Received: 28 April 2019 | Revised: 26 October 2019 | Accepted: 7 November 2019

DOI: 10.1111/1755-0998.13118

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

Widespread prevalence but contrasting patterns of intragenomic rRNA polymorphisms in nematodes: Implications for phylogeny, species delimitation and life history inference

Xue Qing[1,2] | Holly Bik[3] | Timur M. Yergaliyev[4,5] | Jianfeng Gu[6] | Pamela Fonderie[1] | Sigal Brown-Miyara[2] | Amir Szitenberg[4] | Wim Bert[1]

[1]Nematology Research Unit, Department of Biology, Ghent University, Ghent, Belgium

[2]Department of Entomology, Nematology and Chemistry Units, Agricultural Research Organization (ARO), Volcani Center, Rishon LeZion, Israel

[3]Department of Nematology, University of California-Riverside, Riverside, CA, USA

[4]Dead Sea and Arava Science Center, Dead Sea Branch, Masada National Park, Tamar Regional Council, Tel Aviv, Israel

[5]A. Baitursynov Kostanay State University, Kostanay, Kazakhstan

[6]Technical Center of Ningbo Customs (Ningbo Inspection and Quarantine Science Technology Academy), Ningbo, China

Correspondence
Xue Qing, Nematology Research Unit, Department of Biology, Ghent University, Ghent, Belgium.
Email: xueqing4083@gmail.com

Funding information
Bijzonder Onderzoeksfonds, Grant/Award Number: UGent 01N02312

Abstract

Ribosomal RNA genes have long been a favoured locus in phylogenetic and metabarcoding studies. Within a genome, rRNA loci are organized as tandem repeated arrays and the copies are homogenized through the process of concerted evolution. However, some level of rRNA variation (intragenomic polymorphism) is known to persist and be maintained in the genomes of many species. In nematode worms, the extent of rRNA polymorphism (RP) across species and the evolutionary and life history factors that contribute to the maintenance of intragenomic RP is largely unknown. Here, we present an extensive analysis across 30 terrestrial nematode species representing a range of free-living and parasitic taxa isolated worldwide. Our results indicate that RP is common and widespread, ribosome function appears to be maintained despite mutational changes, and intragenomic variants are stable in the genome and neutrally evolving. However, levels of variation were varied widely across rRNA locus and species, with some taxa observed to lack RP entirely. Higher levels of RP were significantly correlated with shorter generation time and high reproductive rates, and population-level factors may play a role in the geographic and phylogenetic structuring of rRNA variants observed in genera such as *Rotylenchulus* and *Pratylenchus*. Although RP did not dramatically impact the clustering and recovery of taxa in mock metabarcoding analyses, the present study has significant implications for global biodiversity estimates of nematode species derived from environmental rRNA amplicon studies, as well as our understanding of the evolutionary and ecological factors shaping genetic diversity across the nematode Tree of Life.

KEYWORDS

concerted evolution, *Halicephalobus*, metabarcoding, Nematoda, phylogeny, rRNA secondary structure

1 | INTRODUCTION

Nuclear genes encoding the ribosomal subunits (commonly referred to as rRNA) have been widely used for the estimation of evolutionary history and the taxonomic assignment of individual organisms. Ribosomal RNA genes are present in all cells across all kingdoms and are highly conserved (Smit, Widmann, & Knight, 2007; Woese, Kandler, & Wheelis, 1990). In eukaryotic

Qing and Bik contribute equally to this work.

organisms, rRNA genes are maintained as large clusters of tandem repeats in chromatids ranging from 30 to 30,000 copies (Ide, Miyazaki, Maki, & Kobayashi, 2010; Prokopowich, Gregory, & Crease, 2003).

The rRNA genes are thought to evolve in a concerted manner, such that tandem repeats are not independent of one another but instead are homogenized by different mechanisms. Concerted evolution is usually a very effective process (Arnheim et al., 1980; Eickbush & Eickbush, 2007), but significant differences between copies of rRNA genes in single organisms have been discovered across all three domains of life and in all three subunits of rRNA genes (Alper, Frenette, & Labrie, 2011; Cui, Zhou, Oren, & Liu, 2009; Pei et al., 2010; Xu et al., 2015). Within Nematoda, the polymorphisms have been found in 28S rRNA genes and noncoding internal transcribed spacer (ITS) regions based on PCR amplification (Pereira & Baldwin, 2016; Van Den Berg et al., 2016; Yoshiga, Kuwata, Takai, & Nishimura, 2014; Zhan, Matafeo, Shi, & Zheng, 2011) and across all rRNA genes based on genome analysis (Bik, Fournier, Sung, Bergeron, & Thomas, 2013). However, the overall knowledge of rRNA polymorphisms in nematodes remains limited, as only a few species have been investigated and evolutionary and ecological analyses are lacking. Current nematode taxonomy and phylogeny have largely relied on rRNA genes to infer evolutionary relationships (e.g. Bert, Leliaert, Vanfleteren, & Borgonie, 2008; Holterman et al., 2006), and the presence of widespread rRNA polymorphisms across nematode species may increase uncertainty with respect to species identification and delimitation in this framework. Moreover, since the estimation of species' composition and abundance is a major goal of environmental studies, understanding the intragenomic and intraspecific rRNA variations will also be critical for informing high-throughput analyses of metabarcoding and metagenomic data sets (Pereira & Baldwin, 2016).

In the present study, we aimed to quantify, locate and understand the evolutionary forces contributing to both intra- and interspecific variation across ribosomal RNA genes in the Phylum Nematoda. We hypothesized that an intragenomic rRNA variation is a rare event that appears in relation to specific life strategies or geographic distributions. Our sampling therefore focused on wide-range nematode taxa with contrasting life cycles and representing distinct geographic distributions. The rRNA polymorphisms were analysed using the following approaches: (a) quantifying the degree of intragenomic rRNA variation across nematode species; (b) mapping polymorphic sites onto rRNA secondary structures and evaluating their impact on the molecule; (c) carrying out functional analyses to screen for potential pseudogenes; (d) evaluating processes of concerted evolution, (e) analysing phylogenetic relationships; and (f) assessing the impact of life strategy or geographic distribution. Furthermore, the implications of rRNA polymorphism are discussed in the context of environmental metabarcoding studies where definitions of molecular 'species' and taxonomic assignments primarily rely on measures of sequence similarity and clustering.

2 | MATERIALS AND METHODS

2.1 | Sample collection and extraction

Samples were collected from various locations and habitats (a total of 30 species); details are listed in Table 1. In general, bulk soil samples were taken by homogenizing 5–10 individual samples (collected in scoop, 20 cm depth) of approximately 100 g each. Nematodes were extracted from 50 g subsamples taken from the bulk sample using a Baermann tray. Samples collected outside Belgium were fixed with the DESS solution (Yoder et al., 2006) at room temperature. For *Halicephalobus* spp., 73 sites in Belgium were first checked for the presence of *Halicephalobus* spp. in compost, soil or manure heaps. In order to evaluate the impact of life strategies to rRNA polymorphisms, we collected further samples from positive sites covering varied habitats, including fresh horse dung, and rectal and maxillary granulomatous from infected horses. The *Halicephalobus gingivalis* specimens from the infected horses were directly isolated from the maxillary tissue (Fonderie et al., 2013), and the *H. gingivalis* from other habitats were extracted by incubating the samples at 30°C on a Petri dish containing 2% bacteriological agar enriched with cholesterol.

Only single nematode specimens were used for DNA extractions in each population except for *H. gingivalis* and *Ruehmaphelenchus quercophilus* where 3–5 individuals were used to compare the variation among individuals. Single nematodes were extracted by worm lysis buffer following Yoder et al. (2006).

2.2 | PCR Amplification and cloning

Three rRNA loci (18S gene, 28S gene and ITS1-5.8S-ITS2 region) were amplified via PCR for polymorphism analysis. Two mitochondrial coding genes, cytochrome oxidase *c* subunit 1 (COI) and NADH dehydrogenase subunit 4 (ND4), were also amplified in *Halicephalobus* spp. for inclusion in phylogenetic analyses. Taxa-specific primer pairs are listed in Table S1. PCRs followed previous protocols for the corresponding primer sets (Bert et al., 2008; Múnera Uribe et al., 2010; Qing, Decraemer, Claeys, & Bert, 2017a; Ye, Giblin-Davis, Braasch, Morris, & Thomas, 2007). The PCR products from the *H. gingivalis* populations (Hgin4 - Hgin18) were directly used for sequencing. Polymorphic rRNA sites were revealed by the presence of ambiguous nucleotides in nematode chromatograms with the method of Nadler et al. (2003). For other species and populations, PCR products were cloned using pGEM®-T easy vector systems (Promega) and subsequently sequenced in two primer directions to minimize possible errors. Resulting contigs were assembled using Geneious R6.1.8.

2.3 | Genome analysis

Five species were used for genome analysis: *Heterorhabditis bacteriophora* (1 sample), *Rotylenchulus reniformis* (2 samples), *Meloidogyne incognita* (8 samples), *Acrobeloides nanus* (1 sample), *and Globodera*

