

# The mitochondrial genome of the soybean cyst nematode, *Heterodera glycines*

Tracey Gibson, Daniel Farrugia, Jeff Barrett, David J. Chitwood, Janet Rowe, Sergei Subbotin, and Mark Dowton

**Abstract:** We sequenced the entire coding region of the mitochondrial genome of *Heterodera glycines*. The sequence obtained comprised 14.9 kb, with PCR evidence indicating that the entire genome comprised a single, circular molecule of approximately 21–22 kb. The genome is the most T-rich nematode mitochondrial genome reported to date, with T representing over half of all nucleotides on the coding strand. The genome also contains the highest number of poly(T) tracts so far reported (to our knowledge), with 60 poly(T) tracts  $\geq 12$  Ts. All genes are transcribed from the same mitochondrial strand. The organization of the mitochondrial genome of *H. glycines* shows a number of similarities compared with *Radopholus similis*, but fewer similarities when compared with *Meloidogyne javanica*. Very few gene boundaries are shared with *Globodera pallida* or *Globodera rostochiensis*. Partial mitochondrial genome sequences were also obtained for *Heterodera cardiolata* (5.3 kb) and *Punctodera chalconensis* (6.8 kb), and these had identical organizations compared with *H. glycines*. We found PCR evidence of a minicircular mitochondrial genome in *P. chalconensis*, but at low levels and lacking a noncoding region. Such circularised genome fragments may be present at low levels in a range of nematodes, with multipartite mitochondrial genomes representing a shift to a condition in which these subgenomic circles predominate.

**Key words:** mitochondrial, nematode, gene rearrangement, *Punctodera*, Punctoderinae, Heteroderidae, *Heterodera cardiolata*.

**Résumé :** Les auteurs ont séquencé complètement la région codante du génome mitochondrial de l'*Heterodera glycines*. La séquence obtenue mesure 14,9 kb et des évidences PCR indiquent que le génome entier serait composé d'une seule molécule circulaire d'environ 21–22 kb. Ce génome est le génome mitochondrial le plus riche en T jamais rapporté chez les nématodes, les T constituant plus de la moitié de tous les nucléotides sur le brin codant. Le génome compte également le nombre le plus élevé de suites poly(T) vu à ce jour (au meilleur de leur connaissance), avec 60 suites comptant plus de 12 T. Tous les gènes sont transcrits à partir du même brin mitochondrial. L'organisation du génome mitochondrial chez l'*H. glycines* affiche de nombreuses similarités avec celui du *Radopholus similis*, mais moins de similitude lorsque comparé à celui du *Meloidogyne javanica*. Très peu de frontières entre les gènes sont en commun avec le *Globodera pallida* ou le *Globodera rostochiensis*. Des séquences partielles du génome mitochondrial ont également été obtenues pour le *Heterodera cardiolata* (5,3 kb) et le *Punctodera chalconensis* (6,8 kb), et ceux-ci présentaient une organisation identique à celle du *H. glycines*. Les auteurs ont trouvé des évidences PCR d'un génome mitochondrial circulaire miniature chez le *P. chalconensis*, mais à de faibles niveaux et sans région non-codante. Des tels fragments de génome circularisés sont possiblement présents en faible nombre chez plusieurs nématodes, les génomes mitochondriaux multipartites constituant une transition vers une condition où ces cercles subgénomiques prédominent.

**Mots-clés :** mitochondrial, nématode, réarrangement de gènes, *Punctodera*, Punctoderinae, Heteroderidae, *Heterodera cardiolata*.

[Traduit par la Rédaction]

## Introduction

*Heterodera glycines* Ichinohe (the soybean cyst nematode) is a plant parasitic nematode belonging to the family Hetero-

deridae. It is a major pest of soybean crops both in Asia and the USA (Subbotin et al. 2001a). A number of races (Riggs and Schmitt 1988) or types (Niblack et al. 2002) of *H. glycines* have been identified, primarily based on their ability to

Received 10 February 2010. Accepted 3 February 2011. Published at [www.nrcresearchpress.com/gen](http://www.nrcresearchpress.com/gen) on 11 July 2011.

Paper handled by Associate Editor Linda Bonen.

**T. Gibson, D. Farrugia, J. Barrett, and M. Dowton.** Centre for Medical Bioscience, School of Biological Sciences, Wollongong University, Wollongong, NSW, 2522, Australia.

**D.J. Chitwood.** Nematology Laboratory, United States Department of Agriculture, Agricultural Research Service, Bldg 011A, BARC-W, Beltsville, MD 20705, USA.

**J. Rowe.** Nematode Interactions Unit, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, United Kingdom.

**S. Subbotin.** Plant Pest Diagnostics Center, California Department of Food and Agriculture, 3294 Meadowview Road, Sacramento, CA 95832-1448, USA.

**Corresponding author:** Mark Dowton (e-mail: [mdowton@uow.edu.au](mailto:mdowton@uow.edu.au)).

infest a number of indicator soybean cultivars. Genetic approaches for distinguishing *H. glycines* populations could provide a means of rapidly distinguishing these types. However, such genetic markers have not yet been developed. The mitochondrial genome of *H. glycines* may provide such genetic markers. For example, intergenic regions in the mitochondrial genome of *Apis mellifera* have been used to distinguish different populations (Cornuet et al. 1991).

The mitochondrial genome of animals is normally a circular, double-stranded molecule containing 36–37 genes: 12–13 protein-coding genes, 2 rRNA genes, and 22 tRNA genes (Boore 1999). Mitochondrial genomes can provide two types of characters for the resolution of phylogeny. The first type is the more traditional set of aligned sequences that can be analyzed by parsimony or likelihood; mitochondrial genomes can provide datasets that are typically 10 000–16 000 characters long. The second type of character is extracted from the organization of the mitochondrial genome (Dowton et al. 2002). It has been argued that large-scale genome changes occur rarely and that reversions or secondary losses will be even rarer (Boore et al. 1998; Rokas and Holland 2000). As a result, mitochondrial gene rearrangements have been used to infer phylogeny between a number of animal groups (Boore et al. 1998; Boore and Staton 2002; Dowton 1999; Dowton et al. 2002; Knudsen et al. 2006; Kurabayashi and Ueshima 2000; Smith et al. 1993). The Heteroderidae represent a group in which mitochondrial gene rearrangement characters may be informative. Extensive gene rearrangements have occurred during the evolution of the Hoplolaimoidea (to which the Heteroderidae belong), as evident from the very different genome organizations of *Meloidogyne javanica* (from the Meloidogynidae) (Okimoto et al. 1991), *Radopholus similis* (Jacob et al. 2009), *Globodera pallida* (Armstrong et al. 2000; Gibson et al. 2007b), and *Globodera rostochiensis* (Gibson et al. 2007a).

