

# Non-homogenized ITS regions in the parasitic nematode *Cooperia oncophora*

T. H. M. MES\* and A. W. C. A. CORNELISSEN

Department of Parasitology and Tropical Veterinary Medicine, Institute of Infectious Diseases and Immunology, Utrecht University, PO Box 80165, 3508 TD Utrecht, The Netherlands

(Received 11 July 2003; revised 12 December 2003 and 21 January 2003; accepted 21 January 2004)

## SUMMARY

Here, the validity of the assumption of concerted evolution of ribosomal regions in larval and adult *Cooperia oncophora* was assessed. In each of 4 individuals of this parasitic nematode, at least 78% of the sequences comprised different ITS variants†. This implies that concerted evolution is not acting, which is corroborated by the scarcity of signatures of gene conversion and recombination. Mis-incorporation of nucleotides and illegitimate PCR-induced recombination turned out to be unlikely, and positions with substantial frequencies of alternative nucleotides corresponded to ambiguous positions in published ITS2 sequences of this and other *Cooperia* species based on direct sequencing. The ITS regions of each individual *C. oncophora* displayed a significant excess of unique mutations in agreement with expansion of the ribosomal gene family. Interesting corollaries of the inferred size changes of this gene family are genomic rearrangements that occur during larval development such as multiple rounds of endoduplication (in Rhabditidae), chromatin diminution (in *Ascaris*), and non-compensatory mutations on the secondary structure of the ITS2. It is yet unknown which process is important in trichostrongylids. Finally, although it can not be rigorously assessed in *Cooperia*, the ITS polymorphisms can readily be envisioned to affect phylogenetic reconstructions of closely related nematodes.

† Nucleotide sequence data are available in the DDBJ/EMBL/GenBank databases under the Accession numbers AJ544390–AJ544465.

Key words: *Cooperia oncophora*, ribosomal, evolution, recombination, gene conversion.

## INTRODUCTION

Some multigene families, notably the nuclear ribosomal family, are renowned for processes that reduce intra-genomic variability. The main reason for the phylogenetic utility of the ribosomal regions is the high copy number and the variation in evolutionary rates of different portions of the ribosomal regions. As such, a spectrum of different evolutionary questions can be addressed. The ribosomal regions typically undergo rapid concerted evolution, which entails the homogenization of ribosomal variants within individuals through unequal crossing-over. In *Onchocerca volvulus* (Higazi *et al.* 2001), concerted evolution is predominant and may lead to an extremely rapid and complete homogenization of ribosomal arrays. Signatures of gene conversion through unequal crossing-over typically are approximately 100 bp long and, as a result, gene conversion can be identified by the presence of nucleotide stretches that result in different phylogenetic estimates. In contrast, signatures that result from recombination between ribosomal variants leave traces that defy any phylogenetic hypothesis. Gene conversion and recombination rely on the same molecular mechanisms, and may interact through a rapid expansion of

variants through the genome, sometimes with a replacement of other ITS variants (Hillis *et al.* 1991). In theory, incomplete homogenization of ITS copies may affect phylogenetic analysis, because identification of orthologues may be compromised. If species have sufficiently diverged or, alternatively, share sufficient numbers of phylogenetically informative mutations, the presence of polymorphisms in partially homogenized ITS gene families may still allow identification of the correct species phylogeny. However, the high level of ITS diversity within individuals (e.g., Stevenson, Chilton & Gasser, 1996; Powers *et al.* 1997; Szalanski *et al.* 1997; Gasser *et al.* 1998; Newton *et al.* 1998; Hugall, Stanton & Moritz, 1999; Nadler *et al.* 2000; Chilton *et al.* 2001) and between individual plant-parasitic and animal-parasitic nematodes (e.g. Powers *et al.* 1997; Adams, Burnell & Powers, 1998; Beckenbach, Blaxter & Webster, 1999; Subbotin *et al.* 2000a; Subbotin, Waeyenberge & Moens, 2000b; Elbadri *et al.* 2002) and the close relationships between at least some nematode species indicate that it is possible that phylogenetically informative mutations are overwhelmed by polymorphisms that occur within and between nematodes of a single species. If so, this means that in practice, the reconstruction of phylogenetic relationships among these species may be compromised if only a few representatives are used. In this study, aspects of the diversity of ribosomal ITS arrays of

\* Corresponding author. Tel: +31 30 2532559/1117. Fax: +31 30 2540784. E-mail: t.mes@vet.uu.nl

*Cooperia oncophora* have been examined, with an emphasis on the genomic processes that may be associated with the diversity among ITS copies and on the consequences of the ITS diversity on phylogenetic reconstruction among closely related nematodes.

#### MATERIALS AND METHODS

##### *Collection of sequence data*

Artificial infections with laboratory strains of *C. oncophora*, which have been maintained for many years (e.g., Kloosterman, Albers & van den Brink, 1978), are regularly performed in our laboratory to infer their impact on local immune responses and to assess aspects of the epidemiology of nematode infections (e.g., Kanobana *et al.* 2001). For this study, *C. oncophora* was collected after natural or artificial infections. The field strains of *C. oncophora* were collected by veterinarians on commercial farms as part of a study that aims to identify regions of the cattle genome that are involved in resistance to nematode infections. In addition, nematodes of a laboratory strain of *C. oncophora* were used. These nematodes shed their eggs *via* the host's faeces. On pasture, they develop into infective 3rd-stage larvae (L3), which may enter a new host. To allow development of eggs into L3 larvae, the faecal samples were mixed with sawdust and stored for 10 days at 27 °C (Roberts & O'Sullivan, 1950).

