Istituto di Nematologia Agraria CNR - 70126 Bari, Italy[1]
Dipartimento di Biochimica e Biologia Molecolare, University of Bari - 70126 Bari, Italy[2]

SEQUENCE ANALYSIS OF THE *CUT-1* GENE AND OF ITS FLANKING REGIONS IN *MELOIDOGYNE ARTIELLIA*

by
F. De Luca[1], C. De Giorgi[2], M. Di Vito[1] and F. Lamberti[1]

Summary. The cuticlin-encoding gene has been isolated from a genomic library of the root-knot nematode, *Meloidogyne artiellia*. The organization revealed a typical eukaryotic-like structure. The sequence of the coding region was similar to the corresponding regions of the free living nematode *Caenorhabditis elegans*, strongly indicating the conservation of this gene during evolution. A detailed sequence analysis of both untranslated regions, 5'-UTR and 3'-UTR, showed the presence of regulatory signals.

It has already been reported that proteins, other than collagens, participate in the formation of the cuticle. In particular, these components are left as an insoluble residue after treating cuticles with strong detergent and disulphide reducing agents (Fujimoto and Kanaya, 1973; Reddigari *et al.*, 1986). Among these proteins, the cuticlin-1 gene (*cut-1*) has been fully studied in the free living nematode *Caenorhabditis elegans*, where it is expressed in the cuticle of dauer larvae (Sebastiano *et al.*, 1991).

In this paper, we report the isolation, characterization and sequence analysis of the *cut-1* gene and its flanking regions in the plant parasitic nematode *Meloidogyne artiellia* Franklin.

Materials and methods

Nematode total DNA was extracted from purified eggs of *M. artiellia* by the liquid nitrogen grinding technique. The phenol extraction and ethanol precipitation were carried out as described by Maniatis *et al.*, (1982). The DNA was visible in suspension and precipitated out around a sterile plastic rod. It was then dissolved in TE buffer (10 mM Tris pH 7.5, 0.1 mM EDTA) and stored at 4 °C.

The cloning of genomic fragments has been greatly facilitated by commercially available genomic cloning systems. The λGem-11 vector (Promega, Madison, WI, USA) was cleaved with *Xho*I and partially filled in, so that the vector arms have CT and TC protruding ends and can only be ligated with the partially filled in ends of genomic DNA fragments. The *M. artiellia* DNA was partially digested with the restriction enzyme *Sau*3A. The product was centrifuged on a discontinuos NaCl gradient at 5-10-15-20-25%. Fractions of 200 µl were collected and aliquots of each fraction analysed on agarose gels in parallel with a molecular weight marker (Fig. 1). The fragments ranging in size from 15 to 23 kb were pooled and partially filled in with *Sau*3A so to have GA and AG as protruding ends.

The library was plated and screened by using a portion of the *C. elegans cut-1* gene. Two positive phages, showing the same restriction map,

were isolated. As the two recombinant phages turned out to be identical, only one phage was fully characterised and sequenced by using the Sanger method (Maniatis *et al.*, 1982).

Results and discussion

The study of plant parasitic nematodes at molecular level can help to understand various aspects of their basic biology. Many genes in plant parasitic nematodes are likely to have a functional counterpart in *C. elegans* similar enough to be cloned by sequence homology (Riddle and Georgi, 1990). Nematode cuticular proteins are of interest because they are involved in protein-protein interactions which confer the mechanical properties and fine architecture of the cuticle. Moreover, it is well known that there is a dramatic change in the cuticular structure among different life stages of nematodes (Bird and Bird, 1991).

Using a portion of the *cut-1* gene isolated from *C. elegans* as probe, the genomic library of *M. artiellia* was screened and thus the *cut-1* containing fragment was isolated from this plant parasitic nematode. This fragment has been fully sequenced and revealed the presence of the entire *cut-1* gene and of its untranslated regions. The overall organization is reported in Fig. 2.