In this study, we report the complete coding sequence of the mitochondrial genome of *H. glycines*. Although a portion of the noncoding region could not be sequenced, our data allow us to identify the relative position of each of the 36 mitochondrial genes commonly found in nematodes and to compare the organization of the *H. glycines* mitochondrial genome with other nematodes. In addition, we sequence partial mitochondrial genomes of *Heterodera cardiolata* and *Punctodera chalcoensis*, to obtain a more complete understanding of the history of mitochondrial gene rearrangement in the Heteroderidae.

## Materials and methods

### Genomic DNA extraction

Cysts from *H. glycines* (USA), *H. cardiolata* (= *Heterodera cynodontis*, Pakistan), and *P. chalcoensis* (Mexico) were collected and stored in 95% ethanol at 4 °C. Genomic DNA from single cysts was prepared as described (Sunnucks and Hales 1996). We verified the identity of each species by amplifying and sequencing the ITS1 and ITS2 regions, using the primers described in Subbotin et al. (2001b). For these reactions, we used the Bio-X-Act long PCR system (Biolone, Australia).

### Amplification and sequencing of mitochondrial fragments

We designed primers to anneal to conserved regions of nematode mitochondrial genes and amplified a number of internal gene fragments from each of the three nematodes studied (primer sequences available upon request). We then designed precisely matching primers, using the sequence data obtained, and amplified a number of multigenic fragments. All amplifications were performed as previously described, using the Bio-X-Act long PCR system (Gibson et al. 2007a). Owing to the presence of poly(T) variation in the amplicons (Riepsamen et al. 2008), most multigenic fragments were then cloned directly into the pGEM-T Easy vector system (Promega, Madison, Wisconsin) and sequenced by primer walking. As this poly(T) variation has been found to be primarily an artefact of amplification of templates with extensive poly(T) regions (Riepsamen et al. 2011), we sequenced multiple clones to infer the biological length of these poly(T) regions.

### Step-out PCR amplification

An approximately 7-kb fragment of the *H. glycines* mitochondrial genome was amplified and sequenced using long PCR (described above). This fragment spanned the *cox2*–*cob* genes. Step-out PCR (Burger et al. 2007) was then used to amplify the flanking regions of this fragment. In the first round of step-out PCR, genomic DNA was amplified using a single outward-facing, species-specific primer. After an initial denaturation at 94 °C for 2 min, the cycling conditions for cycles 1–5 were as follows; 94 °C for 20 s, annealing (temperature set at 5 °C below the melting temperature of the primer) for 20 s, and 68 °C for 8 min. The use of a high annealing temperature is designed to increase the concentration of the desired template by linear (single stranded) amplification. In cycle 6, the conditions remain the same, but the annealing temperature is dropped to 30 °C (for 2 min) to facilitate annealing of the species-specific primer to other regions of the desired template. Cycles 7–22 comprise a touch-down PCR, in which the cycling conditions remain the same, but the annealing temperature is increased to 65 °C (for 20 s), dropping by 1 °C each cycle. The touch-down approach is designed to amplify only those double-stranded molecules produced during cycle 6 (reducing the amount of nonspecific amplification). In cycles 23–40, the annealing temperature is raised to that used during cycles 1–5. The extension time is increased by 30 s in each of cycles 23–40, to offset any reduction in enzyme activity that has occurred. A final incubation at 68 °C for 12 min is included, to complete the synthesis of any partial amplicons.

In the second round of the step-out procedure, two primers were used. One of these was the species-specific primer used in round 1, while the second was another species-specific primer designed from previously acquired sequencing data downstream of the species-specific primer. Amplification using this seminested approach was optimized using an annealing temperature gradient. Generally, single high molecular mass amplicons were only produced using relatively high annealing temperatures. These high molecular mass products were purified by agarose gel electrophoresis and cloned into the pGEM-T Easy vector system. Once the step-out product was sequenced, precisely matching primers were designed and the data were verified by long PCR of genomic DNA,

followed by sequencing. As data accumulated, successive rounds of step-out PCR were performed to obtain the entire coding sequence of the mitochondrial genome.

### Genome annotation

ORFinder ([www.ncbi.nlm.nih.gov/gorf/gorf.html](http://www.ncbi.nlm.nih.gov/gorf/gorf.html)) was used to identify protein-coding genes. The tRNAscan-SE 1.21 (Lowe and Eddy 1997) search server was used to identify tRNA genes using the following settings: source, Nematode Mito; genetic code for tRNA isotype prediction, Invertebrate Mito; and cove score cutoff, 10. Although the authors of tRNAscan recommend a cove cutoff of 20 for organelle tRNA genes (to avoid false positives), we set the cutoff to 10 to screen all putative tRNA genes — this strategy has been useful for the detection of a range of insect tRNA genes (Castro et al. 2006; Downton 1999). Nematode mitochondrial tRNA genes display more plastic secondary structure (and presumably lower cove scores) than most other sources of tRNA (Wolstenholme et al. 1987). We employed an additional criterion to identify tRNA genes: that the tRNA gene should not lie within a protein-coding gene, but lie in the small intergenic regions between genes (or in the longer non-coding regions). rRNA genes were identified by rRNA BLAST search ([www.psb.ugent.be/rRNA/blastrrna.html](http://www.psb.ugent.be/rRNA/blastrrna.html)). Partial mitochondrial genome sequences were deposited in GenBank, accession numbers HM640928–HM640930.

## Results and discussion

### Sequencing of the mitochondrial-coding regions of *H. glycines*

Using primers designed to anneal to conserved regions of nematode mitochondrial genes, we amplified a number of internal gene fragments. We then designed precisely matching primers, using the sequence data obtained, and amplified a number of multigenic fragments. This produced a contiguous fragment of approximately 7 kb, spanning the *nad5*–*cob* genes. However, attempts to amplify the remainder of the genome by long PCR (using outwardly facing primers) were unsuccessful. Instead, we employed step-out PCR (Burger et al. 2007) to obtain the remainder of the coding region of the genome. Each step-out product was cloned and sequenced, and the organization of the genome was checked by designing PCR primers with the newly obtained sequence data. The step-out products were validated by amplification of genomic DNA, using primers that amplified fragments that bridged the previous and newly acquired sequence.

With this approach, we amplified and sequenced 14.9 kb of the *H. glycines* mitochondrial genome. Bioinformatic analysis of this fragment indicated that it contained the 36 mitochondrial genes normally found in nematode mitochondrial genomes. However, the entire noncoding region proved very difficult to amplify. Utilizing outward-facing primers that annealed within the control region, we amplified a faint product of 6–7 kb, but this product was resistant to cloning. This may have been due to the presence of a repeats in the control region. From this information, we estimate that the entire mitochondrial genome of *H. glycines* is 21–22 kb, broadly consistent with an estimate of 22.5–23.5 kb after characterization of the restriction fragments produced from purified *H. glycines* mitochondrial DNA (Radice et al. 1988).