DNA was extracted either from single infective larvae (L3 stage), single adult worms or from a population of L3 larvae. After removal of all fluid on the outside of the nematode, they were subsequently washed in 200 µl of sterile water. DNA was extracted using the following protocol. The larvae were overlaid with 10 µl of 0.25 M NaOH, and stored for 3–14 h. After this incubation, the solution was kept at 92 °C for 5 min. After a short centrifugation to collect all fluid at the bottom of an Eppendorf tube, 9.5 µl of 2% Triton X-100, 0.5 M Tris-HCl pH 7.5, and 1 M HCl were added. Again, this solution was kept at 92 °C for 5 min (Floyd *et al.* 2002). After centrifugation, the solution was thoroughly mixed, and centrifuged. Then 1 µl was used for PCR using Supertaq DNA polymerase (ShaeroQ) in the recommended PCR buffer with 15 mM MgCl<sub>2</sub>. Twenty-five PCR cycles with 20 sec at 92 °C, 30 sec at 55 °C, and 45 sec at 72 °C were used to amplify the non-coding regions between 18S and 25S ribosomal genes using standard primers NC2 and NC5 (Gasser *et al.* 1993). Upon amplification, PCR products were extracted using chloroform, and subsequently ligated and cloned in pGEM-T vectors following the recommendations of the manufacturer (Promega Inc.). PCR amplification without template consistently lacked PCR products. Bidirectional sequence reactions were carried out on Li-Cor automated sequencers (Westburg Genomics) using fluorescent

SP6 and T7 primers. For cloning, DNA from an adult worm and a population of approximately 10 000 L3 larvae, both of which derive from a laboratory strain of *C. oncophora* was used. In addition, 2 larvae from commercial dairy farms in Zuidermeer and Zoeterwoude, The Netherlands (code 1474 and 2409, respectively) were used for cloning and sequencing of ITS regions.

When analysing individual larvae or adult nematodes, a correct taxonomic identification is critical. In addition, DNA of related nematode taxa should be absent from DNA preparations. This is most critical for field samples, because frequently these isolates are mixtures of other Trichostrongylid nematode genera of such as *Haemonchus*, *Trichostrongylus*, *Ostertagia*, and *Oesophagostomum*. We took the following precautions to eliminate contaminating DNA and to obtain a correct taxonomic identification. First, the population sample and the adult were taken from a well-characterized laboratory population of *C. oncophora*. Secondly, nematodes were individually transferred to a microscopic slide for taxonomic identification based on morphological characteristics. Subsequently, the nematodes were rinsed with large amounts of water, before they were transferred to Eppendorf tubes. In these tubes, all fluid was removed once more from the larvae and the adult nematode, and they were rinsed again with 200 µl of sterile water (see above). Because L3 larvae and adults differ in DNA content, the two life-stages serve as a check for the impact of contaminating DNA. Because of the larger amount of DNA in adult nematodes, DNA extracted from adult worms is probably less impacted by the presence of contaminating DNA compared to other nematode species that are present in field isolates (see also Discussion section). Taxonomic identification of the larvae was based on morphological features such as conspicuous oval bodies or a bright band between buccal cavity and the oesophagus, the shape of the head, and the length of the sheath of the tail (MAFF, 1977). To confirm the taxonomic identity, the genetic diversity of the internal transcribed intergenic spacer regions of the nuclear ribosomal regions was evaluated against ITS diversity in a well-characterized lab-strain of *C. oncophora* (this study) and against published ITS2 sequences. Finally, the mitochondrial ND4 gene, which can be used to differentiate many species of the trichostrongylids (Blouin *et al.* 1998), was used for evaluation of the taxonomic identification. Combined with the stringent washing steps (see above), the various checks of the divergence and diversity of mitochondrial and ribosomal DNA allowed taxonomic classification of individual nematodes from field samples. A potential caveat of the taxonomic identification, based on molecular data, is that nematodes may have ingested eggs of other nematode species. This source of error has not been corrected for in this study.

Checks for PCR errors involved calculation of the expected number of mutant ITS variants that accumulate during PCR, under the assumption that each mutation results in a new ITS variant and assuming that the PCR started from a single template DNA molecule. For this calculation, published error rates of DNA polymerases were used. The calculation involved an exponential increase of the total number of non-mutant DNA copies during PCR, and an exponential increase of the number of mutant copies, which comprised all mutant sequences that had accumulated in previous PCR cycles. Another way to evaluate the impact of mutation and recombination (see below) during PCR is to consider the ITS copies resulting from PCR amplification as a population of unlinked gene variants. A thousand simulations of the diversity using a mutation rate proportional to the amount of diversity observed in the sample were conducted (Rozas & Rozas, 1999). Comparison of the observed and simulated levels of diversity and recombination enables determination of the significance of these processes under various levels of recombination. For the simulations, the same number and length of sequences were used as in the collected ITS data. The impact of *in vitro* mutation during PCR was further evaluated using PCR amplification of individual ITS clones, followed by direct sequencing. In this way, polymorphisms that are due to PCR errors are visible as multiple peaks on chromatograms. This allows us to establish whether double peaks in the ITS 2 region in *C. oncophora* (Newton *et al.* 1998) are genuine or are PCR mediated. Finally, the diversity in ITS sequences obtained from amplification using DNA polymerases with (Pwo; Roche Inc.) and without (Supertaq; SphaeroQ) proofreading was evaluated.