TABLE 1 The list of analysed species in the present study

Species	Sample origin	FT	c-p value	Pop.	Gene	No.	Length	Polymorphism
Acrobeloides sp.	Yangling, China	B	2	Asp	18S	11	962	No
Aphelenchoides subtenius	The Netherlands	F/P	2	Asub	18S	4	839–840	Yes
					28S	4	743	Yes
					ITS	21	770–780	Yes
Bicirronema hamiguitanense	Mt. Hamiguitan, Philippines	B	2	Bham	28S	6	620	No
Cephalenchus cephalodiscus[a]	Wyoming, USA	LP	(2)	Ccep	ITS	5	609–638	Yes
Cephalenchus cylindricus[a]	Ensenada, Mexico	LP	(2)	Ccyl2	28S	8	699–730	Yes
				Ccyl1	ITS	9	609–638	Yes
Cephalenchus daisuce[a]	Vancouver, Canada	LP	(2)	Cdai2	28S	6	729–730	Yes
				Cdai1	ITS	6	598–638	Yes
Cephalenchus nemoralis[a]	Cuc Phuong, Vietnam	LP	(2)	Cnem	28S	7	721–735	Yes
Cephalenchus sp.1[a]	Jaguaruna, Brazil	LP	(2)	Csp1	28S	10	680–684	Yes
				Csp2	ITS	9	544–561	Yes
Coslenchus turkeyensis	Ghent, Belgium	LP	(2)	Ctu	18S	16	980	No
Deladenus sp.	USA	IA	2	Dsp	ITS	9	1,555–1,558	Yes
Teratolobus sp.	Mt. Hamiguitan, Philippines	B	2	Tsp	18S	11	962	No
					28S	5	653	No
Filenchus andrassyi	San José, Costa Rica	F/LP	(2)	Fand	18S	7	956	No
Filenchus discrpans	Mt. Qinling, China	F/LP	(2)	Fdis	18S	10	959	No
					28S	6	516	No
Filenchus sp.	Mt. Qinling, China	F/LP	(2)	Fsp	18S	8	921	No
					28S	8	651	No
Geocenamus brevidens	Israel	P	(3)	Gbre	18S	6	1,622	No
					28S	6	681	No
Halicephalobus gingivalis[b]	Deerlijk, Belgium	B/VP	1	Hgin1	18S	7	900–901	Yes
	Dentergem, Belgium	B/VP	1	Hgin2	18S	3	809–811	Yes
	Riverside, USA	B/VP	1	Hgin3	18S	3	902–903	Yes
	East Flanders, Belgium	B/VP	1	Hgin4	18S/28S/COI/ND4	–	836/427/514/445	Yes
	West Flanders, Belgium	B/VP	1	Hgin5		–		Yes
	ILVO, Merelbeke, Belgium	B/VP	1	Hgin6		–		Yes
	ILVO, Merelbeke, Belgium	B/VP	1	Hgin7		–		Yes
	ILVO, Merelbeke, Belgium	B/VP	1	Hgin8		–		Yes
	West Flanders, Belgium	B/VP	1	Hgin9		–		Yes
	West Flanders, Belgium	B/VP	1	Hgin10		–		Yes
	West Flanders, Belgium	B/VP	1	Hgin11		–		Yes
	West Flanders, Belgium	B/VP	1	Hgin12		–		Yes
	West Flanders, Belgium	B/VP	1	Hgin13		–		Yes
	West Flanders, Belgium	B/VP	1	Hgin14		–		Yes
	Antwerp, Belgium	B/VP	1	Hgin15		–		Yes
	Antwerp, Belgium	B/VP	1	Hgin16		–		Yes
	Orlando, Canada	VP	1	Hgin17		–		Yes
	Ghent, Belgium	VP	1	Hgin18		–		Yes

(Continues)

TABLE 1 (Continued)

Species	Sample origin	FT	c-p value	Pop.	Gene	No.	Length	Polymorphism
Halicephalobus sp. 1	West Flanders, Belgium	B/VP	1	Hsp1	18S	6	903	No
					28S	5	543–544	Yes
Halicephalobus brevicauda.	USA RGD892	IA	1	Hbre	18S	6	910	No
Halicephalobus sp. 2	USA RGD838	IA	1	Hsp2	18S	5	910	No
Halicephalobus mephisto	South Africa	B	1	Hmep	18S	3	1,602	No
Halicephalobus similigaster[b]	Berlin, Germany	B	1	Hsim	18S	3	906	No
Malenchus acarayensis	Groenendaal, Belgium	LP	(2)	Maca2	18S	5	912–913	Yes
				Maca1	28S	7	663–667	Yes
Malenchus pachycephalus	Poeke, Belgium	LP	(2)	Mpac1	18S	4	914–915	Yes
	Poeke, Belgium			Mpac2	18S	6	912–914	Yes
Pratylenchus brachyurus	East Gonja, Ghana	P	(3)	Pbra1	18S	7	1,769–1,780	Yes
					28S	16	776–780	Yes
	Imo, Nigeria			Pbra2	18S	8	1,781–1,783	Yes
					28S	8	813–836	Yes
Pratylenchus capsici	Zofar, Israel	P	(3)	Pcap	18S	6	1,775	No
					28S	10	775–778	Yes
					ITS	8	860–872	Yes
Pratylenchus mediterraneus	Shefaa, Israel	P	(3)	Pmed	18S	5	1,763	No
					28S	8	775	No
					ITS	8	927	No
Rotylenchulus macrosoma	Shefaa, Israel	P	(3)	Rmac	18S	10	1,300–1,764	Yes
					28S	11	756–758	Yes
					ITS	12	903–909	Yes
Rotylenchulus reniformis[a]	Alabama, USA	P	(3)	Rren	18S	19	1,717–1,721	Yes
Ruehmaphelenchus quercophilus	Łopuchówko, Poland	F	2	Rque1	18S	27	682–688	Yes
					28S	42	736–745	Yes
				Rque2	18S	35	994–1,003	Yes
Xiphinema parachambersi	Japan	P	(5)	Xpar	ITS	4	1,306–1,313	Yes

Note: FT = feeding type (B = bacteria feeder, *F* = fungi feeder, P = higher plant parasitic, LP = lower plant feeder, including algal, lichen, moss feeders, IA = insect associated, VP = vertebrate parasites). c-p value: a continuum from colonizers to persisters (r- to K-strategists), proposed by Bongers and Bongers (1998) to integrate the nematodes feeding group and life strategy. Values in parentheses refer to plant feeding taxa with a separate plant parasite indexing system. Pop. = population abbreviation used in this study. Gene = the sequenced gene regions (ITS = ITS1-5.8S-ITS2). No. = number of sequenced clones from the single nematode. Length = sequence length in bp.

[a]The polymorphic sequences for *Cephalenchus cephalodiscus* (15T22A13), *C. cylindricus* (ITS: 1T11F09, 28S: 4T03F09), *C. daisuce* (ITS: 1T09H13, 28S: 3T09I13), *Cephalenchus nemoralis* (7T13G10), *Cephalenchus* sp.1 (ITS: 2T17C09, 28S: 1T17C09) and *Rotylenchulus reniformis* (SSU13B) were retrieved from previous studies available in GenBank (Nyaku et al., 2013; Pereira & Baldwin, 2016).

[b]The sequences of *H. gingivalis* population Hgin4 - Hgin18 are acquired from direct sequencing (without cloning). Polymorphisms in 18S rRNA were revealed by the presence of ambiguous nucleotides in nematode chromatograms detailed in Nadler et al. (2003).

rostochiensis (4 samples). The rRNA genes were extracted from the NCBI Sequence Read Archives (Table S2). The species-specific reference sequences for rRNA genes were obtained from NCBI GenBank and the manually curated SILVA database (Quast et al., 2012). For *M. incognita*, the complete rRNA was identified in the whole-genome sequence assembly and reannotated using BLAST 2.6 and MAFFT 7.407 (Katoh & Standley, 2013). The accession numbers and coordinates of the rRNA gene sequences used as reference are available in Table S3. To normalize the SNP count across reference sequences of different lengths, polymorphisms were expressed as SNPs per 50 bp, as in Bik et al. (2013). Other details on genome analysis are provided in the supplementary methods. All generated data in above analysis are included in the GitHub repository (https://doi.org/10.5281/zenodo.2619226).

2.4 | Analysis of sequences and phylogeny reconstruction

Alignments were constructed using the Q-INS-i algorithm implemented in MAFFT v. 7.205 (Katoh & Standley, 2013). For the

phylogenetic analysis of *Halicephalobus* spp., a concatenated alignment of all genes (18S, 28S, ITS1-5.8S-ITS2, COI and ND4) was generated using Geneious R6.1.8. Genetic distance was evaluated by Kumar, Stecher, and Tamura (2016) using the K2P method. The pairwise differences were compared between nematode groups (feeding type and c-p value) or genes to polymorphic variation (genetic distance and maximum pairwise differences) or rRNA secondary structure composition, using ANOVA implied in R version 3.5.2 (R Development Core Team).

For secondary structure analysis, the sequences were extracted and prealigned using the secondary structure from previous studies (Bert et al., 2008; Qing, Slos, Claeys, & Bert, 2017b; Subbotin et al., 2007). Variable regions were aligned and folded in LocARNA (Smith, Heyne, Richter, Will, & Backofen, 2010) using an energy minimization approach. Helices were named by following the universal model of eukaryotic organisms (Wuyts, Van de Peer, & De Wachter, 2001).

For detecting selection effects, neutrality tests were conducted using DNAsp 5.0 (Librado & Rozas, 2009) according to Tajima's method (Tajima, 1989), and Fu and Li's method (Fu & Li, 1993). To identify potential pseudogenes, evolutionary rate constancy was examined. The likelihood ratio test (LRT) of the molecular clock (Felsenstein, 1981; Sorhannus & Bell, 1999) was used to evaluate a likelihood ratio for the null and alternative hypothesis. The details for these tests are provided in the supplementary methods.

Phylogenetic trees were constructed in order to examine the intraspecific patterns of polymorphic rRNA sequences within individual nematode specimens. A maximum-likelihood (ML) analysis was performed with 1,000 bootstrap replicates under the GTRCAT model using RAxML 8.1.11 (Stamatakis, Hoover, & Rougemont, 2008). A Bayesian inference (BI) was carried out with the GTR + I + G model using MrBayes 3.2.3 (Ronquist et al., 2012). The ML and BI analyses were performed at the CIPRES Science Gateway (Miller, Pfeiffer, & Schwartz, 2010).

2.5 | Simulated metabarcoding analysis

The simulated analyses using PCR clones were carried out to assess the potential impact of rRNA polymorphisms on the clustering of Operational Taxonomic Units (OTUs) and nematode taxonomy assignments derived from DNA sequences. PCR clone sequences were first aligned using the default parameters of the SINA online aligner (Pruesse, Peplies, & Glöckner, 2012). Two mock metabarcoding data sets were exported from the trimmed 18S rRNA sequence alignments: a ~380 bp region approximately spanning the V1/V2 hypervariable region, and a ~390 bp region approximately spanning the V8/V9 hypervariable region. Subsequently, mock metabarcoding sequences were subjected to open-reference OTU clustering in the QIIME v1.9.1, using a 99% similarity cut-off. Taxonomic assignments of the resulting metabarcoding OTUs were completed using the naive Bayesian RDP classifier (Wang, Garrity, Tiedje, & Cole, 2007) trained on the QIIME-formatted SILVA 132 database (reference sequences clustered at 99%). Other details are included in the supplementary methods.