### Nucleotide content of the *H. glycines* mitochondrial genome

The nucleotide composition of the coding strand was extremely biased, with T representing more than half of the nucleotides (58% T, 24% A, 10% G, and 7% C). This trend (T > A > G > C) has been reported for a number of other nematodes (Hu et al. 2003b, 2003c; Jacob et al. 2009) (Table 1). Indeed, a survey of all completely sequenced nematode mitochondrial genomes indicated that chromadoreans have a T-content of between 45% and 58% (although *Caenorhabditis briggsae* is a notable exception, with a T-content of only 30%). The highest T-content is found in *H. glycines* (58%), although *Brugia malayi* (54%), *Dirofilaria immitis* (55%), *Onchocerca volvulus* (54%), and *R. similis* (53%) have comparable levels. By comparison, enoplean nematodes tend to have slightly lower T-contents, with a range of 27% (*Trichinella spiralis*) to 44% (*Agamermis* sp.). A comparison with other pseudocoelomates (from the phyla Acanthocephala and Rotifera) indicates that they also have elevated T-contents (Table 1). A high T-content appears to be a feature of pseudocoelomates, with the highest levels evident in the nematode orders Spirurida and Tylenchida (Table 1).

An unusual feature of the genome is the presence of long stretches of T nucleotides, hereafter referred to as poly(T) tracts. This has been noted for other nematodes (Hu et al. 2003b; Jacob et al. 2009; Riepsamen et al. 2008). For example, *Strongyloides stercoralis* has 16 poly(T) tracts  $\geq$  12 Ts, with the longest tract 23 nucleotides long (Hu et al. 2003b). In comparison, *H. glycines* has 60 poly(T) tracts  $\geq$  12 Ts, with the longest tract also 23 nucleotides long (Table 2). In *S. stercoralis*, the longest poly(T) tract lies within the *16S* gene, but in *H. glycines* it lies within the *nad2* gene. To our knowledge, *H. glycines* has the most extreme T-content and the highest number of poly(T) tracts amongst the nematodes.

### Codon usage and amino acid bias

The high T-content is reflected in a biased use of synonymous codons and a biased amino acid content. For example, phenylalanine can be encoded by TTT or TTC; the occurrence of TTT is more than 20 times that of TTC, while nearly one-third of all amino acids are phenylalanine (encoded by the most T-rich codon; Table 3). For every amino acid, the most T-rich synonymous codon outnumbers the sum of all other synonymous codons (Table 3). Our data were not consistent with the genetic code change suggested for *R. similis*, in which it has been proposed that the UAA stop codon is reassigned to tyrosine (Jacob et al. 2009). The conclusion of codon reassignment was based on the presence of premature stop codons in each of the 12 protein-coding genes (Jacob et al. 2009). Although we have reported the occurrence of premature stop codons in other heteroderid nematodes (Gibson et al. 2007b; Riepsamen et al. 2008), this was later found to be due to amplification errors (Riepsamen et al. 2011). Long poly(T) tracts are inaccurately amplified, with single T deletions being the most common error (Riepsamen et al. 2011). These T deletions introduce artefactually premature stop codons in the sequence obtained. *Radopholus similis* also has long poly(T) tracts (Table 2), and we suspect that amplification introduces T deletions in this species and that this is the underlying reason for the observation of premature stop codons in *R. similis*. An alternative explanation

**Table 1.** Nucleotide composition of the coding strand of nematode mitochondrial genomes.

Taxon	%T	%A	%G	%C	GenBank
Phylum Nematoda					
Class Chromadorea					
<b>Order Ascaradida</b>					
<i>Anisakis simplex</i>	48.3	22.9	19.0	9.8	AY994157
<i>Ascaris suum</i>	49.8	22.2	20.4	7.7	X54253
<i>Toxocara canis</i>	46.7	21.9	22.0	9.4	AM411108
<b>Order Oxyurida</b>					
<i>Enterobius vermicularis</i>	49.5	21.7	22.6	6.2	EU281143
<b>Order Rhabditida</b>					
<i>Ancylostoma caninum</i>	48.5	29.0	16.1	6.5	FJ483518
<i>Bunostomum phlebotomum</i>	50.1	26.9	16.7	6.2	FJ483517
<i>Necator americanus</i>	49.2	27.4	17.1	6.3	AJ417719
<i>Angiostrongylus cantonensis</i>	48.7	24.5	20.6	6.1	NC_013065
<i>Metastrongylus pudendotectus</i>	51.4	26.1	16.7	5.8	GQ888714
<i>Syngamus trachea</i>	47.8	26.0	18.8	7.4	GQ888718
<i>Chabertia ovina</i>	46.7	29.6	16.3	7.4	GQ888721
<i>Oesophagostomum dentatum</i>	47.1	28.6	17.1	7.2	GQ888716
<i>Cylicocyclus insignis</i>	46.0	30.6	16.6	6.8	GQ888712
<i>Strongylus vulgaris</i>	47.2	29.3	16.8	6.7	GQ888717
<i>Cooperia oncophora</i>	47.3	30.1	16.1	6.5	AY265417
<i>Haemonchus contortus</i>	44.7	33.4	15.4	6.5	EU346694
<i>Mecistocirrus digitatus</i>	45.7	34.0	14.4	5.9	GQ888722
<i>Teladorsagia circumcincta</i>	46.1	31.0	15.6	7.3	GQ888720
<i>Trichostrongylus axei</i>	45.3	31.1	17.0	6.6	GQ888719
<i>Steinernema carpocapsae</i>	50.8	25.5	14.4	9.3	AY591323
<i>Heterorhabditis bacteriophora</i>	47.0	28.8	16.1	8.1	EF043402
<i>Caenorhabditis elegans</i>	44.8	31.4	14.9	8.9	X54252
<b>Order Spirurida</b>					
<i>Brugia malayi</i>	53.9	21.6	16.8	7.7	AF538716
<i>Dirofilaria immitis</i>	54.9	19.3	19.3	6.6	AJ537512
<i>Onchocerca volvulus</i>	54.0	19.3	19.8	6.9	AF015193
<b>Order Tylenchida</b>					
<i>Radopholus similis</i>	52.9	32.5	10.2	4.4	FN313571
<i>Globodera pallida</i> (circle 1)	46.2	20.0	20.5	13.2	AJ249395
<i>Heterodera glycines</i>	58.5	24.2	10.4	6.9	HM640930
Class Enoplea					
<b>Order Mermithida</b>					
<i>Agamermis</i> sp.	43.7	36.8	9.9	9.6	DQ665656
<i>Hexamermis agrotis</i>	37.4	41.0	10.9	10.7	EF368011
<i>Romanomermis culicivorax</i>	38.0	41.3	10.4	10.3	EF154459
<i>Strelkovimermis spiculatus</i>	42.5	36.0	11.3	10.1	DQ520859
<i>Thaumamermis cosgrovei</i>	39.3	32.1	14.7	13.9	DQ520857
<i>Trichinella spiralis</i>	26.6	40.6	9.8	23.0	AF293969
Phylum Acanthocephala					
<i>Leptorhynchoides thecatus</i>	44.6	26.8	19.9	8.4	AY562383
Phylum Rotifera					
<i>Rotaria rotatoria</i>	43.7	29.5	17.4	9.5	GQ304898
<i>Brachionus plicatilis</i>	36.0	27.0	20.7	16.3	NC_010484

**Note:** Entire mitochondrial genome data was accessed from GenBank on 20th April 2010. Only a single species was included for genera with multiple representatives.

is that +1 or +2 frameshifts correct the reading frame and remove the premature stop codons. During translation +1 frameshifts involve the slippage of the mRNA such that the reading frame is shifted by a single nucleotide — such frameshifts are usually preceded by rare codons that may mediate the slippage and have been reported in the mitochondrial

genes of birds, turtles (Mindell et al. 1998), and ants (Beckenbach et al. 2005). If such translational slippage occurs in *H. glycines*, it would have to involve both +1 and +2 frameshifts in order for all genomic sequences to give rise to functional products. In addition, the poly(T) tracts of *H. glycines* were not preceded by a rare codon.