The sequence analysis revealed that the coding regions are highly conserved between *C. elegans* and plant parasitic nematodes (De Luca *et al.*, 1994). In contrast, both the intron-exon organization and the flanking regulatory regions are different from the *C. elegans cut-1* gene. This observation is in agreement with recent data supporting the hypothesis that 5' and 3'

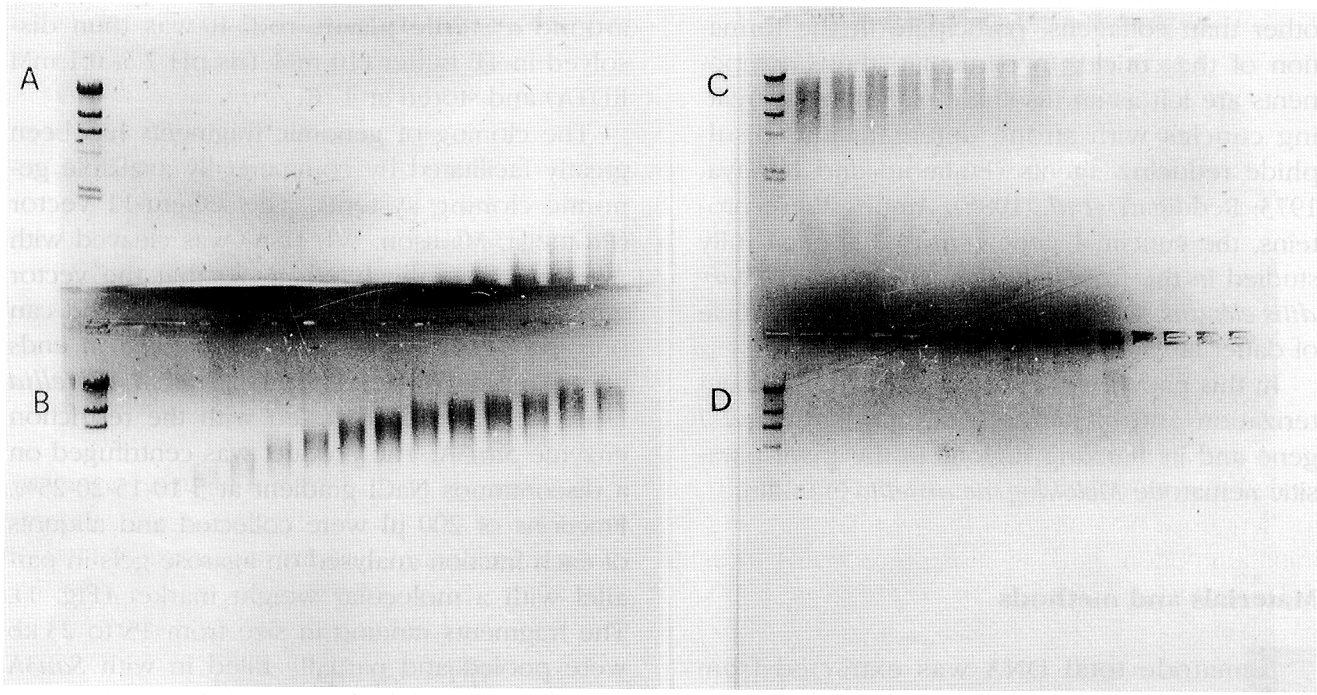

Fig. 1 - *Meloidogyne artiellia* fragment size distribution analysed by using 1% agarose gels. After centrifugation on NaCl gradient, 200 µl fractions were collected. From these fractions 15 µl aliquots were electrophoresed in the different lanes of the gel. Panel A represents the top gradient fractions, panel B and C the intermediates ones, panel D the bottom gradient fractions. The left end lane of each panel contains λDNA fragments as marker.

