	68	1	UNIS
	Nem_COI-8091r						
6816-8956	Omyri_6816 for	45	53	90	68	0.8	UNIS
	Omyri_8956 rev						
8703-10244	Nem_COI-8703f	45	53	90	72	0.3	UNIS
	Nem_COII-10244r						
8766-10244	Nem_COI-8766f	45	53	90	68	1.1	UNIS
	Nem_COII-10244r						
8766-10844	Nem_COI-8766f	45	53	90	72	0.6	UNIS
	Nem_16S-10844r						
8766-10844	Nem_COI-8766f	45	51	120	68	0.8	TaKaRa
	Nem_16S-10844r						
10767-12250	Nem_16S-10767f	45	53	120	72	1.2	UNIS
	Nem_ND5-12250r						
12314-14507	Omyri-12316f	45	53	120	68	0.4	TaKaRa
	AT2-r 14507						
12314-33	Nem_Pro-33 (rev)	45	57	150	68	0.6	GXL
	Omyri_12316 for						
12706-33	Nem_ND5_12706f	45	55	150	68	0.5	TaKaRa
	Nem_Pro-33r						

3.4 Gel electrophoresis

The quantity and actual size of the PCR amplified fragments were verified by gel electrophoresis on 1.5% agarose gel. The gel was prepared by dissolving 3 g of agarose in 200 mL 1x TAE buffer (prepared from stock 50x TAE buffer: 121 g TRIS; 28.55 mL acetic acid; 50 mL 0.5M EDTA at pH 8.0 and 421.45 mL H_2O) in microwave for approximately 3 min. Dissolved agarose solution was let to cool down under cold tap water till ca. 50°C to speed up the solidification of the gel. Before the cooled gel was poured into the prepared form, 15 μL of ethidium bromide (EtBr, 0.5 μg/mL) was added and the liquid gel was slightly mixed and finally poured into the form. Plastic comb was partially inserted into the gel from the top to create small chambers for further loading of the PCR product. The gel solidified roughly 30 min in the dark to avoid bleaching out of EtBr. The stiff gel was stored in buffer that was made of and stored at 4°C for future use.

Onto the prepared gel 3 μL of PCR product mixed with 1.5 μL of loading dye (700μL ddH_2O, 300 μL glycerol and 0.5 g bromophenol blue) was loaded. For approximate determination of the size of PCR product 4 μL of size marker Lambda DNA/EcoRI+Hind III (FERMENTAS) was mixed with 1 μL of bromophenol blue and loaded onto the same gel so that the product fragments could be compared with fragments of DNA present in the size marker. The electrophoresis was run at 120V for 45 min. The separated fragments were observed under UV transilluminator. As an example of the PCR product separation serves Figure 3.

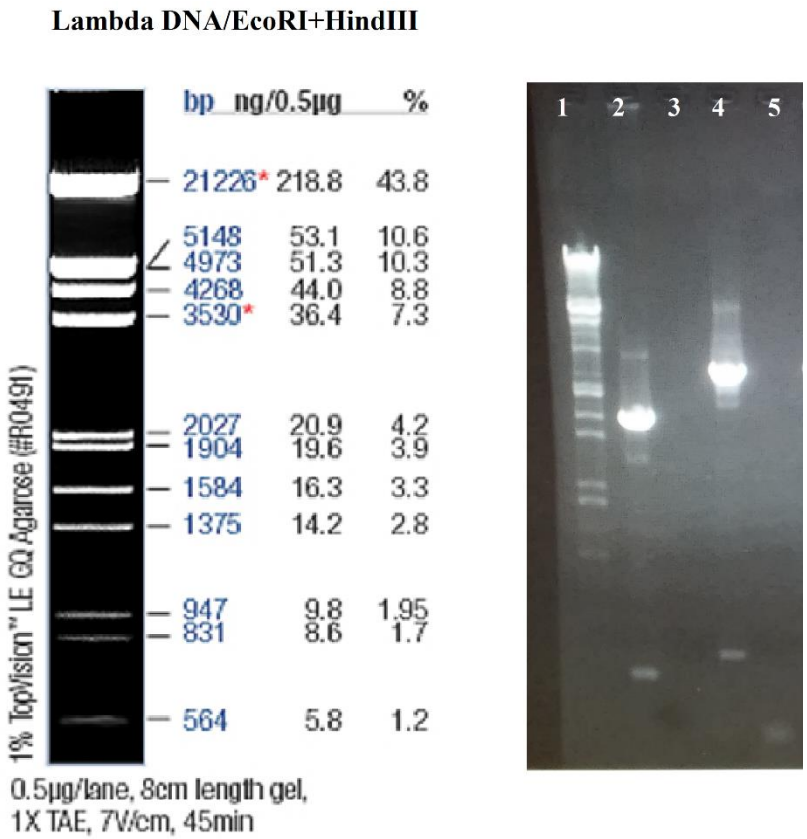

Figure 3: Left: Lambda DNA/EcoRI+HindIII size marker; right: picture of PCR product separation, 1: Lambda DNA/EcoRI+HindIII, 2: successfully amplified fragment (3872-53497); 3 and 5: negative controls; 4: successfully amplified fragment (3872-6383).

3.5 Purification of PCR products

After verification of successful amplification of mtDNA fragment it was needed to be purified from unwanted rests from the reaction (e.g. free nucleotides), which was done with the use of 2 enzymes. To each PCR product of approximately 9 µL, mixture of 0.5 µL Exo (Exonuclease I (20 U/ µL, Fermentas) and 1 µL of FastAP (Thermosensitive Alkaline Phosphatase; 1 U/ µL; Fermentas). The mixture was incubated in thermocycler at 37°C for 30 min followed by enzyme inactivation at 80°C for 15 min.

3.6 Sequencing reaction

Purified PCR products were directly used for sequencing. The sequencing reaction was made of 5 μL of the purified product mixed with 5 μL of respective primer (5 μM) in 1.5 mL Eppendorf tube. If the size of the product exceeded 1.5 kbp, two sequencing reactions were prepared, each with one of the amplification primers (forward and reverse) in order to gain as long fragment from the sequencing reaction as possible. After that samples were sent to GATC Biotech Company for further processing.

3.7 Biostatistics and annotation

3.7.1 Sequence alignment

The quality of the sequencing reaction was checked by analysis of the chromatograms in Bioedit 7.2 software (Hall, 1999; available at http://www.mbio.ncsu.edu/BioEdit/bioedit.html), in which the mtDNA bases must show proper separation of bases represented by distinct peaks. Subsequently, the nucleotide sequences were verified to belong to the Nematoda phylum by using BLAST search (Basic Local Alignment Search Tool; available at https://blast.ncbi.nlm.nih.gov/Blast.cgi) in the GeneBank NCBI database.

Continuous mtDNA sequence (contig) from the sequenced mtDNA fragments was created using a reference mitogenome sequence of *O. chongmingensis* (GenBank accession number KP257594). The alignment and necessary adjustments of sequence were made in SeqManII (DNAStar ver. 4.0; DNAStar, Inc.), where newly sequenced fragments were assembled with the reference sequence.

3.7.2 Annotation

Basic mitogenome annotation was done using of MITOS WebServer (Bernt et al., 2013; available at http://mitos.bioinf.uni-leipzig.de/index.py) with the invertebrate mitochondrial genetic code, which was followed by manual verification of all the genes. Protein coding genes (PCGs) were confirmed in DOGMA (Wyman et al., 2004; available at https://dogma.ccbb.utexas.edu) and by alignment of PCGs to those of *Rhabditidae* nematodes.

Genes coding for tRNAs and rRNAs were checked in the same way. To verify the predicted tRNAs and their secondary structure, the more specific software ARWEN (Laslett and Canbäck, 2008; available at http://mbio-serv2.mbioekol.lu.se/ARWEN/) was used. Translation initiation (START) and termination (STOP) codons of PCGs were identified in ORF Finder (Open Reading Frame Finder; available at https://www.ncbi.nlm.nih.gov/orffinder/). Base composition of the mitogenome was analysed in MEGA 7.0.21 (Kumar et al., 2016; available at http://www.megasoftware.net).

Graphical representation of circular mtDNA was constructed in GenomeVx (Conant and Wolfe, 2007; available at http://wolfe.ucd.ie/GenomeVx/).

Gene order of *O. myriophila* was compared to order of genes of four related Rhabditida species (*Oscheius chongmingensis* (KP257594), *Caenorhabditis elegans* (NC_001328), *Caenorhabditis briggsae* (NC_009885), *Caenorhabditis tropicalis* (KM403565)) and one less related Rhabditida species (*H. bacteriophora).*

Following 6 complete nematode mitochondrial genomes were used for phylogenetic reconstructions: *O. chongmingensis*, *C. elegans*, *C. briggsae*, *C. tropicalis*, *H. bacteriophora* and *Ascaris suum* (NC_001327) as an outgroup. Concatenated nucleotide sequences of protein-coding genes were analysed. Phylogenetic reconstruction was conducted in in MEGA 7.0.21. (Kumar et al., 2016) using Neighbor-Joining (NJ), Maximum Likelihood (ML) and Maximum Parsimony (MP) methods. The robustness of the trees was assessed by bootstrapping (1000 replications). NJ and ML methods were based on DNA distance model. The best model found for NJ was Tajima-Nei (Gamma distribution (+G), Pairwise deletion). ML method was based on Tamura-Nei model (Maximum Composite Likelihood, Gamma distribution with Invariant sites (G+I), Complete deletion. The third tree was built based on the Maximum Parsimony (MP) method (no DNA model, Subtree-Pruning-Regrafting algorithm, Complete deletion). All the trees were rooted with the outgroup *A. suum*.