**Table 2.** Length and frequency of poly(T) tracts in nematode mitochondrial genomes with high T-content.

Length of poly(T) region (nt)	Number of occurrences			
	<i>Strongyloides stercoralis</i>	<i>Radopholus similis</i>	<i>Heterodera glycines</i>	<i>Globodera pallida</i> scmtI
23	1	0	1	0
22	0	0	0	0
21	0	0	1	0
20	0	0	0	0
19	0	0	1	1
18	0	0	4	1
17	0	0	1	0
16	2	0	4	1
15	1	0	5	0
14	1	0	9	1
13	3	4	9	1
12	8	3	25	4
11	9	8	39	4
10	14	17	41	5
9	28	30	50	26
8	42	39	79	32
7	62	60	78	41
6	93	71	109	37
≥12	16	7	60	9
T-content	56%	53%	58%	46%

### Mitochondrial genome organization in *H. glycines*

Figure 1 shows the organization of genes on the 14.9 kb of sequence data obtained. The genome contains a major non-coding region (between *trnG* and *trnN*), which we estimate to be 7–8 kb long. There are also smaller noncoding tracts between the *a6* and *nad5* genes (330 nucleotides), the *trnR* and *trnD* genes (335 nucleotides), and between the *nad4L* and *trnG* genes (130 nucleotides). These noncoding tracts represent regions that could be amplified from a range of *H. glycines* populations, to investigate whether they might be used to distinguish *H. glycines* populations. A similar approach has been used to distinguish various populations of *A. mellifera* (Cornuet et al. 1991).

Our sequence data are consistent with *H. glycines* possessing a conventional, single circle mitochondrial genome, with the 36 genes normally present on nematode mitochondrial genomes present on the one fragment. We found no evidence for the multipartite structure reported for *G. pallida* and *G. rostochiensis* (Armstrong et al. 2000; Gibson et al. 2007a), in which these genes are distributed on multiple circles with overlapping content.

As is the case for other nematodes from the class Chromadorea, all 36 genes are transcribed in the same direction (e.g., Hu et al. 2003a, 2002, 2003b, 2003c; Okimoto et al. 1992). The two ribosomal genes (*12S* and *16S*) are not found adjacent to each other. This is also the case for other nematodes (He et al. 2005), although *T. spiralis* is a notable exception (Lavrov and Brown 2001).

### Comparison of mitochondrial genome organization with other nematodes from the Hoplolaimoidea

Table 4 indicates those members of the Hoplolaimoidea for

which some mitochondrial genome organization is available, together with their classification — this provides a framework for comparisons. Some gene boundaries are shared with other nematodes from the superfamily Hoplolaimoidea. The organization of the mitochondrial genome of *R. similis* (Jacob et al. 2009) shows the highest level of similarity to that of *H. glycines* (Fig. 2). If tRNA genes are ignored, a large block containing six protein-coding and two ribosomal RNA genes are identically organized in the two (Fig. 2). Similarly, *nad6* and *nad4L* are adjacent in both genomes, as are *nad4* and *cob*, and *nad2* and *cox3*. A similar comparison of *H. glycines* with *M. javanica* (He et al. 2005; Okimoto et al. 1991) reveals many fewer shared gene boundaries (if tRNA genes are ignored), with the largest block being three genes long (*cox2–16S–nad3*) (see Supplementary data,<sup>1</sup> Table S1). The higher level of organizational similarity between *H. glycines* and *R. similis* is consistent with the close relationship between the Pratylenchidae (to which *Radopholus* belongs) and the Heteroderidae, while Meloidogynidae and Heteroderidae are more distantly related (Bert et al. 2008). Our mitochondrial genome organization data are also consistent with recent molecular analyses of the Tylenchida. Based on an analysis of the 28S rRNA gene, Subbotin et al. (2006) reported that the data did not support the traditional sister grouping of Meloidogynidae and Heteroderidae. Although the level of support was not convincing (in terms of bootstrap support or posterior probabilities), the Heteroderidae were most consistently recovered in a clade together with the Hoplolaimidae, Rotylenchulidae, and *Radopholus*. The data were sufficient to reject the traditional hypothesis (Heteroderidae + Meloidogynidae).

We then compared the genome organization of *G. pallida* (Armstrong et al. 2000; Gibson et al. 2007b) and *G. rosto-*

<sup>1</sup>Supplementary data are available with the article at [www.nrcresearchpress.com/gen](http://www.nrcresearchpress.com/gen)

**Table 3.** Number of times each codon is used for all protein-coding genes of the mitochondrial genome of *Heterodera glycines*.

AA	Codon	Number
F	TTT	1029
	TTC	45
L <sub>2</sub>	TTA	299
	TTG	55
L <sub>1</sub>	CTT	94
	CTC	3
	CTA	24
	CTG	7
I	ATT	239
	ATC	17
M	ATA	92
	ATG	8
V	GTT	132
	GTC	10
	GTA	24
	GTG	6
S <sub>2</sub>	TCT	105
	TCC	17
	TCA	21
	TCG	5
P	CCT	30
	CCC	2
	CCA	21
	CCG	2
T	ACT	44
	ACC	9
	ACA	10
	ACG	2
A	GCT	35
	GCC	2
	GCA	10
	GCG	2
Y	TAT	113
	TAC	7
H	CAT	43
	CAC	1
Q	CAA	18
	CAG	12
N	AAT	122
	AAC	10
K	AAA	75
	AAG	16
D	GAT	51
	GAC	7
E	GAA	43
	GAG	15
C	TGT	20
	TGC	2
W	TGA	41
	TGG	17
R	CGT	20
	CGC	2
	CGA	5
	CGG	1
S <sub>1</sub>	AGT	78
	AGC	3
	AGA	26
	AGG	13
G	GGT	87
	GGC	3
	GGA	23
	GGG	13

*chiensis* (Gibson et al. 2007a) to that of *H. glycines*. Although the two *Globodera* species are more closely related to *H. glycines* than is *R. similis* (*Globodera* and *Heterodera*

are both members of the Heteroderidae), a comparison of the relative positions of protein-coding and ribosomal genes reveals no shared gene boundaries. An examination of the positions of tRNA genes indicates that there are just two regions of similarity (Table S1). One of these is a cluster of 4–5 tRNA genes that are nearly identical between the two, while the *nad4* gene is surrounded by the *T* and *H* genes in both. The lack of similarity between the *Globodera* and *Heterodera* mitochondrial genomes is likely owing to the multipartite nature of the *Globodera* mitochondrial genome (Armstrong et al. 2000; Gibson et al. 2007a; Riepsamen et al. 2008).

