Repeated sequence sets in mitochondrial DNA molecules of root knot nematodes (*Meloidogyne*): nucleotide sequences, genome location and potential for host-race identification

Ronald Okimoto, Helen M. Chamberlin, Jane L. Macfarlane and David R. Wolstenholme*
Department of Biology, University of Utah, Salt Lake City, UT 84112, USA

Received November 16, 1990; Revised and Accepted February 22, 1991

EMBL accession nos X57625, X57626

ABSTRACT

Within a 7 kb segment of the mtDNA molecule of the root knot nematode, *Meloidogyne javanica*, that lacks standard mitochondrial genes, are three sets of strictly tandemly arranged, direct repeat sequences: approximately 36 copies of a 102 ntp sequence that contains a *TaqI* site; 11 copies of a 63 ntp sequence, and 5 copies of an 8 ntp sequence. The 7 kb repeat-containing segment is bounded by putative tRNA^{asp} and tRNA^{f-met} genes and the arrangement of sequences within this segment is: the tRNA^{asp} gene; a unique 1,528 ntp segment that contains two highly stable hairpin-forming sequences; the 102 ntp repeat set; the 8 ntp repeat set; a unique 1,068 ntp segment; the 63 ntp repeat set; and the tRNA^{f-met} gene. The nucleotide sequences of the 102 ntp copies and the 63 ntp copies have been conserved among the species examined. Data from Southern hybridization experiments indicate that the 102 ntp and 63 ntp repeats occur in the mtDNAs of three, two and two races of *M. incognita*, *M. hapla* and *M. arenaria*, respectively. Nucleotide sequences of the *M. incognita* Race-3 102 ntp repeat were found to be either identical or highly similar to those of the *M. javanica* 102 ntp repeat. Differences in migration distance and number of 102 ntp repeat-containing bands seen in Southern hybridization autoradiographs of restriction-digested mtDNAs of *M. javanica* and the different host races of *M. incognita*, *M. hapla* and *M. arenaria* are sufficient to distinguish the different host races of each species.

INTRODUCTION

Metazoan mitochondrial (mt-) genomes are, with rare exception, single circular DNA molecules that contain the same set of genes for 2 rRNAs, 22 tRNAs, and 12 or 13 proteins all concerned with oxidative phosphorylation (refs in 1,2). None of the metazoan mt-genes contain introns and there are very few or no nucleotides between genes. However, all metazoan mtDNA

molecules contain an apparently non-coding region that varies in size between species from 121 ntp to approximately 20 kb (3–5). As this region has been found in vertebrate and *Drosophila* mtDNAs to contain the molecule's origin of replication (6), and in mammalian mtDNAs to contain transcription promoter sequences (7), it has been designated the control region.

Repeated segments have been found in the mtDNA molecules of a number of metazoan species. Tandemly arranged, repeated sequences occur in the control regions of the mtDNA molecules of some *Drosophila* species (470 ntp, 1 to 5 copies; 8), a cricket, *Gryllus firmus* (220 ntp, 1 to 7 copies; 9), three weevils, *Pissodes* species (800–2,000 ntp, various numbers of copies, 5), lizards of the genus *Cnemidophorus* (64 ntp, 3 to 9 copies; 10) and two fishes, *Alosa sapidissima* (1,500 kb, 1 to 3 copies; 11); *Acipenser transmontanus* (82 ntp, 1 to 4 copies; 12). Two copies of a non-tandemly arranged, direct repeat occur in the control region of the mtDNA molecule of *Xenopus laevis* (13). MtDNA from the scallop, *Placopecten magellanicus* contains between 2 and 8 copies of a 1,442 ntp direct repeat, but the location within the molecule is not known (14). For each of the above, except *X. laevis*, repeat copy number variation occurs between individuals of a species and in some cases, within individuals (heteroplasmy).

MtDNAs that include duplicated segments of heterogeneous lengths comprising various portions of the control region and adjacent rRNA, tRNA and protein genes have been isolated from individuals of different *Cnemidophorus* species (15,16). A single duplication of a sequence containing the large rRNA, small rRNA, ND1 and ND2 genes has been reported in mtDNA from newts (*Triturus cristatus*; 17). Also, a segment of mtDNA that may contain coding sequences is both directly and inversely repeated in mtDNA of *Romanomermis culicivorax*, a parasitic nematode (18).

Of the above mentioned mtDNA repeats, nucleotide sequence information has been obtained only for those of *G. firmus*, *A. transmontanus*, *P. magellanicus* and *X. laevis* (9,12–14).

In this paper we report the finding of three sets of repeat sequences of 102 ntp, 63 ntp and 8 ntp, in mtDNA molecules

* To whom correspondence should be addressed

of the plant parasitic root knot nematode, *Meloidogyne javanica*. The 102 ntp and 63 ntp repeats have been analyzed in regard to their nucleotide sequences, copy number, genome location, and occurrence and variation in different *Meloidogyne* species and host races.

MATERIAL AND METHODS

Origins of species and host races

Eggs of *Meloidogyne javanica* (NCSU #7-2), *M. incognita*-Race 1 (NCSU #68), -Race 3 (NCSU #E1135) and -Race 4 (NCSU #401), *M. arenaria*-Race 1 (NCSU #352) and -Race 2 (NCSU #480), and *M. hapla*-Race A (NCSU #86) and -Race B (NCSU #48), produced by worms grown on eggplant (*Solanum melongena*; 19), were obtained from Michael A. McClure, Department of Plant Pathology, University of Arizona, Tucson, Arizona.