4. Results

In this work, almost complete mitogenome of *O. myriophila* was obtained. Of the presumed 15 kbp, 12 986 bp were sequenced and annotated. The analysed partial mitogenome was compared with five related species; the most related one from *Oscheius* genus (*O. chongmingensis*), two species from *Caenorhabditis* genus (*C. elegans*, *C. briggsae*) and one species from *Heterorhabditis* genus (*H. bacteriophora*). The phylogenetic reconstruction was done using the five related Rhabditida species with an outgroup species from *Ascaris* genus (*Ascaris suum*).

4.1 Mitochondrial genome of *Oscheius myriophila*

Sequence of mitochondrial genome of *O. myriophila* was obtained by PCR amplification and by "primer walking" method. The whole continual sequence of mtDNA was assembled in program SeqManII (DNAStar ver. 4.0; DNAStar, Inc.). Total length of mitogenome of *O. myriophila* might be estimated to approximately 15.4 kbp according to the mitogenome length of closely related *O. chongmingensis* (Jarošová et al., 2015) that served as a reference sequence. The size of mitogenome that was successfully sequenced contains 12 986 bases, segments that could not be analysed are approximately 2.5 kbp long. However, 11 protein coding genes, 16 tRNAs and 2 ribosomal RNAs were identified.

The graphical representation of the mitogenome is shown on Figure 4. From the picture it can be seen that two regions could not be sequenced: from the end (ca. 12 987 bp) till the beginning of the mtDNA (ca. 780) and between *tRNA-(Cys)* (ca. 9 410 bp) and *COX2* gene (ca. 9689). These regions were coloured in grey, named as the hypothetical genes. The region at the interface of the beginning and the end of the molecule should contain *tRNA-(Pro)*, *tRNA-(Val)* and *ND6* genes respectively. Transfer RNAs should occur in the region between *tRNA-(Cys)* and *COX2* gene. The region at the end of mtDNA should contain *tRNA-(Ala)* and "AT-rich" region. All the successfully identified genes are on the same DNA strand and are transcribed in the same direction (clockwise).

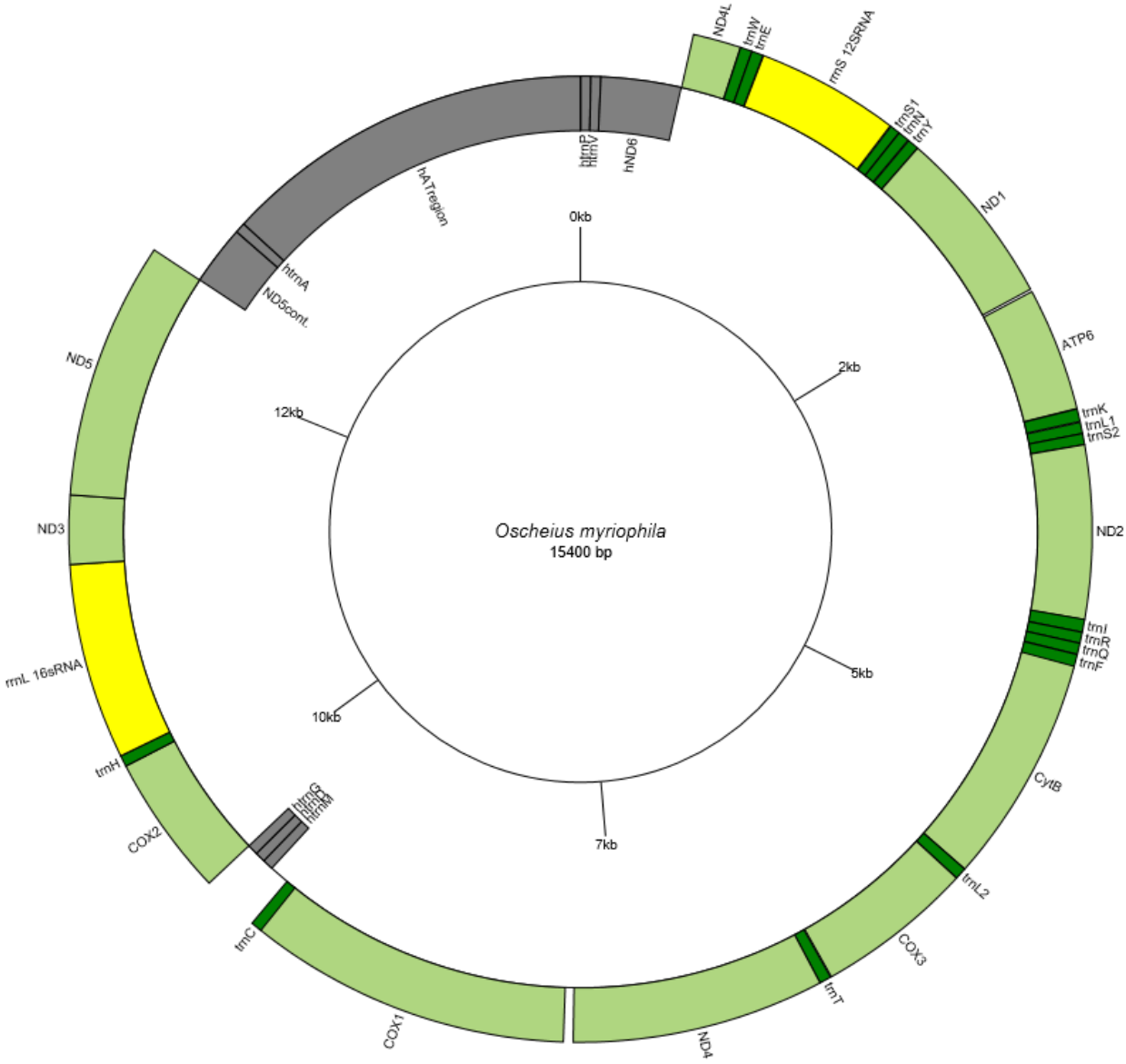

Figure 4: Circular map of mtDNA of *O. myriophila;* tRNAs coloured in dark green, PCGs in light green, rRNAs in yellow, unsuccessfully sequenced fragments (named as hypothetical h-regions) in grey. All genes are encoded in the same clockwise direction.

4.1.1 Base composition of the mitogenome

The base composition of the whole sequence was statistically evaluated by MEGA 7.0.21 program (Kumar et al., 2016; available at http://www.megasoftware.net) using the Sequence Data Explorer. It was found that the mitogenome of *O. myriophila* consists of 45.7% T, 30.2% A, 15.4% G and 8.7 % C, hence it is strongly A-T rich.

4.1.2 Protein coding genes (PCGs)

After annotation in MITOS (Bernt et al., 2013; available at http://mitos.bioinf.uni-leipzig.de/index.py) it was shown that 11 PCGs should be present in total. Subsequently by comparison of the predicted PCGs with those of the *O. chongmingensis*, all of the predicted PCGs were verified (Table 5). *ND5* gene was detected partially (because of low quality resolution of respective chromatograms), thus it was excluded from further analysis. *COX1* and *CytB* were sequenced fully but their chromatograms showed certain discrepancies too, hence they were not used for phylogenetic reconstruction. The gene that was not completely detected is *ND6*, which occurred in the "problematic" region at the beginning of the sequence.

Translation initiation (START) and termination (STOP) codons of PCGs were verified in ORF Finder (available at https://www.ncbi.nlm.nih.gov/orffinder/).

Four types of START codons were detected. Five PCGs (*ND4L*, *ATP6*, *ND4*, *COX2*, *ND3* and *ND5)* started with ATT codon, three PCGs (*ND1*, *ND2* and *CytB*) started with TTG, one gene (*COX3*) with ATA, and TTA was a start codon of *COX1* gene.

Most of the STOP codons started with TAA codon, *ND4L* had TAG codon, *ND1* had TTA codon, *ND2* used TAC stop codon and *COX1* might stop with TTA or TTG codon. Stop codon of *ND5* could not be resolved because of the partial sequencing of this gene.

Table 5: Verified PCGs found in the mitogenome of *O. myriophila*, *h-ND6* (in grey) marked as hypothetical gene, "?" means the insecure stop codons.

Name	Start	Stop	Strand	Length	Start codon	Stop codon
h-ND6	119	541	+	423	ATT	TAA
ND4L	546	779	+	234	ATT	TAG
ND1	1763	2634	+	872	TTG	TTA
ATP6	2647	3250	+	604	ATT	TAA
ND2	3428	4273	+	846	TTG	TTA
CytB	4503	5618	+	1116	TTG	TAA
COX3	5677	6444	+	768	ATA	TAA
ND4	6506	7735	+	1230	ATT	TAA
COX1	7782	9363	+	1582	TTA	TAG/TTA
COX2	9690	10382	+	693	ATT	TAA
ND3	11394	11729	+	336	ATT	TAA
ND5	11729	12986	+	1258	ATT	?

4.1.3 tRNA

The prediction of tRNAs in MITOS program was verified in ARWEN (Laslett and Canbäck, 2008; http://mbio-serv2.mbioekol.lu.se/ARWEN/). Out of presumed 22 tRNAs, 18 tRNAs were detected by MITOS, but only 16 of them were confirmed by ARWEN (Table 6). The exact positions of *tRNA-(Asp)* and *tRNA-(Gly)* could not be proved since their positions were considerably shifted and did not correspond to their positions in the reference sequence of *O. chongmingensis*. *tRNA-(Pro), tRNA-(Val), tRNA-(Met)* and *tRNA-(Ala)* were not detected at all.