#### Partial mitochondrial genome organization in *H. cardiolata*

We next amplified and sequenced a number of multigenic mitochondrial fragments from *H. cardiolata*. *Heterodera cardiolata* and *H. glycines* are relatively divergent members of the Heteroderinae (Ma et al. 2008; Subbotin et al. 2001b). A comparison of their genomes would, thus, provide an indication of the level of variability of organizational structure within the Heteroderinae.

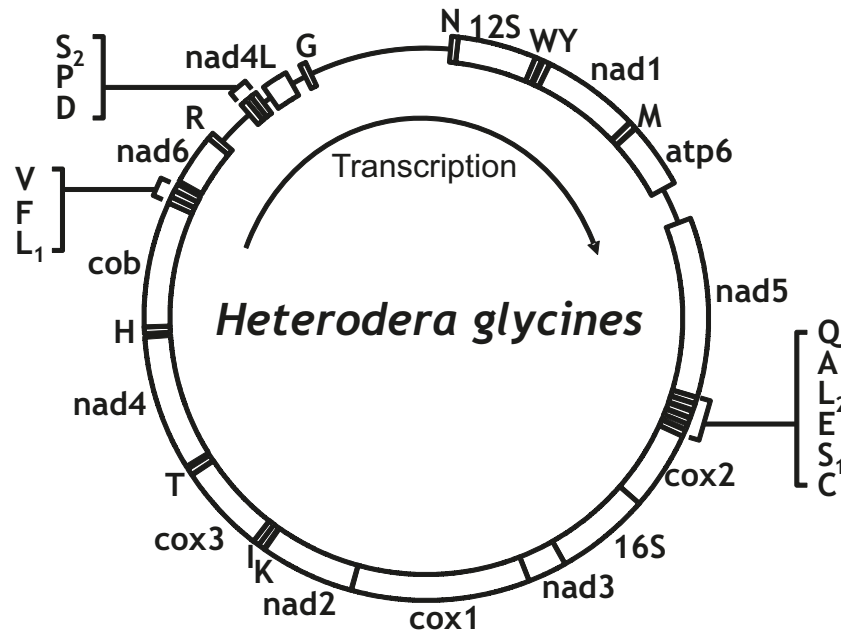
Using primers designed to anneal to conserved regions of the nematode *nad4* and *cox2* genes, we amplified a 5.5-kb mitochondrial fragment from *H. cardiolata*. Similarly, using *cox1* and *nad4* primers, we were able to amplify a *nad4*–*cox1* fragment (3 kb), and we also amplified a *cox1*–*cox2* fragment (2.5 kb) using *cox1* and *cox2* primers. We cloned and sequenced these two latter fragments, as they presumably represent portions of the larger 5.5-kb amplicon. Sequence data indicated the partial genome organization shown in Fig. 2. This indicated that the organization of genes in these mitochondrial fragments was identical when *H. glycines* and *H. cardiolata* are compared.

#### Partial mitochondrial genome organization in *P. chalconensis*

*Punctodera chalconensis* is a member of the Punctoderinae, and thus represents a distinct subfamily within the Heteroderidae (compared with the Heteroderinae) and a distinct lineage within the Punctoderinae (compared with *Globodera*) (see Table 4) (Subbotin et al. 2001b). We amplified three overlapping, multigenic fragments from *P. chalconensis*. Complete sequencing indicated that the fragments were precisely overlapping and comprised a contig of 6.8 kb. The organization of genes within this 6.8 kb was identical to that of *H. glycines* and *H. cardiolata* (Fig. 2). Most attempts to amplify the remainder of the mitochondrial genome, using outwardly facing primers, were unsuccessful (but see below). Step-out PCR is likely necessary to obtain the remainder of this genome, but was beyond the scope of the present investigation.

However, one amplification using outwardly facing primers yielded a very faint product, much shorter than expected (primer positions shown in Fig. 3). Amplification with outward facing *nad4* and *cox1* primers yielded an amplicon of 1 kb. Cloning and sequencing of this fragment yielded a sequence that precisely matched some of the sequence obtained from the *nad4* and *cox1* sequences from the larger fragment, but both genes were incomplete; *nad4* was truncated at the 3' end, while *cox1* was incomplete at the 5' end. A single A residue sat between these two fragments. Such truncated and rejoined genes were observed in the multipart-

**Fig. 1.** Mitochondrial genome organization of *Heterodera glycines*. Genes are abbreviated according to Boore (1999), with the tRNA genes indicated by their corresponding one-letter amino acid code. All genes are encoded on the same strand.



**Table 4.** Members of the Hoplolaimoidea for which some mitochondrial genome organization is available.

Taxon	Reference
Hoplolaimoidea	
Meloidogynidae	
Meloidogyninae	
<i>Meloidogyne javanica</i>	He et al. 2005; Okimoto et al. 1991
Pratylenchidae	
Radopholinae	
<i>Radopholus similis</i>	Jacob et al. 2009
Heteroderidae	
Heteroderinae	
<i>Heterodera glycines</i>	Present study
<i>Heterodera cardiolata</i>	Present study
Punctoderinae	
<i>Punctodera chalconensis</i>	Present study
<i>Globodera pallida</i>	Armstrong et al. 2000; Gibson et al. 2007b
<i>Globodera rostochiensis</i>	Gibson et al. 2007a

tite mitochondrial genomes of *G. rostochiensis* (Gibson et al. 2007a); small circular mitochondrial DNA III from *G. rostochiensis* (Gro-scmtDNA III) contains a partial *nad4* joined directly to a partial *nad1*. In *G. rostochiensis*, the junction of the two partial genes contained 17 nucleotides that could be aligned with both genes from *G. pallida*. These data were consistent with a repeat-mediated recombination event leading to the deletion of a large portion of Gro-scmtDNA III (Gibson et al. 2007a). We propose that the *nad4-cox1* fragment in *P. chalconensis* is the result of a similar event. Amplification of genomic DNA with inward-facing *nad4* and *cox1* primers produced a product of 3.9 kb, consistent with there being low levels of minicircular genomes in *P. chalconensis* (Fig. 3C). However, this minicircle does not contain a non-coding region and may not be able to replicate.

