CHAPTER VI

DISENTANGLING TAXONOMY WITHIN THE *RHABDITIS (PELLIODITIS) MARINA* (NEMATODA, RHABDITIDAE) SPECIES COMPLEX USING MOLECULAR AND MORPHOLOGICAL TOOLS

Accepted with minor revision for publication in Zoological Journal of the Linnean Society as:

Derycke S, Fonseca G, Vierstraete A, Vanfleteren J, Vincx M, Moens T. Disentangling taxonomy within the *Rhabditis (Pellioditis) marina* (Nematoda: Rhabditida) species complex using molecular and morphological tools.

ABSTRACT

Correct taxonomy is a prerequisite for biological research, but currently it is undergoing a serious crisis, resulting in the neglect of many highly diverse groups of organisms. In nematodes, species delimitation remains problematic due to their high morphological variability. Evolutionary approaches using DNA sequences can potentially overcome the problems caused by morphology, but they are also affected by flaws. A holistic approach with a combination of morphological and molecular methods can therefore produce a straightforward delimitation of species.

The present study investigates the taxonomic status of some highly divergent mitochondrial haplotypes in the *Rhabditis (Pellioditis) marina* species complex by using a combination of molecular and morphological tools. We used concordance among phylogenetic trees of three molecular markers (COI, ITS, D2D3) to infer molecular lineages. Subsequently, morphometric data from nearly all lineages were analysed with multivariate techniques. The results showed that highly divergent genotypic clusters were accompanied by morphological differences, and we created a graphical polytomous key for future identifications. This study indisputably demonstrates that *R. (P.) marina* and *R. (P.) mediterranea* belong to a huge species complex and that biodiversity in free-living marine nematodes may seriously be underestimated.

INTRODUCTION

Nematodes have high species diversity as well as high abundances in marine, freshwater and terrestrial environments. Species delimitation in nematodes remains problematic mainly due to the high morphological variability among populations which reduces the number of diagnostic characters (Coomans 2002, Nadler 2002, Powers 2004). Molecular techniques and phylogenetic analyses may overcome this problem, and barcoding seems a promising tool to assess biodiversity in free-living nematodes (Floyd et al. 2002, Blaxter et al. 2005, Bhadury et al. 2006). However, it is difficult to decide when individuals are sufficiently distinct to discern them as different species based on sequence divergence. This is mainly due to the lack of a straightforward relationship between genetic divergence and reproductive isolation (Ferguson 2002), to the occurrence of theorethical observations (like incomplete lineage sorting, incongruence between gene and species trees Avise 1995, Nadler 2002), and to discrepancies between morphological and molecular data. Many examples of morphological stasis despite substantial genetic differentiation have been observed in nematode genera (e.g. *Caenorhabditis* Butler et al. 1981, *Globodera* and *Heterodera* Bakker & Bouwman-Smits 1988), while morphological differentiation between genetically similar species has also been reported (De Ley et al. 1999).

The problems of either morphological or molecular species delimitation can be resolved by applying a holistic approach, in which analyses of several independently evolving molecular markers circumvent the theorethical observations of the molecular method. Subsequently, the observed phylogenetic lineages can be used to aim more precisely for diagnostic morphological characters between nematode lineages (Coomans 2002, see also Chapter 5).

In a recent study on the phylogeny and systematics of the Rhabditidae, Sudhaus & Fitch (2001) considered *Pellioditis* Dougherty (1953) as one of the 15 subgenera within the genus *Rhabditis* Dujardin (1845). The subgenus comprises 18 species (Andrassy 1983, Sudhaus & Nimrich 1989, Gagarin 2001), only four of which inhabit the marine environment. *Rhabditis* (*Pellioditis*) *marina* Bastian, 1865 has been reported most frequently (Inglis & Coles 1961, Sudhaus & Nimrich 1989). The large intraspecific variability within *R.* (*P.*) *marina* is reflected in the description of a number of varieties, all but one of them later having been considered as synonyms of *R. (P.) marina* (Inglis & Coles 1961). A recent study based on mitochondrial and

nuclear DNA sequences revealed at least four cryptic species within *R. (P.) marina*, all of which were sympatrically distributed on a fairly small geographical scale (100 km) (Derycke et al. 2005) and were morphologically distinguishable (See Chapter 5). In addition, a temporal survey in which more than 1600 individuals were analysed led to the discovery of specimens with highly divergent DNA sequences (referred to as the Z lineages), of which the taxonomic position and phylogenetic relationships with the other lineages remained unclear (Derycke et al. 2006).

The present study aims to elucidate the phylogenetic and taxonomic uncertainties in the *R.* (*P.*) *marina* species complex through a combination of molecular and morphological methods. We performed phylogenetic analyses on three genes (mitochondrial COI, nuclear ITS and D2D3 regions) and used concordant tree topologies between these genes as evidence for independent evolutionary histories. We subsequently used multivariate analyses of morphological characters to investigate whether the observed genetic differences were accompanied by morphological differences and created a polytomous key for future identifications.

MATERIAL AND METHODS

SAMPLE COLLECTION AND PROCESSING

A detailed description of the sampling strategy and isolation protocol of *R. (P.) marina* has been described in Derycke et al. (2006). From the 1615 individuals analysed in that study, 11 individuals from Blankenberge, a coastal location situated in the northern part of the Belgian coastline (51° 19' N, 3° 8' E), possessed highly divergent mitochondrial COI haplotypes (called Z, Z2 and Z3). Prior to molecular analysis, each of the 1615 specimens were transferred to an embryo dish containing sterile artificial seawater, which was briefly heated to 60 °C to kill the nematodes. Each nematode was transferred in a drop of sterile distilled water on a glass slide and photographed digitally under a Leica DMR microscope equipped with a Leica DC 300 camera. These pictures served as a morphological back-up. Subsequently, each nematode was preserved in an Eppendorf reaction tube of 0.5 ml filled with acetone. Morphological and molecular data were thus obtained from the same specimens.

For the present study, we additionally used specimens collected in the frame of an ongoing larger-scale phylogeographic study of *R. (P.) marina* (see Chapter 9). Nematodes with Z haplotypes were collected in South Africa (Ngazi estuary) and eastern Mexico (Playa del Carmen, Yucatan). Collection sites for all lineages are summarized in Table 6.1. The morphological back-up of these nematodes was created by randomly picking 5 - 10 adult specimens from each location and mounting them into glycerin slides according to Vincx (1996). The remaining specimens were preserved on acetone for molecular analyses. Here, morphological and molecular data were thus obtained from different specimens. The link between both datasets was maintained because each location contained only one molecular lineage.

Morphological back-up	Digital pictures								Glycerin slides							
Lineage	PmI	PmII	PmIII	PmIV	Z	Z2	Z3	Z4	PmI	PmII	PmIII	PmIV	Z	Z2	Z3	Z4
Location																
Belgium - Blankenberge		X	X		X	X	X									
Belgium - Nieuwpoort	X															
The Netherlands - Westerschelde	X	X	X						X							
The Netherlands - Oosterschelde	X		X													
The Netherlands - Lake Grevelingen		X		X								X				
UK - Scotland (Westroy)										X						
USA - Massachusetts (Boston)											X					
Mexico - Yucatan																X
South Africa - Ngazi estuary													X	X		

Table 6.1: *Rhabditis (Pellioditis) marina*. Collection of specimens from each lineage. Specimens from pictures were collected in Belgium and The Netherlands (100 km), while specimens in slides were collected worldwide.

