

Horizontal transfer of a bacterial gene involved in polyglutamate biosynthesis to the plant-parasitic nematode *Meloidogyne artiellia*

Pasqua Veronico^a, John Jones^b, Mauro Di Vito^c, Carla De Giorgi^{a,*}

^aDipartimento di Biochimica e Biologia Molecolare, Università di Bari, Via Orabona 4, 70126 Bari, Italy

^bMycology, Bacteriology and Nematology Unit, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, UK

^cIstituto di Nematologia Agraria CNR, Via Amendola 165/A, 70126 Bari, Italy

Received 13 July 2001; revised 26 October 2001; accepted 30 October 2001

First published online 9 November 2001

Edited by Takashi Gojobori

Abstract Analysis of a genomic fragment from the plant parasitic nematode *Meloidogyne artiellia* revealed the presence of a gene which, in bacteria, is involved in the formation of polyglutamate capsule. Searching of various databases, including the *Caenorhabditis elegans* genome sequence and the large EST datasets from a variety of parasitic nematodes, showed that no similar genes have been identified in other nematodes or in any other eukaryotic organisms. The *M. artiellia* gene has a typical eukaryotic structure and its mRNA is present in the intestine. The gene is expressed in all life cycle stages tested. These findings demonstrate horizontal gene transfer may be important in catalyzing the diversification of nematode lineages. © 2001 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Gene expression; Horizontal transfer; In situ hybridization; Nematode; Xenology

1. Introduction

Nematodes exist in virtually every habitat and many species exist solely as obligate parasites of plants or animals. Plant parasitic nematodes have been estimated to cause damage valued in excess of 70 billion US\$ each year [1]. Much effort has therefore been made to understand the molecular nature of the interactions between plants and nematodes both in terms of the plant response to nematodes [2,3] and the mechanisms used by nematodes that allow them to parasitize plants [4]. The most studied plant parasitic nematodes are the cyst (*Globodera* and *Heterodera* spp.) and root knot nematodes (*Meloidogyne* spp.). These nematodes hatch into the soil as infective juveniles, but to complete their life style they must locate and enter the roots of a host plant where they induce the formation of highly specialized structures inside the host vascular cylinder. The nematodes remain at this feeding site for the remainder of their life cycle and depend on it to provide the food required for development to the adult stage.

Parasitic nematodes have evolved specializations in order to acquire a successful parasitic behavior and these adaptations

have been built on a framework of basic nematode anatomy. The availability of the entire sequenced genome from the free-living nematode *Caenorhabditis elegans* therefore represents a remarkable resource for understanding the organismal biology of other nematodes [5].

Cellulose and pectin are essential components of plant cells, and the production of enzymes able to degrade plant cell wall is of critical importance for phytoparasitic nematodes. Genes encoding cellulases and pectinase have been recently identified in various species of plant parasitic cyst nematodes including *Globodera rostochiensis*, *Heterodera glycines* and *Meloidogyne incognita* [6–8]. Another parasitic-specific gene encoding a potentially secreted chorismate mutase, a key enzyme in the biosynthetic pathway leading to phenylalanine and tyrosine in plants, has been recently studied in both *Meloidogyne javanica* [9] and *G. pallida* [10]. While all these genes are endogenous to the nematodes, they are more similar to microbial than eukaryotic (where present) enzymes. This unexpected resemblance may suggest conservation in the nematode genome of ancestral bacterial genes or, more likely, it may reflect past entry of bacterial genes into the nematode genome by horizontal gene transfer [11]. Horizontal gene transfer is a mechanism that allows the acquisition of novel capability within a single generation and it is widely recognized to be an important mechanism in divergence and adaptation of bacterial populations [12]. Horizontal gene transfer has been generally accepted as operating in eukaryotes [13], and it has been proposed that microbes in the digestive system of nematodes may have been the source of the cellulase and other genes which, in the course of evolution, have helped the nematode to convert to a parasitic habit [14,15].

This paper describes another gene, presumably of bacterial origin, present in the genome of a plant-parasitic nematode *Meloidogyne artiellia* (*Mt*). Although its role is not fully understood, the gene is expressed throughout the life cycle of the nematode, and the mRNA derived from this gene is present specifically in the intestine. Our data strongly indicate that this gene is not derived through bacterial contamination and has entered the genome of *Mt* by horizontal transfer.