DNA preparation

MtDNA was isolated from eggs of *M. javanica*, *M. incognita*-Race 3 and *M. hapla*-Race A, as follows. Between 2 and 3 ml packed volume of eggs were suspended on ice in 10–12 ml of 0.2 M mannitol, 0.07 M sucrose, 0.05 M Tris-HCl (pH 7.5), 0.01 M EDTA and 200 µg/ml proteinase K (20), and broken using a 15 ml (pestle A) Dounce homogenizer. From a mitochondrial pellet, obtained by differential sedimentation, DNA was purified by phenol and chloroform extraction, and ethanol precipitation (21). Covalently-closed circular mtDNA molecules were isolated using CsCl-ethidium bromide centrifugation (21,22).

Total cell DNA was extracted from 50–100 µl packed volume of eggs of *M. incognita*-Races 1 and 4, *M. arenaria*-Races 1 and 2, and *M. hapla*-Races A and B, using proteinase K digestion, SDS lysis, phenol and chloroform extraction and ethanol precipitation (21).

Restriction enzyme digestions and cloning

Conditions used for restriction enzyme digestions were those recommended by the manufacturers. When partial *TaqI* digestion of mtDNA was required, 0.2–0.5 µg DNA in a 25 µl reaction mixture containing 1 U enzyme was incubated at 37°C (rather than 65°C) for 30 min, cooled on ice or frozen. Restriction fragments of *Meloidogyne* mtDNAs were cloned into pUC9 or pUC12, or bacteriophages M13mp18 or M13mp19 and amplified in *Escherichia coli* strains JM101 and DH5αF' (Bethesda Research Laboratories (BRL)). Other details regarding electrophoresis, cloning and purification of single-stranded M13 DNAs are given or referred to in ref. 23.

DNA probe labeling

Whole *M. incognita* mtDNA was ³²P-labeled by nick translation (21). ³²P-labeled probes were made from mtDNA-containing M13 clones by extension synthesis using the Klenow fragment of *E. coli* DNA polymerase I and [α-³²P] dATP. A synthetic oligonucleotide sequence (the complement of nt 6,441–6,470, Fig. 3, synthesized using an Applied Biosystems Synthesizer 380B) was end-labeled using [γ-³²P]ATP, and T4 polynucleotide kinase (24). A 123 ntp ladder (BRL) was ³²P-end labeled using T4 polymerase (BRL).

Nucleic acid hybridizations

Capillary transfer of DNAs from agarose gels to the hybridization support, Gene Screen Plus, was as given in New England Nuclear

Catalog No. NEF0976 (Jan. 1984). Aqueous DNA:DNA blot hybridizations were carried out as in ref. 25 except that prehybridization and hybridization were both at 55°C when mtDNA-containing M13 probes were used, and at 50°C and 45°C, respectively, when the oligonucleotide probe was used.

Sequencing

DNA sequences were obtained (26: but using [α-P-³⁵S]dATP) from sets of deletion clones (27). These clones contained overlapping sequences representing the entire sequences of both complementary strands of the DNA segment shown in Fig. 3, except for the region containing the 102 ntp repeat (see below). Other details concerning sequencing, and computer assembly and analysis of sequences are given in ref. 23.

RESULTS

Data from restriction analyses of mtDNAs isolated from eggs of *Meloidogyne javanica* and *M. incognita*-Race 3 indicated that the mt-genomes of each of these organisms is a single molecule of 20.5 kb and 19.5 kb, respectively. The approximately 1 kb difference in size between these molecules is mainly due to differences within a single *XbaI* fragment: 7.94 kb in *M. javanica* and 6.88 kb in *M. incognita*-Race 3. Using electron microscopy (22) it was shown that the *M. incognita*-Race 3 mtDNA molecule is circular with a contour length approximately equal to that estimated from restriction analysis. Circularity of the *M. javanica* mtDNA molecule was confirmed by sequencing (see Fig. 3).

A *TaqI* site-containing 102 ntp repeated sequence in *M. javanica* and *M. incognita* mtDNA molecules

We inadvertently exposed *M. javanica* mtDNA to the *TaqI* enzyme at 37°C (rather than the optimum 65°C). Examination of an autoradiograph of a blot of the electrophoresed *TaqI* partial digestion product after it had been hybridized with whole, ³²P-labeled *M. javanica* mtDNA revealed a ladder of bands (Fig. 1A and B). Given the non-optimum temperature of the *TaqI* digestion, this band pattern suggested that *M. javanica* mtDNA includes at least 28 copies of a tandemly arranged, directly repeated, 100 ntp sequence that contains a *TaqI* site. A similar result was obtained for *M. incognita*-Race 3 mtDNA digested with *TaqI* at 37°C, (Fig. 1A and B). For both species, the control experiment in which mtDNA was digested with *TaqI* at 65°C revealed 10 corresponding bands, a band (1.77 kb) unique to *M. javanica* mtDNA, and a band (0.88 kb) unique to *M. incognita*-Race 3 mtDNA (Fig. 1A and B). The migration distances and stoichiometry of the 10 corresponding bands in *M. javanica* and *M. incognita*-Race 3 mtDNAs were approximately those expected for *TaqI* fragments greater in size than 200 ntp, as later determined from the nucleotide sequence of the *M. javanica* mtDNA molecule. However, bands containing fragments less than 200 ntp that would include the postulated 100 ntp repeat and 10 other small (18 ntp–180 ntp) *TaqI* fragments were not visible, possibly due to insufficient transfer of small fragments in this experiment. Both the *M. javanica* 1.77 kb band and the *M. incognita*-Race 3, 0.88 kb band are sub-stoichiometric, suggesting that our preparations of *M. javanica* and *M. incognita*-Race 3 mtDNAs each include a minor population of a sequence variant. The fragments contained in each of these bands could have resulted either from a deletion or insertion, or from the presence of an extra *TaqI* site in the minor, relative to the major population of mtDNA molecules of the respective species. If the

latter were the case, then the second extra band expected may be too light to discern, or may be of a size that is poorly transferred, as discussed above.