Molecular data

The DNA-extraction protocol, PCR-amplification, screening of genetic variation in the mitochondrial cytochrome oxidase c subunit 1 gene (COI) with the SSCP method and primer sequences are described in Derycke et al. (2005). The COI gene was amplified from 1 µl of genomic DNA, and with primers JB3 and JB5, and all samples with different SSCP band mobility patterns were sequenced with the ABI 3130XL capillary DNA sequencer. The sequencing reaction was performed with the BigDye Terminator v 3.1Mix (PE Applied Biosystems) under the following conditions: an initial denaturation of 2 min at 98 °C was followed by 40 cycles of denaturation at 98 °C for 10 s, annealing at 50 °C for 5 s and extension at 60 °C for 60 s. Both strands were sequenced using the amplification primers. DNA samples were stored at -80 °C so that multiple loci could be amplified from the same specimens.

Subsequently, we created a subset of individuals (n = 28) based on the COI topology and sequenced two nuclear loci. The highly variable ribosomal internal transcribed spacer region (ITS1, 5.8S and ITS2) was amplified as described in Derycke et al. (2005). The D2D3 expansion segments of the conserved 28S ribosomal DNA were amplified using primers D2A (5' ACAAGTACCGTGAGGGAAAGTTG 3') and D3B (5' TCCTCGGAAGGAACCAGCTACTA 3'). Amplification of this fragment started with a denaturation at 94 °C for 5 min, followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 58 °C for 30 s and extension at 72 °C for 60 s, and was terminated by a final extension period of 10 min at 72 °C. Both nuclear fragments were amplified from 1 µl genomic DNA, and both strands were sequenced with the amplification primers. New COI, ITS and D2D3 sequences are submitted in GenBank (Accession numbers: AM398819 – AM398833; AM399037 – AM399068).

Morphological data

Morphological variability in males and females containing the Z haplotypes was compared with that of the four lineages PmI, PmII, PmIII and PmIV (Chapter 5) in two ways. First, a detailed investigation was performed on specimens mounted into glycerin slides. These specimens were collected worldwide (Table 6.1) and were measured by video capture with the Leica Q500+MC software. A total of 29 morphological characters were considered, 11 of which were shape parameters (Table 6.2). A detailed description of all morphological characters can be found in Chapter 5, Appendix 5.1.

Morphometric characters	abbreviation	slides	pictures
Body length	L	X	X
Body width	W	X	X
Pharynx length	Ph	X	X
Pharynx corresponding body diameter	Phcbd	X	X
Position of the mid-bulb from the anterior end	Mid-bulb	X	
Nerve ring	nr	X	X
Midbulb diameter	M bulb diam	X	
Bulb diameter	Bulb diam	X	
Position of the anus	anus	X	X
Tail length	tail	X	X
Testis length	testis	X	X
Buccal cavity length	bc L	X	X
Buccal cavity width	bc W	X	X
Head length	head	X	X
Spicule length	spic	X	X
	Pos-intest	X	X
Anal body diameter	abd	X	X
Vulva	v	X	X
Shape parameters			
L/W	a	X	X
L/Ph	b	X	X
L/Tail	c	X	X
	c´	X	X
	spic/abd	X	X
	V%	X	X
	Pos-Int/abd	X	X
	testis/L	X	X
	nr%	X	X
	bcL/w	X	X
	bcL/head	X	X

Table 6.2: Morphometric characters and shape variables used for morphological identification of the Z lineages. Characters measured on specimens in slides and pictures are indicated with a cross.

Second, to compare the degree of morphological differentiation due to geographical variation, results from the first dataset were compared with those of

measurements from specimens collected in populations from Belgium and The Netherlands. These measurements were performed on a subset of characters (those that were used in the discriminant analysis, see next section) on the digital pictures (Table 6.2). Morphological and molecular data for this second analysis were from the same set of individuals. Drawings were made with a Leica DMLS microscope (Appendix 6.2).

DATA ANALYSES

Molecular data

Sequences of each locus (COI, ITS, D2D3) were aligned in ClustalX v.1.81 (Thompson et al. 1997) using default parameter settings (gap opening/gap extension costs of 15/6.66). We also amplified COI, ITS and D2D3 sequences from *R. (R.) nidrosiensis*, which was isolated from decomposing algae in The Netherlands (Derycke et al. 2005), and from *R. (P.) mediterranea* (New Zealand). Deeper phylogenetic relationships between our *R. (P.) marina* sequences and sequences of *R.(R.) nidrosiensis* and *R. (P.) mediterranea* were inferred from the nuclear dataset, which was rooted with sequences from the nematodes *Ancylostoma caninum* (D2D3: AM039739; ITS: DQ438079) and *Necator americanus* (D2D3: AM039740; ITS: AF217891) obtained from GenBank. Both species belong to the same order (Rhabditida) as *R. (P.) marina*.

An unambiguous alignment was obtained from the COI sequences, while indels were observed in both nuclear loci, especially for the ITS region. Hence, each of the two nuclear alignments was checked for unreliable positions in SOAP 1.2.a4 (Löytynoja & Milinkovitch 2001), using the following Clustalw parameter range: gap penalties were allowed to range between 11 and 19 with a two-step increase, and extension penalties ranged between 3 to 11, also with a two-step increase. We used a threshold level of 90 % for the D2D3 locus, which resulted in the removal of 17 unreliable positions. The threshold level for the ITS alignment was created as follows: first, we removed the outgroup sequences *N. americanus* and *A. caninum*. At the 90 % level, 713 out of 913 sites appeared unreliable. However, manual inspection of the alignment showed that many of these 'unreliable sites' did not contain much variation among sequences. Therefore, we lowered the threshold level until all indel events remained excluded. This was at the 60 % level. Second, we also excluded *R.(R.)*

nidrosiensis from the dataset, which resulted in the exclusion of 'only' 277 out of 903 positions at the 90 % level. Hence, the alignment of ITS sequences within *Pellioditis* was highly reliable at the 90 % level, and the threshold for the ITS alignment including *N. americanus*, *A. caninum* and *R. nidrosiensis* was set at 60 %.

Prior to phylogenetic analysis, the appropriate model of evolution for each locus was determined with Modeltest 3.7 (Posada & Crandall 1998) using the Akaike Information Criterion (AIC) (Posada & Buckley 2004). For each dataset, the overall transition/transversion ratio was calculated using the values from Modeltest. The COI dataset was screened for saturation at first, second and third codon positions by calculating the uncorrected pairwise distances and corrected maximum likelihood distances for each codon position in Paup. A linear relationship between both distances indicates that no saturation has occurred. Phylogenetic relationships were calculated for each locus separately according to three methods: most parsimonious (MP) and maximum likelihood (ML) trees were calculated in Paup 4.0 beta 10 (Swofford 1998) using heuristic searches and a tree-bisection-reconnection branch swapping algorithm (10 000 rearrangements), and a random stepwise addition of sequences in 100 replicate trials. One tree was held at each step. Robustness of the obtained trees was tested by bootstrapping with 1000 replications for MP and 100 replications for ML and 10 replicate trials of sequence addition. Gaps were treated as missing data. In addition, a Bayesian analysis was performed in Mr Bayes v 3.1.2 (Huelsenbeck & Roncquist 2005). Four independent Markov chains were run for 500 000 generations and a tree was saved every 10^{th} generation. The first 10 000 trees were discarded as burn-in. The best model for Bayesian analysis of the three loci was determined with MrModeltest 2.2 (Nylander 2004) using the Akaike Information Criterion (AIC).