2. Materials and methods

2.1. Biological material

Mt is a sedentary endoparasitic nematode, which reproduces both by cross-fertilization and facultative meiotic parthenogenesis. Each female lays approximately 500 eggs in a gelatinous sac deposited at the root surface. The eggs contain embryos at different developmental stages. Embryogenesis consists of four larval stages with the first stage

*Corresponding author. Fax: (39)-80-544 3317.

E-mail addresses: jjones@scri.sari.ac.uk (J. Jones), nemamd03@area.ba.cnr.it (M. Di Vito), c.degiorgi@biologia.uniba.it (C. De Giorgi).

Abbreviations: *Mt*, *Meloidogyne artiellia*; *pgsA*, polyglutamate synthetase gene A; SL1 and -2, *trans*-spliced sequence leader 1, and 2

(J1) occurring in the egg. The nematode moults to the second juvenile stage (J2) before the hatching from the egg, migrates into the soil, invades the root and continues its life cycle through J3 and J4 stages at a feeding site within the roots. Finally, a moult to the adult stage occurs. Males leave the roots and seek out the adult females which remain attached to the roots but whose bodies emerge through the root surface, providing the males with a means of access for mating to occur.

For experiments, adult female nematodes were collected by hand from infected roots. To obtain eggs, the gelatinous sac was collected and eggs were freed as previously described [16]. Hatched J2s were obtained by incubating eggs, collected as above, in water at 25°C for several days. No other nematode stages were available for experimentation due to the endoparasitic life style of the nematode (above). All nematode material was frozen in liquid nitrogen and stored at −80°C before use, except J2s used for in situ hybridization, which were fixed immediately after collection.

2.2. Sequencing and computational analysis

The *Mt* polyglutamate synthetase gene A (*Mt-pgsA*) gene was identified during sequencing of a lambda genomic clone undertaken during another project [17]. The entire clone (17065 nt) was sequenced and found to contain three protein-coding sequences: the *Mt-pgsA* gene, a chitin synthase gene [17], and a small open reading frame. These potential coding regions were identified using TestCode [18] and GeneFinder (<http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html>) programs and were confirmed using RT-PCR (below). Similarity searches were performed using WU-BLASTX2 (<http://dove.embl-heidelberg.de/blast2/>), and sequence alignments were generated using Multalin software [19]. The sequences reported in this study have been submitted to EMBL nucleotide sequence database under the accession numbers AJ310212 (*Mt-pgsA*) and AJ310213 (small open reading frame). The chitin synthase gene has been previously submitted under the accession number AY013285.

2.3. RT-PCR

RT-PCR was conducted on total RNA isolated from *Mt* eggs, J2s and adult females. First-strand cDNA was synthesized with an oligo d(T)₁₆ primer and AMV Reverse Transcriptase (Roche Diagnostics).

Subsequent amplifications were performed using gene-specific primers which annealed to sequences flanking introns, in order to allow PCR products generated from cDNA and genomic DNA to be differentiated. PCR products were separated on agarose gels using standard protocols [20]. Where genomic DNA was used for PCR (in control reactions), it was extracted from frozen eggs using previously described methods [16].

In all PCR reactions 10 ng template DNA was used for amplification in 100 µl reactions containing 0.5 µM each primer, 200 µM dNTPs and 2.5 U Taq DNA Polymerase (Roche Diagnostics) in the buffer provided by the supplier.

2.4. In situ hybridization

A 240 nt fragment of the *Mt-pgsA* cDNA was amplified from plasmid DNA template using as the primers CGGAATCAATGAGAAC-GA (forward) and CCTTCGCAATTTGCCGAC (reverse). This PCR product was used as a template to synthesize digoxigenin-labelled single-stranded sense or antisense probes which were used in whole-mount in situ hybridization reactions as previously described [21].

3. Results

Sequencing of the genomic clone containing a chitin synthase gene [17] revealed the presence of another gene which was very similar to a sequence present in *Mycobacterium tuberculosis* [22] and in other bacterial species (Fig. 1) including the *capA* gene from *Bacillus anthracis*. This gene is one of three, *capA*, *capB* and *capC*, involved in the polymerization of D-glutamic acid, an essential step in the encapsulation process in this organism [23]. Similarly, the homologous gene sequence from *Bacillus subtilis*, named polyglutamate synthetase gene A (*pgsA*), together with *pgsB* and *pgsC*, is a component of the polyglutamate synthesis system in this organism [24]. The gene that we have isolated from *Mt* has therefore been named *Mt-pgsA*, in keeping with the nomenclature sug-