Fragments resulting from partial *TaqI* digestion (37°C) that collectively included all of the *M.javanica* 7.94 kb *XbaI* fragment, and all of the *M.incognita* Race-3 6.88 kb *XbaI* fragment (and the remainder of each of these mtDNAs) were cloned using an M13mp19 and *E.coli* DH5 α F' (RecA⁻) vector-host combination. The ends of some of these *M.javanica* and *M.incognita*-Race 3 mtDNA cloned inserts were sequenced and found to contain between one and ten copies of a 102 ntp *TaqI* site-containing sequence. (The sequences of up to four repeats could be read, but the presence of more repeats in some gels could be inferred from a repeated banding pattern in the upper part of the gel.) The 102 ntp sequences in these *M.javanica* and *M.incognita*-Race 3 mtDNA clones were identical both within and between species (with the exceptions discussed below) and were tandemly arranged without intervening nucleotides.

Purified mtDNAs from *M.javanica* and *M.incognita*-Race 3 were cleaved with *XbaI* and, separately, with *TaqI* at 65°C, electrophoresed and blot-transferred. To the blot was hybridized a ³²P-labeled M13 probe (1×102R) containing a single copy of the 102 ntp repeat sequence. In the lanes containing *XbaI* digested *M.javanica* and *M.incognita*-Race 3 mtDNA this probe hybridized strongly to the 7.94 kb and the 6.88 kb fragments,

respectively (Fig. 1C). Also, in each of these lanes the probe hybridized weakly to a fragment of unexpected size (5.2 kb, *M.javanica* and 4.5 kb, *M.incognita*-Race 3; Fig. 1C) again suggesting that each mtDNA contains a low frequency sequence variant. These data indicate that the 102 ntp, *TaqI* site-containing repeat is limited to the 7.94 kb and 6.88 kb *XbaI* mtDNA fragments of *M.javanica* and *M.incognita*-Race 3 mtDNAs, respectively. In the lanes that contained *M.javanica* mtDNA and *M.incognita*-Race 3 mtDNA completely cleaved with *TaqI* (65°C), only a single band was observed at the approximate position expected for a 102 ntp fragment (Fig. 1C). This finding further supports the view that the 102 ntp, *TaqI* site-containing sequence occurs in both *M.javanica* and *M.incognita*-Race 3 mtDNAs mainly, if not exclusively, in a directly repeated arrangement.

Organization of the *M.javanica* mtDNA molecule

Using M13 clones (grown in *E.coli* DH5 α F') containing *TaqI* and *Sau3A* fragments of the 7.94 kb *XbaI* fragment of *M.javanica* mtDNA, and M13 clones (grown in *E.coli* JM101) containing various restriction fragments that collectively contained the remainder of the *M.javanica* mtDNA molecule, we obtained the nucleotide sequence of the *M.javanica* mtDNA molecule (R. Okimoto, J.L.Macfarlane, H.M.Chamberlin and D.R.Wolstenholme, in preparation). This molecule (Fig. 2)

Figure 1. Autoradiographs of Southern hybridization experiments (1% agarose gels) that provide evidence for the presence of a 102 ntp, *TaqI* restriction site-containing, directly repeated sequence in mtDNA molecules of *Meloidogyne javanica*, and *M.incognita*-Race 3. The lanes in panels A, B, and C contain the following: Pj, and Pi, *M.javanica* and *M.incognita* mtDNAs, respectively, partially digested (37°C) with *TaqI*; Cj and Ci, *M.javanica* and *M.incognita* mtDNAs, respectively, digested to completion (65°C) with *TaqI*; Xj and Xi, *M.javanica* and *M.incognita* mtDNAs, respectively, digested to completion with *XbaI*; 123, ³²P-labeled 123 ntp ladder; L, ³²P-end labeled *HindIII* digestion products of bacteriophage lambda: 23.1 kb, 9.4 kb, 6.7 kb, 4.6 kb, 2.3 kb, 2.0 kb, 0.56 kb. Panel A is a shorter exposure of the two left most lanes in Panel B. Panel B was probed with whole *M.javanica* mtDNA (mt), ³²P-labeled by nick translation. Panel C was probed with an M13 clone (1×102R; ³²P-labeled by a synthesis reaction) containing a single copy of the 102 ntp *TaqI* site-containing mtDNA repeat. The arrowhead indicates the 102 ntp monomers. Dots indicate the band unique to *M.javanica* mtDNA, and the band unique to *M.incognita* mtDNA.