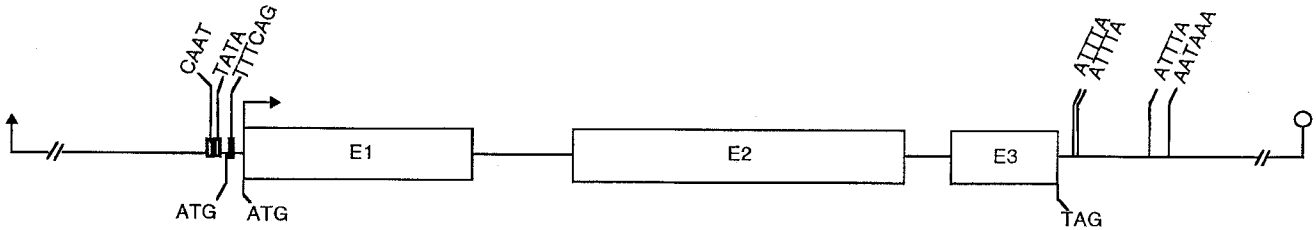

Fig. 2 - Overall organization of *cut-1* gene and of its flanking regions. White boxes indicate the coding regions. The arrow indicates the ATG translation start codon, while TAG indicates the translation stop codon. The symbols at the ends of the maps represent the *Bam*HI and *Xba*I restriction sites used for the subcloning in the plamid vector.

untranslated regions may play a specific role in the regulation of transcription and translation (Pesole *et al.*, 1994).

The upstream region of the *M. artiellia cut-1* gene sequenced is 1766 nt long. The 5'-UTR encompasses 300 nt upstream from the ATG start codon. In this portion no open reading frames were detected by computer analysis, while many regulatory elements have been identified. In particular, the eukaryotic promoter sequences TATA and CAAT boxes have been found at positions - 106 and - 132, respectively (Fig. 2). Two possible ATG start codons were detected, separated 71 nt from each other. The internal ATG codon has been assumed as the real translation initiation codon on the basis of the following considerations: it is in frame with the desumed polypeptide which is very similar to that of *C. elegans* and, on the other hand, it well matches the consensus ribosome-binding site for initiation (Kozak, 1991).

The sequence of the genomic *cut-1* DNA revealed the presence of a consensus splice acceptor sequence TTTCAG at -49 upstream from the initiation codon. This signal is preceded by three polypyrimidines tracts like those occurring in trypanosomes *trans*-spliced mRNAs (Vassella *et al.*, 1994). These findings also suggest that *M. artiellia cut-1* mRNA could be *trans*-spliced like many nematodes mRNAs. In particular the *C. elegans cut-1* gene is clearly *trans*-spliced. *Trans*-splicing consists in the addition of 22 nt sequence leader (SL1) to mRNA at the *trans*-splicing acceptor sequence (Krause and Hirsh, 1987). Indeed, Southern blot experiments (data not showed) clearly indicated that SL1 or its precursor do exist in the *M. artiellia* genome providing additional evidence that *trans*-splicing processing occurs in this plant parasitic nematode. To demonstrate the existence of *trans*-spliced mRNA, the combined techniques of "reverse transcriptase" and "polymerase chain reaction" (RT-PCR) were used (Frohman *et al.*, 1988). The single-strand cDNA was synthesised by reverse transcription using a gene-specific primer, while the amplification was carried out by using the same gene specific primer and the 22 nt sequence leader as a second primer (Fig. 3). The absence of any amplified fragment demonstrates that the *cut-1* mRNA is not *trans*-spliced and that the acceptor splice signal is silent, probably being an evolutionary relic (De Giorgi *et al.*, 1995).

Another peculiar motif of the *cut-1* 5'-UTR is the presence of homoguanines stretches no longer than 5 nt and homocytidines stretches which are variable in length. These homocytidines stretches are also present at the 3'-UTR of *M. artiellia cut-1* and internally at the *cut-1* introns. It has been suggested it might have a possible role as a binding site for regulatory factors involved in the accurate maturation of mRNAs (Singh *et al.*, 1995). Interestingly, these sequences seem to be absent in the corresponding region of *C. elegans*.

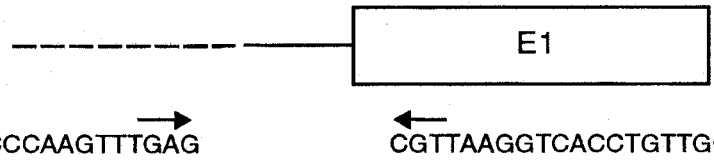

Fig. 3 - Primer sequences used in the RT-PCR experiments. The upstream primer of the PCR is the SL1 primer, while the other is the gene-specific primer. The dashed line indicates the presumptive mRNA region of annealing of the SL1 primer.