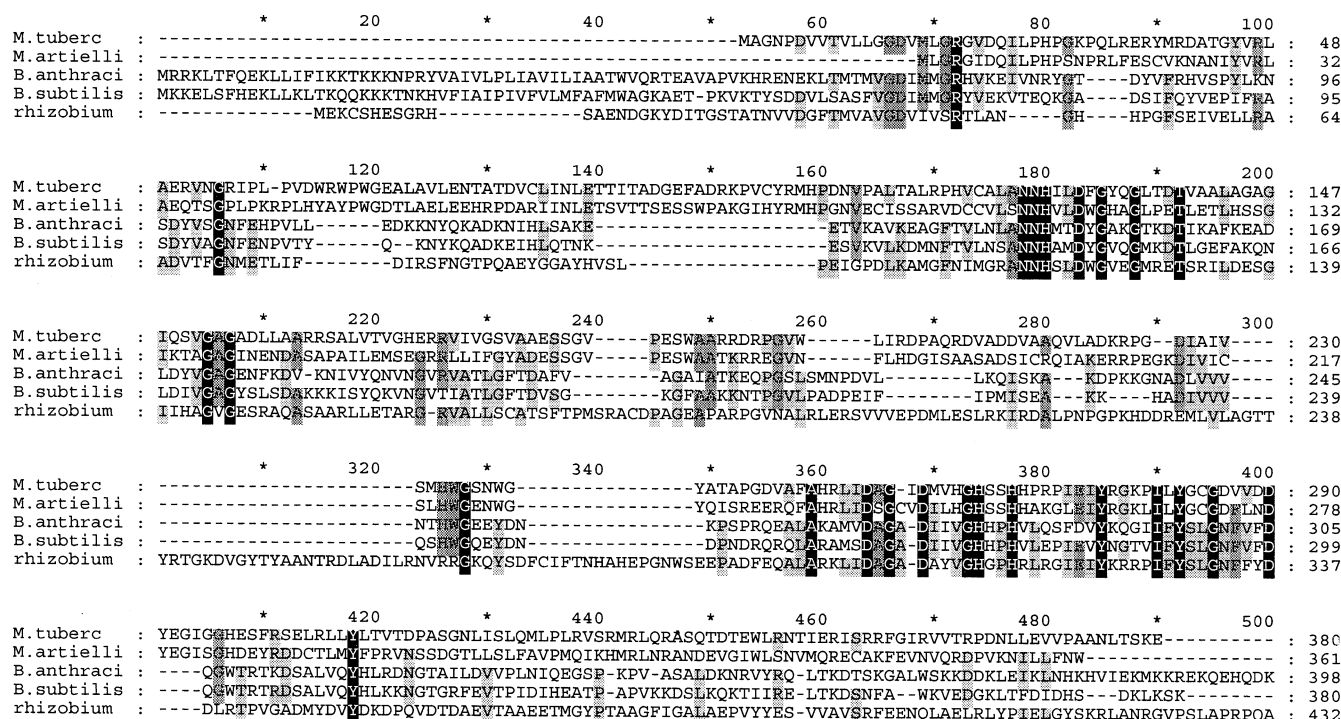


Fig. 1. Multiple alignment of deduced amino acid sequence of the *Mt-pgsA* gene and sequences present in bacterial species *M. tuberculosis* (AC: AL021941), *B. anthracis* (AC: P19579), *B. subtilis* (AC: H70069) and *Rhizobium* spp. (AC: P55651). Black shading indicates identity of a particular residue in all sequences, dark gray indicates identity in four of five sequences and light gray indicates identity in three of the five sequences.

gested for genes identified from parasitic nematodes [25] and in order to reflect its proposed function. Searching of a variety of different databases, including the animal parasite EST dataset and the completed *C. elegans* genome sequence, revealed that no similar genes have been identified in other nematodes or indeed in any other eukaryotic organism. However, the gene is present and conserved in different strains and species of bacteria, as shown in Fig. 1. The gene from *Mt* is 65% similar (at the amino acid level) to the *Meloidogyne tuberculosis* gene and this high level of similarity is particularly significant when compared against the 67.7% similarity found within the *Bacillus* genus. We also compared this figure with the levels of similarity between other genes present in both *M. tuberculosis* and nematodes (*C. elegans*). Despite the fact that such genes (almost by definition) have extremely conserved functional roles, the similarity between the nematode and bacterial genes was between 23% (aspartyl tRNA-synthase) and 32% (glucose 1 dehydrogenase) (not shown), considerably less than the figure for the *Mt-pgsA* gene.