3 | RESULTS

3.1 | Intragenomic polymorphisms are present across species and life strategies

In this study, we analysed 445 de novo PCR amplicon sequences belonging to 24 species collected in 13 countries, and analysed 60 PCR amplicon sequences (6 species) and the genomes of 5 species (16 samples, each from one individual or from group of genetically identical individuals depends on genome sources) from previous studies. Our results show that rRNA polymorphisms are widely present in different species of nematodes associated with various life strategies. Among them, we detected polymorphisms in 22 out of 32 species. The 18S gene in the plant-parasitic species *Rotylenchulus macrosoma* had the highest rRNA variation (29.9% maximum pairwise differences) due to a *c.a.* 400-bp deletion (two sequences for 1,300 bp vs. ten sequences for about 1764 bp), followed by ITS1-5.8S-ITS2 in the free-living *Cephalenchus* sp.1 (18.6% maximum pairwise differences), while other species have variation less than 10% (Figure 1, Table 1, Table S4). Within a nematode genus, the presence of polymorphisms varied greatly among species, ranging from widely present (at least one gene in *Cephalenchus*, *Pratylenchus*, *Malenchus* and *Rotylenchulus*), to relatively less common (two out of six species in *Halicephalobus*), to no polymorphisms detected at all (*Filenchus*).

Among the three analysed loci, ITS1-5.8S-ITS2 shows the highest average variation of 7.7%, while 18S and 28S share similar 4.5% and 4.8% nucleotide differences. However, differences across loci were not significant (p = .381; Figure 1f). Similarly, the maximum polymorphism percentage shows no difference between nematode feeding types and trophic groups (c-p value). We also detected significant differences between trophic groups and genetic distance (p = .046), and a post hoc analysis suggested that free-living nematodes with a short generation time and a high reproduction rate (c-p = 2) have more divergent rRNA gene copies compared to species with similar reproductive strategies that are root epidermal feeders (plant feeder c-p = 2). These results indicate that nematode life strategies may have a strong impact on the development and maintenance of rRNA polymorphisms. Genetic distances are positively correlated with percentage differences except for a few cases; for example, *Cephalenchus* sp.1 shows the highest intragenomic variability (0.0719), followed by 18S in *R. macrosoma* (0.0465). Subsequent analysis revealed that polymorphic sites were mostly randomly distributed in ITS1-5.8-ITS2 and 28S region, or concentrated in specific regions in the 18S gene (e.g. 78% of the polymorphic sites in 18S rRNA of *H. gingivalis* are concentrated in two distinctive regions).

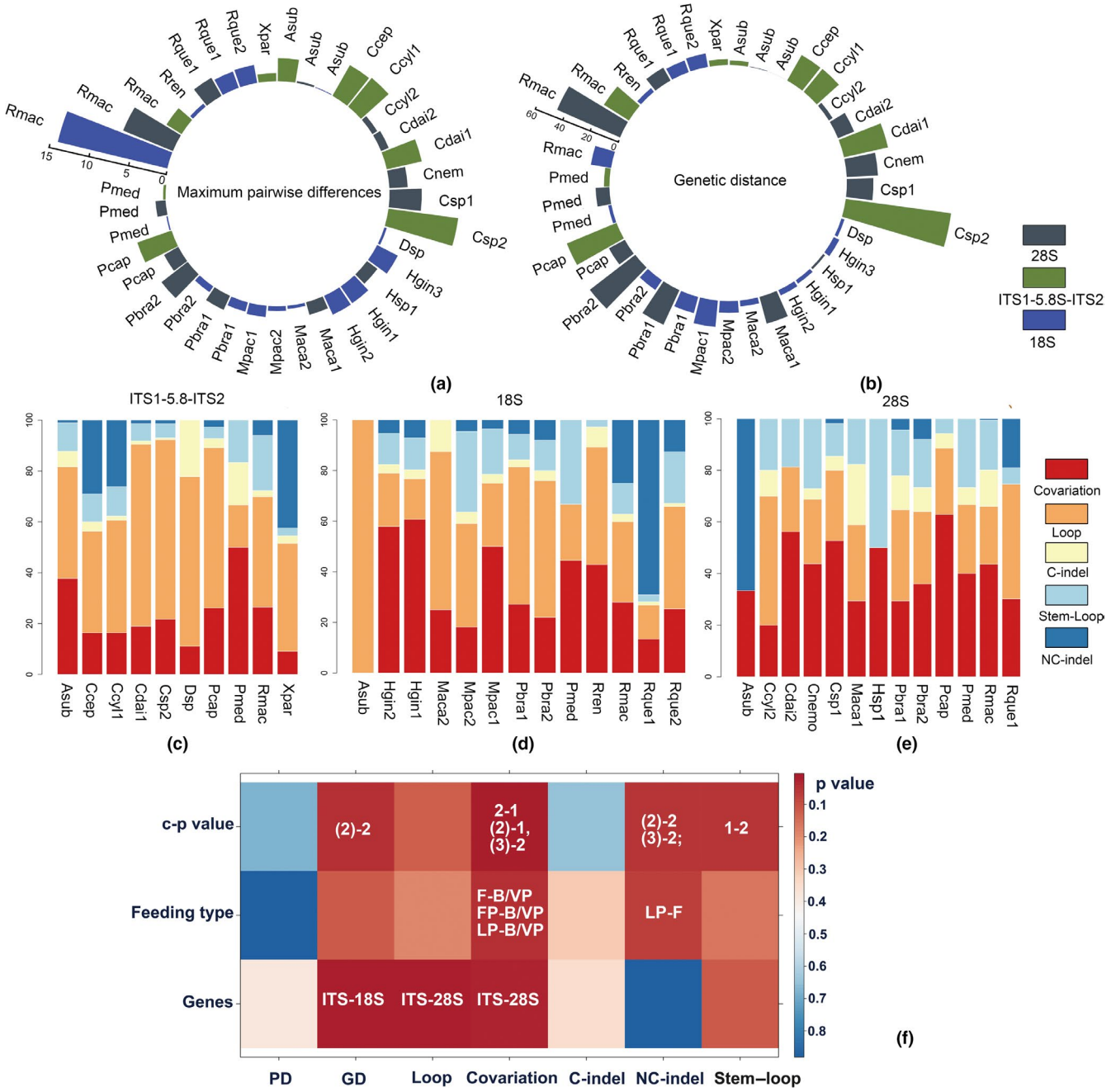

FIGURE 1 The intragenomic rRNA polymorphic variation measured in maximum pairwise differences (a), genetic distances (b) and the compositions of polymorphic mutation in rRNA secondary structure (c–e), heatmap of *p*-value in ANOVA, compared between nematode life strategies or genes and polymorphism measured by different factors (f). The maximum pairwise differences provide the highest variation one species can reach, while genetic distances consider the variation among polymorphic copies. The conserved mutation types that do not alter base pairing and structure were defined as (1) covariation: located in a stem, but causing GC:GU conversions or covariation resulting in no change in base pairing mismatch; (2) loop: mutations are in the loop; or (3) C-indel: conserved indel due to the deletion of a single unpaired nucleotide in stem. The nonconserved mismatch that changes the structure was classified as a (1) stem–loop: stem-to-loop transition caused by several deletions or insertions; or (2) NC-indel: nonconserved indel introduced by single insertion. The colour indicates the significance level from significantly different (red) to no difference (blue). Post hoc tests were conducted when $p < .1$ and revealed significant factor pairs are given as text at heatmap. GD, genetic distance; ITS, ITS1-5.8S-ITS2; PD, percentage of difference [Colour figure can be viewed at wileyonlinelibrary.com]

We further compared intragenomic rRNA variation among individuals within populations of *H. gingivalis and R. quercophilus*. The results showed that one individual did not always possess all polymorphisms, indicating either that the presence of polymorphisms differs across individual nematodes, or alternatively that our limited sampling strategy did not capture the full range of polymorphisms present in each specimen.

3.2 | rRNA secondary structure reveals specific polymorphism sites

The mutations are not randomly distributed but more concentrated in few sites. Among all helixes, 23/e1 and 23/e4 (in V4 region) are most variable (e.g. Figures S1 and S2, concurring with Bert et al. (2008)). The mapping of polymorphic rRNA sites shows that most mutations are conservative changes that maintain the structure of functional domains. In nonconservative changes, most changes were single paired–unpaired changes and none of them appeared to cause loss of a helix or structural rearrangements (Figure 1c–e). Comparisons of mutation types among genes suggest that ITS1-5.8-ITS2 gene harbours significant more polymorphic sites in loop regions (p = .015) and in the type of covariation changes (p = .031) than 28S gene. The covariation and nonconserved indel mutations are significantly different (vs. feeding types) among feeding types (p = .042 to covariation, p = .069 to nonconserved indel) and trophic groups (p = .005 to covariation, p = .055 to nonconserved indel), while stem-to-loop transition is only significantly different (p = .052) in trophic groups (Figure 1f).

3.3 | Intraspecific polymorphism, a specific trait independent of habitat or geography

Halicephalobus gingivalis is a free-living bacterivorous nematode, which is also capable of facultative parasitism in equines. Among 73 examined locations, *Halicephalobus* spp. were found in compost, soil or manure heaps at eight locations, and we subsequently collected fresh horse dung and equine rectal samples from selected locations. Thirty-one out of 56 fresh horse dung samples and two out of 11 equine rectal samples revealed the presence of *Halicephalobus* spp. A total of 18 populations of *H. gingivalis* were identified via morphometric (30 individuals were measured for each population; see Table S5) and molecular data, and subsequently used for analysis together with five additional *Halicephalobus* species (Figure 2). Our results showed that rRNA polymorphisms were present in all 18 populations of *H. gingivalis* isolated from soil, manure heaps, compost heaps, horse dung and infected horse tissue, while polymorphic sites are absent in five other species from the same genus (Table 1, Figure 2). This suggests that the presence of rRNA polymorphisms is more likely to represent a species instead of a population-level character associated with specific habitats or population-specific life strategies.

The analysis based on two *Pratylenchus brachyurus* populations isolated from Ghana and Nigeria reveals that rRNA polymorphisms can be prevalent across wide geographic ranges with only slight differences in polymorphism diversity (see the π value in Table S6). Thus, both *P. brachyurus* and *H. gingivalis* showed similar polymorphism levels among geographically distinctive populations (Ghana and Nigeria populations for *P. brachyurus*, and Belgium and US populations for *H. gingivalis*), suggesting the presence and level of polymorphisms seem to be independent of geography.