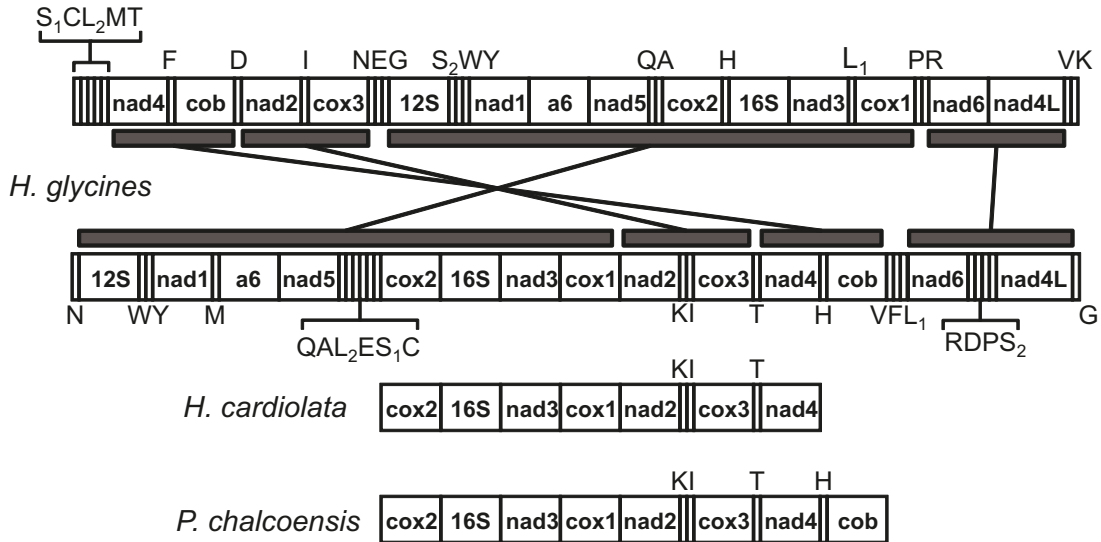
Our data are insufficient to judge whether the mitochondrial genome of *P. chalconensis* is multipartite. The 6.8-kb

contig (Fig. 3A) contains more genes and coding sequence than any of the fully characterized subgenomes of *G. pallida* (Gibson et al. 2007b). If our sequence data from *P. chalconensis* are found on a subgenome, it must be larger than any of those characterized in *G. pallida*. The failure to amplify the remainder of the *P. chalconensis* genome using outward facing primers is further evidence that this genome is larger than the multipartite mitochondrial genomes of *G. pallida* and *G. rostochiensis*.

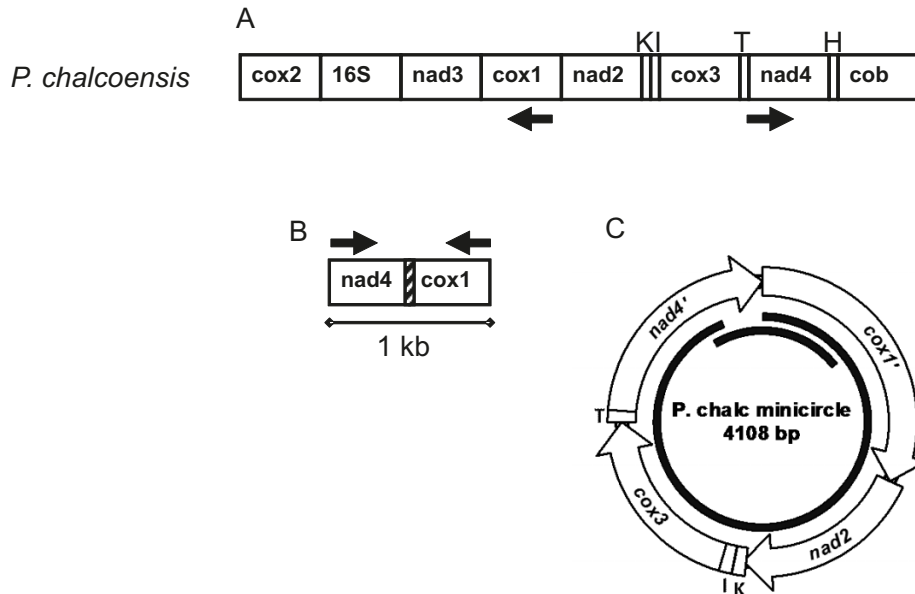
Our data do suggest that *P. chalconensis* contains low levels of mitochondrial subgenomes, as evident from the *nad4-cox1* fragment (Fig. 3C). Such circularised genome fragments may be present at low levels in a range of nematodes (and have been documented in humans; (Kajander et al. 2000)), with multipartite mitochondrial genomes representing a shift to a condition in which these subgenomic circles predominate. In humans, such circularised genome fragments are most preva-

**Fig. 2.** Partial mitochondrial genome organization of *Heterodera cardiolata* and *Punctodera chalcoensis* and a comparison of genome organization with *Heterodera glycines* and *Radopholus similis*. Bars joined by lines indicate regions of conserved genome organization.

*R. similis*



**Fig. 3.** (A) Partial genome organization of *Punctodera chalcoensis*, showing the positions of outward facing primers used in an attempt to amplify the remainder of the genome. (B) The weakly amplified product of the outward amplification. (C) The minicircle that would be produced from the inward and outward amplifications using *nad4* and *cox1* primers.



lent in tissues subject to high oxidative stress (Kajander et al. 2000). Similarly, plant parasitic nematodes are exposed to high levels of oxidative stress, owing to the generation of reactive oxygen species by the plant, as a response to infection (Apel and Hirt 2004; Waetzig et al. 1999). It is tempting to speculate that an increased prevalence of subgenomic circles results from an increased exposure to reactive oxygen species, but there is no evidence that *Globodera* (in which subgenomes are most prevalent) are exposed to higher levels of reactive oxygen species than are other plant parasitic nematodes.

### The evolution of gene rearrangement in the Hoplolaimoidea

We found only minor differences in mitochondrial genome organization between *R. similis*, *H. glycines*, *H. cardiolata*, and *P. chalcoensis*. In stark contrast, there were very few gene boundaries shared when either *Globodera* species was compared with their closest relative examined in this study, *P. chalcoensis*. This indicates that dramatic reorganization of the mitochondrial genome occurred during the divergence of *Punctodera* and *Globodera*. Timing the age of this diver-



gence is difficult, owing to the poor representation of nematodes in the fossil record. However, based on an analysis of both geological and biological information (Picard et al. 2008) the *Punctodera* and *Globodera* last shared a common ancestor between 60–173 million years ago.