Table 6: Overview of detected tRNAs in the sequence

tRNAs	Start	End	Strand	Length	Anticodon
tRNA-(Arg)	4335	4388	+	54	ACG
tRNA-(Asn)	1648	1703	+	56	GTT
tRNA-(Cys)	9354	9409	+	56	GCA
tRNA-(Gln)	4389	4443	+	55	TTG
tRNA-(Glu)	838	894	+	57	TTC
tRNA-(His)	10382	10436	+	55	GTG
tRNA-(Ile)	4272	4334	+	63	GAT
tRNA-(Leu1)	3319	3373	+	55	TAA
tRNA-(Leu2)	5619	5676	+	58	TAG
tRNA-(Lys)	3253	3314	+	62	TTT
tRNA-(Phe)	4445	4512	+	68	GAA
tRNA-(Ser1)	1598	1647	+	50	TGA
tRNA-(Ser2)	3374	3427	+	54	TCT
tRNA-(Thr)	6449	6505	+	57	TGT
tRNA-(Trp)	781	837	+	57	TCA
tRNA-(Tyr)	1707	1762	+	56	GTA

The secondary cloverleaf structures of the tRNAs of *O. myriophila* showed characteristic arrangement of the loops they are consisted of. As Jühling et al. (2012) has mentioned in his research, almost all the species of Chromadorea class are typically absent of T-arm (except D-armless serine tRNAs) within the tRNA sequences.

There were few exceptions from this fact, where structures of *tRNA-(Ile)*, *tRNA-(Lys)* and *tRNA-(Phe)* of *O. myriophila* showed cloverleaf structures with four stems (Figure 5).

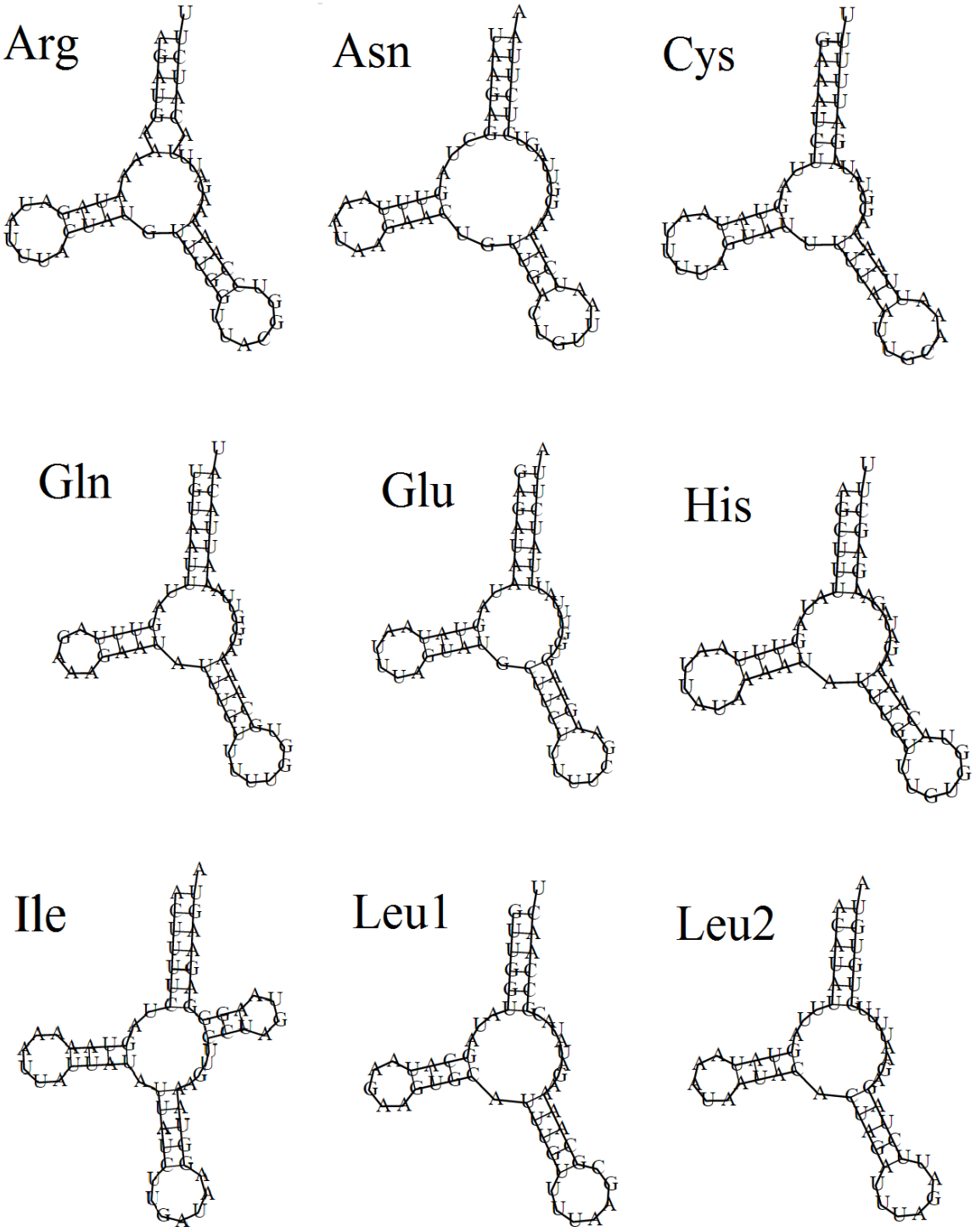

Figure 5: Cloverleaf structures of tRNAs (MITOS)

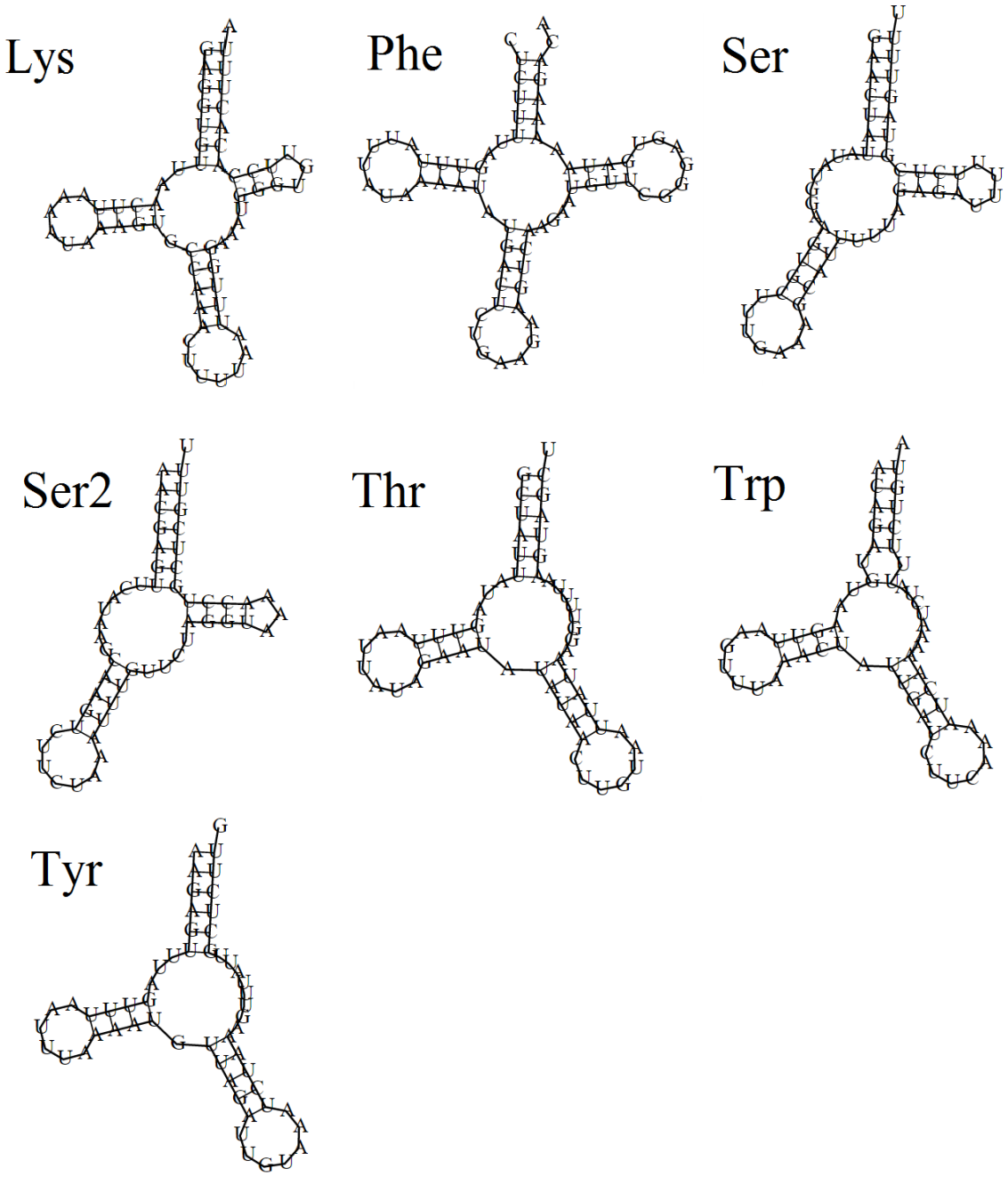

Figure 5 (cont.): Cloverleaf structures of tRNA (MITOS)

4.1.4 rRNA

In the sequence there were two rRNAs: rrnS (12S) at position 897 with the length of 1593 bp and rrnL (16S) at 10 437 bp that was 965 bp long. According to the reference sequence of *O. chongmingensis* both subunits of rRNA of *O. myriophila* should be complete.