3.4 | Intraspecific polymorphism strongly impacts phylogenetic patterns

The phylogenies of rRNA polymorphisms inferred from Bayesian analyses are shown in Figures S3–S36. The analysed polymorphic sequences exhibited two different patterns: (a) polymorphic sequences are polyphyletic, forming two well-supported clades that appear as sister clades to another species (Figure 3a,b), and (b) polymorphic sequences form a single well-supported clade, clearly distinct from sister species (Figure 3c,d). This second scenario is the most commonly observed pattern and found in all species except for the genus *Rotylenchulus*. The polyphyletic polymorphic sequences in *Rotylenchulus* suggest that this phylogeny is based on paralogs that located in different rRNA gene clusters or chromosomes, while the monophyletic placements suggest that other phylogenies are based on orthologous sequences.

The influence of polymorphisms on phylogeography patterns appears to be variable. In the 18S rRNA gene phylogeny, the Ghana populations of *Pratylenchus brachyurus* formed a relatively well-supported clade with respect to unresolved Nigeria sequences (Figure 3d), whereas phylogeographic patterns are absent in the 28S rRNA phylogeny (Figure 3c). Conversely, both according to 28S and ITS1 sequences, *R. macrosoma* sequences tend to be clustered by their geographic origin although clades are not always well supported (Figure 3a,b). Compared with *P. brachyurus* where each of the recovered 28S copies is different (24 sequences), the *R. quercophilus* has less diversity in 28S polymorphism, where 10 out of 42 copies are identical (Figure S35). Previous studies have shown that rRNA polymorphisms form few (two or three) well-separated clades and each of these clades includes all similar sequences (Muir, Fleming, & Schlötterer, 2001; Van Den Berg et al., 2016; Zhan et al., 2011). The multiple lineages observed suggested the rRNA polymorphism in *P. brachyurus* is very high and our limited clone sampling was not able to recover all the existing rRNA polymorphisms.

3.5 | rRNA polymorphisms show neutral rather than concerted evolution

Our results suggest that most analysed nematodes have a negative value of Tajima's D (23 out of 30 tested populations and genes), Fu and Li's D* (25 out of 30 tested populations and genes), and Fu and Li's F* (23 out of 30 tested populations and genes) (Table S6), but most of them are not significantly supported by p-values from coalescent simulations. Therefore, rRNA polymorphisms appear to have evolved neutrally in most of the analysed nematode species. These polymorphic copies are relatively stable in the genome with no ongoing concerted evolution. Conversely, a few species show negative values with significant support (p < .1) in all three tests, that is the 28S of *Cephalenchus* sp.1, the ITS1-5.8S-ITS2 of *C. cephalodiscus*, *Deladenus* sp., the 28S of *H. gingivalis*, the 18S of *Rotylenchulus reniformis* and 18S of *R. quercophilus* population Rque1, suggesting that the number of rRNA tandem arrays may still be expanding in

these species. In addition, neutrality tests for some species are not in agreement with each other; for example, the 28S of *R. macrosoma* has a significant negative Tajima's D value ($p < .02$) but not for Li's D and Fu and Li's F* value, while the opposite is true for 28S of *R. quercophilus* where it is significantly negative for Li's D ($p < .02$) and Fu and Li's F* ($p < .05$) but not for Tajima's D. Despite the limited number of analysed sequences and the uncertainty of neutrality test values, we consider the rRNA gene copies of *R. macrosoma* and *R. quercophilus* to be neutrally evolving.

3.6 | Different evolutionary rates indicate that pseudogenes or dysfunctional rRNA variants may exist

Assuming that all rRNA copies have maintained their function, they should be exposed to similar evolutionary constraints and thus show similar rates of evolution. If some of them have lost function (pseudogenes), then they are expected to show an elevated rate of evolution. To discriminate between these two alternative hypotheses, we conducted the LRT of global and local molecular clocks. The results are presented in Table S7. Most of the analysed species (26 of 36 tested populations and genes) show significantly variable substitution rates in rRNA polymorphism copies ($p < .1$), suggesting different evolutionary rates within the same individual. This is likely introduced by contrasting selective pressures and pseudogenes, or dysfunctional rRNA subunits may thus exist in those fast-evolving gene copies. Ten populations or genes belonging to eight species (*A. subtenuis*, *R. reniformis*, *X. parachambersi*, *P. mediterraneus*, *M. pachycephalus*, *M. acarayensis*, *C. cylindricus* and *C. daisuce*; Table S7) have rRNA copies that evolved with more constant rates among lineages, suggesting these genes are equally functional and pseudogenes may not exist.

3.7 | Genome analysis reveals polymorphisms are significantly different among species

The genome analysis of five nematode species (16 samples) suggests that the polymorphisms are intraspecifically stable among sequenced samples, although more variations are found in ITS1 and ITS2 of *M. incognita*. The highest polymorphism levels were observed in the plant-parasitic nematodes *R. reniformis* and *M. incognita*, in which SNP counts were substantially higher in all five rRNA

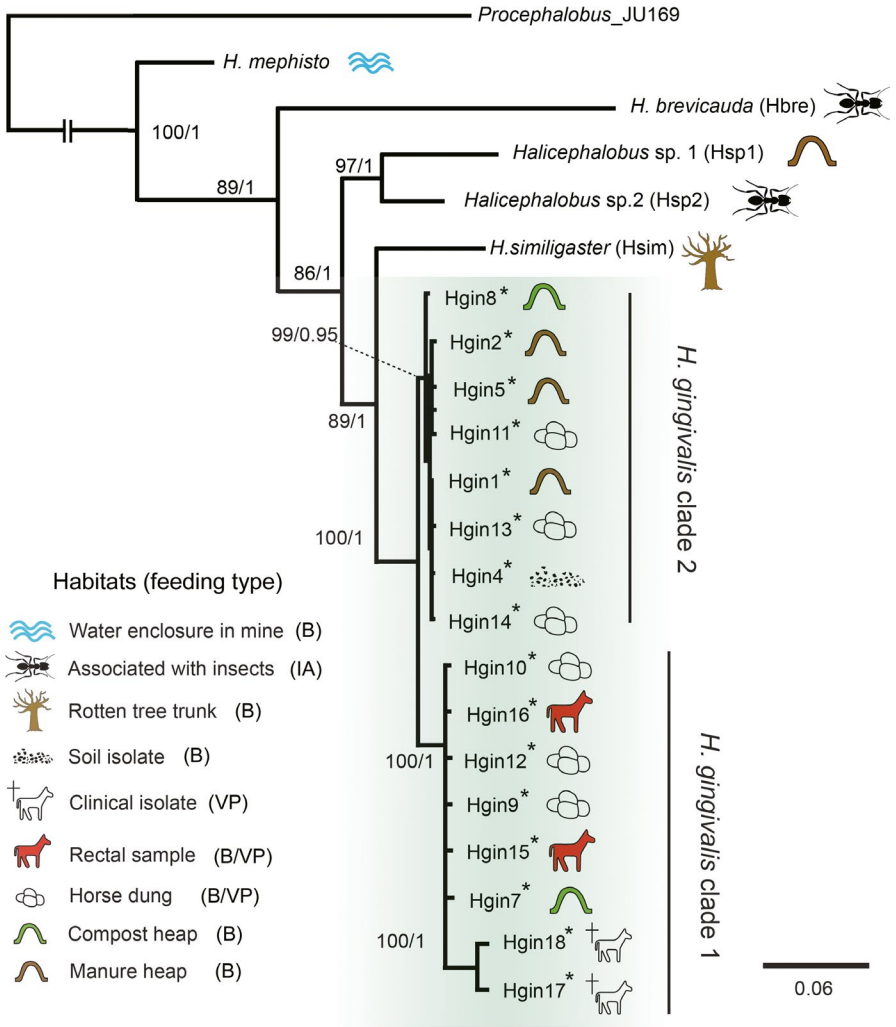

FIGURE 2 Partitioned Bayesian inference and maximum-likelihood phylogeny based on a concatenated data set of four genes (18S, 28S, ND4 and COI). Values at nodes are presented as RAxML bootstrap support/Bayesian posterior probabilities. The origin of each isolate is schematically represented. Isolates originally containing polymorphic positions in 18S sequences are indicated with an asterisk (*). The abbreviations of feeding types in parentheses are corresponding to Table 1. The morphometric data for *Halicephalobus* spp. used in the phylogeny are listed in Table S5. Scale bar denotes nucleotide substitutions per site [Colour figure can be viewed at wileyonlinelibrary.com]