#### Statistical analyses

Descriptive statistics such as the number of ITS variants, the mean number of absolute pairwise differences, and the number of segregating sites (S) were calculated with PAUP\* version 10 (Swofford, 1998). Maximum parsimony cladograms were constructed using PAUP\*. The general setting included  $n=1000$ ,  $nchuck=25$ ,  $chuckscore=10$ ,  $nreps=5$ ,  $addseq=random$ ,  $steepest=yes$ . Sequence divergence among the closely related ITS sequences was based on the proportion of polymorphic sites (Swofford, 1998). Overall support for trees was assessed using unconstrained permutation tail probabilities based on 1000 permutations, and confidence statistics such as the consistency indices and the decay index. The Partition-Homogeneity Test with 100 replicates was used to determine whether the ITS1 and ITS2 gave different phylogenetic estimates. To assess whether the rate heterogeneity and complex substitution models were implicated in the evolution of the ITS regions, Modeltest version 3.06

(Posada & Crandall, 1998) was used. The importance of processes such as gene conversion and recombination was assessed independently. Stretches of gene conversion were detected using GENECONV based on 10000 permutations and Bonferroni-corrected scores (Sawyer, 1989, 1999). Signatures of recombination (i.e. linkage disequilibrium) and gene conversion were detected using DNASP version 3.51 (Rozas & Rozas, 1999), and they were used to examine whether the diversity in the ITS variants could be faithfully represented by a phylogenetic tree. The frequency distribution of sites in a population sample of sequences is expected to follow Tajima's distribution under stationarity (i.e. under no population size changes and no selection; Tajima, 1989). Deviation was assessed through Tajima's D statistic, which contrasts the diversity in a sample based on the number of polymorphic sites and the diversity based on the average pairwise differences between sequences. If many mutations are unique, these will contribute more to estimates based on polymorphic sites, and they will have less impact on the average pairwise difference in the sample. Under these conditions, Tajima's D will be negative. Significantly negative values of Tajima's D are typically interpreted as resulting from population size expansions or a deviation from neutrality (Mes, 2003).

#### RESULTS

The individual larvae used for amplification of the ribosomal regions had all the characteristics of *C. oncophora*, i.e. they had a clearly distinguishable head and neck, they were approximately 100  $\mu\text{m}$  long, with a long tail relative to other parasitic nematode species. In addition, the individual nematodes had conspicuous oval bodies. ND4 sequences of the same nematodes as used for the ITS regions showed 96% identity to the mitochondrial genome of *C. oncophora* in Genbank (AY265417), and much less to other trichostrongylid species (e.g. *Mazamastrongylus ocoidei* 86% identity). The former level of diversity is well within the range of divergence observed within other nematode species of livestock (Blouin *et al.* 1998). However, the lack of comparative ND4 data of other *Cooperia* species makes it difficult to exclude the possibility that the sequences derive from a closely related *Cooperia* species. ITS sequences of the population sample, the adult nematode, and the 2 individual L3 larvae from the field were very similar, and they match published ITS sequences. For example, sequence identity between the ITS sequences of this study and ITS sequences of *Cooperia* species (Accession numbers 2261370, 5019425, 5019424) deposited in Genbank ranged from 97 to 99%. Nevertheless, these features and the geographical origin of the samples (*C. punctata*, a species which can be confused with *C. oncophora*, is common in warm temperate and subtropical areas) support the

Table 1. Molecular descriptives of the ITS region of *Cooperia oncophora*

(S: number of segregating sites. Unique: number of unique sites. k: average number of pairwise differences. Ti/Tv: transition–transversion ratio.)

Sample	No. sequences	No. variants	S	Unique	k	Ti/Tv	Rate heterogeneity	Nucleotide frequencies			
								A	C	G	T
Field											
1474 (single L3 larvae)	18	16	41	30	7.20	1.86	No	0.30	0.17	0.20	0.33
2409 (single L3 larvae)	23	18	44	39	4.97	6.20	No	0.30	0.17	0.20	0.33
Laboratory											
Population of L3 larvae	24	20	53	34	6.73	15.76	No	0.30	0.16	0.20	0.34
Single adult <i>C. oncophora</i>	23	22	52	48	5.83	2.91	No	0.30	0.17	0.20	0.33

Table 2. Phylogenetic descriptives of the ITS variants of *Cooperia oncophora*

(MPTs: Most Parsimonious Trees with a length indicated in steps. CI: consistency index excluding uninformative characters. PTP: permutation-tail probabilities based on 100 permutations. An asterisk indicates a consistency index excluding uninformative positions.)

Sample	No. MPTs	Length	CI*	Length of MPTs of random trees	PTP
Field					
1424 (single L3 larvae)	11	46	0.85	54	0.001
2409 (single L3 larvae)	67	51	0.75	62	0.001
Laboratory					
Population of L3 larvae	23	61	0.86	74	0.001
Single adult <i>C. oncophora</i>	1	56	1.00	61	0.001
All	676	201	0.72	392	0.001

classification of these nematodes as *C. oncophora* based on the morphological features of the larval and adult stages.

The number of sequences, segregating sites, and haplotypes, and the average divergence among ITS copies is very similar within individuals and between individuals (Table 1). Although the transition – transversion ratio varies somewhat across samples, the composition and mode of evolution of these ITS sequences were also similar. The nucleotide substitution models for each data set had equal rates among sites (i.e. no significant rate heterogeneity; Table 1). The cladistic analyses resulted in most parsimonious trees (MPTs) that had a high consistency index and a strong hierarchical structure as judged from random trees (Table 2). These random trees were much longer than the MPTs of the actual data sets. Likewise, permutation-tail probabilities based on 1000 replicates suggested a highly significant hierarchical structure in the ITS data sets (Table 1). Partition-Homogeneity tests indicated that there was no significant difference between the ITS1 and the ITS2 data sets (not shown;  $P=0.001$ ).