Figure 2. Gene map of the circular *Meloidogyne javanica* mtDNA molecule. The identities and arrangements of the various features shown were determined from nucleotide sequence studies. ORF indicates an unidentified open reading frame. Designation of the 5' end of the large rRNA (dashed line) is tentative. The locations of the tRNA^{asp} (D), tRNA^{f-met} (M) and tRNA^{his} (H) genes (Figs. 3 and 4) are shown. Hatched regions indicate other sequences tentatively interpreted as containing tRNA genes. The direction of transcription of all identified genes is indicated. Blacked-in areas indicate the locations of tandemly arranged, directly repeated sequences: 102R, approximately 37 copies of a 102 ntp sequence; 8R, 5 copies of an 8 ntp sequence; 63R, 11 copies of a 63 ntp sequence. The continuous sequence of the segment between the dotted lines in the 102R region has not been determined (Fig. 3). SL101 and SL37 identify potential, highly stable, stem and loop forming sequences. A partial restriction map of the *M.javanica* mtDNA molecule is shown inside the gene map (A, *Ara3*; E, *EcoRI*; H, *HindIII*; Hg, *HgiA1*; Mb, *MboI*; R, *Rsa1*; X, *XbaI*). The arc on the outside of the map (*M.i*) identifies a sequenced region of the *M. incognita*-Race 3 mtDNA molecule.

contains genes homologous to the two rRNA genes and 12 protein genes found in all other metazoan mtDNAs sequenced to date. Between the ends of some protein and rRNA genes are sequences that we have tentatively interpreted as tRNA genes of the sort found in *Caenorhabditis elegans* and *Ascaris suum* mtDNAs (24): that is, structures in which the T ψ C arm and variable loop are together replaced with a simple loop of nucleotides (Figs. 3 and 4). All of the protein, rRNA, and tRNA genes so far identified within the *M.javanica* mtDNA molecule are transcribed in the same direction and are contained within a sequence of 13,565 ntp bounded by a putative tRNA^{asp} gene and a putative tRNA^{f-met} gene (Fig. 2). Within the remaining, approximately 7 kb segment of the molecule is the 102 ntp repeat set, and sets of directly repeated sequences of 63 ntp and 8 ntp (Figs. 2 and 3). The lack of genes, the presence of repeats, two highly stable stem and loop structures (Figs. 2 and 3), and bracketing tRNA genes are consistent with the interpretation that the 7 kb segment is the control region of the molecule.

The 102 ntp repeat sequence set

The 102 ntp repeat set is separated from the tRNA^{asp} gene by 1,528 ntp and from the tRNA^{f-met} gene by 1,758 ntp (Fig. 3). Due to the high copy number of this repeat and to almost perfect sequence conservation among the copies (see below), we were unable to sequence through the repeat set. However, sequences were obtained of the 3.5 copies at the tRNA^{asp} gene-proximal end of the repeat set and of 3.3 copies at the tRNA^{f-met} gene-proximal end of the repeat set (Fig. 3). The series of tRNA^{asp}

Table 1. Variation in nucleotide sequence among 102 ntp repeats of *M.javanica* and *M.incognita*-Race 3 mtDNAs.

Repeat	Species	Nucleotide Number ^b						
		10	20	61	62	81	82	88
102R-L1 ^a (partial:50 nt)	<i>M. javanica</i>	—	—	T	C	T	Ⓞ	A
	<i>M. incognita</i>	—	—	T	C	C	A	Ⓞ
102R-L2	<i>M. javanica</i>	G	G	A	C	C	A	A
	<i>M. incognita</i>	G	G	A	C	T	A	A
102R-L3	<i>M. javanica</i>	G	G	A	C	C	A	A
	<i>M. incognita</i>	G	Ⓞ	A	C	C	A	A
102R-L4	<i>M. javanica</i>	Ⓞ	G	A	C	C	A	A
	<i>M. incognita</i>	G	G	A	C	C	A	A
Standard (Internal) 102R ^c		G	G	A	C	C	A	A
Variant <i>M. incognita</i> 102R ^d		G	G	A	C	C	A	Ⓞ
102R-R1	<i>M. javanica</i>	G	G	A	T	C	A	A
102R-R2	<i>M. javanica</i>	G	G	A	T	C	A	A
102R-R3	<i>M. javanica</i>	G	G	A	T	C	A	A
102R-R4 (partial:32 nt)	<i>M. javanica</i>	G	G	—	—	—	—	—

a. Designation of repeats is that given in Fig. 3

b. Nucleotide numbers in each 102 ntp repeat correspond to those of repeat 102R-L2 of *M.javanica* mtDNA (Fig. 3), beginning with the first nucleotide of the *TaqI* site (TCGA; nt 1,729, Fig. 3) of that repeat.

c. The internal standard is the nucleotide sequence that was found to be identical for 21 *M.javanica* and 9 out of 10 *M.incognita* 102 ntp repeats in randomly selected clones that resulted from partial *TaqI* digests of mtDNA of the respective species.

d. The only *M.incognita* 102 ntp repeat from randomly selected clones of partial *TaqI* mtDNA digests that contained a sequence variant.

gene-proximal 102 ntp repeats (102R-L1 to 102R-L4) begins 50 ntp before the first *TaqI* site contained in these repeats (102R-L1, Fig. 3), but the series of tRNA^{f-met}-gene proximal repeats (102R-R1 to 102R-R4, Fig. 3) ends at a different location within a 102 ntp repeat: 28 ntp after the last 102 ntp repeat-containing *TaqI* site (ntp 5,330, Fig. 3). Therefore, it is not possible from the sequence information to define the boundaries of the original monomeric 102 ntp repeated sequence. For discussion, the monomeric unit is defined as beginning with the first nucleotide of the *TaqI* site and ending with the nucleotide that precedes the next *TaqI* site.