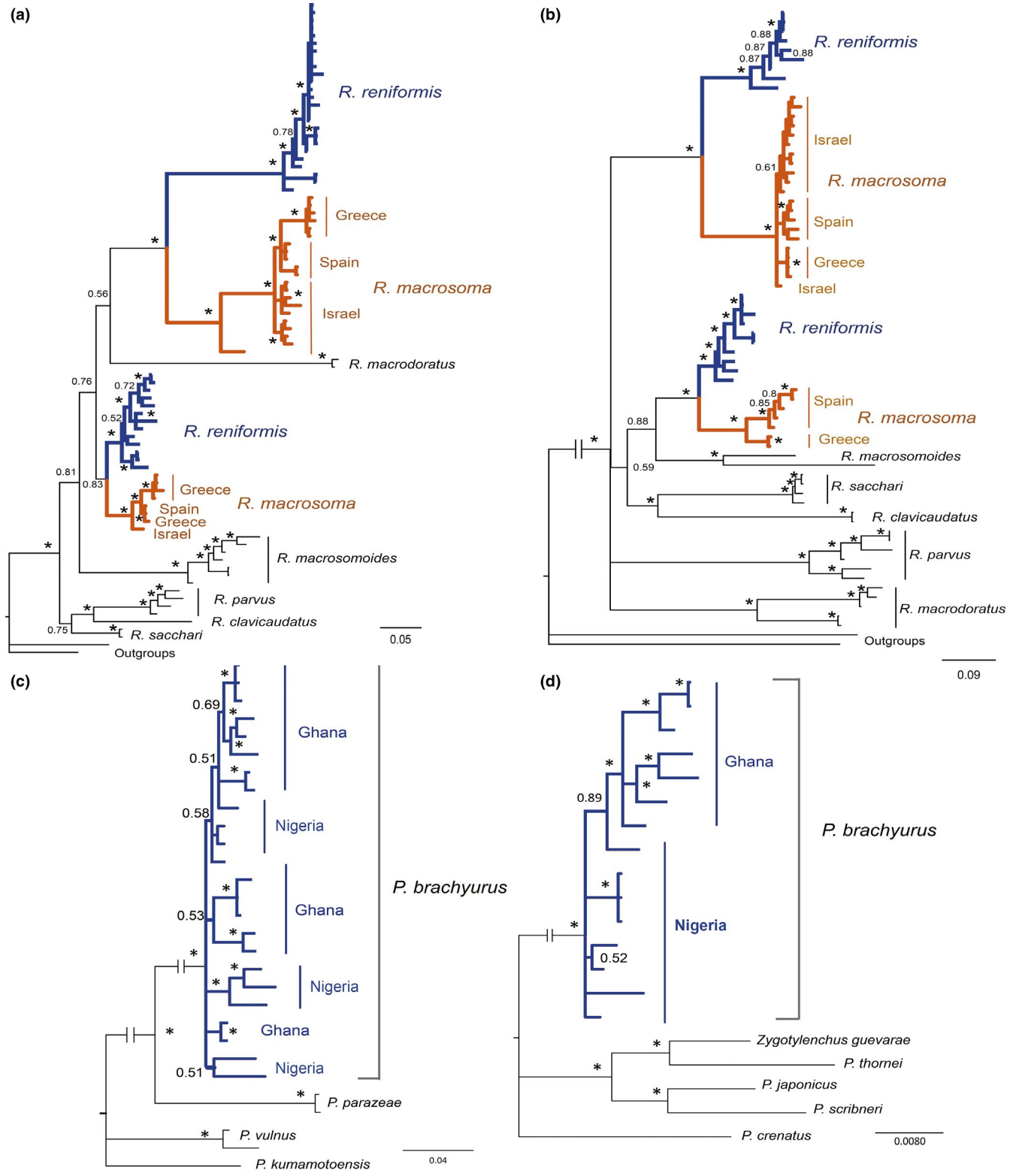

FIGURE 3 The two types of phylogenetic placements for polymorphic sequences. (a, b) polymorphic sequences are polyphyly, species are mixed. (c, d) All polymorphic sequences are monophyly, species are well separated. (a, c) 28S, (b) ITS1, (d) 18S. Scale bar denotes nucleotide substitutions per site. Bayesian posterior probabilities are indicated at the node or marked as an asterisk (*) when the value higher than 0.9 [Colour figure can be viewed at wileyonlinelibrary.com]

regions in comparison with other species (Figure 4). In contrast, the entomoparasitic *Heterorhabditis bacteriophora* had very few SNPs, which were filtered out due to low coverage (<10% of the reads). The plant-parasitic nematode *Globodera rostochiensis* had an intermediate SNP count, but this varied significantly among rRNA genes: polymorphism levels were comparable to *M. incognita* in 28S rRNA

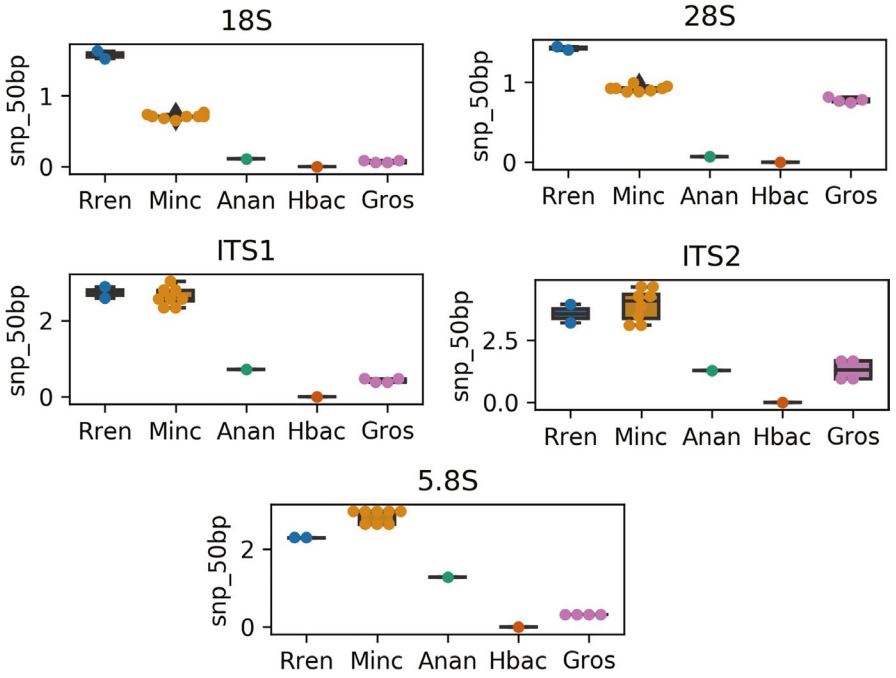

FIGURE 4 Intra- and interspecies diversity of SNP counts in rRNA regions. Boxplots represent the distribution of SNP numbers per 50 bp in different rRNA regions (18S, ITS1, 5.8S, ITS2 and 28S). Each dot denotes the SNP count per 50 bp in one sample. Anan, *Acrobeloides nanus*; Gros, *Globodera rostochiensis*; Hbac, *Heterorhabditis bacteriophora*; Minc, *Meloidogyne incognita*; Rren, *Rotylenchulus reniformis* [Colour figure can be viewed at wileyonlinelibrary.com]

but much lower in ITS regions and other subunits (Figure 4). The relatively high level of SNP count variability among isolates of some of the species (*M. incognita*, *G. rostochiensis*) persisted in spite of rigorous SNP filtering. During this step, positions in which at least 90% of the reads had the same character state were skipped to increase the confidence in the results, at the expense of detection sensitivity. Out of five rRNA regions, the 18S and 28S genes were most conserved among samples of the same species, while ITS regions and the 5.8S gene were more variable, in agreement with our PCR amplicon-based analysis.

3.8 | rRNA polymorphisms have a minor impact on simulated metabarcoding analysis

Mock metabarcoding analyses were carried out on 18S rRNA clone sequences, comprising 109 sequences for the V1/V2 region (representing 8 nematode species) and 89 sequences for the V8/V9 region (representing 7 nematode species). For the majority of nematode species examined, sequence reads were clustered into a single 99% OTU corresponding to the correct species and this OTU did not share sequence reads with any other nontarget nematode species. Some nematode species were recovered as having two (Hgin1 in the V1/V2 data set, and Rmac, Rque1 and Pbra1 in the V8/V9 data set) or three (Rque2, Pbra1, Pbra2 in the V1/V2 data set, and Pbra2 in the V8/V9 data set) 99% OTUs presumably indicative of intragenomic rRNA polymorphisms, although we note that sequence reads within variant OTUs were typically not shared with any other species. In some cases, an OTU had sequence reads assigned to two separate species belonging to the same nematode genus, which is not entirely surprising considering the slow evolution of the 18S rRNA locus and lack of nucleotide differences between closely related species (Floyd, Abebe, Papert, & Blaxter, 2002). Some OTUs were given high-level, uninformative taxonomic assignments (e.g. 'Metazoa' or assignments to the nematode orders Tylenchida or Rhabditida), while other OTUs were correctly assigned down to the species level (*H. gingivalis*, *R. reniformis*, *A. subtenuis*). In some cases, OTUs were assigned down to the wrong species in the correct genus (e.g. *Pratylenchus*, *Rotylenchulus*), indicating that the quality and accuracy of rRNA-derived taxonomy could vary widely even within the same genus of nematodes. Variations in taxonomic assignments are likely related to persistent taxonomic gaps in the SILVA database (e.g. where the global genetic diversity of soil nematodes is vastly underrepresented and many nematode lineages lack a reference sequence) and the particulars of metabarcoding taxonomy assignment algorithms (which require a robust training set of reference sequences for optimal performance). The choice of 18S metabarcoding region (V1/V2 vs. V8/V9) did not impact the observed OTU clustering or taxonomic assignments, and the same patterns were observed across both regions.

4 | DISCUSSION

4.1 | Origin and possible function of rRNA polymorphisms

The eukaryotic rRNA array consists of several hundred tandemly repeated copies but homogenized to nearly identical by concerted evolution (Arnheim, 1983). However, exceptions to this rule have emerged. For example, the flatworm *Dugesia mediterranea* contains two different types of SSU rRNA within single organisms with an overall sequence divergence of 8% (Carranza, Giribet, Ribera, & Riutort, 1996). In Nematoda, rRNA polymorphisms have also been reported in different taxa, for example in *Meloidogyne*

(Blok, Phillips, & Fargette, 1997; Hugall, Moritz, & Stanton, 1999; Zijlstra, Lever, Uenk, & Van Silfhout, 1995), *Heterodera* (Szalanski, Sui, Harris, & Powers, 1997), *Belonolaimus* (Cherry, Szalanski, Todd, & Powers, 1997), *Haemonchus* (Heise, Epe, & Schnieder, 1999), *Globodera* (Subbotin, Perry, Warry, & Halford, 2000), *Nematodirus* (Heise et al., 1999; Nadler, Hoberg, Hudspech, & Rickard, 2000) and *Bursaphelenchus* (Cardoso, Fonseca, & Abrantes, 2012).

Our results confirm that rRNA polymorphisms are widely present across nematode lineages, and intragenomic rRNA variation is commonly found in many species regardless of life history strategies. In agreement with previous studies, lower levels of polymorphism are observed in 18S and 28S rRNA gene regions compared to noncoding ITS rRNA (Bik et al., 2013), and the conserved nature of most mutations recovered from 18S/28S subunit genes emphasizes the need for species to maintain the structure of rRNA functional domains. We observed no major nonconservative changes that would cause loss of an rRNA helix or another type of structural rearrangement, suggesting that the appearance and maintenance of intragenomic rRNA polymorphisms in nematodes do not generally impact the cellular functions of ribosome molecules. It is currently unknown whether some rRNA variants could be indicative of pseudogenes. The presence of rRNA pseudogenes has been noticed in different species (Muir et al., 2001; Pei et al., 2010; Xu et al., 2015). Although the exact reasons for this remain unclear, some studies have shown that in addition to ribosome biosynthesis, the rRNA and nucleolus may also play prominent roles in evolutionary processes of ageing, gene regulation and retaining the integrity of the genome (Ide et al., 2010; Kobayashi, 2008; Kobayashi & Ganley, 2005; Sinclair & Guarente, 1997). It has also been demonstrated that the mRNAs of some pseudogenes are able to regulate the expression of other genes (Saka, Ide, Ganley, & Kobayashi, 2013; Tam et al., 2008). Although the secondary structure and evolutionary rate analysis of rRNA variant sequences in the present study support the existence of functional ribosomal subunits in nematodes, it is possible that our PCR clone-based approach did not capture the full spectrum of rRNA variants present in any given genome. Further investigation is necessary to clarify whether the rRNA pseudogenes are present in the genomes of the nematode species studied to date.