Signatures of the action of concerted evolution in these ITS sequences were scarce. Only a single

nucleotide stretch with an aberrant sequence composition that may be typical for gene conversion was detectable using GENECONV (Sawyer, 1999). This stretch comprised nucleotide position 5 through 17 in sequence 2409-24 (Fig. 1). Because this nucleotide stretch is located on the borders of the ITS region, it is simply a region with a lower than average similarity without strongly supported recombination breakpoints. This region has only marginal effects on the relationships among ITS variants (Fig. 1), because only a single informative base substitution occurs in this region. These findings suggest that gene conversion is not a dominant process in the ITS sequences of *C. oncophora*.

Given the large number of ITS variants, it is appropriate to consider the potential impact of polymerase errors during PCR. There are several indications that the different ITS copies found after PCR do not result from artifacts. First, although an initially high error rate may be propagated in the course of the PCR, the error rate of DNA polymerases is especially important for analyses of single cells or single DNA molecules, and not for these multicellular nematodes that have several hundreds of cells and nuclei. In addition, the number of mutations is far too high to be caused by PCR errors.



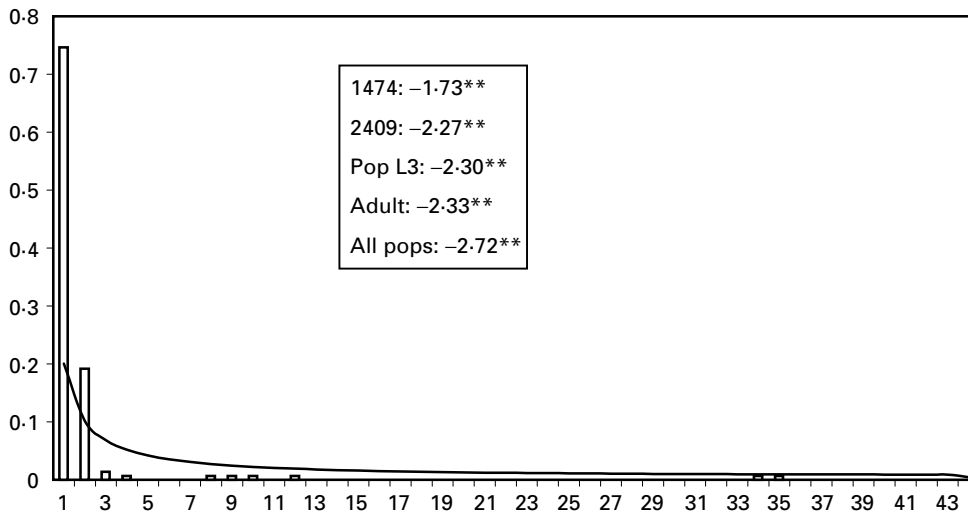


Fig. 2. Distribution of the numbers of segregating sites in 88 ITS sequences of *Cooperia oncophora* (bars). The significance of deviation from the neutral expectation (line; Tajima, 1989) is tested by coalescent simulations based on Tajima's D under no recombination. Values and significance of Tajima's D are displayed for all ITS data sets as well as for all data sets together. \*\*  $P < 0.01$ .

the probability that  $ZZ < -0.0491$  was only 0.0070. Higher levels of recombination resulted in higher ZZ values with still lower probabilities. In conclusion, recombination can be neglected as a cause for the observed ITS diversity, because even a single recombination event would readily lead to higher estimates of ZZ than observed.

The ITS sequences amplified using Pwo also displayed several base substitutions, and the most polymorphic position in the ITS sequences amplified using Taq polymerase was also polymorphic using Pwo (pos. 117; cf. Newton *et al.* 1998). Coalescent simulations also indicated that the number of ITS variants are not likely caused by errors of the DNA polymerases Supertaq and Pwo, because these enzymes with and without proofreading gave similar levels of diversity. The level of diversity as determined by the pairwise number of differences among ITS sequences was not significantly different with and without proofreading ( $P = 0.90$ ). Likewise, the number of haplotypes was also not significantly different from the initial survey (5 of 6 *vs.* 70 of 88;  $P = 0.30$ ). Three Pwo amplified sequences were identical to the group of sequences comprising 25R (Box in Fig. 1), whereas the three others represented novel variants (not shown). Interestingly, 6 of 8 polymorphic sites were also polymorphic in the large data set of Fig. 1. One of these, the highly polymorphic site at position 117 reported by Newton *et al.* (1998) and in the large sample of ITS sequences (Fig. 1) was also highly polymorphic in this small sample (3 clones had A and 3 clones had G). PCR amplification of 3 ITS clones followed by direct sequencing gave no peaks that comprised multiple nucleotides and, as such, there is no evidence that polymorphic sites observed using direct sequencing resulted from polymerase errors. Also

the polymorphic sites in Newton *et al.* (1998) were monomorphic in these sequences.

Next, the impact of the ITS polymorphisms on phylogenetic reconstruction of a few species of *Cooperia* is considered. In spite of the large number of different ITS variants, the level of variability is low (average distance across all pairwise comparisons is 0.010933 based on the Jukes-Cantor correction). Alignment of our and Newton *et al.*'s (1998) sequences under a conservative mode of scoring (ambiguous sites of Newton *et al.* (1998) were excluded) indicated that the *C. oncophora* and *C. surnabada* sequence of Newton *et al.* (1998) is identical to the group of sequences including 2R (Box in Fig. 1). A single mutation distinguishes *C. punctata* and *C. cuticei* from the other sequences of the *Cooperia* species (Newton *et al.* 1998). Examination of the stretch of bases comprising this mutation showed that it is unique to species of *Cooperia*. No other ITS sequence of nematode species shows significant similarity in this stretch of the ITS2. As such, it can not be established whether or not this site has undergone additional mutations in a wider sample of trichostrongylids and whether *C. oncophora*/*C. surnabada* or *C. punctata*/*C. cuticei* comprises the root of the *Cooperia* phylogeny. Clearly, these data do not allow an assessment of the impact of intra-individual and inter-individual polymorphisms on phylogenetic reconstruction in *Cooperia*.