In addition to the tRNA^{asp} gene-proximal and tRNA^{f-met} gene-proximal 102 ntp repeats, we have sequenced a total of 21 copies of the 102 ntp repeat contained in 11 clones (between one and four copies per clone) derived from partial *TaqI* digestion (37°C) of *M.javanica* mtDNA. As the sequences of all 21 of these repeats (presumed to be mainly from the internal region of the 102 ntp repeat set) are identical, this common sequence is referred to as the standard sequence (Table 1). It is noted that although the repeat-containing clones were chosen at random, we do not know the extent to which a single repeat from the 102 ntp set might have multiple representation in this collection of 21 sequences. Minor variations in nucleotide sequences among the *M.javanica* tRNA^{asp} gene-proximal and tRNA^{f-met} gene-proximal 102 ntp repeats, relative to the standard (internal) repeat, are summarized in Table 1.

We have made use of the *MboI* (Sau3A) sites that bracket the set of 102 ntp repeats (Figs. 2 and 3) to estimate the total number of 102 ntp repeats. *M.javanica* mtDNA was digested with *MboI*,

Figure 3. The nucleotide sequences of 2,034 ntp and 2,280 ntp segments of the *Meloidogyne javanica* mtDNA molecule that lie between the 3' end of the ND4 gene and the 5' end of the ATPase6 gene. The predicted amino acid sequences of the ND4 and ATPase6 gene segments are shown, and three asterisks indicate the termination codon of the ND4 gene. Sequences of three putative tRNA genes are identified by dotted overlines within which brackets identify the anticodons. The nucleotide sequence shown is the (5'-3') sense strand for the 5 genes, all of which are transcribed (arrows) in the same direction. Regions containing the 102 ntp, *TaqI* site (underlined)-containing direct repeats, the 8 ntp direct repeats, and the 63 ntp direct repeats are identified by broken overlines. The 102 ntp repeats are marked off as sequences between *TaqI* sites (but see text), beginning at the ND4 gene-proximal end of the 102 ntp sequence-containing region and designated 102R-L1 to 102R-L4 (left 1-4) and 102R-R1 to 102R-R4 (right 1-4). The 63 ntp and 8 ntp repeats are also marked off, and labeled 63R-1 to 63R-11 and 8R-1 to 8R-5, respectively. Nucleotides within individual repeats of each kind that are variant relative to the majority are indicated by inverted arrowheads below the sequence. The d below the 8R-2 sequence indicates a deleted T nucleotide relative to the remaining 8R repeats. Arrow pairs beneath the sequence identify inverted repeat sequences. Two large stem and loop structures (SL101 and SL37) are also identified by continuous overlines. The *MboI* sites used to estimate the copy number of the 102 ntp repeat are shown. The sequence (ntp 6,441-6,470) to which a complementary synthetic oligonucleotide was made is indicated by a wavy underline. All of the sequences in this figure, except ntp 1,762-2,034 and ntp 4,993-5,298, were determined by sequencing overlapping fragments of both complementary strands.

electrophoresed, blotted and hybridized with a ^{32}P -labelled M13 probe ($2 \times 102\text{R}$) containing two copies of the 102 ntp repeat. In the resulting autoradiograph (Fig. 5A) one major band was observed (plus a minor band that might result from a sequence variant as discussed above) at a position expected for a 5,200 ntp fragment. The *MboI* sites that must have been cleaved to produce this fragment lie 967 ntp and 516 ntp from the tRNA^{asp} gene-proximal and tRNA^{f-met} gene-proximal ends of the 102 ntp repeat set, respectively (Fig. 3). This indicates that the total length of the 102 ntp repeat set is approximately 3,717 ntp (5,200 - [967 + 516]), equivalent to 36.4 copies of the 102 ntp repeat.

The 8 ntp and 63 ntp repeat sequence sets

Beginning 5 ntp from the tRNA^{f-met} gene-proximal end of the 102 ntp repeat set, are five copies of an 8 ntp directly repeated sequence, again tandemly arranged (8R-1 to 8R-5, Fig. 3). In the second copy (8R-2), the fourth nucleotide is deleted, and the fifth copy (8R-5) contains a single nucleotide substitution.

Beginning 1,069 ntp from the tRNA^{f-met} gene-proximal end of the 102 ntp repeat set are 11 copies of a 63 ntp sequence (63R-1 to 63R-11, Fig. 3). These copies are in a perfect tandem arrangement. The first ten copies are identical, except for a single

substitution in 63R-9. The eleventh copy (63R-11) contains the same substitution as is found in 63R-9, overlaps the predicted tRNA^{f-met} gene by 3 ntp, and ends in a T rather than an A.

Open reading frames in the 102 ntp and 63 ntp repeat sets

All ATN codons and TTG appear to be used as translation initiation codons among metazoan mt-protein genes (1). In the direction of gene transcription (Figs. 2 and 3), there are two open reading frames (ORF) in the *M.javanica* 102 ntp repeat set. One of these ORFs, that could begin with either an ATT or an ATA, would traverse the entire repeat set, assuming that only standard repeats (Table 1) occur in the unsequenced portion of this set. The second ORF, that could begin with either ATT or TTG,

Figure 4. The three putative tRNA genes found at the boundaries of the protein and rRNA gene-containing segment and the repeat sequence-containing segment of the *M.javanica* mtDNA molecule. Each gene is shown in the presumed secondary structural form of the corresponding tRNA in which the TyC arm and variable loop are replaced with a loop of between 4 and 6 nt.

runs through the tRNA^{asp} gene-proximal copies of the 102 ntp repeat region (and the standard (internal) repeats), but the C → T substitution in each of the tRNA^{f-met} gene-proximal 102 ntp repeats creates a stop codon in this reading frame. A single ORF, that could begin with any ATN codon or TTG, traverses the reverse complement of the sequence containing the 102 ntp repeat set.