Contrasting phylogenetic patterns and geographic structuring of rRNA variant sequences, as well as statistically significant differences between nematode trophic groups and genetic distance, suggest that some factors relating to nematode ecology or reproduction could be playing a strong role in the development and maintenance of rRNA polymorphisms in some species. Higher rRNA divergence was characteristic of free-living nematodes with short generation time and high reproduction rate (c-p = 2). Increased levels of intragenomic rRNA variation in these species may contribute to rapid growth rates and ecological strategies that facilitate the rapid exploitation of environmental resources (Elser et al., 2000), thus conferring a competitive advantage in natural ecosystems. Evolutionary tests indicated that the majority of rRNA polymorphisms are neutrally evolving, and rRNA variants are stable in the genome and not subjected to ongoing concerted evolution. However, we obtained conflicting results in neutrality tests of some species (e.g. *Rotylenchulus* spp.), and negative values recovered may imply positive selection or an increase in rRNA copy number is occurring. Furthermore, contrasting geographic patterns were recovered in the phylogenies of 18S vs. 28S rRNA variant sequences in *P. brachyurus* (Figure 3), suggesting population-level factors may strongly influence the persistence and divergence of specific rRNA variants in nematode genomes. Our limited sample set does not allow for rigorous assessment of these distinct patterns, and further targeted studies are needed to evaluate the evolutionary forces that may be at play.

4.2 | rRNA patterns in the nematode genus *Halicephalobus*

Ribosomal RNA polymorphisms in *H. gingivalis* had been previously suspected in the 28S sequence of *H. gingivalis* (Nadler et al., 2003); however, their use of pooled individuals did not allow the authors to determine whether these copies represented intraindividual heterogeneity or variation between individuals. In our study, the occurrence of rRNA polymorphisms in both sequenced strands of individual *H. gingivalis* specimens reflects genuine sequence variants. Theoretically, the analyses of additional clones of each isolate could reveal additional rRNA sequence variants and determine their corresponding proportion.

PCR-mediated recombination could be playing a role in generating the observed rRNA polymorphisms. However, this is usually present as a low proportion of sequences (Xu et al., 2015) and is unlikely to explain polymorphisms in present study for two reasons: first, clones from other species in the same experiment have shown to exhibit significant homogeneity (all 28S clones of *H. gingivalis* and 18S clones of *H. mephisto* were completely identical, and 18S clones from *A. subtenuis* exhibited only a single nucleotide deletion); and second, instead of being randomly distributed, the mutation sites tend to be located in the variable helix of rRNA secondary structure revealed by previous studies. Thus, our data suggest that rRNA polymorphisms are biologically valid results and intragenomic rRNA variation is a key feature of nematode genomes.

The rRNA repeats can escape concerted evolution in several circumstances. An important assumption for evolution under the concerted evolution model is the organization of genes in tandem arrays (Pillet, Fontaine, & Pawlowski, 2012). Recombination and gene conversion occurs less frequently between sequences on heterologous chromosomes than on homologous chromosomes (Ironside, 2013). Therefore, in the rare case that rRNA repeats are not strictly organized in tandem arrays and are instead dispersed throughout the genome, concerted evolution acts less strongly, as was suggested for the 5.8S rRNA heterogeneity in the amphibian *Xenopus laevis* (Peterson, Doering, & Brown, 1980) and the loach fish *Misgurnus fossilis* (Mashkova et al., 1981). DNA heterogeneity can also be caused by polyploidization, which has on several occasions been described in nematodes, for example in *Meloidogyne* (Triantaphyllou & Hirschmann, 1997). In the case of polyploidization, concerted evolution is restricted to each set of chromosomes

that encode for a different ribotype, as was suggested for the microsporidian *Nosema bombi* (O'Mahony, Tay, & Paxton, 2007). Other proposed explanations for rRNA heterogeneity include interspecific hybridization events, as suggested for *Meloidogyne* (Fargette et al., 2010; Lunt, 2008) and the foraminiferan *E. macellum* (Pillet et al., 2012).

Although the origin of the rRNA heterogeneity in *Halicephalobus* is uncertain, it is remarkable that only the isolates of the *H. gingivalis* clade, including facultative parasites, exhibit polymorphic regions in their 18S rRNA genes. In *Plasmodium*, the expression of two distinct types of 18S rRNA has been shown to be developmentally regulated (i.e. correlated with discrete stages of sporozoite development), linked to different parasitic life stages (i.e. one predominantly expressed in the mosquito host and the other in the mammalian host (Gunderson et al., 1987; Li, Wirtz, McConkey, Sattabongkot, & McCutchan, 1994; Rogers et al., 1996)), and possibly related to functional differences between ribosome types or mechanisms of transcriptional control. Thus, it is possible that the presence of varying life histories in *Halicephalobus* is the key to understanding the presence or absence of rRNA variation and heterogeneity. For example, in *H. gingivalis* rRNA polymorphisms could facilitate the survival of nematode populations in extremely different habitats, and consequently enabling individuals in this species to be opportunistically parasitic.

4.3 | Impact on metabarcoding analyses

Although rRNA polymorphisms may complicate attempts to definitively link molecular OTUs with biological species, the most pressing challenges for high-throughput metabarcoding studies are sparse eukaryotic databases and an overall lack of reference DNA barcodes for most nematode species. PCR clone sequences from a single nematode species or geographic population were largely grouped into one OTU representing the target species, with no sequence reads being shared with other species or OTUs. Nematode species in the genera *Rotylenchulus* and *Pratylenchus* were commonly split into 2–3 OTUs, but in most cases, the distribution of sequence reads followed a characteristic 'Head-Tail' pattern showing a dominant OTU containing the most sequence reads; this pattern is commonly observed in nematode metabarcoding studies (Porazinska et al., 2010; Porazinska, Giblin-Davis, Powers, & Thomas, 2012) and can be leveraged in bioinformatic workflows to separate species barcodes (e.g. the dominant rRNA variant typically recoverable through PCR and Sanger sequencing) from minor intragenomic variants. However, OTU patterns reflective of rRNA polymorphism may be significantly impacted by the choice of rRNA metabarcoding locus. Although we did not observe any distinct patterns between OTUs clustered from the V1/V2 and V8/V9 hypervariable regions of the 18S gene, our secondary structure analysis suggested that the V4 hypervariable region may be a hot spot of polymorphism and sequence divergence in nematode 18S rRNA. The V4 region is a common alternative target locus for eukaryotic and metazoan metabarcoding studies (Decelle, Romac, Sasaki, Not, & Mahé, 2014; Stoeck et al., 2010), and thus, higher rRNA polymorphisms in this gene region could potentially inflate species diversity estimates even further. The choice of rRNA metabarcoding locus and the impact of intragenomic rRNA variation on the recovery of OTUs across different hypervariable 18S gene regions are topics worthy of further investigation.

In contrast to OTU clustering patterns, taxonomic assignments in the mock metabarcoding data set were generally poor regardless of whether or not an OTU represented a dominant or minor rRNA variant copy. Species-level taxonomic assignments were correct in only four out of 19 mock 18S metabarcoding samples (Figure S39), indicating that reference databases did not contain a sufficient number of nematode sequences to make robust low-level assignments or, alternatively, that the RDP classification algorithm was performing poorly on our eukaryotic data set where taxonomic hierarchies may be conflicting or unrepresentative of the underlying molecular diversity within a morphologically defined genus. Order-level taxonomic assignments (Tylenchida, Rhabditida) make it almost impossible to partition molecular OTUs into putative biological species without using sequence read number or relative abundance as a guide, and the level of rRNA polymorphism across species is likely to complicate matters further. More accurate metabarcoding taxonomic assignments could partially alleviate some issues related to bioinformatic species delimitation, for example by flagging genera or species known to exhibit high levels of rRNA polymorphisms such as *H. gingivalis* isolates identified in the present study. Regardless of the accuracy of taxonomic assignments, the presence of intragenomic polymorphisms will always result in inflated species estimates when relying on clustered metabarcoding OTUs derived from common rRNA loci. The number of OTUs recovered from any given nematode specimen is known to be highly variable and not related to body size or taxonomy (Porazinska et al., 2010, 2012); such intragenomic rRNA variation confounds biodiversity estimates across all microbial eukaryote groups, and currently cannot be overcome with simple bioinformatic workarounds such as the use of different pairwise identity cut-offs (Thiéry et al., 2016; Thornhill, Lajeunesse, & Santos, 2007). The move towards 'Amplicon Sequence Variants' (essentially 100% OTUs; Callahan, McMurdie, & Holmes, 2017) and an increasing use of phylogenetic approaches in metabarcoding studies (Washburne et al., 2019) may soon help to overcome some of the inherent challenges faced in linking specimen-level data with patterns observed in environmental metabarcoding studies.

5 | CONCLUSIONS

A growing body of evidence now suggests that intragenomic rRNA variation is a key feature of nematode genomes. However, the presence of polymorphisms can be highly variable across taxa and species within a single genus can show highly contrasting patterns. Our study lends support to the hypothesis that geography and nematode life history factors may play a key role in the generation and maintenance of rRNA polymorphisms. However, further research

and systematic investigations of specific nematode taxa are needed in order to clearly elucidate the biological and evolutionary explanations of the observed rRNA variations. Regardless of the driving mechanisms for intragenomic rRNA polymorphisms and copy number variation, the existence of minor variant gene copies presents substantial challenges for molecular-based biodiversity estimates and the analysis of environmental metabarcoding and metagenomic data sets. Further investigation is needed across a variety of metazoan taxa in order to evaluate the potential impact of rRNA variation on high-throughput studies of natural ecosystems.

ACKNOWLEDGMENTS

This work was supported by a special research fund UGent 01N02312. We thank Dr. Yao Kolombia for providing *Pratylenchus*, Dr. Toon Janssen for his supports in phylogeny analysis and Andy Vierstraete for his help in sequencing.

AUTHOR CONTRIBUTIONS

W.B. and X.Q. designed the experiment. P.F., J.G. and S.B. provided materials, and conducted sampling, PCR and cloning. X.Q., T.Y. and A.S. analysed data. All authors discussed the results. X.Q., H.B. and W.B. wrote the manuscript.