The ITS sequences have many unique and rare sites (Table 1) and the number of these categories of sites are significantly different from the number expected under stationarity (Tajima, 1989). This holds true for each of the 4 individual data sets as well as for the 4 data sets together (Fig. 2). The negative estimates of Tajima's D indicates that the 18S-26S ribosomal gene family has expanded.

## DISCUSSION

*Possibility of PCR-induced errors*

These analyses agree with the literature on the fidelity of DNA polymerases in which the usual error rates are considered to be 'insufficient to generate a diverse library of variant sequences, especially over a region shorter than 1000 nucleotides' (e.g. Cadwell & Joyce, 1994). Also manufacturers consider it unlikely that – largely independent of the error rate of the polymerase – many unique sequence variants result solely from polymerase errors when short amplicons up to 1 kb are used (e.g. Stratagene Co.). There are other indications that mutation is not the prime factor. If polymerase-induced mutations are predominant, their distribution is *a priori* expected to be largely uniform across the amplicon. However, the patterns of diversity that resulted from a sliding window approach on the ITS data (not shown) indicate lower levels of polymorphism in the 5.8S ribosomal gene, which is located between the ITS regions. Further, the observed diversity in the ITS sequences is similar between enzymes with different error rates, under conditions recommended by the manufacturer (standard reaction buffer with 2 mM MgSO<sub>4</sub>) and comprise positions that had already been identified as ambiguous using direct sequencing (Newton *et al.* 1998). The MgCl<sub>2</sub> and MgSO<sub>4</sub> concentrations are generally considered to be the most critical for fidelity of DNA polymerases (Liang, Chen & Fulco, 1995) and these concentrations are commonly used for normal PCR. A strong reduction of the polymorphism as implicated by the 10-fold lower error rate of Pwo (Roche Inc.) that is expected if PCR-induced errors were in play, is evidently not at hand. In spite of the fact that sample sizes were rather different for the two polymerases, these results indicate that both polymerases detect multiple and similarly diverged ITS regions in these parasitic nematodes. Another possibility, that the ITS diversity is PCR-induced via *in situ* recombination between ITS copies (e.g., Cronn *et al.* 2002) can also be discarded based on the simulations.

Similar evidence pointing to real ITS polymorphisms in *Cooperia* comes from a study by Newton *et al.* (1998), who used direct sequencing of ITS2 regions of 4 *Cooperia* species, i.e. *C. oncophora*, *C. surnabada*, *C. punctata*, and *C. curticei*. The ITS sequences of these species exhibited 10 ambiguous positions out of 21 variable (polymorphic plus ambiguous) positions. A comparison of the ambiguous positions of Newton *et al.* (1998) and this study reveals that 5 of the 10 ambiguous positions show multiple nucleotides in our sample. Furthermore, the ambiguity codes in Newton *et al.* (1998) agree with the alternative polymorphic nucleotides found in this study. Interestingly, in our study a few sites have alternative nucleotides at considerable frequencies, so that they can be identified by direct

sequencing protocols as employed by Newton *et al.* (1998). For example, positions 89, 96, and 117 of Newton *et al.* (1998) have minority nucleotide frequencies of 9, 11 and 39%, respectively. These polymorphic positions nearly completely explain Newton's data, because only a single ambiguous site in the 3 closely related *Cooperia* species (*C. oncophora*, *C. punctata*, and *C. surnabada*; *C. curticei* is more distantly related) was not polymorphic in our survey. One adult of each of these 3 *Cooperia* species was used for DNA isolation (Newton *et al.* 1998). This again indicates that the observed ITS diversity is real. It also suggests that cloned ITS fragments greatly increase the power to detect different ribosomal variants.

*Cause of ITS heterogeneity*

The presence of highly polymorphic ITS variants in *C. oncophora* is well established and is not surprising given that polymorphic ITS variants were reported for other nematodes (see Introduction section). Although evaluation of the causes of this heterogeneity is complicated by a lack of information on the development of parasitic trichostrongylids, in plants and higher animals clear evidence has been obtained that polymorphic ITS variants are associated with polyploidy (e.g. Campbell *et al.* 1997). Although limited, the available evidence also suggests a similar pattern for nematodes of the Rhabditidae (Hugall *et al.* 1999). A number of observations suggest that dynamic and complex genome rearrangements may be associated with the ribosomal heterogeneity. Two kinds of processes may be outlined. The first are successive rounds of endoreduplication (Flemming *et al.* 2000) which results in polyploid nuclei in the hypodermal syncytium in *C. elegans*. At the late L4 stage most nuclei are tetraploid, and in adults the ploidy levels may have risen to decaploid levels in some nematode species. The degree of endoreduplication may also differ from cell to cell and among others occurs in hypodermal cells (Flemming *et al.* 2000). Although most of the developmental data are available for free-living nematodes (Triantaphyllou, 1991; Flemming *et al.* 2000), this process likely also occurs in parasitic nematodes, which virtually all have syncytial cells (Chitwood & Chitwood, 1972; Flemming *et al.* 2000). The duplications may involve entire genomes, entire chromosomes, or portions of chromosomes. Considering the link between ITS diversity and ploidy level in plant-parasitic nematodes (*Meloidogyne*; Hugall *et al.* 1999) and plants (see above), these genome dynamics in *C. oncophora* may also be associated with the high ITS diversity, and they may very well have impacted the diversity of ITS variants as observed in individual nematodes. In another potentially relevant process, chromatin diminution (Jentsch, Tobler & Muller, 2002), large amounts of predominantly