We subsequently performed an incongruence length difference (ILD) test (Mickevich & Farris 1981) in Paup to investigate whether the different gene fragments could be combined in one analysis.

Morphological data

Morphological differences among the molecular lineages were analysed using backward stepwise discriminant function analyses (DFA) in Statistica 6.0 (Statsoft 2001). DFA determines which variables are best to discriminate between a priori defined groups. In our study, we defined eight groups based on the molecular COI

data (PmI, PmII, PmIII, PmIV, Z, Z2, Z3 and Z4). We only had information on one specimen for haplotype Z3, and hence it was removed from the dataset. Variables which were correlated with each other above the 0.8 level were omitted. This threshold was determined after calculation of the correlation between variables that are expected to be correlated (e.g. length and width, length and tail length, tail and anal body diameter). Morphological characters for which means and variances were correlated, were log transformed (body length and body length/body width in females; body length, body length/pharynx length and position of the nerve ring in males). Missing data were replaced by the average value in a particular lineage.

Since specimens from the different localities were preserved by different methods (pictures or permanent slides), morphological data from each method were analysed and interpreted separately. A first DFA analysis involved all specimens (females and males, n = 46 and n = 26, respectively) from the seven lineages which had been prepared in slides (Table 6.1). This yielded morphological information obtained from a vast geographical scale (Europe, Africa, USA). Subsequently, females and males were analysed separately so that sexually dimorphic and gender specific variables could be included in the DFA. We performed a third DFA which involved six lineages from a fairly small geographical area (ca. 100 km) in Belgium and The Netherlands that had been photographed digitally (Table 6.1). Lineage Z4 has not been observed in Belgium and The Netherlands, and hence, this lineage was not included in this last analysis. In addition, no males from Z and Z2 from Belgium and The Netherlands were available, so this last DFA was restricted to females.

No single morphometric character could unambiguously separate the species. Therefore, we created a polytomous key in which species are identified graphically by a combination of characters. Characters are chosen in accordance with the number of different frequency peaks found in their distribution range. The best characters to use at each step of the key have the highest number of peaks (= the highest variation) (Fonseca et al. 2006).

RESULTS

MOLECULAR DATA: PHYLOGENETIC ANALYSIS OF COI

The three methods of phylogenetic inference (MP, ML, BA) showed highly concordant tree topologies and divided the 58 mitochondrial COI sequences of *R. (P.) marina* into seven lineages and one terminal branch (Fig 6.1). The only difference between MP, ML and BA was the inclusion of the Z2 haplotypes within the PmII lineage in BA, which explains the low bootstrap support of the PmII lineage (Fig 6.1). Within lineages, little or no substructure was observed. All Z haplotypes were pooled into three distinct lineages (Z, Z2 and Z4) and one terminal branch (Z3) with high bootstrap support and which were highly divergent from the known cryptic lineages PmI, PmII, PmIII and PmIV (Table 6.3). The positioning of the sister species *R. (P.) mediterranea* remained unresolved, as were the deeper phylogenetic nodes. The clade containing *R. (P.) mediterranea*, PmII, Z, Z2, Z3 and Z4 contained one amino acid substitution (valine changed to leucine). Calculations of the transition/transversion ratio indicated that

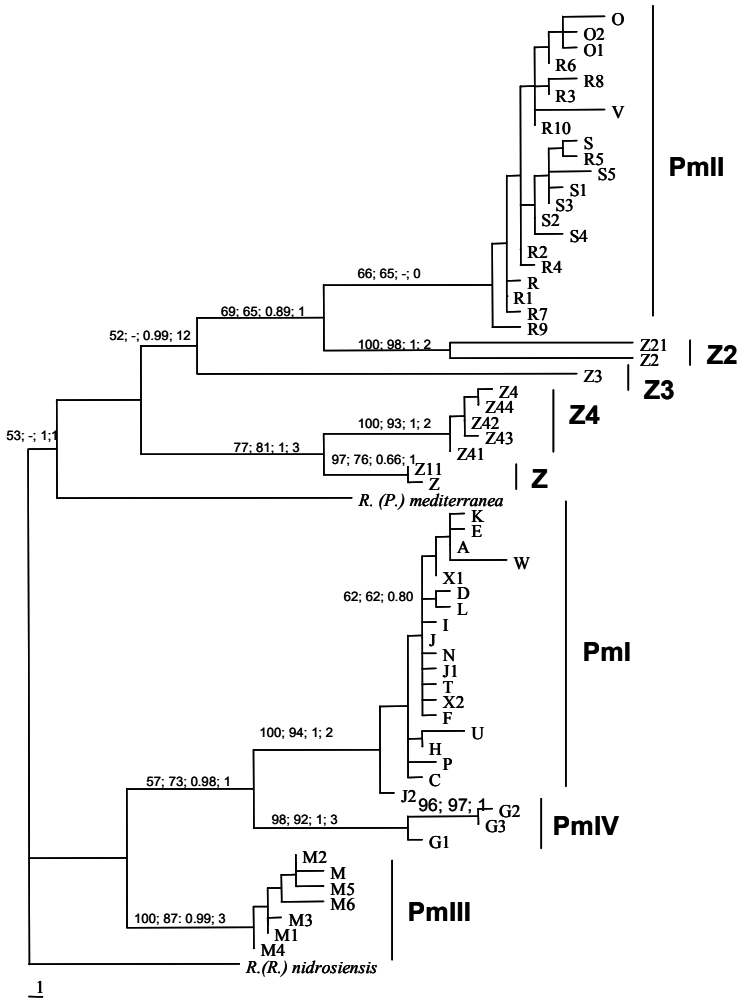


Fig. 6.1: *Rhabditis* (*Pellioditis) marina.* One of the 46 most parsimonious trees based on 396 bp of the mitochondrial COI gene. Values above branches are bootstrap supports from MP, ML, posterior probability values of BA and the number of fixed differences for each branch. Only bootstrap values above 50 are indicated. Lineages are indicated next to each branch. A dash indicates the absence of a branch in the respective analysis.

transitions vastly outnumbered transversions (Table 6.4). Plotting the uncorrected pairwise distances against the ML distances for each codon position separately indicated that saturation occurred at the third codon position of the COI gene (data not shown). The number of fixed differences for each lineage is indicated above branches (Fig. 6.1). Only the PmII lineage did not contain any fixed differences. Divergence ranges were lower within lineages (0.2 – 2.3 %) than between lineages (3.5 – 10.6 %) (Table 6.3).