4.1.5 Non- coding regions

The non-coding regions (so called intergenic nucleotides; IGN) were found in the mitogenome. These regions are not so large, the biggest IGNs occurred between *ND4* and *COX* I genes (7735-7782) and *ND1* and *ATP6* (2634-2647). Between certain genes there was an overlap of the sequences. The biggest overlaps occurred between *tRNA-(Phe)* and *CytB*, and *COX1* and *tRNA-(Thr)*. All the non-coding regions can be found in Table 7.

Non-coding repeats region ("AT-rich") could not be sequenced, however, its position is assumed to occupy the sequence at position from approximately 12 987 till 15 400, depending on the total actual size of the mtDNA molecule.

Table 7: Hypothetical intergenic nucleotides (h-IGN); positive values indicate gaps between the genes, negative values indicate overlapping of base between genes, "?" means the insecure region of non-verified tRNAs.

Genes	Start	End	IGN
ND4L	546	779	1
tRNA-(Trp)	781	837	0
tRNA-(Glu)	838	894	2
rrnS (12S-RNA)	897	1593	4
tRNA-(Ser1)	1598	1647	0
tRNA-(Asn)	1648	1703	3
tRNA-(Tyr)	1707	1762	0
ND1	1763	2634	12
ATP6	2647	3250	2
tRNA-(Lys)	3253	3314	4
tRNA-(Leu1)	3319	3373	0
tRNA-(Ser2)	3374	3427	0
ND2	3428	4273	-2
tRNA-(Ile)	4272	4334	0
tRNA-(Arg)	4335	4388	0
tRNA-(Gln)	4389	4443	1
tRNA-(Phe)	4445	4512	-10
CytB	4503	5618	0
tRNA-(Leu2)	5619	5676	0
COX3	5677	6444	4
tRNA-(Thr)	6449	6505	0
ND4	6506	7735	46
COX1	7782	9363	-10
trnC(gca)	9354	9409	?
COX2	9690	10382	-1
tRNA-(His)	10382	10436	0
rrnL (16s-RNA)	10437	11393	0
ND3	11394	11729	-1
ND5	11729	12986	0

4.2 Phylogenetic comparison with related species

4.2.1 Gene order

The overall order of genes in the mitogenome of *O. myriophila* remains the same as in the related species (Figure 6). There were only certain regions where the amplification of mtDNA was not successful thus the gene content at this segments was not established. As already mentioned, it comprises the region from the end of the sequence (from approximately 12 987 base) till the beginning of the sequence (approximately 780 base). Moreover, it is obvious that closer to the end of the sequence, there is a region of tRNAs *(Met, Asp, Gly)* that was not detected by any annotation tools even though this region (at position approximately 9410-9689 bp) was indicated as successfully sequenced. Position of *tRNA-(Ala)* could not be determined exactly, however, it should occur in the vicinity of the disputable "AT-rich" region.

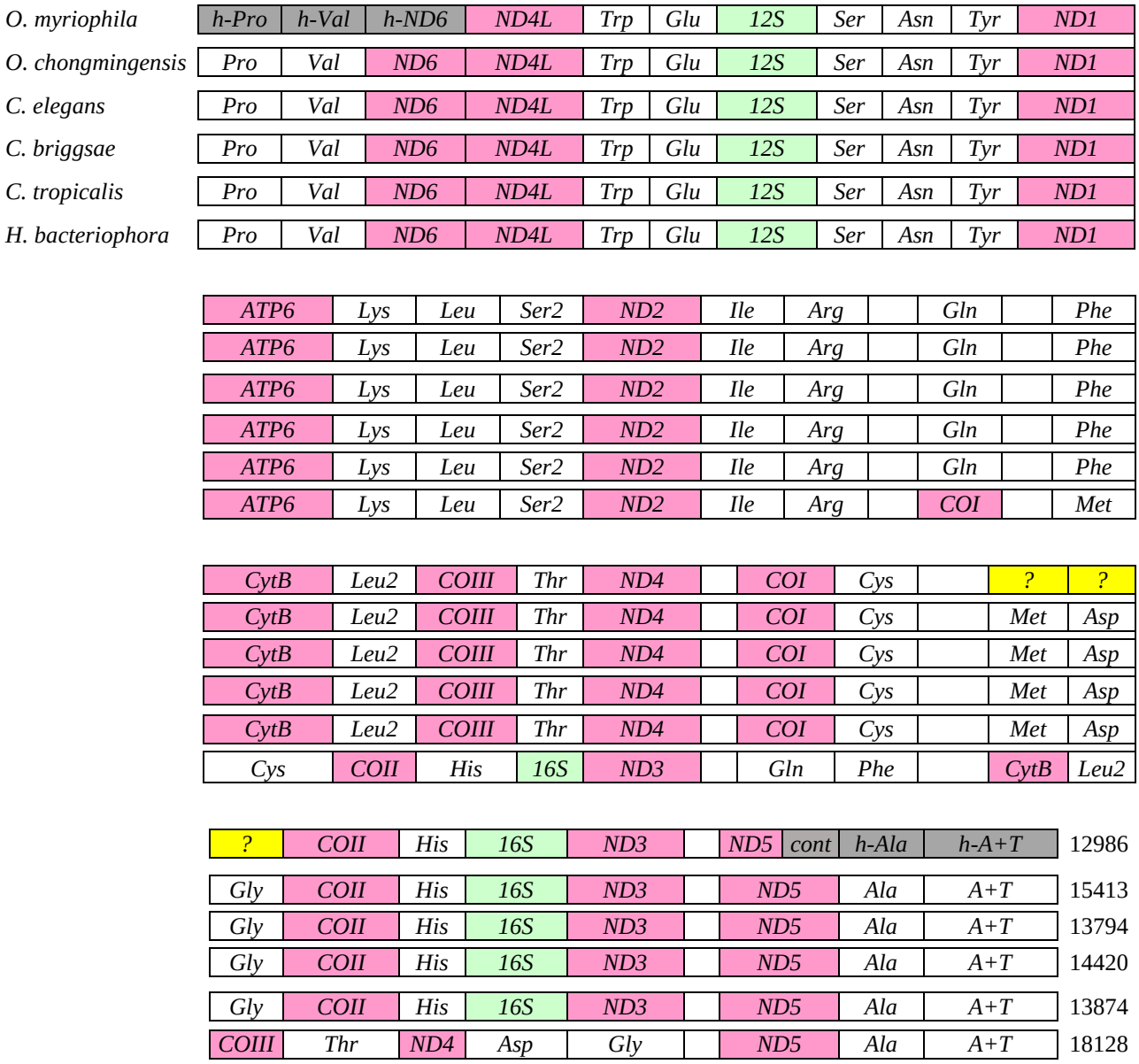

Figure 6: Gene order in mitogenome of *O. myriophila* and its related species; segments in pink stand for PCGs, green stand for rRNAs, white segments signify tRNAs, regions in yellow show insecure region of tRNA genes, segments in grey unsuccessfully sequenced (hypothetical h-genes).

4.2.2 Phylogenetic reconstruction

The phylogenetic reconstruction was done using the more conserved genes, the protein coding genes. For this purpose, sequenced PCGs (without *ND5, COX1* and *CytB)* from mitogenomes of related Rhabditidae species (*C. tropicalis, C. elegans, C. briggsae, H. bacteriophora* and *O. chongmingensis*) and more distinct *Ascaris suum* were separately aligned with PCGs of *O. myriophila* by Muscle (Codons) algorithm. After that, concatenated file of 8 PCGs of all nematode species was created.

Prior to the further analyses, concatenated file was used to find the best DNA model for phylogenetic reconstruction. Three types of phylogenetic trees were built; two of them were based on the detected model for phylogenetic reconstruction: Neighbor-Joining (NJ) Tree based on Tajima-Nei model with Gamma distribution (+G), and Maximum Likelihood (ML) Tree for which Tamura-Nei model with discrete Gamma distribution and evolutionary Invariant value (G+I) was used. The third tree was built with the Maximum Parsimony (MP) method based on Subtree-Pruning-Regrafting algorithm. All the phylogenetic trees were built with bootstrapping of 1000 replications and rooted with *A. suum* as the outgroup. Since the trees showed the same topology and almost identical percentages of clustering the sequences, only NJ phylogenetic tree is presented on the Figure 7. The other two trees (ML and MP) are shown in Appendix (Figures 10 and 11).