telomeric DNAs are deleted at specific times during development. Presumably, this process has a role in gene regulation (Muller & Tobler, 2000) and is – as far as known – not acting in free-living Rhabditidae. In *Ascaris lumbricoides*, chromatin diminution leads to elimination of the germline variant of a ribosomal protein (S19), while its divergent somatic homologue is maintained leading to ribosomal heterogeneity (Etter *et al.* 1994). Again, in spite of the lack of developmental information, the role of chromatin diminution in parasitic trichostrongylids is unknown – the polyploid nature of nematode hypodermal nuclei and the large number of reports of ITS heterogeneity (see Introduction section) suggest that a link with nematode development is more than circumstantial. These results do not conflict with EST studies of parasitic nematodes such as *Haemonchus contortus* and *Ostertagia ostertagi*, which have not identified many ribosomal genes. This is because the ribosomal genes are not targeted by the library generation protocols (Makedonka Dautova, personal communication). Other organisms also have highly dynamic numbers of ribosomal rRNA genes during development. Such size changes of the ribosomal gene family are especially well documented for *Tetrahymena thermophila* (e.g. Ward *et al.* 1997), in which the ribosomal genes are excised and amplified from two to many thousands of copies. After amplification, a substantial fraction of the expanded ribosomal gene family is transposed, resulting in many smaller ribosomal gene families on different chromosomes. There is little evidence that the ribosomal loci of *C. oncophora* comprise a limited number of loci on different chromosomes, which, in principle, can be invoked to explain the relatively high level of heterogeneity. If so, groups of divergent ITS variants and, most importantly, positive values of Tajima's D (Tajima, 1989) would be expected. Both of these attributes are not observed.

An alternative explanation for the high ITS diversity is an increased mutation rate. In this respect, it is potentially important to distinguish nematode germ line cells and somatic cells, because during differentiation of *C. elegans*, somatic cells undergo extensive chromatin modulation (Shin & Mello, 2003) that does not occur in germ line cells. Interestingly, also transposon activity occurs only in somatic cell lineages, because germ line cells are silenced (Sijen & Plasterk, 2003). Presumably, germ line cells are protected against the potentially deleterious effects of such genome rearrangements (Shi & Blackwell, 2003). To be able to interpret the significance of the findings of elevated levels of ITS diversity reported here, it would be advantageous to analyse ITS diversity in germ line and somatic cells separately in order to address spatial and temporal heterogeneity of the ribosomal gene family. This might be possible using laser dissection (Bernsen *et al.* 1998). Alternatively,

eggs and sperm could be analysed for ITS diversity. These approaches would be especially fruitful when combined with techniques such as *in situ* hybridization using ribosomal probes to track the chromosomal position and approximate copy number of ribosomal genes.

If either gene family size or mutation rate has increased in *C. oncophora*, the ITS diversity may impact the efficiency of RNA transcript processing (e.g. ITS 2; Chilton *et al.* 1998). In *C. oncophora*, the frequency of polymorphic sites across ITS 1 (102 polymorphic of 422 sites), 5.8S gene (38 polymorphic of 152 sites), and ITS 2 (24 polymorphic of 116 sites) is very similar. Also the distribution of polymorphic sites is similar across the ITS 1, the 5.8S gene, and the ITS 2 is very similar, and conserved stretches longer than about 10 nucleotides are lacking. Finally, insertion and deletions are very short (1 to 2 base pairs) and these mutations were relatively rare (2, 0, and 3, in the ITS 1, the 5.8S gene, and the ITS 2, respectively). If many of the mutations are deleterious, they can be expected to affect the processing of rRNA transcripts. As such, they can be expected to affect RNA folding, which, in parasitic trichostrongylids, is important for processing. In *Trichostrongylus*, most mutations in the ITS 2 were in loops and bulges and many of the mutations in helices were completely or partially compensatory (94%; Chilton *et al.* 1998). The ITS 2 sequence divergence between *Trichostrongylus* species and *C. oncophora* is too large to simply evaluate the polymorphic positions of *C. oncophora* using the secondary structure of *Trichostrongylus* as a template. Similarity searches using Genbank accessions only detected ITS 2 sequences of *Cooperia* species. Folding of the ITS 2 sequence and portions of the 5.8S and 26S ribosomal rRNA gene of *C. oncophora* using the same algorithm as Chilton *et al.* (1998) gave a very similar secondary structure relative to *Trichostrongylus*. All 7 helices of *Trichostrongylus* were retrieved (not shown), although domain VII was much larger in *C. oncophora* than in *Trichostrongylus* and comprised more bulges (cf. Fig. 2 in Chilton *et al.* 1998). Interestingly, 24 of the 46 polymorphisms that occurred in helices were fully or partially compensatory, whereas incompatible base pairings occurred at 17 positions (complex mutations at base pairings were not considered). Thus, a much smaller fraction of the mutations in helices (52%) were completely or partially compensatory in *C. oncophora* relative to *Trichostrongylus* (Chilton *et al.* 1998). Unlike the ITS 2 sequence of *Trichostrongylus*, only a small fraction of the mutations were in loops and bulges (19 of 68). Whether these findings imply a higher deleterious mutation rate or different dynamics of the ribosomal gene families in the latter species is currently unknown.