	PmI	PmII	PmIII	PmIV	Z	Z2	Z3	Z4	*R. (P.) mediterranea*	*R. (R.) nidrosiensis*
PmI	0.2 - 1.7	5 - 5.7	13.8 - 14.9	0.7 - 1.1	15.0 - 15.1	4.4	5.1 - 5.2	14.4 - 14.6	11.8 - 12.0	24.1 - 24.2
PmII	7.3 - 10.3	0.2 - 2.3	14.8 - 15.4	4.7 - 5.7	15.5 - 15.8	1.3 - 1.8	3.1 - 3.7	14.8 - 15.2	12.3 - 12.7	23.4 - 23.6
PmIII	6.8 - 8.3	7.8 - 10.3	0.2 - 1.3	13.8 - 15.0	12.9 - 13.8	14.6 - 14.8	14.7 - 14.9	12.9 - 13.4	10.1 - 10.6	23.9 - 24.2
PmIV	5.3 - 7.1	7.8 - 9.6	6.6 - 7.3	0.2 - 1.3	15.4 - 15.7	4.4 - 4.7	5.2 - 5.6	14.8 - 15.1	11.9 - 12.2	24.0 - 24.1
Z	8.8 - 10.6	6.3 - 8.1	6.6 - 7.1	7.8 - 8.3	0.2	15.5	15.8	5.0 - 5.2	11.3	24.9
Z2	9.3 - 9.8	4.1 - 5.5	8.5 - 9.1	9.1 - 9.3	7.3	0.4	2.7	14.8 - 14.9	11.8	23.4
Z3	8.8 - 9.6	6.3 - 8.5	8.8 - 9.8	9.3 - 9.6	8.1 - 8.3	8.5 - 8.8	-	15.2 - 15.6	12.2	23.5
Z4	8.5 - 9.6	6.6 - 8.5	6.8 - 7.8	8.8 - 9.6	3.5 - 4.3	7.3 - 7.5	9.3 - 9.8	0.2 - 0.4	10.8	24.2 - 24.3
R. (P.) mediterranea	9.3 - 10.6	7.5 - 9.1	7.1 - 7.8	9.1 - 9.6	6.3 - 6.5	8.5	9.8	8.3 - 8.5	-	24.6
R. (R.) nidrosiensis	7.3 - 8.5	7.5 - 9.6	7.5 - 8.1	8.3 - 8.5	8.8 - 9.1	9.1	11.1	7.8 - 8.1	8.8	-

Table 6.3: Sequence divergence among the molecular lineages in *R. (P.) marina*, and among *R. (P.) mediterranea* and *R. (R.) nidrosiensis*. Below diagonal are divergences based on COI, above diagonal based on ITS. Values on the diagonal are intralineage divergence for the COI gene.

	COI	ITS		D2D3	ITS-D2D3
		90%	60%		
# taxa	62	32	32	33	30
Sequence length	396	669 - 858	669 - 858	579 - 589	1248 - 1603
Alignment length	396	913	913	597	1646
# unreliable positions	0	707	395	24	418
# parsimony informative sites	76 (19%)	24 (12%)	121 (23%)	41 (7%)	162 (13%)
Substitution model	K81uf + I + G	SYM + G	GTR + G	GTR + I + G	GTR + I + G
Tree length	247	77	297	156	455
# trees	46	15	3	3	4
Ts/Tv	2.5	1.34	1.92	3.17	2.24

Table 6.4: *Rhabditis (Pellioditis) marina.* Summary of phylogenetic analyses for each gene separately and for the combined ITS-D2D3 dataset. Sequences of *Necator americanus* and *Ancylostoma caninum* are not considered in these calculations. Percentages indicate the threshold level used in SOAP for the ITS data.

MOLECULAR DATA: PHYLOGENETIC ANALYSES OF THE NUCLEAR ITS AND D2D3 REGIONS

MP, ML and BA of both nuclear genes were highly concordant and the ILD test allowed us to combine them into one dataset (p = 1, Fig. 6.2). The nuclear tree generally gave the same topology as the mitochondrial COI gene, the only difference was caused by the inclusion of the Z specimen within the Z4 lineage in the nuclear dataset, while it was a strongly supported monophyletic branch in the COI dataset.

Divergences between Z and Z4 were relatively low (Table 6.3, and 0 – 0.4 % in D2D3). Removing the Z specimen from the dataset yielded a non significant ILD test between the mitochondrial and nuclear dataset (p = 0.28). The deeper nodes in the tree were well resolved in the nuclear tree, which supported the monophyly of the subgenus *Pellioditis*. Within the 29 *Pellioditis* sequences, the PmI, PmII, PmIII and PmIV lineages are again clearly separated and well supported (bootstrap > 90), except for lineage PmIV. The Z4 haplotypes are more closely related to the PmIII lineage and to *R. (P.) mediterranea* than to the other *R. (P.) marina* lineages. In addition, Z2 and Z3 form a monophyletic clade with the PmII lineage. They are, however, as divergent from each other as they are from the other lineages within the *Pellioditis* group (Table 6.3). Finally, the PmI and PmIV lineage are considered sister taxa.

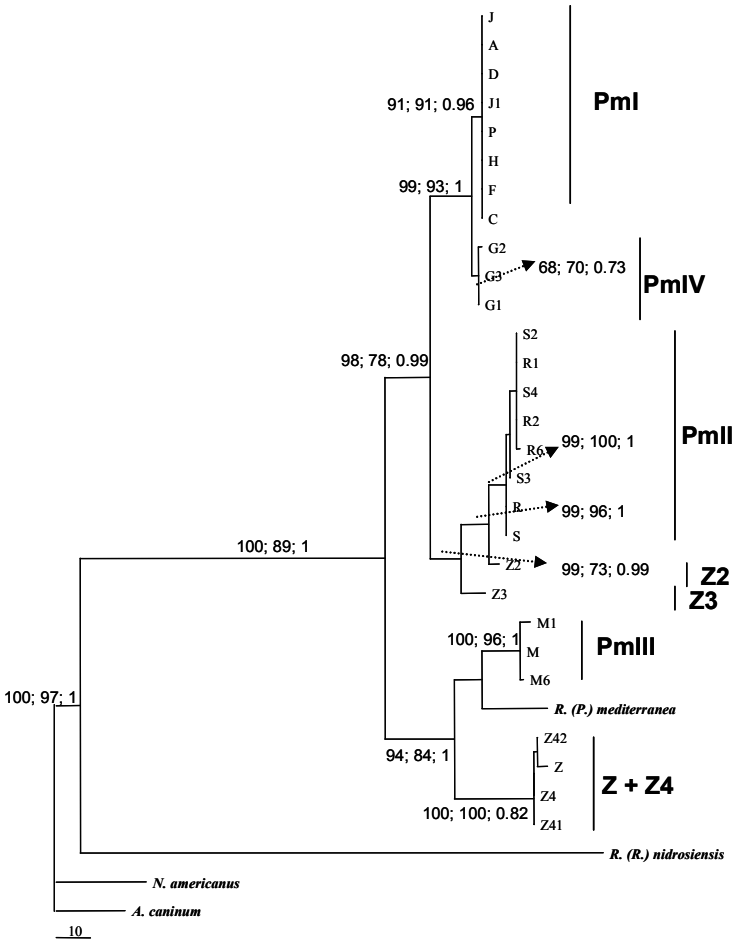


Fig. 6.2: *Rhabditis* (*Pellioditis) marina*. One of the 7 most parsimonious trees of the combined nuclear ITS and D2D3 expansion segments. Values above branches (or indicated by arrow) are bootstrap support from MP, ML, and posterior probability values from BA. Only bootstrap values > 50 are indicated. Lineages are indicated next to each branch.