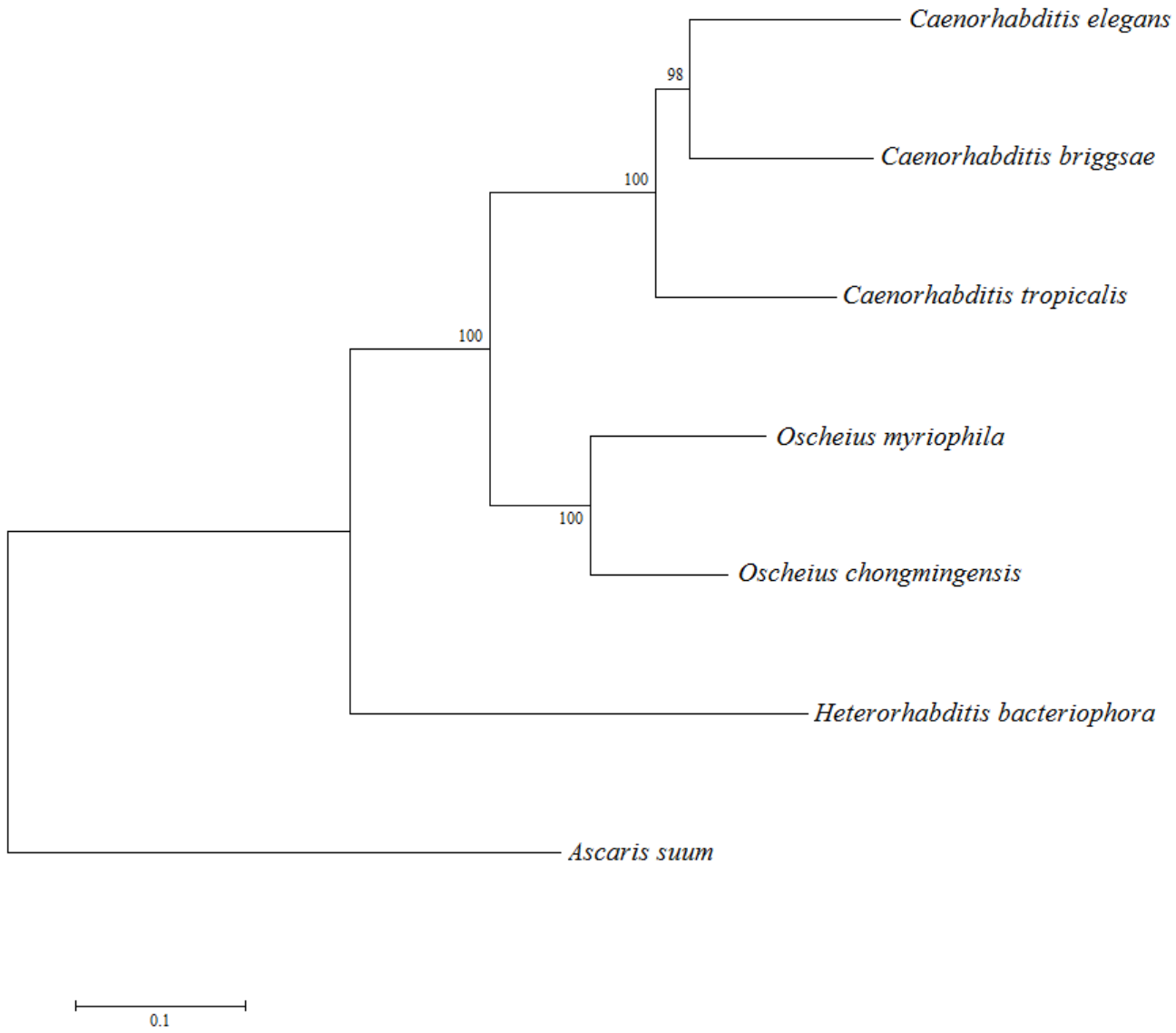

Figure 7: Neighbor-Joining phylogenetic tree of the selected nematode species using *A. suum* as an outgroup (Tajima-Nei model, Bootstrap 1000x, Gama distribution (+G), Pairwise deletion). Bootstrap values are next to the branches. The value below the scheme represents a distance scale (0.1: 10% difference between sequences). The tree was based on 8 complete PCGs.

5. Discussion

In the presented work, almost complete mitogenome of *O. myriophila* was described and analysed. However, approximately 2.5 kbp remained unresolved, although several different types of Taq polymerase and altered PCR profiles were used. The missing part should contain the "AT-rich" region, which is usually composed of approximately 80% of A and T bases. Such parts of DNA are generally very problematic for successful amplification. At the strings of AAAs or TTTs in the sequences the polymerase tends to slide over the DNA strand, which means that the amplification stops – both in PCR as well as sequencing reactions. Moreover, these regions are strongly attracted and might form unwanted secondary structures.

Also, there was other problematic part – between *tRNA-(Cys)* and *COX2* gene where the amplification seemed unsuccessful, although this region was assumed to be completely sequenced.

5.1 Size of the mitogenome

It was successfully sequenced 12 986 bp out of approximately 15 400 bp. The approximate total size of the mitogenome of *O. myriophila* was estimated according to the mitogenome of related species *O. chongmingensis*, that consists of 15 413 bp.

However, the size of the mitogenome of Rhabditidae species varies quite widely. Their size ranges from 13 794 bp (*Caenorhabditis elegans*) till 18 128 bp (*Heterorhabditis bacteriophora*). Hence, the actual size of complete mitogenome of *O. myriophila* might be even smaller or bigger than those of *O. chongmingensis*; especially since the variability is due to the repeated short motifs in the "AT- rich" region.

5.2 Base composition of the mitogenome

The base composition of the mtDNA of *O. myriophila* was found to be strongly A-T biased. Similar nucleotide composition was found in related Rhabditidae species. For instance, mtDNA of *O. chongmingensis* is composed of 33.44% A, 45.45% T, compared to *O. myriophila* with its 30.2% A and 45.7% T; however, the small difference could be due to the missing "AT-rich" region.

5.3 Protein coding genes (PCGs)

Out of 12 predicted PCGs in total, 11 of them were identified but only 8 PCGs were used for phylogenetic reconstruction. The PCGs that could not be resolved completely were *COX1, CytB, ND5* and *ND6*. *COX1* and *CytB* were localized at right positions, however, their sequences showed too many discrepancies in the chromatograms. Therefore, they could not be taken in account for reconstruction of relationships between nematodes as they would lead to incorrect phylogeny. The *ND5* gene was found by MITOS program within the sequence, however, it was found out that its sequence is partial (1258 bp out of approximately 1580 bp). The unresolved part of *ND5* gene (approximately 320 bases) interfered with the "AT-rich" region, thereby lowering the quality of chromatograms that caused exclusion of the *ND5* gene from phylogenetic analysis. Hence exact sizes and compositions of *COX1, CytB* and *ND5* need to be verified. On the other hand, *ND6* was completely not detected because of its close proximity to "A-T rich" region which was not sequenced successfully.

The absence of *ATP8* gene in mitogenome of *O. myriophila* is typical for all the nematodes, so it is not presumed to be found in the missing part. In accordance with other nematodes, all the genes were transcribed in the same direction on one strand.

Four types of start codons were detected The most of them were ATT start codons which are the most frequent start codons in PCGs of *O. chongmingensis* and other related Rhabditida as well. ATA start codons are observed at PCGs of both *Oscheius* species only once, but differently positioned (*O. chongmingensis* at *COX2*, whereas *O. myriophila* at *COX3)*. Positions of two TTG start codons are in accordance for both *Oscheius* species, both starting in *ND1* and *ND2* genes. *O. myriophila* possessed another TTG start codon in *CytB* gene. TTA was a start codon of *COX1* gene of *O. myriophila*, but TTA start codon in *O. chongmingensis* occurred in *CytB* gene.

As in PCGs of related species, most of the stop codons of *O. myriophila* were TAA. PCGs of *O. myriophila* also showed another three types of stop codons; TAG, TTA and TAC from whose only TAG stop codon is widely observed in related species. Presence of other types of stop codons might seem to be unusual, since they have not been observed in other related species. It was also unsure, it the *COX1* stops with TTA or TAG which corresponds to the

discrepancies that were observed in the sequence of *COX1* gene (Figure 8). Stop codon of *ND5* gene could not be resolved, because of non-sequenced ending part of this gene.

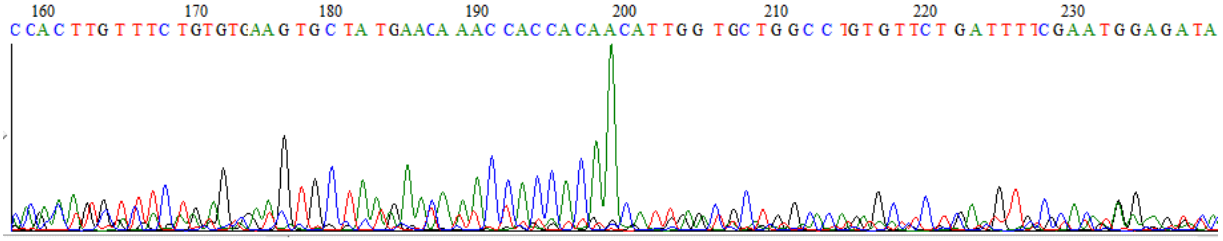

Figure 8: Example of low quality chromatogram; fragment of a sequence positioned at approximately 8250 bp (within *COX1* gene).

In order to clarify these improperly sequenced fragments, PCR product would need to be purified from interfering subbands, best by cloning method. This was not accomplished because of time and finance demands that cloning requires.

5.4 tRNA

With the use of two softwares 16 tRNAs out of 22 in total were verified. The exact positions of 6 tRNAs (*tRNA-(Ala), tRNA-(Gly), tRNA-(Val), tRNA-(Pro), tRNA-(Asp)* and *tRNA-(Met)*) could not be determined. The missing tRNAs occur probably at close proximity to the problematic AT-rich region.

It is also possible that some kind of translocation appeared in the region of *tRNA-(Met) tRNA-(Asp)* and *tRNA-(Gly)* (9410-9689 kbp), thus the actual position and structure of these tRNAs might not correspond to the predictions. That might be more specified once the sequencing of the mitogenome is complete.

The structures of tRNAs which were observed in *O. myriophila* are characteristic for nematode species with the T-armless-shape. Serine tRNAs (*tRNA-(Ser1)* and *tRNA-(Ser2))* showed D-armless structure which is also very typical for metazoan species. Figure 9 serves for comparison of four-stem secondary structures of human tRNA versus three-stem tRNA structures of metazoan organisms (nematodes), D and T-armless secondary structures are possible in enoplean nematodes as well.