ORCID

Xue Qing https://orcid.org/0000-0002-0355-9956

DATA AVAILABILITY STATEMENT

The authors confirm that the data supporting the findings of this study are available in the supplementary materials. Other raw data are available from the first author on request.

REFERENCES

Alper, I., Frenette, M., & Labrie, S. (2011). Ribosomal DNA polymorphisms in the yeast *Geotrichum candidum*. *Fungal Biology*, *115*(12), 1259–1269. https://doi.org/10.1016/j.funbio.2011.09.002

Arnheim, N. (1983). Concerted evolution of multigene families. In N. Arnheim, M. Nei, & R. K. Koehn (Eds.), *Evolution of genes and proteins* (pp. 38–61). Sunderland, MA: Sinauer Associates.

Arnheim, N., Krystal, M., Schmickel, R., Wilson, G., Ryder, O., & Zimmer, E. (1980). Molecular evidence for genetic exchanges among ribosomal genes on nonhomologous chromosomes in man and apes. *Proceedings of the National Academy of Sciences of the United States of America*, *77*(12), 7323–7327. https://doi.org/10.1073/pnas.77.12.7323

Bert, W., Leliaert, F., Vierstraete, A. R., Vanfleteren, J. R., & Borgonie, G. (2008). Molecular phylogeny of the Tylenchina and evolution of the female gonoduct (Nematoda: Rhabditida). *Molecular Phylogenetics and Evolution*, *48*(2), 728–744. https://doi.org/10.1016/j.ympev.2008.04.011

Bik, H. M., Fournier, D., Sung, W., Bergeron, R. D., & Thomas, W. K. (2013). Intra-genomic variation in the ribosomal repeats of nematodes. *PLoS ONE*, *8*(10), e78230. https://doi.org/10.1371/journal.pone.0078230

Blok, V. C., Phillips, M. S., & Fargette, M. (1997). Comparison of sequences from the ribosomal DNA intergenic region of *Meloidogyne mayaguensis* and other major tropical root-knot nematodes. *The Journal of Nematology*, *29*(1), 16–22.

Bongers, T., & Bongers, M. (1998). Functional diversity of nematodes. *Applied Soil Ecology*, *10*(3), 239–251.

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*(12), 2639–2643. https://doi.org/10.1028/ismej.2017.119

Cardoso, J. M. S., Fonseca, L., & Abrantes, I. (2012). Genetic diversity of ITS sequences of *Bursaphelenchus xylophilus*. *Genetics and Molecular Research*, *11*(4), 4508–4515. https://doi.org/10.4238/2012.October.15.1

Carranza, S., Giribet, G., Ribera, C., & Riutort, M. (1996). Evidence that two types of 18S rDNA coexist in the genome of *Dugesia* (Schmidtea) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida). *Molecular Biology and Evolution*, *13*(6), 824–832. https://doi.org/10.1093/oxfordjournals.molbev.a025643

Cherry, T., Szalanski, A. L., Todd, T. C., & Powers, T. O. (1997). The internal transcribed spacer region of *Belonolaimus* (Nemata: Belonolaimidae). *The Journal of Nematology*, *29*(1), 23–29.

Cui, H.-L., Zhou, P.-J., Oren, A., & Liu, S.-J. (2009). Intraspecific polymorphism of 16S rRNA genes in two halophilic archaeal genera, *Haloarcula* and *Halomicrobium*. *Extremophiles*, *13*(1), 31–37. https://doi.org/10.1007/s00792-008-0194-2

Decelle, J., Romac, S., Sasaki, E., Not, F., & Mahé, F. (2014). Intracellular diversity of the V4 and V9 regions of the 18S rRNA in marine protists (Radiolarians) assessed by high-throughput sequencing. *PLoS ONE*, *9*(8), e104297. https://doi.org/10.1371/journal.pone.0104297

Eickbush, T. H., & Eickbush, D. G. (2007). Finely orchestrated movements: Evolution of the ribosomal RNA genes. *Genetics*, *175*(2), 477–485. https://doi.org/10.1534/genetics.107.071399

Elser, J. J., Sterner, R. W., Gorokhova, E., Fagan, W. F., Markow, T. A., Cotner, J. B., … Weider, L. W. (2000). Biological stoichiometry from genes to ecosystems. *Ecology Letters*, *3*(6), 540–550. https://doi.org/10.1111/j.1461-0248.2000.00185.x

Fargette, M., Berthier, K., Richaud, M., Lollier, V., Franck, P., Hernandez, A., & Frutos, R. (2010). Crosses prior to parthenogenesis explain the current genetic diversity of tropical plant-parasitic *Meloidogyne* species (Nematoda: Tylenchida). *Infection, Genetics and Evolution*, *10*(6), 806–813. https://doi.org/10.1016/j.meegid.2009.04.013

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, *17*(6), 368–376. https://doi.org/10.1007/BF01734359

Floyd, R., Abebe, E., Papert, A., & Blaxter, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology*, *11*(4), 839–850. https://doi.org/10.1046/j.1365-294X.2002.01485.x

Fonderie, P., de Vries, C., Verryken, K., Ducatelle, R., Moens, T., van Loon, G., & Bert, W. (2013). Maxillary granulomatous inflammation caused by *Halicephalobus gingivalis* (Nematoda) in a *Connemara Mare* in Belgium. *Journal of Equine Veterinary Science*, *33*(3), 186–190. https://doi.org/10.1016/j.jevs.2012.05.066

Fu, Y.-X., & Li, W.-H. (1993). Statistical tests of neutrality of mutations. *Genetics*, *133*(3), 693–709.

Gunderson, J., Sogin, M., Wollett, G., Hollingdale, M., de la Cruz, V., Waters, A., & McCutchan, T. (1987). Structurally distinct, stage-specific ribosomes occur in *Plasmodium*. *Science*, *238*(4829), 933–937. https://doi.org/10.1126/science.3672135

Heise, M., Epe, C., & Schnieder, T. (1999). Differences in the second internal transcribed spacer (ITS-2) of eight species of gastrointestinal nematodes of ruminants. *The Journal of Parasitology*, *85*(3), 431–435. https://doi.org/10.2307/3285774

Holterman, M., van der Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., … Helder, J. (2006). Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution*, *23*(9), 1792–1800. https://doi.org/10.1093/molbev/msl044

Hugall, A., Stanton, J., & Moritz, C. (1999). Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic *Meloidogyne*. *Molecular Biology and Evolution*, *16*(2), 157–164. https://doi.org/10.1093/oxfordjournals.molbev.a026098

Ide, S., Miyazaki, T., Maki, H., & Kobayashi, T. (2010). Abundance of ribosomal RNA gene copies maintains genome integrity. *Science*, *327*(5966), 693–696. https://doi.org/10.1126/science.1179044

Ironside, J. E. (2013). Diversity and recombination of dispersed ribosomal DNA and protein coding genes in microsporidia. *PLoS ONE*, *8*(2), e55878. https://doi.org/10.1371/journal.pone.0055878

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kobayashi, T. (2008). A new role of the rDNA and nucleolus in the nucleus—rDNA instability maintains genome integrity. *BioEssays*, *30*(3), 267–272. https://doi.org/10.1002/bies.20723

Kobayashi, T., & Ganley, A. R. D. (2005). Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science*, *309*(5740), 1581–1584. https://doi.org/10.1126/science.1116102

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*(7), 1870–1874. https://doi.org/10.1093/molbev/msw054

Li, J., Wirtz, R. A., McConkey, G. A., Sattabongkot, J., & McCutchan, T. F. (1994). Transition of *Plasmodium vivax* ribosome types corresponds to sporozoite differentiation in the mosquito. *Molecular and Biochemical Parasitology*, *65*(2), 283–289. https://doi.org/10.1016/0166-6851(94)90079-5

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11), 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Lunt, D. H. (2008). Genetic tests of ancient asexuality in root knot nematodes reveal recent hybrid origins. *BMC Evolutionary Biology*, *8*, 194. https://doi.org/10.1186/1471-2148-8-194

Mashkova, T. D., Serenkova, T. I., Mazo, A. M., Avdonina, T. A., Timofeyeva, M. Y., & Kisselev, L. L. (1981). The primary structure of oocyte and somatic 5S rRNAs from the loach Misgurnus fossilis. *Nucleic Acids Research*, *9*(9), 2141–2152. https://doi.org/10.1093/nar/9.9.2141

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). *Creating the CIPRES Science Gateway for inference of large phylogenetic trees*. Paper presented at IEEE 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA.

Muir, G., Fleming, C. C., & Schlötterer, C. (2001). Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. *Molecular Biology and Evolution*, *18*(2), 112–119. https://doi.org/10.1093/oxfordjournals.molbev.a003785

Múnera Uribe, G. E., Bert, W., Vierstraete, A. R., de la Pena, E., Moens, M., & Decraemer, W. (2010). Burrowing nematodes from Colombia and their relationship with *Radopholus similis* populations, *R. arabocoffeae* and *R. duriophilus*. *Nematology*, *12*(4), 619–629. https://doi.org/10.1163/138855410X12628646276005

Nadler, S. A., Carreno, R. A., Adams, B. J., Kinde, H., Baldwin, J. G., & Mundo-Ocampo, M. (2003). Molecular phylogenetics and diagnosis of soil and clinical isolates of *Halicephalobus gingivalis* (Nematoda: Cephalobina: Panagrolaimoidea), an opportunistic pathogen of horses. *International Journal for Parasitology*, *33*(10), 1115–1125. https://doi.org/10.1016/S0020-7519(03)00134-6

Nadler, S. A., Hoberg, E. P., Hudspeth, D. S. S., & Rickard, L. G. (2000). Relationships of *Nematodirus* species and *Nematodirus battus* isolates (Nematoda: Trichostrongyloidea) based on nuclear ribosomal DNA sequences. *The Journal of Parasitology*, *86*(3), 588–601. https://doi.org/10.2307/3284877

Nyaku, S. T., Sripathi, V. R., Kantety, R. V., Gu, Y. Q., Lawrence, K., & Sharma, G. C. (2013). Characterization of the two intra-individual sequence variants in the 18S rRNA gene in the plant parasitic nematode. *Rotylenchulus reniformis*. *PLoS ONE*, *8*(4), e60891.