MORPHOLOGICAL ANALYSES

The DFA carried out on the complete dataset from slides (females + males) without sexual dimorphic (body length/body width) and gender-specific characters (spicule length, position of the vulva) separated most lineages in the first two roots (Fig 6.3). Root 1 was best explained by body length and separated three clusters: Z-Z2-PmIII, Z4-PmII and PmI-PmIV. Each lineage within these clusters was separated along root 2, except for lineages PmII-Z4 and Z-Z2. All interlineage squared

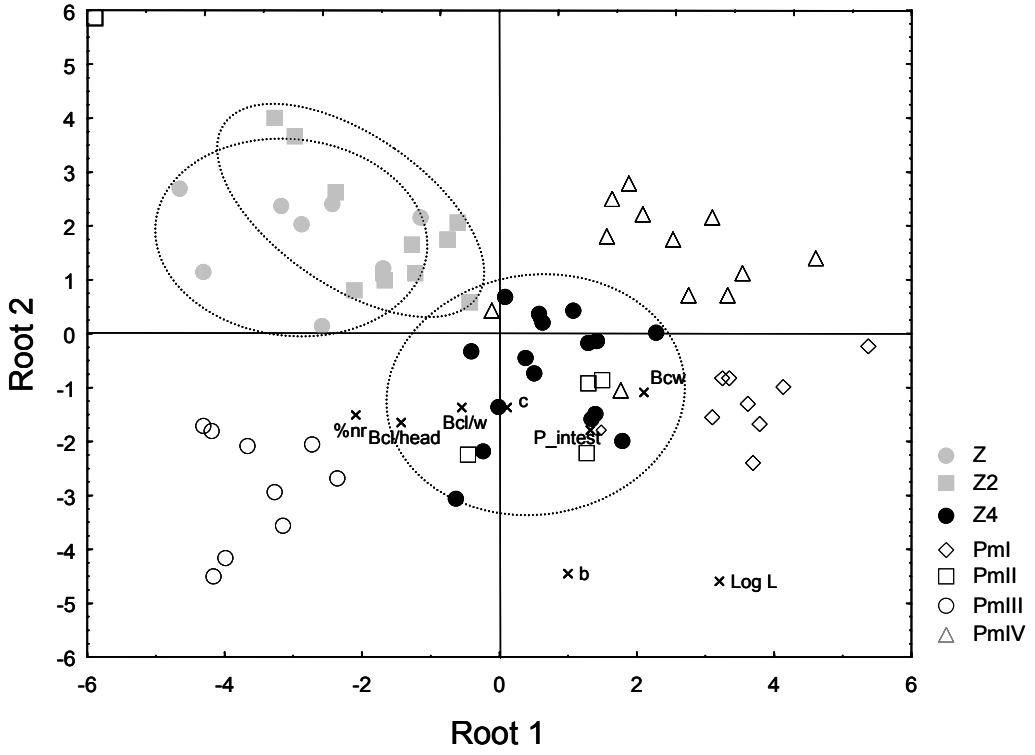


Fig. 6.3: *Rhabditis (Pellioditis) marina.* Canonical scatterplot along the first two roots of morphological measurements in males and females which have been mounted in glycerine slides. The areas occupied by lineages Z, Z2 and Z4 are encircled. Variables included in the model are indicated with crosses. Abbreviations are as in Table 6.2.

Mahalanobis distances (D^2- values) were significantly different from zero ($p < 0.01$ for all pairwise comparisons) except for lineages Z-Z2 ($p = 0.05$). D^2-values ranged between 2.1 (Z-Z2) and 30.3 (PmIII-IV). When sexually dimorphic and gender-specific characters were included in the DFA, the canonical biplot of females separated lineages Z and Z2 from each other and from all other lineages along the first root (Fig. 6.4a). Z4 specimens clustered again with PmII, and D^2-values between PmII - Z4, PmII - PmI and PmII - PmIV were non significant at the $p < 0.05$ level (D^2 = 24.9, p = 0.14; D^2 = 26.9, p = 0.15; D^2 = 30.3, p = 0.08, respectively). However, this result should be interpreted with caution, as only two specimens of lineage PmII were available. All other D^2 values were highly significant ($p < 0.001$, except for PmI-PmIV where p = 0.03 and for PmII-Z2, where p = 0.009) and ranged between 13.1 - 191.5. Based on measurements in males, all lineages were clearly separated in the first two roots of the canonical biplot (data not shown). D^2-values were high among all lineages and ranged between 45.3 - 721.5. They were non-significant only between Z-Z4 and Z-PmIII ($p = 0.3$ and $p = 0.1$, respectively). However, this is most likely due to the small number of males (n = 2) analysed in these lineages.