Of the six possible reading frames that traverse the 63 ntp repeat set, only one (that could begin with ATC) is open, in the direction of transcription.

The largest ORFs within the 1,528 ntp sequence between the tRNA^{asp} gene and the 102 ntp repeat set, the 1,068 ntp sequence that separates the 102 ntp and 63 ntp repeat sets, and the complements to these sequences, that could begin with ATG, ATT or ATC, range from 53 to 93 codons.

We have been unable to find an identity for any of the amino acid sequences predicted from the above mentioned ORFs.

Secondary structures in the repeated sequences and flanking sequences

The 102 ntp repeat includes two regions with the potential to form stem and loop structures. These are indicated in 102R-L2 (Fig. 3): nucleotides 1,746–1,766 can fold into a stem of 5 ntp and a loop of 11 nt; nucleotides 1,817–1,828 can fold into a stem of 4 ntp (that includes two GC pairs) and a loop of 4 nt. Also, the *TaqI* site is within a 6 ntp sequence of dyadsymmetry (5' TTCGAA; an *NspV* site).

Near the 5' end of the 63 ntp repeat is a sequence (nucleotides 6,402–6,434, Fig. 3) that can fold into a stem of 9 ntp (including one G·T pair) and a loop of 15 nt.

Within the 1,528 ntp segment between the tRNA^{asp} gene and the 102 ntp repeat set are two sequences of 101 ntp and 37 ntp that can be folded into highly stable stem and loop structures (Figs. 2 and 3, SL101 and SL37).

Figure 5. Autoradiographs resulting from Southern hybridization experiments to determine the distribution of the 102 ntp and 63 ntp directly repeated sequences in mtDNAs of different *Meloidogyne* species and races. Lanes L contain ³²P-labeled *Hind*III digestion products of bacteriophage lambda; see Fig. 1. All other lanes contain *Mbo*I digestion products of the following: j, i3 and hA, mtDNAs of *M.javanica*, *M.incognita*-Race 3 and *M.hapla*-Race A, respectively; i1, i4, hB, a1 and a2, whole cell DNAs of *M.incognita*-Race 1, *M.incognita*-Race 4, *M.hapla*-Race B, *M.arenaria*-Race 1 and *M.arenaria*-Race 2, respectively. Panel A was probed with an M13 clone (2×102R; ³²P-labeled by a synthesis reaction) containing two copies of the 102 ntp *TaqI* site-containing repeat. Panel B was probed with a ³²P-end labelled 30 nt oligomer (63R) complementary to a sequence (nt 6,441–6,470, Fig. 3) within the *M.javanica* 63 ntp repeat region.

Sequences that include 102 ntp repeats in *M.incognita*-Race 3 mtDNA

We sequenced a total of ten 102 ntp repeats, contained in five clones (between one and three copies per clone) derived from a partial *TaqI* digestion (37°C) of *M.incognita*-Race 3 mtDNA. Of these, nine were identical to the standard, *M.javanica* 102 ntp repeat and one contained a single nucleotide deletion (Table 1). We also sequenced a 629 ntp segment of the *M.incognita*-Race 3 mtDNA molecule that is homologous to nucleotides 1,410–2,038 of the *M.javanica* sequence (Figs. 2 and 3) and includes the tRNA^{asp} gene-proximal 3.5, 102 ntp repeats and the immediately adjacent 269/270 ntp unique sequence. Minor variations in these 3.5, 102 ntp repeats, are shown in Table 1. The corresponding 269/270 ntp *M.javanica* and *M.incognita*-Race 3 mtDNA sequences that are continuous with the 102 ntp repeat region differed by only a single nucleotide substitution and a single insertion/deletion, (0.74% divergence).

Repeats in mtDNAs of different species and host races of *Meloidogyne*

The ³²P-labelled, *M.javanica* 2×102R probe was hybridized to blots of electrophoresed, *MboI* digests of mtDNA of *M.hapla*-Race A, and whole worm DNA of *M.incognita*-Races 1 and 4, *M.hapla*-Race B and *M.arenaria*-Races 1 and 2. The results (Fig. 5A) clearly indicate that the mtDNA of each race of each species tested contains the 102 ntp sequence. The differences in migration distance for the single bands observed for *M.hapla*-Race A, *M.hapla*-Race B and *M.arenaria*-Race 2, could have resulted from differences in either repeat number or location of *MboI* sites relative to one or both ends of the repeat set. From consideration of the sizes of multiple bands observed for *M.incognita*-Races 1, 3 and 4, and *M.arenaria*-Race 1, it seems likely that the mtDNAs of each of these races include sequence variants. These variants could again represent differences in repeat numbers and/or *MboI* site locations. The band patterns in Fig. 5A do not suggest cross-contamination between the three races of *M.incognita*, between the two races of *M.hapla*, or between the two races of *M.arenaria*. Therefore, in spite of the apparent presence of sequence variants among some of the mtDNAs tested, the band patterns are diagnostic for the different host races within a species. Also, the band patterns of the three *M.incognita* races, *M.hapla*-Race B and *M.arenaria*-Race 1, can be distinguished from those of all other races tested.