Both *G. pallida* and *G. rostochiensis* have a multipartite mitochondrial genome organization (Armstrong et al. 2000; Gibson et al. 2007a), in which the mitochondrial genes are distributed amongst a number of subgenomic circles. Each circle has a substantial noncoding region, which presumably contains the information for replication and transcription. It is perhaps not surprising that gene rearrangement is extensive in *Globodera*, given that multipartite mitochondrial genomes are mosaics of each other and are likely formed by intermitochondrial recombination (Gibson et al. 2007b). Thus, the generation of a multipartite structure involves gene rearrangement. However, it is premature to speculate as to whether *Punctodera* and *Heterodera* also have multipartite mitochondrial genomes. We consider it unlikely that *H. glycines* has a multipartite mitochondrial genome, based on our mapping of all 36 nematode mitochondrial genes to a single fragment. The situation is less clear for *P. chaltoensis*, in which we find evidence of subgenomes with coding sequence, but no noncoding region (and presumably no or reduced capacity for replication). Although it is tempting to speculate that substantial gene rearrangement coincides with the origin of a multipartite structure, our data are currently insufficient to draw this conclusion. A complete characterization of the mitochondrial genome of a representative of the *Punctodera* is the most direct way of assessing this possibility.

## Acknowledgments

Thanks are extended to M.A. Maqbool for providing cysts of *H. cardiolata* and to V. Blok for helpful discussions on *Globodera* biology. This work was supported by grants from the Australian Research Council (DP0556520) and the Australian Academy of Science.

## References

- Apel, K., and Hirt, H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.* **55**(1): 373–399. doi:10.1146/annurev.arplant.55.031903.141701. PMID:15377225.
- Armstrong, M.R., Blok, V.C., and Phillips, M.S. 2000. A multipartite mitochondrial genome in the potato cyst nematode *Globodera pallida*. *Genetics*, **154**(1): 181–192. PMID:10628979.
- Beckenbach, A.T., Robson, S.K.A., and Crozier, R.H. 2005. Single nucleotide +1 frameshifts in an apparently functional mitochondrial cytochrome *b* gene in ants of the genus *Polyrhachis*. *J. Mol. Evol.* **60**(2): 141–152. doi:10.1007/s00239-004-0178-5. PMID:15785844.
- Bert, W., Leliaert, F., Vierstraete, A.R., Vanfleteren, J.R., and Borgonie, G. 2008. Molecular phylogeny of the Tylenchina and evolution of the female gonoduct (Nematoda: Rhabditida). *Mol. Phylogenet. Evol.* **48**(2): 728–744. doi:10.1016/j.ympev.2008.04.011. PMID:18502668.
- Boore, J.L. 1999. Animal mitochondrial genomes. *Nucleic Acids Res.* **27**(8): 1767–1780. doi:10.1093/nar/27.8.1767. PMID:10101183.
- Boore, J.L., and Staton, J.L. 2002. The mitochondrial genome of the sipunculid *Phascolopsis gouldii* supports its association with Annelida rather than Mollusca. *Mol. Biol. Evol.* **19**(2): 127–137. PMID:11801741.
- Boore, J.L., Lavrov, D.V., and Brown, W.M. 1998. Gene translocation links insects and crustaceans. *Nature*, **392**(6677): 667–668. doi:10.1038/33577. PMID:9565028.
- Burger, G., Lavrov, D.V., Forget, L., and Lang, B.F. 2007. Sequencing complete mitochondrial and plastid genomes. *Nat. Protoc.* **2**(3): 603–614. doi:10.1038/nprot.2007.59. PMID:17406621.
- Castro, L.R., Ruberu, K., and Downton, M. 2006. Mitochondrial genomes of *Vanhornia eucnemidarum* (Apocrita: Vahorniidae) and *Primeuchroeus* spp. (Aculeata: Chrysididae): evidence of rearranged mitochondrial genomes within the Apocrita (Insecta: Hymenoptera). *Genome*, **49**(7): 752–766. doi:10.1139/G06-030. PMID:16936784.
- Cornuet, J.-M., Garnery, L., and Solignac, M. 1991. Putative origin and function of the intergenic region between COI and COII of *Apis mellifera* L. mitochondrial DNA. *Genetics*, **128**(2): 393–403. PMID:1649072.
- Downton, M. 1999. Relationships among the cyclostome braconid (Hymenoptera: Braconidae) subfamilies inferred from a mitochondrial tRNA gene rearrangement. *Mol. Phylogenet. Evol.* **11**(2): 283–287. doi:10.1006/mpev.1998.0580. PMID:10191072.
- Downton, M., Castro, L.R., and Austin, A.D. 2002. Mitochondrial gene rearrangements as phylogenetic characters in the invertebrates: the examination of genome ‘morphology’. *Invertebr. Syst.* **16**(3): 345–356. doi:10.1071/IS02003.
- Gibson, T., Blok, V.C., and Downton, M. 2007a. Sequence and characterization of six mitochondrial subgenomes from *Globodera rostochiensis*: multipartite structure is conserved amongst close nematode relatives. *J. Mol. Evol.* **65**(3): 308–315. doi:10.1007/s00239-007-9007-y. PMID:17674076.
- Gibson, T., Blok, V.C., Phillips, M.S., Hong, G., Kumarasinghe, D., Riley, I.T., and Downton, M. 2007b. The mitochondrial subgenomes of the nematode *Globodera pallida* are mosaics: evidence of recombination in an animal mitochondrial genome. *J. Mol. Evol.* **64**(4): 463–471. doi:10.1007/s00239-006-0187-7. PMID:17479345.
- He, Y., Jones, J., Armstrong, M., Lamberti, F., and Moens, M. 2005. The mitochondrial genome of *Xiphinema americanum sensu stricto* (Nematoda: Enoplea): considerable economization in the length and structural features of encoded genes. *J. Mol. Evol.* **61**(6): 819–833. doi:10.1007/s00239-005-0102-7. PMID:16315110.
- Hu, M., Chilton, N.B., and Gasser, R.B. 2002. The mitochondrial genomes of the human hookworms, *Ancylostoma duodenale* and *Necator americanus* (Nematoda: Secernentea). *Int. J. Parasitol.* **32**(2): 145–158. doi:10.1016/S0020-7519(01)00316-2. PMID:11812491.
- Hu, M., Chilton, N.B., Abs El-Osta, Y.G., and Gasser, R.B. 2003a. Comparative analysis of mitochondrial genome data for *Necator americanus* from two endemic regions reveals substantial genetic variation. *Int. J. Parasitol.* **33**(9): 955–963. doi:10.1016/S0020-7519(03)00129-2. PMID:12906879.
- Hu, M., Chilton, N.B., and Gasser, R.B. 2003b. The mitochondrial genome of *Strongyloides stercoralis* (Nematoda) - idiosyncratic gene order and evolutionary implications. *Int. J. Parasitol.* **33**(12): 1393–1408. doi:10.1016/S0020-7519(03)00130-9. PMID:14527522.
- Hu, M., Gasser, R.B., Abs El-Osta, Y.G., and Chilton, N.B. 2003c. Structure and organization of the mitochondrial genome of the canine heartworm, *Dirofilaria immitis*. *Parasitology*, **127**(1): 37–51. doi:10.1017/S0031182003003275. PMID:12885187.
- Jacob, J.E., Vanholme, B., Van Leeuwen, T., and Gheysen, G. 2009. A unique genetic code change in the mitochondrial genome of the parasitic nematode *Radopholus similis*. *BMC Res. Notes*, **2**: 192. doi:10.1186/1756-0500-2-192. PMID:19778425.