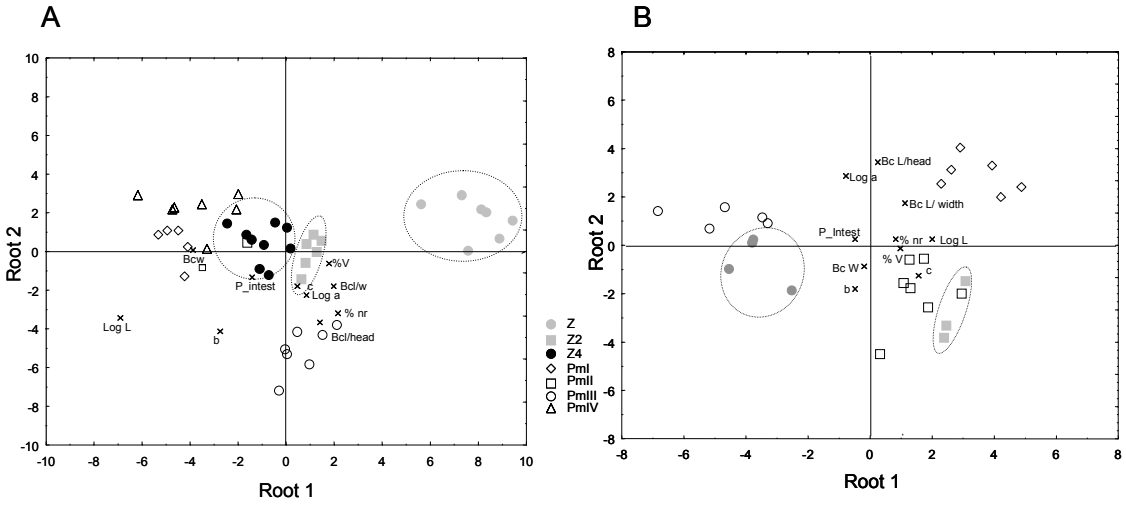


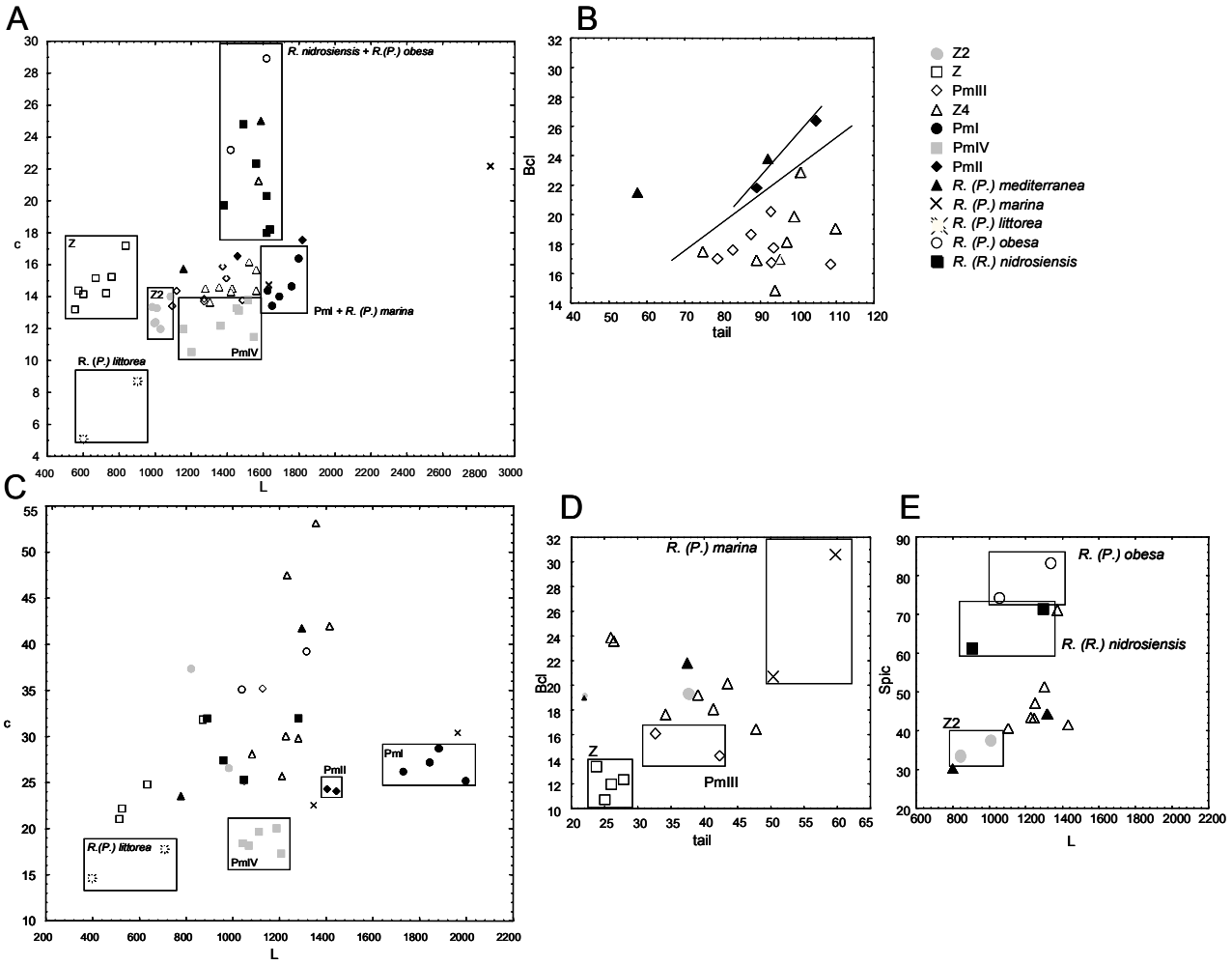
Fig 6.4: *Rhabditis (Pellioditis) marina.* Canonical scatterplots along the first two roots of morphological measurements in females. A) Females mounted in slides and collected worldwide. B) Females photographed digitally and collected in Belgium and The Netherlands, without Z4 and PmIV. Z lineages are encircled. Variables included in the model are indicated with crosses. Abbreviations are as in Table 6.2.

Finally, we compared female morphometric data from pictures to infer variation in the observed morphological differentiation between lineages on a smaller geographical scale (100 km). For this analysis, we only considered populations between which gene flow was known to occur from a previous population genetic study (Derycke et al. 2006). The canonical biplot clearly separated lineages PmI, PmII and PmIII, while lineage Z clustered with lineage PmIII (D^2 = 15.4, p = 0.27) and lineage Z2 clustered with lineage PmII (D^2 = 11.5, p = 0.52, Fig. 6.4b).

We subsequently compared our morphometric data from slides with data from the literature on rhabditid nematodes that have been observed on decomposing seaweeds (see Appendix 6.1). For *R. (R.) nidrosiensis*, morphometric data were available from several specimens, while we had minimum and maximum values for *R. (P.) marina*, *R. (P.) meditteranea*, *R. (P.) littorea* Sudhaus & Nimrich 1989 and *R. (P.) obesa* Gagarin 2001. The graphical polytomous key based on a combination of five characters (body length, tail length, buccal cavity length, body length/tail length and spicule length) unambiguously separated several species depending on the gender analysed (Fig. 6.5). For females, six species were clearly differentiated. The separation of PmII and *R. (P.) mediterranea* was less obvious, but in general, PmII specimens had a larger body length and a longer tail. Differences between the PmIII and Z4 specimens were absent in the first two steps of the key, but clear differences in buccal cavity width were observed (minimum-maximum values of 3 – 5 µm vs 5 – 7 µm, for PmIII and Z4 respectively). In addition, females of PmIII had a sharp conical

tail, while females of Z4 had a rounded tail tip (Appendices 5.4 and 6.2). For males, seven species could be differentiated with the first to steps of the key (Fig. 6.5 C, D). We have no data on the buccal cavity length of *R. (R.) nidrosiensis* and *R.(P.) obesa*, and consequently, both species are absent in Fig 6.5 D. Spicule length separated the remaining species, except for one outlier specimen of Z4 and *R. (P.) mediterranea* (Fig 5.5 E). Males from the latter species are distinguishable from Z2 and Z4 (and from the other lineages) by the absence of a structured bursa.

Fig 6.5: Graphical polytomous key for identification of species within the *R. (P.) marina* species complex. A) Females from all species, body length vs. body length/ tail length; B) Females from the clustered species in A, tail length vs. buccal cavity length; C) Males from all species, body length vs. body length/tail length; D) males from the clustered species in C, tail length vs. buccal cavity length; E) Males from the clustered species in D, body length vs. spicule length.



DISCUSSION

MOLECULAR RESULTS

The phylogenetic analyses of three molecular loci (COI, ITS, D2D3) show highly concordant tree topologies with respect to the subdivision of *R. (P.) marina* individuals into several deeply divergent lineages. The few inconsistencies between the mitochondrial and the nuclear dataset are caused either by saturation effects (Dolphin et al. 2000) or by conflicting phylogenetic signals in both datasets (Sanderson & Shaffer 2002). Saturation (multiple substitutions at the same sites) masks the true levels of sequence divergence and obscures the deeper phylogenetic relationships among sequences (Arbogast et al. 2002). Several observations do in fact indicate that saturation is present in our mitochondrial COI data: 1) the inability of the COI dataset to infer deeper phylogenetic nodes, 2) the high number of transitions with respect to transversions at the third codon position, 3) the high bootstrap support situated only at the tips of the branches and 4) the differences between MP and ML bootstrap values (Page et al. 2005). In the present study, the principle cause of the conflicts between the nuclear and mitochondrial dataset are most likely differences in phylogenetic signal: after identifying the conflicting partition (the COI gene) and the problematic taxa (Z haplotype) by the "conditional combinability" method (Bull et al. 1993), a separate analysis of mitochondrial and nuclear fragments appeared the best approach for our data. In this way, we could infer recent phylogenetic relationships with inclusion of all taxa from the mitochondrial DNA, while the deeper nodes in the tree were resolved in the nuclear dataset.