Future research may be directed at strengthening the link between the genomic consequences of



developmental processes and the diversity of multi-gene families. For example, determination of the diversity in individual cells at different stages of development may improve our understanding of this link. In *C. oncophora*, the results suggest that the ribosomal gene family increased in copy number in individual nematodes. The expansion is also detectable in populations of nematodes. Considering the average tract length resulting from gene conversion (e.g. *Drosophila*; Betrán *et al.* 1997), signatures of the gene conversion would be readily detectable in the ITS sequences of *C. oncophora*. The lack of these signatures suggests a (nearly) complete absence of gene conversion. Although the observed ITS heterogeneity may well affect phylogenetic reconstructions of closely related nematode species, this could not be rigorously documented due to a lack of comparative ITS sequence data of *Cooperia*. Many studies of the phylogenetic relationships of closely related nematodes exploit levels of ITS variability that are comparable to the variability within *C. oncophora* (e.g., Stevenson *et al.* 1995; Stevenson, Gasser & Chilton, 1996; Adams *et al.* 1998; Nadler *et al.* 2000). As such, surveying ITS heterogeneity of additional nematode species is expected to be helpful in assessing the validity of ITS regions for phylogenetics of closely related parasitic nematode species.

This research was supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.

## REFERENCES

- ADAMS, B. J., BURNELL, A. M. & POWERS, T. O. (1998). A phylogenetic analysis of *Heterorhabditis* (Nematoda: Rhabditidae) based on internal transcribed spacer 1 DNA sequence data. *Journal of Nematology* **30**, 22–39.
- BECKENBACH, K., BLAXTER, M. L. & WEBSTER, J. M. (1999). Phylogeny of *Bursaphelenchus* species derived from analysis of ribosomal internal transcribed spacer DNA sequences. *Nematology* **1**, 539–548.
- BERNSEN, M. R., DIJKMAN, H. B. P. M., DE VRIES, E., FIGDOR, C. G., RUITER, D. J., ADEMA, G. J. & VAN MUIJEN, G. N. P. (1998). Identification of multiple mRNA and DNA sequences from small tissue samples isolated by laser-assisted microdissection. *Laboratory Investigation* **78**, 1267–1273.
- BETRÁN, E., ROZAS, J., NAVARRO, A. & BARBADILLA, A. (1997). The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. *Genetics* **146**, 89–99.
- BLOUIN, M. S., YOWELL, C. A., COURTNEY, C. H. & DAME, J. B. (1998). Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Molecular Biology and Evolution* **15**, 1719–1727.
- CADWELL, R. C. & JOYCE, G. F. (1994). Mutagenic PCR. *PCR Methods and Applications* **3**, S136–S140.
- CAMPBELL, C. S., WOJCIECHOWSKI, M. F., BALDWIN, B. G., ALICE, L. A. & DONOGHUE, M. J. (1997). Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae). *Molecular Biology and Evolution* **14**, 81–90.
- CHILTON, N. B., HOSTE, H., NEWTON, L. A., BEVERIDGE, I. & GASSER, R. B. (1998). Common secondary structures for the second internal transcribed spacer pre-rRNA of two subfamilies of trichostrongylid nematodes. *International Journal for Parasitology* **28**, 1765–1773.
- CHILTON, N. B., NEWTON, L. A., BEVERIDGE, I. & GASSER, R. B. (2001). Evolutionary relationships of Trichostrongylid nematodes (Strongylida) inferred from ribosomal DNA sequence data. *Molecular Phylogenetics and Evolution* **19**, 367–386.
- CHITWOOD, B. G. & CHITWOOD, M. B. (1972). *Introduction to Nematology*. University Park Press.
- CRONN, R., CEDRONI, M., HASELKORN, T., GROVER, C. & WENDEL, J. F. (2002). PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* **104**, 482–489.
- ELBADRI, G. A. A., DE LEY, P., WAEYENBERGE, L., VIERSTRAETE, A., MOENS, M. & VANFLETEREN, J. (2002). Intraspecific variation in *Radopholus similis* isolates assessed with restriction fragment length polymorphism and DNA sequencing of the internal transcribed spacer region of the ribosomal RNA cistron. *International Journal for Parasitology* **32**, 199–205.
- ETTER, A., BERNARD, V., KENZELMANN, M., TOBLER, H. & MULLER, F. (1994). Ribosomal heterogeneity from chromatin diminution in *Ascaris lumbricoides*. *Science* **265**, 954–956.
- FLEMMING, A. J., SHEN, Z. Z., CUNHA, A., EMMONS, S. W. & LEROI, A. M. (2000). Somatic polyploidization and cellular proliferation drive body size evolution in nematodes. *Proceedings of the National Academy of Sciences, USA* **97**, 5285–5290.
- FLOYD, R., ABEBE, E., PAPERT, A. & BLAXTER, M. (2002). Molecular barcodes for soil nematode identification. *Molecular Ecology* **11**, 839–850.
- GASSER, R. B., CHILTON, N. B., HOSTE, H. & BEVERIDGE, I. (1993). Rapid sequencing of rDNA from single worms and eggs of parasitic helminths. *Nucleic Acids Research* **21**, 2525–2526.
- GASSER, R. B., ZHU, X. Q., CHILTON, N. B., NEWTON, L. A., NEDERGAARD, T. & GULDBERG, P. (1998). Analysis of sequence homogenization in rDNA arrays of *Haemonchus contortus* by denaturing gradient gel electrophoresis. *Electrophoresis* **19**, 2391–2395.
- HIGAZI, T. B., KATHOLI, C. R., MAHMOUD, B. M., BARAKA, O. Z., MUKHTAR, M. M., AL QUBATI, Y. & UNNASCH, T. R. (2001). *Onchocerca volvulus*: Genetic diversity of parasite isolates from Sudan. *Experimental Parasitology* **97**, 24–34.
- HILLIS, D. M., MORITZ, C., PORTER, C. A. & BAKER, R. J. (1991). Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science* **251**, 308–310.
- 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**, 157–164.
- JENTSCH, S., TOBLER, H. & MULLER, F. (2002). New telomere formation during the process of chromatin diminution in *Ascaris suum*. *International Journal of Developmental Biology* **46**, 143–148.
- KANOBANA, K., VERVELDE, L., VAN DER VEER, M., EYSKER, M. & PLOEGER, H. W. (2001). Characterization of host