Each of the DNAs represented in Fig. 5A was partially digested (37°C) with *TaqI*, electrophoresed, blotted and hybridized with the *M.javanica* 2×102R probe. In each case a ladder of fragments was observed (data not shown), confirming that multiple tandemly arranged 102 ntp repeats occur in each of the *Meloidogyne* mtDNAs tested.

MboI digested *Meloidogyne* DNAs were probed with a 30 nt oligomer, complementary to the *M.javanica* 63 ntp repeat (Fig. 3; nt 6,441–6,470). Data from the resulting autoradiographs (Fig. 5B and C) indicate that the 63 ntp repeated sequence is also present in the mtDNAs of all of the *Meloidogyne* species and races tested. The similarly located single band in each of the lanes representing *M.incognita*-Race 3, *M.incognita*-Race 1 and *M.arenaria*-Race 2 is consistent with the conclusion that the copy number, sequence arrangement and location relative to *MboI* sites of the 63 ntp repeat is similar in the mtDNAs of these races. Possible sequence variants in regard to the 63 ntp repeat are again

indicated by the double bands visible in each of the lanes containing *M.hapla*-Race A, *M.hapla*-Race B and *M.arenaria*-Race 1 DNAs.

DISCUSSION

The data presented in this paper establish that three sets of tandemly arranged, directly repeated sequences occur within a 7 kb segment of the mtDNA molecule of *M.javanica*: approximately 36 copies of a 102 ntp sequence; 11 copies of a 63 ntp sequence; 5 copies of an 8 ntp sequence. It seems likely that the repeat-containing 7 kb segment is the control region of the *M.javanica* mtDNA molecule. However, none of the *M.javanica* mtDNA repeats have convincing sequence similarity to any of the sequenced, control region-containing repeats found in other metazoan species (9,12–14).

A function for any of the *Meloidogyne* mtDNA repeats remains undetermined. As open reading frames traverse the 102 ntp and 63 ntp repeat sets in *M.javanica* mtDNA, each of these sequences has the potential to encode a protein comprising a repeated amino acid sequence. However, the single nucleotide deletion that occurs in two copies of the *M.incognita*-Race 3, 102 ntp sequence would result in a protein encoded by either strand that is variant and truncated relative to the corresponding, putative *M.javanica* protein. An alternative possibility is that individual 102 ntp and 63 ntp repeats could each encode a short protein. Such a protein might be produced either by proteolytic cleavage of a long repeated polypeptide or from translation of a repeat length transcript. However, at this time we do not have functional evidence to support the view that any of the *Meloidogyne* repeats are expressed: we have been unable to detect RNAs that contain transcripts of either the 102 ntp or the 63 ntp sequences (Okimoto, R. and Wolstenholme, D.R., unpublished data).

Our data provide clear evidence that the 102 ntp and 63 ntp sequences were present, and that at least the 102 ntp sequence was tandemly repeated in an ancestor common to all four of the *Meloidogyne* species tested. Comparisons of corresponding protein gene-containing sequences (four sequences totalling 2,389 ntp) from *M.javanica* and *M.incognita* mtDNAs have indicated a nucleotide divergence in the coding region of the molecule of only 0.08% (R.Okimoto, N.A.Okada, D.R.Wolstenholme, unpublished data), suggesting that the establishment of *M.javanica* and *M.incognita* as distinct species was a relatively recent event. The high degree of sequence similarity between *M.javanica* and *M.incognita*-Race 3 102 ntp repeats is also consistent with this latter view.

In regard to the presence and multiplicity of the 102 ntp and 63 ntp repeats, it is interesting to note that the mode of gene expression in metazoan mtDNAs, which necessitates the use of a minimum of DNA sequence (see ref. 7), together with the small number of genes retained in these molecules, has been reasoned to result from selection for smallness (see discussions in 5,27). If this is indeed the case, then it follows that the repeated sequences in *Meloidogyne* mtDNAs must confer a strong selective advantage to molecules that contain them.

It seems likely that multiple copies of the 102 ntp and 63 ntp repeats could have been generated by a mechanism similar to that recently proposed by Buroker et al. (12) to account for the presence of between four and eight copies of an 82 ntp sequence in the control region of sturgeon (*Acipenser transmontanus*) mtDNA. Operation of this mechanism requires a tandemly duplicated sequence and relies on the asymmetrical mode of

replication peculiar to metazoan mtDNAs. In vertebrate and in *Drosophila* mtDNAs, DNA synthesis of one strand (the H strand) initiates in the control region and continues for a considerable distance around the molecule before synthesis of the second (L) strand is initiated (6). In the Buroker et al. model, synthesis of the H strand proceeds through the duplicated segment and is then displaced by the parental H strand. The duplicated copy that is proximal to the replication origin is stabilized by intramolecular base pairing, whilst the distal copy base pairs with the proximal parental L strand copy. Continued synthesis of the nascent H strand on the distal L strand copy results in the addition of a third copy of the sequence to the nascent H strand. A double-stranded molecule containing three copies is completed in the next replication cycle. Whether or not the *Meloidogyne* mtDNA repeats could have attained their present level of multiplicity by this proposed mechanism would depend on the potential for intra-strand pairing in the 102 ntp and 63 ntp repeats (that might involve small hairpin forming sequences (Fig. 3)), and on whether the repeats are actually located in a region that is replicated asymmetrically.