Each lineage contains 2-17 fixed differences, this number differing between gene fragments. The COI gene is generally assumed to reach fixation four times more rapidly than the nuclear genome, because of its maternal inheritance and haploid state (Nadler 2002). From Table 6.5 the number of fixed differences per 100 bases is in most cases 1 – 6 times higher in the mitochondrial COI. Clearly, this number is strongly dependent on the number of individuals analysed in each lineage and further demonstrates the shortcomings of species delimitation based solely on fixed differences (Wiens & Servedio 2000). Sequence divergence is less susceptible to the number of specimens analysed, but seems too variable across taxa to be a good universal predictor for species delimitation (Ferguson 2002, Cognato 2006). Within

	COI	ITS	D2D3
PmI	0.51	0.76	0.00
PmII	0.00	0.38	0.00
PmIII	0.76	1.26	0.34
PmIV	0.76	0.13	0.00
Z	0.25	0.25	0.00
Z2	0.51	0.00	0.34
Z3	3.03	0.50	0.17
Z4	0.51	0.00	0.00
Total	6.31	3.28	0.84

Table 6.5: Number of fixed differences in COI, ITS and D2D3 genes per 100 bp, for each lineage.

the species complex investigated here, the lineages of *R. (P.) marina* are as divergent from each other as they are from their close relatives *R. (P.) mediterranea* and *R. (R.) nidrosiensis*. Divergent molecular lineages are not compatible with species if 1) extremely high rates of evolution are present in both mitochondrial and nuclear DNA, 2) strong balancing selection is acting on the genome, or 3) vicariant events have occurred (Rocha-Olivares et al. 2001). Morphological differences were consistent with molecular results and hence, false conclusions due to high molecular rates can be discarded in our data. With respect to balancing selection, we find it unlikely that highly divergent polymorphisms in two independently evolving genomes would be maintained in the population. Balancing selection in the mitochondrial DNA genome in invertebrates has been associated with sex determination (Quesada et al. 1999), but this is unlikely here as relative frequencies of some lineages are not equally distributed across geographical regions (e.g. PmIV in Lake Grevelingen, Z4 in Mexico) (Rocha-Olivares et al. 2001). Finally, if the deeply divergent lineages are to be explained by vicariant events, they should be able to hybridize once they occur in sympatry. The monophyletic status of the lineages in the nuclear gene trees indicates that they do not hybridize. This is obviously disputable for lineages Z and Z4. Most likely, speciation between both lineages has occurred too recently to be detected in the nuclear genes.

MORPHOLOGICAL RESULTS

The set of morphological variables used in this study clearly demonstrates that the three Z lineages exhibit morphological differences with respect to each other and to the previously described lineages within *R. (P.) marina*. Regardless of which morphological variables are responsible for this differentiation, it shows that molecular lineages in free-living nematodes can be morphologically quite distinct. Similar observations have been made on parasitic nematodes (e.g. Carneiro et al. 1998, Han et al. 2006). Allthough different methodologies were applied to obtain

morphological data, our analyses strongly suggest that the morphological variation is affected by geographical scale, as the differences between some lineages were less pronounced or even disappeared when only specimens from geographically close populations were considered. Similar effects of geography on morphology in parasitic nematodes have been reported (Agudelo et al. 2005, Nguyen et al. 2006) and clearly illustrates the problem of morphological variability in nematodes.

Comparing our measurements from slides with those of *R. (P.) marina* reported in Sudhaus (1974) and of the congeners *R. (P.) mediterranea, R. (P.) ehrenbaumi, R. (P.) obesa* and *R. (P.) littorea* reported in the literature (Sudhaus 1974, Inglis & Coles 1961, Gagarin 2001 and Sudhaus & Nimrich 1989, respectively, Appendix 6.1) shows that our specimens are more similar to *R. (P.) marina* and *R. (P.) mediterranea* than to the other congeners. Moreover, the graphical polytomous key indicates that the combination of four morphometric characters (body length, tail length, buccal cavity length, spicule length) and one shape parameter (body length/tail length) is sufficient to differentiate all species. The three Z lineages show some similarities to, but clearly also differences from the *R. (P.) marina* and *R. (P.) mediterranea* described by Sudhaus (1974). *R. (P.) mediterranea* was initially described as a subspecies of *R. (P.) marina* due to its geographical distribution (Sudhaus 1974), and was later raised to species level mainly based on the female tail shape (Andrássy 1983, Sudhaus & Nimrich 1989). The high divergences between *R. (P.) mediterranea* and the *R. (P.) marina* lineages in both mitochondrial and nuclear fragments support this view.

COMBINING MOLECULAR AND MORPHOLOGICAL RESULTS TO INFER TAXONOMIC STATUS OF THE 'CRYPTIC' LINEAGES WITHIN R. (P.) MARINA

Inferring species status of the Z haplotypes requires a solid framework from which we can conclude whether the observed differences are situated at the intra- or interspecific level. For nematodes, evolutionary approaches are very promising for delimiting species as they produce phylogenetic relationships based on many characters (Adams 1998, 2001). Nevertheless, phylogenetic analyses of DNA sequences can easily lead to misinterpretations of the evolutionary processes underlying the observed patterns (Arbogast 2002, Nadler 2002). These theoretical drawbacks are substantially reduced when several independently evolving molecular

markers are analysed in the same set of individuals (Nadler 2002). We used concordant patterns among different markers as evidence for independent evolutionary histories of the four Z-lineages. The analyses of one mitochondrial and two nuclear genes yielded highly concordant tree topologies, indicating that the divergent phylogenetic lineages are caused by a common evolutionary process, i.e. speciation. Furthermore, at least three of the four lineages are accompanied by morphological differences. Although morphology may be influenced by geography, each of the lineages is differentiated from each other and from *R. (P.) marina* and *R. (P.) mediterranea* by a combination of morphometric characters and morphological observations (Fig 6.5). For example, lineages Z and Z4, which had similar nuclear gene sequences, are morphologically quite distinct. This clearly illustrates the usefulness of combining molecular and morphological data to delineate species. Furthermore, lineages Z and PmIII have been observed in very distant geographical populations (Belgium and South Africa, Belgium and USA, respectively), despite the limited dispersal of *R. (P.) marina*. This wide geographical distribution suggests that *R. (P.) marina* dispersal is not that limited at all or, alternatively, that parallel evolution may be acting in the *R.* (*P.) marina* complex. This clearly needs further research.[23]

[23] See Chapter IX

CONCLUSION

Based on molecular and morphological data, we have identified eight species within the 'morphospecies' *R. (P.) marina*, of which four are new. We here refer to these species as Z, Z2, Z3 and Z4. Although nuclear sequences from Z were very similar to those of Z4, specimens belonging to both lineages were morphologically quite distinct. Our molecular data also confirms the species status of *R. (P.) mediterranea*. Most importantly, our results indicate that the true level of biodiversity in free-living nematodes is hitherto seriously underestimated. This study further illustrates the usefulness of a holistic approach for identifying species in problematic taxa. Obviously, more species are likely to be present within *R. (P.) marina* species complex, due to its cosmopolitan distribution. In view of this, we are currently collecting samples from over the world in order to further unravel the speciation modes in this cryptic species complex.

ACKNOWLEDGEMENTS

S.D. acknowledges a grant from the Flemish Institute for the Promotion of Scientific- Technological Research (I.W.T.). G.F. is supported by the Brazilian ministry of science and technology (CNPq). T.M. is a postdoctoral fellow with the Flemish Fund for Scientific Research. Further financial support was obtained from Ghent University in BOF – projects 01GZ0705 and 011060002. The authors acknowledge Drs. Daniel Leduc who kindly provided specimens of *R. (P.) mediterranea* and are grateful to the following people for their help in collecting and retrieving samples of *R. (P.) marina*: Prof. Dr. G. Borgonie, Drs. S. Vandendriessche, G. Van Hoey and K. Houthoofd, Ms. H. Wydaeghe and Mr. T. Volant.