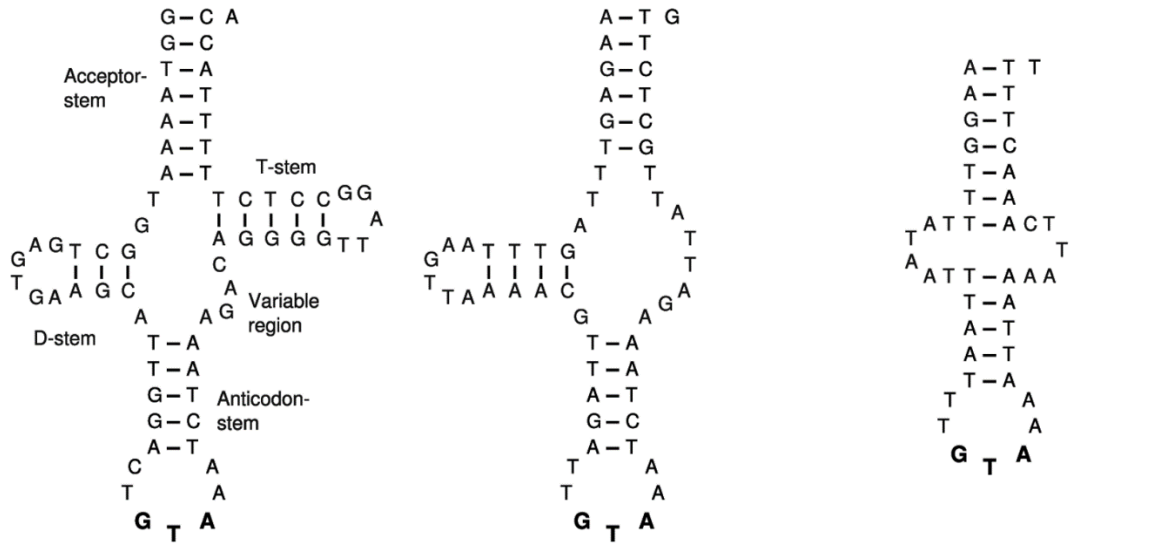

Figure 9: Human mitochondrial *tRNA-(Tyr)* has a canonical cloverleaf structure forming four stems. Its orthologue in *C. elegans* lacks the T-arm. In *R. nielseni* both arms are missing (taken from Jühling et al., 2012).

There were few exceptions from the typical composition of metazoan tRNAs. The cloverleaf structures of *tRNA-(Ile)*, *tRNA-(Lys)* and *tRNA-(Phe)* consisted of four stems, which is an unusual observation, since all the tRNAs (except two serine tRNAs) of metazoan species should miss the T-arm. The presence of full cloverleaves might be caused by their inexact identification and wrong prediction by the softwares.

5.5 rRNA

Two types of ribosomal RNAs were found; srRNA (697 bp) and lrRNA (956 bp). After comparison of rRNAs of *O. myriophila* with those of related *O. chongmingensis* it was found, that both the ribosomal subunits should be complete.

5.6 Non- coding regions

The non-coding regions were found quite frequently in the mitogenome. Their extent appeared to be very similar to those in related species. Size of non-coding region could not be determined in the insecure region of tRNAs (9410-9689 kbp). Since the assembled sequence

contained a lot of unresolved bases in these parts, these regions should be sequenced and proof-read.

5.7 Gene order in comparison with related Rhabditidae

The order of successfully amplified genes in the mitogenome of *O. myriophila* was the same as in four related (*Oscheius* and *Caenorhabditis*) species. *H. bacteriophora* is more distinct species; hence the gene order did not correspond to gene order of the other species. The order of genes at the non-sequenced region could not be determined exactly. Positions of genes in this region were only estimated after comparison with related *O. chongmingensis*.

5.8 Phylogenetic reconstruction

All the methods of phylogenetic reconstruction showed that *O. myriophila* is phylogenetically most closely related to *O. chongmingensis* since the bootstrap support of the cluster with two *Oscheius* species was calculated to 100. According to the strong evidence that *Caenorhabditis* branch stands in one clade with *Oscheius* branch (again, the highest bootstrap support), both the *Oscheius* species are more closely related to *Caenorhabditis* species. This observation is in contradiction with Blaxter et al. (1998) who have determined that *Oscheius* species share a common ancestor with *Heterorhabditis*, while *Caenorhabditis* species do not, meaning that *Oscheius* genus might be more distant. Since both presented *Oscheius* species belong to the Insectivora group, more research focused also on species from Dolichura group should resolve this dilemma.

6. Conclusion

Almost complete mitogenome of roundworm *Oscheius myriophila* was achieved by *de novo* sequencing. 11 out of 12 presumed protein coding genes were identified, both complete genes coding for small and large ribosomal subunit and 16 tRNAs. Several tRNAs showed some discrepancies of secondary structures which need to be improved in further research. The order of newly sequenced genes approved the order of genes of four related species of Rhabditidae family (*O. chongmingensis, C. elegans, C. briggsae* and *C. tropicalis*).

The phylogenetic analysis showed that *O. myriophila* is closely related to *O. chongmingensis* which is indicated as potential entomopathogenic nematode, thereby confirming that *O. myriophila* might be the potential insect parasite using the symbiotic pathogenic bacteria as well. Furthermore, both *Oscheius* species are probably more closely related to *Caenorhabditis* species than to *H. bacteriophora*, which is in contradiction to previous analysis by Blaxter et al. (1998).

7. References

ABEBE, E., JUMBA, M., BONNER, K., GRAY, V., MORRIS, K. (2010). An entomopathogenic *Caenorhabditis briggsae*. *J. Exp. Biol.* 213, 3223–3229.

ABHISHEK, A., BAVISHI, A., CHOUDHARY, M. (2011). Bacterial genome chimaerism and the origin of mitochondria. *Can. J. Microbiol.* 57, 49–61.

ALBERTS, B. (2001). Základy buněčné biologie: úvod do molekulární biologie buňky. Translated by Arnošt Kotyk. Ústí nad Labem: Espero Publishing, xxvi, Pp. 630, 18, 62, 30 p., ISBN 8090290604.

AQUINALDO, A. M. A., TURBEVILLE, J. M., LINFORD, L. S., RIVERA, M. C., GAREY, J. R., RAFF, R. A., LAKE, J. A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* 387, 489-93.

AVISE, J. C., ARNOLD, J., BALL, R. M., BERMINGHAM, E., LAMB, T., NEIGEL, J. E., REEB, C. A., SAUNDERS, N. C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge between Population Genetics and Systematics. Annual Review of Ecology and Systematics, Palo Alto, California: *Annu. Rev.* 18(1): 489-522.

BEDDING, R. A., MOLYNEUX, A. S. (1982). Penetration of insect cuticle by infective juveniles of *Heterorhabditis* spp. (Heterorhabditidae: Nematoda). *Nematologica* 28, 354–359.

BENDICH, A. J. (1993). Circular-mapping mitochondrial genomes may actually exist as multimeric linear molecules in vivo. *Curr. Genet*. 24, 279-290.

BLAXTER, M. L., DE LEY, P., GAREY, J. R., LIU, L. X., SCHELDEMAN, P., VIERSTRAETE, A. J., VANFLETEREN, J. R., MACKEY, L. Y., DORRIS, M., FRISSE, L. M., VIDA, J. T., THOMAS, W. K. (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature.* 392, 71–75.

BOEMARE, N. E., LAUMOND, C., MAULÉON, H. (1996). The nematode-bacterium complexes: biology, life cycle and vertebrate safety. *Biocontrol Science and Technology* 6, 333-345.

BOEMARE, N. E., AKHURST, R. J., MOURANT, R. G. (1993). DNA relatedness between *Xenorhabdus* spp (Enterobacteriaceae), symbiotic bacteria of entomopathogenic nematodes, and a proposal to transfer *Xenorhabdus luminescens* to a new genus, *Photorhabdus* gen. nov. *Int. J. Syst. Bacteriol.* 43, 249–255.

BROWN, W. M., GEORGE, M. JR., WILSON, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. U.S.A* 76, 1967-71.

CAMPBELL, J. F., LEWIS, E. E., STOCK, S. P., NADLER, S., KAYA, H. K. (2003). Evolution of host search strategies in entomopathogenic nematodes. *J. Nematol.* 35, 142–145.

CAVALIER-SMITH, T. (1987b). The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. *Ann NY Acad. Sci.* 503, 55–71.

CLEMENTS, A., BURSAC, D., GATSOS, X., PERRY, A. J., CIVCIRISTOV, S. (2009). The reducible complexity of a mitochondrial molecular machine. *Proc. Natl. Acad. Sci. U.S.A* 106, 15791–15795.

COLEMAN, A., THOMPSON, W. F., GOFF, L. J. (1991). DNA Analysis of Eukaryotic Algal Species. *J. Protozool*. 38, 129.

COX, A. J., HEBERT, P. D. N. (2001). Colonization, extinction and phylogeographic patterning in a freshwater crustacean. *Mol. Ecol.* 10, 331–386.

DELATTRE, M., FÉLIX, M.-A. (2001). Microevolutionary studies in nematodes: a beginning. *Bioessays* 23, 807–819.

DILLMAN, A. R., CHASTON, J. M., ADAMS, B. J., CICHE, T. A., GOODRICH-BLAIR, H., STOCK, P. S., STERNBERG, P. W. (2012) An Entomopathogenic Nematode by Any Other Name. *PLoS Pathog*. 8(3), 1-4.