O'Mahony, E. M., Tay, W. T., & Paxton, R. J. (2007). Multiple rRNA variants in a single spore of the microsporidian *Nosema bombi*. *Eukaryotic Microbiology*, *54*(1), 103–109. https://doi.org/10.1111/j.1550-7408.2006.00232.x

Pei, A. Y., Oberdorf, W. E., Nossa, C. W., Agarwal, A., Chokshi, P., Gerz, E. A., … Pei, Z. (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Applied and Environmental Microbiology*, *76*(12), 3886–3897. https://doi.org/10.1128/AEM.02953-09

Pereira, T. J., & Baldwin, J. G. (2016). Contrasting evolutionary patterns of 28S and ITS rRNA genes reveal high intragenomic variation in *Cephalenchus* (Nematoda): Implications for species delimitation. *Molecular Phylogenetics Evolution*, *98*, 244–260. https://doi.org/10.1016/j.ympev.2016.02.016

Peterson, R. C., Doering, J. L., & Brown, D. D. (1980). Characterization of two *Xenopus* somatic 5S DNAs and one minor oocyte-specific 5S DNA. *Cell*, *20*(1), 131–141. https://doi.org/10.1016/0092-8674(80)90241-X

Pillet, L., Fontaine, D., & Pawlowski, J. (2012). Intra-genomic ribosomal RNA polymorphism and morphological variation in *Elphidium macellum* suggests inter-specific hybridization in Foraminifera. *PLoS ONE*, *7*(2), e32373. https://doi.org/10.1371/journal.pone.0032373

Porazinska, D. L., Giblin-Davis, R. M., Esquivel, A., Powers, T. O., Sung, W. A. Y., & Thomas, W. K. (2010). Ecometagenetics confirm high tropical rainforest nematode diversity. *Molecular Ecology*, *19*(24), 5521–5530. https://doi.org/10.1111/j.1365-294X.2010.04891.x

Porazinska, D. L., Giblin-Davis, R. M., Powers, T. O., & Thomas, W. K. (2012). Nematode spatial and ecological patterns from tropical and temperate rainforests. *PLoS ONE*, *7*(9), e44641. https://doi.org/10.1371/journal.pone.0044641

Prokopowich, C. D., Gregory, T. R., & Crease, T. J. (2003). The correlation between rDNA copy number and genome size in eukaryotes. *Genome*, *46*(1), 48–50. https://doi.org/10.1139/g02-103

Pruesse, E., Peplies, J., & Glöckner, F. O. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, *28*(14), 1823–1829. https://doi.org/10.1093/bioinformatics/bts252

Qing, X., Decraemer, W., Claeys, M., & Bert, W. (2017a). Molecular phylogeny of *Malenchus* and *Filenchus* (Nematoda: Tylenchidae). *Zoologica Scripta*, *46*(5), 625–636. https://doi.org/10.1111/zsc.12236

Qing, X., Slos, D., Claeys, M., & Bert, W. (2017b). Ultrastructural, phylogenetic and rRNA secondary structural analysis of a new nematode (Nematoda: Tylenchomorpha) from mushroom with recovery of intestinal crystals. *Zoologischer Anzeiger*, *269*, 13–25. https://doi.org/10.1016/j.jcz.2017.06.002

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., … Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Rogers, M. J., Gutell, R. R., Damberger, S. H., Li, J., McConkey, G. A., Waters, A. P., & McCutchan, T. F. (1996). Structural features of the large subunit rRNA expressed in *Plasmodium falciparum* sporozoites that distinguish it from the asexually expressed subunit rRNA. *RNA*, *2*(2), 134–145.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., … Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, *61*(3), 539–542. https://doi.org/10.1093/sysbio/sys029

Saka, K., Ide, S., Ganley, A. R. D., & Kobayashi, T. (2013). Cellular senescence in yeast is regulated by rDNA noncoding transcription. *Current Biology*, *23*(18), 1794–1798. https://doi.org/10.1016/j.cub.2013.07.048

Sinclair, D. A., & Guarente, L. (1997). Extrachromosomal rDNA circles—a cause of aging in yeast. *Cell*, *91*(7), 1033–1042. https://doi.org/10.1016/S0092-8674(00)80493-6

Smit, S., Widmann, J., & Knight, R. (2007). Evolutionary rates vary among rRNA structural elements. *Nucleic Acids Research*, *35*(10), 3339–3354. https://doi.org/10.1093/nar/gkm101

Smith, C., Heyne, S., Richter, A. S., Will, S., & Backofen, R. (2010). Freiburg RNA tools: A web server integrating INTARNA, EXPARNA and LocARNA. *Nucleic Acids Research*, *38*(Suppl 2), W373–W377. https://doi.org/10.1093/nar/gkp316

Sorhannus, U., & Bell, C. V. (1999). Testing for equality of molecular evolutionary rates: A comparison between a relative-rate test and a likelihood ratio test. *Molecular Biology and Evolution*, *16*(6), 849. https://doi.org/10.1093/oxfordjournals.molbev.a026169

Stamatakis, A., Hoover, P., & Rougemont, J. (2008). A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, *57*(5), 758–771. https://doi.org/10.1080/10635150802429642

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D. M., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology*, *19*, 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x

Subbotin, S., Perry, R., Warry, A., & Halford, P. (2000). Variations in ribosomal DNA sequences and phylogeny of *Globodera* parasitising solanaceous plants. *Nematology*, *2*(6), 591–604. https://doi.org/10.1163/156854100509484

Subbotin, S. A., Sturhan, D., Vovlas, N., Castillo, P., Tambe, J. T., Moens, M., & Baldwin, J. G. (2007). Application of the secondary structure model of rRNA for phylogeny: D2–D3 expansion segments of the LSU gene of plant-parasitic nematodes from the family Hoplolaimidae Filipjev, 1934. *Molecular Phylogenetics and Evolution*, *43*(3), 881–890. https://doi.org/10.1016/j.ympev.2006.09.019

Szalanski, A. L., Sui, D. D., Harris, T. S., & Powers, T. O. (1997). Identification of cyst nematodes of agronomic and regulatory concern with PCR-RFLP of ITS1. *The Journal of Nematology*, *29*(3), 255–267.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*(3), 585–595.

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., … Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, *453*, 534–538. https://doi.org/10.1038/nature06904

Thiéry, O., Vasar, M., Jairus, T., Davison, J., Roux, C., Kivistik, P. A., … Öpik, M. (2016). Sequence variation in nuclear ribosomal small subunit, internal transcribed spacer and large subunit regions of *Rhizophagus irregularis* and *Gigaspora margarita* is high and isolate-dependent. *Molecular Ecology*, *25*(12), 2816–2832. https://doi.org/10.1111/mec.13655

Thornhill, D. J., Lajeunesse, T. C., & Santos, S. R. (2007). Measuring rDNA diversity in eukaryotic microbial systems: How intragenomic variation, pseudogenes, and PCR artifacts confound biodiversity estimates. *Molecular Ecology*, *16*(24), 5326–5340. https://doi.org/10.1111/j.1365-294X.2007.03576.x

Triantaphyllou, A. C., & Hirschmann, H. (1997). Evidence of direct polyploidization in the mitotic parthenogenetic *Meloidogyne microcephala*, through doubling of its somatic chromosome number. *Fundamental and Applied Nematology*, *20*(4), 385–391.

Van Den Berg, E., Palomares-Rius, J. E., Vovlas, N., Tiedt, L. R., Castillo, P., & Subbotin, S. A. (2016). Morphological and molecular characterisation of one new and several known species of the reniform nematode, *Rotylenchulus* Linford & Oliveira, 1940 (Hoplolaimidae: Rotylenchulinae), and a phylogeny of the genus. *Nematology*, *18*(1), 67–107. https://doi.org/10.1163/15685411-00002945

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*(16), 5261–5267. https://doi.org/10.1128/AEM.00062-07

Washburne, A. D., Silverman, J. D., Morton, J. T., Becker, D. J., Crowley, D., Mukherjee, S., … Plowright, R. K. (2019). Phylofactorization: A graph partitioning algorithm to identify phylogenetic scales of ecological data. *Ecological Monographs*, *89*(2), e01353. https://doi.org/10.1002/ecm.1353

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(12), 4576–4579. https://doi.org/10.1073/pnas.87.12.4576

Wuyts, J., Van de Peer, Y., & De Wachter, R. (2001). Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA. *Nucleic Acids Research*, *29*(24), 5017–5028. https://doi.org/10.1093/nar/29.24.5017

Xu, J., Xu, Y., Yonezawa, T., Li, L., Hasegawa, M., Lu, F., … Zhang, W. (2015). Polymorphism and evolution of ribosomal DNA in tea (*Camellia sinensis*, Theaceae). *Molecular Phylogenetics and Evolution*, *89*, 63–72. https://doi.org/10.1016/j.ympev.2015.03.020

Ye, W., Giblin-Davis, R. M., Braasch, H., Morris, K., & Thomas, W. K. (2007). Phylogenetic relationships among *Bursaphelenchus* species (Nematoda: Parasitaphelenchidae) inferred from nuclear ribosomal and mitochondrial DNA sequence data. *Molecular Phylogenetics and Evolution*, *43*(3), 1185–1197. https://doi.org/10.1016/j.ympev.2007.02.006

Yoder, M., Wm King, I., De Ley, I. T., Mann, J., Mundo-Ocampo, M., Poiras, L., … Blaxter, M. (2006). DESS: A versatile solution for preserving morphology and extractable DNA of nematodes. *Nematology*, *8*(3), 367–376. https://doi.org/10.1163/156854106778493448

Yoshiga, T., Kuwata, R., Takai, H., & Nishimura, K. (2014). Heterogeneity of the large subunit of ribosomal RNA gene sequences in a *Halicephalobus gingivalis* isolate. *Nematology*, *16*(10), 1233–1236. https://doi.org/10.1163/15685411-00002841

Zhan, Y., Matafeo, A., Shi, H.-L., & Zheng, J.-W. (2011). Morphological and molecular characterization and host range of *Rotylenchulus reniformis* population occurring in Hangzhou, Zhejiang, China. *Acta Phytopathologica Sinica*, *41*(1), 37–43.

Zijlstra, C., Lever, A. E. M., Uenk, B. J., & Van Silfhout, C. H. (1995). Differences between ITS regions of isolates of root-knot nematodes *Meloidogyne hapla* and *M. chitwoodi*. *Phytopathology*, *85*(10), 1231–1237.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Qing X, Bik H, Yergaliyev TM, et al. Widespread prevalence but contrasting patterns of intragenomic rRNA polymorphisms in nematodes: Implications for phylogeny, species delimitation and life history inference. *Mol Ecol Resour*. 2020;20:318–332. https://doi.org/10.1111/1755-0998.13118