Appendix 6.1: Summary of morphometric data of all genetic lineages in *R. (P.) marina*. Literature data of *R. (P.) marina* (Sudhaus 1974) and of the congeners *R. (P.) mediterranea* (Sudhaus 1974), *R. (P.) nidrosiensis* (Inglis & Coles 1961), *R. (P.) littorea* (Sudhaus & Nimrich 1989) and *R. (P.) obesa* (Gagarin 2001) are included. Values are given in µm and as minimum – maximum (average).

		L	W	BcL	Ph	tail	a	b	c	%V	Spic	testis
FEMALES	Z	556 - 836 (675)	22 - 36 (29)	11 - 14(13)	127 - 177 (147)	39 - 51 (45)	20.1 - 24.9 (23.2)	4.3 - 4.9 (4.6)	13.2 - 17.2 (14.8)	54 - 57 (56)	-	-
	Z2	985 - 1088 (1018)	45 - 57 (49)	15 - 20 (17)	189 - 226 (202)	73 - 85 (79)	17.2 - 22.9 (20.8)	4.5 - 5.8 (5.1)	11.9 - 14.0 (12.9)	49 - 59 (54)	-	-
	Z4	1282 - 1573 (1447)	56 - 89 (64)	15 - 23 (18)	207 - 234 (221)	73 - 108 (94)	17.5 - 26.9 (22.6)	5.6 - 7.6 (6.5)	13.6 - 21.26 (15.4)	48 - 58 (51)	-	-
	PmI	1626 - 1798 (1705)	76 - 84 (81)	21 - 26 (23)	240 - 277 (256)	109 - 123 (117)	19.9 - 21.9 (20.9)	6.2 - 7.2 (6.7)	13.4 - 16.4 (14.6)	50 - 53 (52)	-	-
	PmII	1457 - 1818 (1638)	71 - 92 (81)	22 - 26 (24)	215 - 326 (270)	88 - 103 (95)	19.8 - 20.4 (20.1)	5.6 - 6.8 (6.2)	16.6 - 17.5 (17.0)	52 - 56 (54)	-	-
	PmIII	1095 - 1514 (1309)	42 - 60 (54)	16 - 20 (18)	183 - 203 (192)	77 - 107 (90)	21.0 - 28.6 (24.1)	5.8 - 7.7 (6.7)	13.4 - 15.9 (14.3)	50 - 53 (51)	-	-
	PmIV	1160 - 1548 (1387)	51 - 76(64)	19 - 24 (21)	220 - 245 (235)	97 - 134 (112)	18.9 - 24.5 (21.8)	5.1 - 6.6 (5.9)	10.5 - 13.8 (12.3)	47 - 53 (50)	-	-
	P. mediterranea	1157 - 1590	45 - 78	22 - 24	197 - 237	56 - 91	18.2 - 25.8	5.2 - 7.0	15.7 - 25.0	51 - 55	-	-
	P. marina	1628 - 2875	69 - 118	30-39	237 - 354	99 - 139	20 - 24.5	5.2 - 8.1	14.7 - 22.0	53 - 57	-	-
	P. ehrenbaumi	1380 - 1640	-	-	-	-	14.9 - 20.3	3.6 - 4.2	18.0 - 24.8	52 - 56	-	-
	P. littorea	599 - 900(731)	34 - 58 (46)	16 - 20 (18)	118 - 156 (137)	89 - 147 (113)	14.4 - 17.8 (15.9)	4.7 - 6.4 (5.3)	5.1 - 8.7 (6.4)	47 - 53 (50)	-	-
	P. obesa	1422 - 1619 (1524)	-	-	361-416 (387)	52-59 (56)	-	16.0 - 20.0 (19)	23.2 - 28.9 (25.6)	57 - 59 (58)	-	-
MALES	Z	515 - 870 (635)	25 - 34 (29)	10 - 13 (12)	115 - 155 (133)	23 - 27 (25)	15.6 - 25.1 (21.5)	4.3 - 5.6 (4.7)	21.0 - 31.8 (24.9)	-	30 - 35 (33)	338 - 669 (461)
	Z2	822 - 985 (904)	43 - 53 (48)	19 - 19 (19)	166 - 179 (173)	22 - 37 (29)	15.4 - 22.7 (19.1)	4.9 - 5.5 (5.2)	26.5 - 37.3 (31.9)	-	34 - 38 (36)	655 - 823 (739)
	Z4	1084 - 1413 (1258)	48 - 71 (55)	16 - 24 (20)	186 - 226 (201)	25 - 47(36)	19.1 - 25.2 (22.9)	5.4 - 7.3 (6.3)	25.7 - 53.1 (36.6)	-	41 - 52 (48)	983 - 1233 (1101)
	PmI	1731 - 1998 (1864)	72 - 87 (79)	24 - 28 (26)	297 - 312 (304)	65 - 79 (69)	21.4 - 27.7 (23.6)	5.7- 6.4 (6.1)	25.2 - 28.7 (26.8)	-	50 - 54 (53)	1205 - 1555 (1316)
	PmII	1403 - 1445 (1424)	62 - 62 (62)	23 - 23 (23)	260 - 280 (270)	57 - 60 (58)	22.6 - 23.16 (22.9)	5.0 -5.6 (5.3)	24.0 - 24.3 (24.2)	-	57 - 64 (61)	1168 - 1238 (1203)
	PmIII	1051 - 1130 (1090)	39 - 32 (36)	14 - 16(15)	179 - 163 (171)	32 - 41 (36)	28.4 - 32.5 (30.5)	6.3 - 6.4 (6.4)	25.1 - 35.2 (30.2)	-	37 - 42 (40)	885 - 948 (917)
	PmIV	1043 - 1210 (1125)	47 - 62 (54)	18 - 20 (19)	202 - 227 (213)	56 - 70 (60)	18.9 - 21.9 (20.6)	4.9 - 5.5 (5.3)	17.3 - 20.0 (18.7)	-	52 - 62 (57)	909 - 1078 (996)
	P. mediterranea	779 - 1298	32 - 49	19 - 22	153 - 200	22 - 37	18.7 - 33.2	4.4 - 6.6	23.6 - 41.7	-	31 - 45	-
	P. marina	1337 - 1978	43 - 71	21 - 31	221 - 291	50 - 59	20.9 - 32.4	4.6 - 7.7	22.4 - 30.2	-	37 - 57	-
	P. ehrenbaumi	890 - 1280	52 - 69	-	217 - 312	27 - 40	17.0 - 18.3	3.2 - 4.1	25.3 - 32.0	-	-	-
	P. littorea	400 - 708 (501)	25 - 45 (33)	15 - 20 (16)	109 - 159 (125)	25 - 33 (28)	13.6 - 19.9 (16.3)	3.5 - 4.8 (3.9)	14.6 - 17.8 (16.2)	-	23 - 30 (26)	201 - 453 (286)
	P. obesa	1039 - 1318 (1116)	-	-	322 - 357 (340)	28 - 37 (32)	-	11.0 - 19.0 (16)	35.1 - 39.2 (37.1)	-	75 - 84 (79)	-

Appendix 6.2. Drawings of the *Rhabditis (Pellioditis) marina* species complex. a-c: Z2; d-f: Z; g-i:Z4