ESSER, C., AHMADINEJAD, N., WIEGAND, C., ROTTE, C., SEBASTIANI, F. (2004). A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* 21, 1643–1660.

FLEMMING, A. J., SHEN, Z.-Z., CUNHA, A., EMMONS, S. W., LEROI, A. M. (2000). Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5285–5290.

FLORES-LARA, Y., RENNECKAR, D., FORST, S., GOODRICH-BLAIR, H., STOCK, P. (2007). Influence of nematode age and culture conditions on morphological and physiological parameters in the bacterial vesicle of *Steinernema carpocapsae* (Nematoda: Steinernematidae). *J. Invert. Pathol.* 95, 110–118.

FRATI, F., SPINSANTI, G., DALLAI, R. (2001). Genetic variation of mtCOII gene sequences in the collembolan *Isotoma klovstadi* from Victoria Land, Antarctica: evidence for population differentiation. *Polar. Biol.* 24, 934–940.

GAUGLER, R., LEWIS, E., STUART, R. J. (1997). Ecology in the service of biological control: the case of entomopathogenic nematodes. *Oecologia* 109, 483–489.

GLAZER, I. (2014). Genetic improvement and breeding of EPN: the race for the “super nematode”. *J. Nematol.* 46, 168.

GLAZER, I. (2015). Improvement of entomopathogenic nematodes: a genetic approach. In: CAMPOS-HERRERA, R. (Ed.), *Nematode Pathogenesis of Insects and Other Pests* (pp. 29–55). Springer.

GOODRICH-BLAIR, H. (2007). They've got a ticket to ride: *Xenorhabdus nematophila-Steinernema carpocapsae* symbiosis. *Current Opinion in Microbiology* 10, 225–230.

GRAY, M. W., LANG, B. F., CEDERGREN, R., GOLDING, G. B, LERNIEUX, C., SANKOFF, D., TURMEL, M., BROSSARD, N., DELAGE, E., LITTLEJOHN, T. G., PLANTE, I., RIOUX, P., SAIN-LOUIS, D., ZHU, Y., BURGER, G. (1998). Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acid Res*. 26, 865.

GRAY, M. W. (1982). Mitochondrial genome diversity and the evolution of mitochondrial DNA. *Can. J. Biochem.* 60, 157–171.

GRAY, M. W. (1983). The bacterial ancestry of plastids and mitochondria. *BioScience* 33, 693–699.

GRAY, M. W., SPENCER, D. F. (1996). Organellar evolution. In: ROBERTS, D., SHARP, P., ALDERSON, G., COLLINS, M. (Eds.), *Evolution of microbial life. Society for General Microbiology Symposium 54* (pp. 109–126). Cambridge University Press, U.K.

GRAY, M. W., BURGER, G., LANG, B. F. (1999). Mitochondrial evolution. *Science* 283, 1476–1481.

GRAY, M. W. (2012). Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* 4, a011403.

GREWAL, P. S., KOPPENHÖFER, A. M., CHOO, H. Y. (2005b). Lawn, turfgrass, and pasture applications. In: GREWAL, P. S., EHLERS, R. -U., SHAPIRO-ILAN, D. I. (Eds.), *Nematodes as Biocontrol Agents, CAB International* (pp. 115–146), Wallingford, U.K.

GYLLENSTEN, U., WHARTON, D., WILSON, A. C. (1985). Material inheritance of mitochondrial DNA during backcrossing of two species of mice. *J. Hered.* 76, 321-24.

GYLLENSTEN, U. WHARTON, D., JOSEFSSON, A., WILSON, A. C. (1991). Paternal inheritance of mitochondrial DNA in mice. *Nature* 352, 255 – 257.

HITPOLD, I. (2015). Prospects in the application technology and formulation of entomopathogenic nematodes for biological control of insect pests. In: CAMPOS-HERRERA, R. (Ed.), *Nematode Pathogenesis of Insects and Other Pests: Ecology and Applied Technologies for Sustainable Plant and Crop Protection* (pp. 187–206). Springer.

JAROŠOVÁ, A., PŮŽA, V., ŽUROVCOVÁ, M. (2015). The complete mitochondrial genome of the facultative entomopathogenic nematode *Oscheius chongmingensis* (Rhabditida: Rhabditidae). *Mitochondrial DNA A DNA Mapp. Seq. Anal.* 27(5), 3109-10.

JÜHLING, F., PÜTZ, J., FLORENTZ, C., STADLER, P. F. (2012). Armless mitochondrial tRNAs in Enoplea (Nematoda). *RNA Biology* 9(9), 1161-1166.

KUNG, S. P., GAUGLER, R., KAYA, H. K. (1991). Effects of soil temperature, moisture and relative humidity on entomopathogenic nematode persistence. *J. Invert. Pathol.* 57, 242–249.

LANG, B. F., LAVROV, D. V., BURGER, G. (2004). Mitochondrial Genome, Evolution. *Encyclopedia of Biological Chemistry*, Volume 2.

MARGULIS, L. (1970). Origin of eukaryotic cells. *Yale University Press,* New Haven, Connecticut.

MÜLLER, M., MENTEL, M., VAN HELLEMOND, J. J., HENZE, K., WOEHLE, C. (2012). Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* 76, 444–495.

POINAR, G. O., JR., THOMAS, G. M. (1967). The Nature of *Achromobacter nematophilus* as an insect pathogen. *J. Invert. Pathol.* 9, 510–514.

POINAR, G. O., JR. (1976). Description and biology of a new insect parasitic rhabditoid, *Heterorhabditis bacteriophora* n. gen. n. sp. (Rhabditida; Heterorhabditidae n. fam.). *Nematologica* 21, 463–470.

POINAR, G. O., JR., THOMAS, G. M., HAYGOOD, M., NEALSON, K. H. (1980a). Growth and luminescence of the symbiotic bacteria associated with the terrestrial nematode, *Heterorhabditis bacteriophora*. *Soil Biol. and Biochem.* 12, 5–10.

POINAR, G. O., JR. (1986). *Rhabditis myriophila* sp. n. (Rhabditidae: Rhabditida), associated with the millipede, *Oxidis gracilis* (Polydesmida: Diplopoda). *Proc. Helminthol. Soc. Wash.* 53, 232–236.

POINAR, G. O., JR. (1993). Origins and phylogenetic relationships of the entomophilic rhabditids, *Heterorhabditis* and *Steinernema*. *Fundamental and Applied Nematology* 16, 333–338.

STEINER, G. (1923). *Aplectana kraussei* n. sp., einer in der Blattwespe, *Lyda* sp. parasitierende Nematodenform, nebst Bemerkungen über das Seitenorgan der parasitischen Nematoden. *Zentralblatt für Bakteriologie, Parasitenkunde, Infektionskrankheiten und Hygiene* 59, 14–18.

SCHNEIDER, A. F. (1866). *Monographie der Nematoden*, Berlin, viii, Pp.357.

SUDHAUS, W. (1976). Vergleichende Untersuchungen zur Phylogenie, Systematik, Oekologie, Biologie und Ethologie der Rhabditidae (Nematoda). *Zoologica* 43, 1- 229.

SUDHAUS, W., SCHULTE, F. (1989). *Rhabditis* (*Rhabditis*) *necromena* sp. n. (Nematoda: Rhabditidae) from South Australian Diplopoda with notes on its sibling *R.myriophila* Poinar, 1986 and *R.caulleryi* Maupas, 1919. *Nematologica* 35, 15- 24.

TABASSUM, K. A, SHAHINA, F., NASIRA, K., ERUM, Y. I., (2016). Description of six new species of Oscheius Andrassy, 1976 (Nematoda: Rhabditida) from Pakistan with a key and diagnostic compendium to species of the genus. *Pakistan J. Nematol.* 34 (2), 109-161.

TORRINI, G., MAZZA, G., STRANGI, A., BARABASCHI, D., LANDI, S., MORI E., MENCHETTI, M., SPOSIMO, P., GIULIANI, C., ZOCCOLA, A., LAZZARO, L., FERRETTI, G., FOGGI, B., ROVERSI, P. F. (2016). *Oscheius tipulae* in Italy: Evidence of an Alien Isolate in the Integral Natural Reserve of Montecristo Island (Tuscany). *J. Nematol.* 48(1), 8-13.

TZAGOLOFF, A., MYERS, A. M. (1986). Genetics of mitochondrial biogenesis. Annu. Rev. Biochem 55, 249-85.

WEISER, J. (1955). *Neoaplectana carpocapsae* n. sp. (Anguillata, Steinernematidae), nový cizopasník housenek obaleče jablečného, *Carpocapsae pomonella. L. Věštník Československé Společnosti Zoologické* 19, 44–52.

YANG, D., OYAIZU, Y., OYAIZU, H., OLSEN, G. J., WOESE, C. R. (1985). Mitochodrial origins. *Proc. Natl. Acad. Sci. U.S.A* 82, 4443– 4447.

YE, W., TORRES-BARRAGAN, A., CARDOZA, Y. J. (2010). *Oscheius carolinensis* n. sp. (Nematoda: Rhabditidae), a potential entomopathogenic nematode from vermicompost. *Nematology* 12, 121-135.

ZHANG, C., LIU, J., SUN, J., YANG, S., AN, X. (2008). *Heterorhabditidoides chongmingensis* gen. nov., sp. nov.; (Rhabditida: Rhabditidae), a novel member of the entomopathogenic nematodes. *J Invert. Pathol* 98, 153–168.