- responder types after a single *Cooperia oncophora* infection: kinetics of the systemic immune response. *Parasite Immunology* **23**, 641–653.
- KLOOSTERMAN, A., ALBERS, G. A. A. & VAN DEN BRINK, R. (1978). Genetic variation among calves in resistance to nematode parasites. *Veterinary Parasitology* **4**, 353–368.
- LIANG, Q. W., CHEN, L. S. & FULCO, A. J. (1995). An efficient and optimized PCR method with high fidelity for site-directed mutagenesis. *PCR Methods and Applications* **4**, 269–274.
- MAFF (1977). *Manual of Veterinary Parasitological Laboratory Techniques*. Technical Bulletin 18. Her Majesty's Stationary Office, London, UK.
- MES, T. H. M. (2003). Demographic expansion of parasitic nematodes of livestock based on mitochondrial DNA regions that conflict with the infinite-sites model. *Molecular Ecology* **12**, 1555–1566.
- MULLER, F. & TOBLER, H. (2000). Chromatin diminution in the parasitic nematodes *Ascaris suum* and *Parascaris univalens*. *International Journal for Parasitology* **30**, 391–399.
- 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. *Journal of Parasitology* **86**, 588–601.
- NEWTON, L. A., CHILTON, N. B., BEVERIDGE, I. & GASSER, R. B. (1998). Genetic evidence indicating that *Cooperia surnabada* and *Cooperia oncophora* are one species. *International Journal for Parasitology* **28**, 331–336.
- POSADA, D. & CRANDALL, K. A. (1998). Modeltest: testing the model of nucleotide substitution. *Bioinformatics* **14**, 817–818.
- POWERS, T. O., TODD, T. C., BURNELL, A. M., MURRAY, P. C. B., FLEMING, C. C., SZALANSKI, A. L., ADAMS, B. A. & HARRIS, T. S. (1997). The rDNA internal transcribed spacer region as a taxonomic marker for nematodes. *Journal of Nematology* **29**, 441–450.
- ROBERTS, F. H. S. & O'SULLIVAN, P. J. (1950). Methods for egg counts and larval cultures for strongyles infecting the gastro-intestinal tract of cattle. *Australian Journal of Agricultural Research* **1**, 99–102.
- ROZAS, J. & ROZAS, R. (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**, 174–175.
- ROZAS, J., GULLAUD, M., BLANDIN, G. & AGUADE, M. (2001). DNA variation at the rp49 gene region of *Drosophila simulans*: evolutionary inferences from an unusual haplotype structure. *Genetics* **158**, 1147–1155.
- SAWYER, S. A. (1989). Statistical tests for detecting gene conversion. *Molecular Biology and Evolution* **6**, 526–538.
- SAWYER, S. A. (1999). GENECONV: a computer package for the statistical detection of gene conversion. Distributed by the author, Dept. of Mathematics, Washington University in St Louis, available at <http://www.math.wustl.edu/~sawyer>.
- SHI, Y. & BLACKWELL, T. K. (2003). A two-tiered transcription regulation mechanism that protects germ cell identity. *Molecular Cell* **12**, 1062–1064.
- SHIN, T. H. & MELLO, C. C. (2003). Chromatin regulation during *C. elegans* germline development. *Current Opinion in Genetics and Development* **13**, 455–462.
- SIJEN, T. & PLASTERK, R. H. A. (2003). Transposon silencing in the *Caenorhabditis elegans* germ line by natural RNAi. *Nature, Genetics* **426**, 310–314.
- STEVENSON, L. A., CHILTON, N. B. & GASSER, R. B. (1995). Differentiation of *Haemonchus placei* from *H. contortus* (Nematoda: Trichostrongylidae) by the ribosomal DNA second internal transcribed spacer. *International Journal for Parasitology* **25**, 483–488.
- STEVENSON, L. A., GASSER, R. B. & CHILTON, N. B. (1996). The ITS-2 rDNA of *Teladorsagia circumcincta*, *T. trifurcata* and *T. davitiani* (Nematoda: Trichostrongylidae) indicates that these taxa are one species. *International Journal for Parasitology* **26**, 1123–1126.
- SUBBOTIN, S. A., HALFORD, P. D., WARRY, A. & PERRY, R. N. (2000a). Variations in ribosomal DNA sequences and phylogeny of *Globodera* parasitising solanaceous plants. *Nematology* **2**, 591–604.
- SUBBOTIN, S. A., WAEYENBERGE, L. & MOENS, M. (2000b). Identification of cyst forming nematodes of the genus *Heterodera* (Nematoda: Heteroderidae) based on the ribosomal DNA-RFLP. *Nematology* **2**, 153–164.
- SWOFFORD, D. L. (1998). *PAUP\*. Phylogenetic Analysis using Parsimony (\* and other Methods)*. Sinauer Associates, Sunderland, Massachusetts.
- 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. *Journal of Nematology* **29**, 255–267.
- TAJIMA, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.
- TRIANTAPHYLLOU, A. C. (1991). Further studies on the role of polyploidy in the evolution of *Meloidogyne*. *Journal of Nematology* **23**, 249–253.
- WARD, J. G., BLOMBERG, P., HOFFMAN, N. & YAO, M. C. (1997). The intranuclear organization of normal, hemizygous and excision-deficient rRNA genes during developmental amplification in *Tetrahymena thermophila*. *Chromosoma* **106**, 233–242.