Root knot, caused by nematodes of the genus *Meloidogyne*, is one of the most economically important diseases of crop plants. Root knot nematodes (about 30 species) infect up to 3,000 plant species that include most of the earth's commercial crops, and are responsible for a worldwide annual yield loss estimated at about 5% (29,30). Most of the root knot damage in the United States of America is caused by *M.javanica*, *M.incognita*, *M.hapla* and *M.arenaria*. Host-specificity of races of the latter three species is the basis for management of root knot disease by crop rotation (31). Although different species of *Meloidogyne* can be identified by morphological and cytological characteristics, race identification depends on lengthy and cumbersome host specificity tests (31).

Restriction enzyme cleavage site differences between mtDNAs of different *Meloidogyne* species and some host races have been demonstrated previously using purified mtDNAs, and it has been suggested that such data might be useful for host race identification (20,32). In the presently reported experiments we were able to distinguish all of the host races of *M.incognita*, *M.hapla* and *M.arenaria* tested by hybridizing a single 102 ntp repeat-containing probe to either restricted mtDNA or restricted whole worm DNAs. Although some of the mtDNAs included sequence variants, it is clear that the procedure used offers the potential for development of a relatively simple test for *Meloidogyne* host race identification. Determination of the real value of such a test must await the availability of data from hybridizations of the 102 ntp repeat-containing probe to restriction enzyme-cleaved DNAs from single worms or from multiple worms descended from a single female.

ACKNOWLEDGEMENTS

We are grateful to Michael A. McClure for providing eggs of root knot nematodes and for many helpful discussions during the course of this study. We thank Kirk Thomas for an oligonucleotide, and Martin C. Rechsteiner, John F. Atkins and Raymond F. Gesteland for comments on the manuscript. This work was supported by NIH Grant No. GM18375 and USDA Grant No. 86-CRCR-1-1994.

REFERENCES

- Okimoto, R. and Wolstenholme, D.R. (1990) *Nucleic Acids Res.* **18**, 6113–6118.
- Desjardins, P. and Morais, R. (1990) *J. Mol. Biol.* **212**, 599–634.
- Jacobs, H.T., Elliot, D.J., Math, V.B. and Farquharson, A. (1988) *J. Mol. Biol.* **202**, 185–217.
- Fauron, C.M.R. and Wolstenholme, D.R. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3623–3627.
- Boyce, T.M., Zwick, M.E., Aquadro, C.F. (1989) *Genetics* **123**, 825–836.
- Clayton, D.A. (1982) *Cell* **28**, 693–705.
- Clayton, D.A. (1984) *Ann. Rev. Biochem.* **53**, 573–594.
- Solignac, M., Monnerot, M. and Mounolou, J.-C. (1986) *J. Mol. Evol.* **24**, 53–60.
- Rand, D.M. and Harrison, R.G. (1989) *Genetics* **121**, 551–569.
- Densmore, L.D., Wright, J.W. and Brown, W.M. (1985) *Genetics* **110**, 689–707.
- Bentzen, P., Leggett, W.C. and Brown, G.G. (1987) *Genetics* **118**, 509–518.
- Buroker, N.E., Brown, J.R., Gilbert, T.A., O'Hara, P.J., Beckenbach, A.T., Thomas, W.K. and Smith, M.J. (1990) *Genetics* **124**, 157–163.
- Roe, B.A., Ma, D.P., Wilson, R.K. and Wong, J.F.-H. (1985) *J. Biol. Chem.* **260**, 9759–9774.
- La Roche, J., Snyder, M., Cook, D.I., Fuller, K. and Zouros, E. (1990) *Mol. Biol. Evol.* **7**, 45–64.
- Moritz, C. and Brown, W.M. (1986) *Science* **233**, 1425–1427.
- Moritz, C. and Brown, W.M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7183–7187.
- Wallis, G.P. (1987) *Heredity* **58**, 229–238.
- Hymen, B.C., Beck, J.L. and Weiss, K.C. (1988) *Genetics* **120**, 707–712.
- McClure, M.A., Cruik, T.H. and Mifaghi, I. (1973) *J. Nematol.* **5**, 230.
- Powers, T.O., Platzer, E.G. and Hyman, B.C. (1986) *J. Nematol.* **18**, 288–293.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Wolstenholme, D.R. and Fauron, C.M.-R. (1976) *J. Cell Biol.* **71**, 434–448.
- Wahleithner, J.A. and Wolstenholme, D.R. (1987) *Curr. Genet.* **12**, 55–67.
- Okimoto, R. and Wolstenholme, D.R. (1990) *EMBO J.* **9**, 3405–3411.
- Helentjaris, T., King, G., Slocum, M., Siederstrand, C. and Wegman, S. (1985) *Plant Mol. Biol.* **5**, 109–118.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Dale, R.M.K., McClure, B. and Houchins, J.P. (1985) *Plasmid* **13**, 31–40.
- Moritz, C., Dowling, T.E. and Brown, W.M. (1987) *Ann. Rev. Ecol. Syst.* **18**, 269–292.
- Dropkin, V.H. (1980) *Introduction to plant nematology*. J. Wiley and Sons, Inc. New York, NY.
- Poinar, G.O. Jr. (1983) *The natural history of nematodes*. Prentice Hall Inc., Englewood, NJ.
- Eisenback, J.D., Hirschmann, H., Sasser, J.N. and Triantaphyllou, A.C. (1981) *A guide to the four most common species of root knot nematodes (Meloidogyne Spp.)*, with a pictorial key. International Meloidogyne project. A publication of the U.S.A.I.D.
- Powers, T.O. and Sandall, L.J. (1988) *J. Nematol.* **20**, 505–511.