- Kajander, O.A., Rovio, A.T., Majamaa, K., Poulton, J., Spelbrink, J. N., Holt, I.J., et al. 2000. Human mtDNA sublineages resemble rearranged mitochondrial genomes found in pathological states. *Hum. Mol. Genet.* **9**(19): 2821–2835. doi:10.1093/hmg/9.19.2821. PMID:11092758.
- Knudsen, B., Kohn, A.B., Nahir, B., McFadden, C.S., and Moroz, L. L. 2006. Complete DNA sequence of the mitochondrial genome of the sea-slug, *Aplysia californica*: conservation of the gene order in Euthyneura. *Mol. Phylogenet. Evol.* **38**(2): 459–469. doi:10.1016/j.ympev.2005.08.017. PMID:16230032.
- Kurabayashi, A., and Ueshima, R. 2000. Complete sequence of the mitochondrial DNA of the primitive opisthobranch gastropod *Pupa strigosa*: systematic implication of the genome organization. *Mol. Biol. Evol.* **17**(2): 266–277. PMID:10677849.
- Lavrov, D.V., and Brown, W.M. 2001. *Trichinella spiralis* mt DNA: a nematode mitochondrial genome that encodes a putative ATP8 and normally structured tRNAs and has a gene arrangement relatable to those of coelomate metazoans. *Genetics*, **157**(2): 621–637. PMID:11156984.
- Lowe, T.M., and Eddy, S.R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**(5): 955–964. doi:10.1093/nar/25.5.955. PMID:9023104.
- Ma, H., Overstreet, R.M., and Subbotin, S.A. 2008. ITS2 secondary structure and phylogeny of cyst-forming nematodes of the genus *Heterodera* (Tylenchida: Heteroderidae). *Org. Divers. Evol.* **8**(3): 182–193. doi:10.1016/j.ode.2007.09.001.
- Mindell, D.P., Sorenson, M.D., and Dimcheff, D.E. 1998. An extra nucleotide is not translated in mitochondrial ND3 of some birds and turtles. *Mol. Biol. Evol.* **15**(11): 1568–1571. PMID:12572620.
- Niblack, T.L., Arelli, P.R., Noel, G.R., Opperman, C.H., Orf, J.H., Schmitt, D.P., et al. 2002. A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *J. Nematol.* **34**(4): 279–288. PMID:19265945.
- Okimoto, R., Chamberlin, H.M., Macfarlane, J.L., and Wolstenholme, D.R. 1991. Repeated sequence sets in mitochondrial DNA molecules of root knot nematodes (*Meloidogyne*): nucleotide sequences, genome location and potential for host-race identification. *Nucleic Acids Res.* **19**(7): 1619–1626. doi:10.1093/nar/19.7.1619. PMID:2027769.
- Okimoto, R., MacFarlane, J.L., Clary, D.O., and Wolstenholme, D.R. 1992. The mitochondrial genome of two nematodes, *Caenorhabditis elegans* and *Ascaris suum*. *Genetics*, **130**(3): 471–498. PMID:1551572.
- Picard, D., Sempere, T., and Plantard, O. 2008. Direction and timing of uplift propagation in the Peruvian Andes deduced from molecular phylogenetics of highland biotaxa. *Earth Planet. Sci. Lett.* **271**(1–4): 326–336. doi:10.1016/j.epsl.2008.04.024.
- Radice, A.D., Powers, T.O., Sandall, L.J., and Riggs, R.D. 1988. Comparisons of mitochondrial DNA from the sibling species *Heterodera glycines* and *H. schachtii*. *J. Nematol.* **20**(3): 443–450. PMID:19290236.
- Riepsamen, A.H., Blok, V.C., Phillips, M., Gibson, T., and Downton, M. 2008. Poly(T) variation within mitochondrial protein-coding genes in *Globodera* (Nematoda: Heteroderidae). *J. Mol. Evol.* **66**(3): 197–209. doi:10.1007/s00239-007-9064-2. PMID:18288437.
- Riepsamen, A.H., Gibson, T., Rowe, J., Chitwood, D.J., Subbotin, S. A., and Downton, M. 2011. Poly(T) variation in heteroderid nematode mitochondrial genomes is predominantly an artefact of amplification. *J. Mol. Evol.* **72**(2): 182–192. doi:10.1007/s00239-010-9414-3. PMID:21161202.
- Riggs, R.D., and Schmitt, D.P. 1988. Complete characterization of the race scheme for *Heterodera glycines*. *J. Nematol.* **20**(3): 392–395. PMID:19290228.
- Rokas, A., and Holland, P.W.H. 2000. Rare genomic changes as a tool for phylogenetics. *Trends Ecol. Evol.* **15**(11): 454–459. doi:10.1016/S0169-5347(00)01967-4. PMID:11050348.
- Smith, M.J., Arndt, A., Gorski, S., and Fajber, E. 1993. The phylogeny of echinoderm classes based on mitochondrial gene arrangements. *J. Mol. Evol.* **36**(6): 545–554. doi:10.1007/BF00556359. PMID:8350349.
- Subbotin, S.A., Peng, D., and Moens, M. 2001a. A rapid method for the identification of the soybean cyst nematode *Heterodera glycines* using duplex PCR. *Nematology*, **3**(4): 365–371. doi:10.1163/156854101317020286.
- Subbotin, S.A., Vierstraete, A., De Ley, P., Rowe, J., Waeyenberge, L., Moens, M., and Vanfleteren, J.R. 2001b. Phylogenetic relationships within the cyst-forming nematodes (Nematoda, Heteroderidae) based on analysis of sequences from the ITS regions of ribosomal DNA. *Mol. Phylogenet. Evol.* **21**(1): 1–16. doi:10.1006/mpev.2001.0998. PMID:11603932.
- Subbotin, S.A., Sturhan, D., Chizhov, V.N., Vovlas, N., and Baldwin, J.G. 2006. Phylogenetic analysis of Tylenchida Thorne, 1949 as inferred from D2 and D3 expansion fragments of the 28S rRNA gene sequences. *Nematology*, **8**(3): 455–474. doi:10.1163/156854106778493420.
- Sunnucks, P., and Hales, D.F. 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Mol. Biol. Evol.* **13**(3): 510–524. PMID:8742640.
- Waetzig, G.H., Sobczak, M., and Grundler, F.M.W. 1999. Localization of hydrogen peroxide during the defence response of *Arabidopsis thaliana* against the plant-parasitic nematode *Heterodera glycines*. *Nematology*, **1**(7): 681–686. doi:10.1163/156854199508702.
- Wolstenholme, D.R., Macfarlane, J.L., Okimoto, R., Clary, D.O., and Wahleithner, J.A. 1987. Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proc. Natl. Acad. Sci. U.S.A.* **84**(5): 1324–1328. doi:10.1073/pnas.84.5.1324. PMID:3469671.