Web sources:

https://blast.ncbi.nlm.nih.gov/Blast.cgi

https://www.ncbi.nlm.nih.gov/orffinder

Online softwares:

BERNT, M., DONATH, A., JÜHLING, F., EXTERNBRINK, F., FLORENTZ, C., FRITZSH, G., PÜTZ, J., MIDDENDORF, M., STADLER, P. F. (2013). MITOS: Improved de novo Metazoan Mitochondrial Genome Annotation [software]. *Mol. Phylogenet. Evol.* 69(2), 313-9. Available at http://mitos.bioinf.uni-leipzig.de/index.py.

CONANT, G. C., WOLFE, K. H. (2007). GenomeVx: Simple web-based creation of editable circular chromosome maps [software]. *Bioinformatics* 15; 24(6), 861-2. Available at http://wolfe.ucd.ie/GenomeVx.

HALL, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95-98. BioEdit updated version 7.2.6 (2017) [software]. Available at http://www.mbio.ncsu.edu/BioEdit/bioedit.html.

KUMAR, S., STECHER, G., TAMURA, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets [software]. *Molecular Biology and Evolution* 33, 1870-1874. Available at http://www.megasoftware.net.

LASLETT, D., CANBÄCK, B. (2008). ARWEN, a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences [software]. *Bioinformatics* 24, 172-175. Available at http://mbio-serv2.mbioekol.lu.se/ARWEN.

WYMAN, S. K., JANSEN, R. K., BOORE, J. L. (2004). DOGMA: Dual Organellar GenoMe Annotator [software]. Automatic annotation of organellar genomes with DOGMA, *Bioinformatics* 2004 20(17), 3252-3255. Available at https://dogma.ccbb.utexas.edu.

8. Appendix

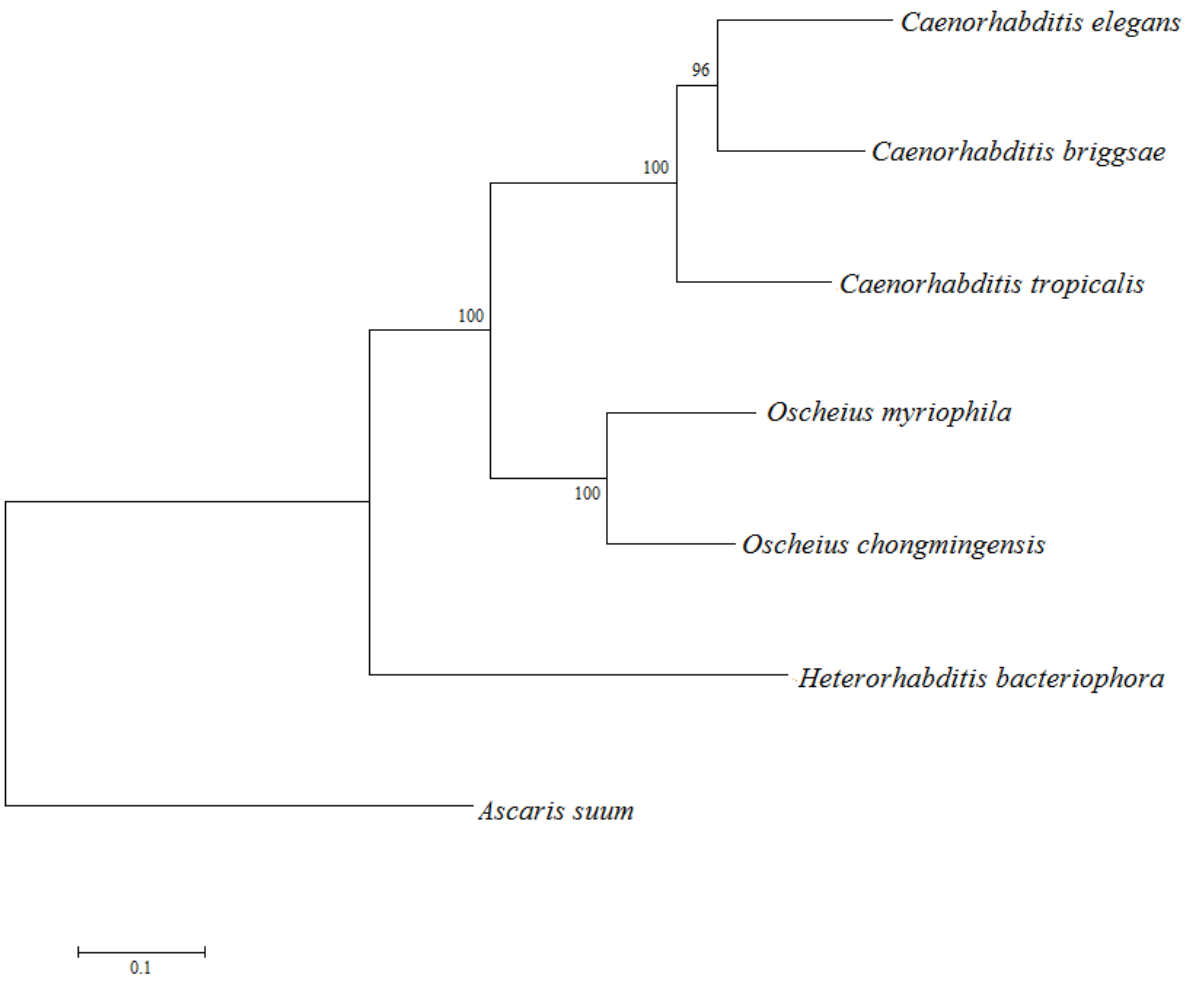

Figure 10: Maximum Likelihood phylogenetic tree of the selected nematode species using *A. suum* as an outgroup (Tamura-Nei model, Bootstrap 1000x, Maximum Composite Likelihood, Gamma distribution with evolutionary Invariant value (G+I), Complete deletion). Bootstrap values are next to the branches. The value below the scheme represents a distance scale (0.1: 10% difference between sequences). The tree was based on 8 complete PCGs.

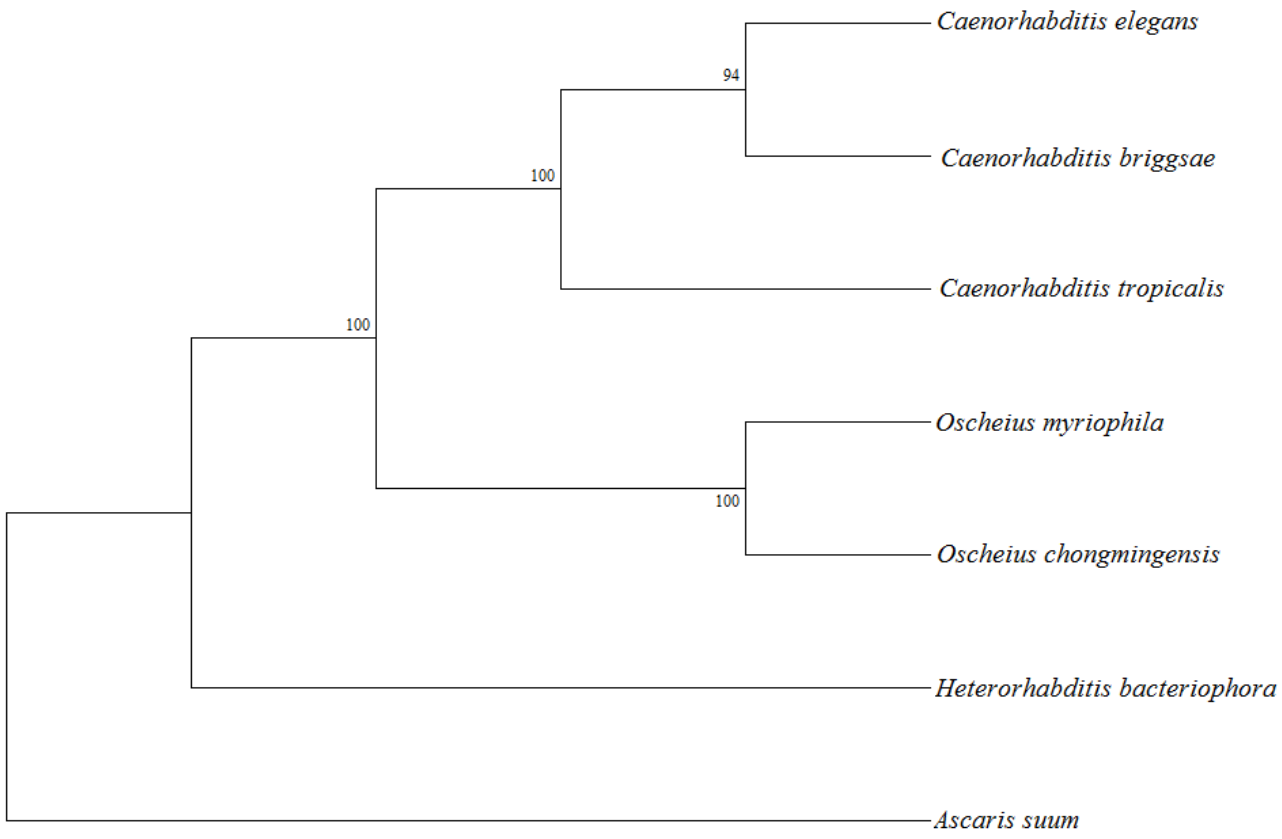

Figure 11: Maximum Parsimony phylogenetic tree of the selected nematode species using *A. suum* as an outgroup (Subtree-Pruning-Regrafting algorithm, Bootstrap 1000x, Complete deletion). Bootstrap values are next to the branches. The tree was based on 8 complete PCGs.