The downstream region of the *cut-1* gene sequenced is 600 nt long. The 3'-UTR is exactly 503 nt long downstream from the translation stop codon. It displays the presence of a peculiar hexanucleotide AATAAA known to regulate the polyadenilation which influences the stability of mRNA. In addition, the sequence ATTTA has been found three times in the *cut-1* 3'-UTR. This element has been already found at the 3'-UTR of lymphokine and cytokine mRNAs and seems to be involved in the regulation of mRNA stability through interaction with specific RNA-binding proteins (Bohjanen *et al.*, 1992).

In conclusion it appears that the *cut-1* coding regions are the same as those of *C. elegans cut-1* suggesting that the cuticlin-1 protein has the same function in the formation of cuticle in plant parasitic nematodes as in *C. elegans*. In contrast to free living nematodes, however, plant parasitic nematodes have evolved different regulatory signals, presumably because of different mechanisms for regulation of transcription and translation.

Literature cited

Bird A. F. and Bird J., 1991. *The structure of nematodes.* Academic Press, San Diego, pp. 56-74.

Bohjanen P. R., Petryniak B., June C. H., Thompson C. B. and Lindsten T., 1992. AU-RNA binding factors differ in their binding specificities and affinities. *J. Biol. Chem.*, *267*: 6302-6309.

De Giorgi C., De Luca F. and Lamberti F., 1996. A silent *trans*-splicing signal in the cuticlin-encoding gene of the plant parasitic nematode *Meloidogyne artiellia*. *Gene* in press.

De Luca F., De Giorgi C. and Lamberti F., 1994. Cuticlin gene organization in *Meloidogyne artiellia*. In: Lamberti F., De Giorgi C. and Bird D. McK (Eds), *Advances in Molecular Plant Nematology*, Plenum Press, New York, pp. 85-89.

Frohman M. A., Dush M. K. and Martin G. R., 1988. Rapid production of full-lenght cdna from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Sci. Usa*, *85*: 8998-9002.

Fujimoto D. and Kanaya S., 1973. Cuticlin: A noncollagen structural protein from *Ascaris* cuticle. *Arch. Biochem. Biophis.*, *157*: 1-6.

Kozak M., 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell. Biol.*, *115*: 887-903.

Krause M. and Hirsh D., 1987. A *trans*-spliced leader sequence on actin mRNA in *C. elegans*. *Cell*, *49*: 753-761.

Maniatis T., Fritsch E. F. and Sambrook J., 1982. *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 479 pp.

Pesole G., Fiormarino G. and Saccone C., 1994. Sequence analysis and compositional properties of untranslated regions of human mRNAs. *Gene*, *140*: 219-225.

Reddigari S. R., Jansma P. L., Premachandran D. and Hussey R. S., 1986. Cuticular collagenous proteins of second-stage juveniles and adult females of *Meloidogyne incognita*: isolation and partial characterization. *J. Nematol.*, *18*: 294-302.

Riddle D. L. and Georgi L. L., 1990. Advances in research on *Caenorhabditis elegans*: application to plant parasitic nematodes. *Annu. Rev. Phytopathol.*, *28*: 247-269.

Sebastiano M., Lassandro F. and Bazzicalupo P., 1991. *cut-1* a Caenorhabditis elegans gene coding for a dauer-specific noncollagenous component of the cuticle. *Dev. Biology*, *146*: 519-530.

Singh R., Valcarcel J. and Green M. R., 1995. Distinct binding specificities and functions of the higher Eukaryotic polypyrimidine tract-binding proteins. *Science*, *268*: 1173-1176.

Vassella E., Braun R. and Roditi I., 1994. Control of polyadenylation and alternative splicing of transcripts from adjacent genes in a procyclin expression site: a dual role for polypyrimidine tracts in trypanosomes? *Nucleic Acids Res.*, *22*: 1359-1364.

Accepted for publication on 25 January 1996.