The *Mt* gene displays a typical eukaryotic gene structure. It is 2518 nt long and is composed of seven exons. The introns, bordered by canonical *cis*-splicing signals, are small and AT rich, all features commonly found in nematode introns [26]. Although the *cis*-splicing donor site of the fourth intron has a non-canonical junction, the sequence of the mRNA transcript demonstrates that all introns are correctly processed. A 5'UTR region of 130 nt is transcribed which fits well with the expected size of eukaryotic untranslated regions [27,28]. Southern blotting showed that a single copy of this gene is present in the genome of *Mt* and provided further evidence that it is not derived from bacterial contamination (Fig. 2). Low stringency Southern blots of genomic DNA from other plant parasitic nematodes probed with the same fragment of the *Mt-pgsA* gene gave only a faintly hybridizing band in

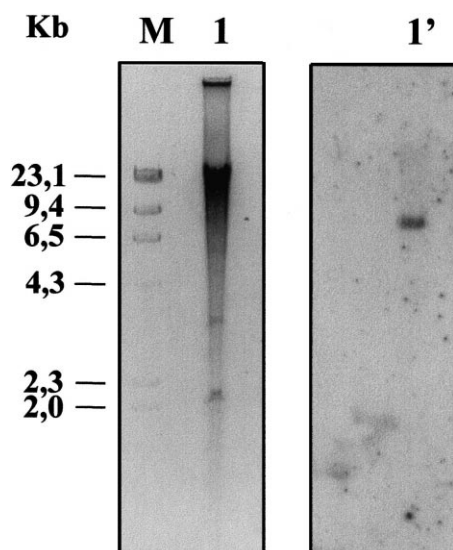


Fig. 2. Southern blotting of *Mt* total DNA cleaved with *Sac*I and hybridized with a portion of the *Mt-pgsA* gene. The probe of 170 bp spans a portion of the first exon. The figure shows that it hybridizes with a single genomic fragment of 9 kb, as predicted from the genomic sequence. The data are consistent with the presence of a single copy of the *Mt-pgsA* gene in *Mt*. Lane 1: Cleaved genomic DNA; lane 1': hybridization of a *Mt-pgsA* fragment with the genomic DNA. M: molecular size marker.

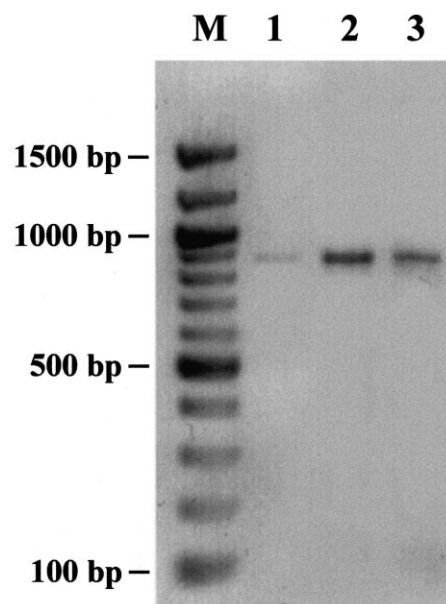


Fig. 3. Results of RT-PCR experiments indicating that the *Mt-pgsA* gene is expressed all *Mt* life cycle stages tested. Lane 1: eggs; lane 2: females; lane 3: J2 juveniles. M: molecular size marker. The expected size of the product amplified from cDNA is 881 bp.

Meloidogyne javanica, while no hybridization was observed with total DNA extracted from *G. rostochiensis*, a more distantly related plant parasitic nematode (data not shown).

The mRNA derived from the *Mt-pgsA* gene could encode a protein of 361 amino acids with a predicted molecular mass of 41 189 Da. This is similar to the predicted molecular masses of the proteins encoded by the two most similar genes in the database (*M. tuberculosis* – 41 547 Da and *Bacillus anthracis* – 41 547 Da). Predictions using the TMDD software (<http://www.cbs.dtu.dk/services/tmhmm-1.0/>) [29] suggested that no transmembrane helices were present, a property shared with the *M. tuberculosis* predicted protein but in contrast to the proteins predicted from *B. anthracis* and *B. subtilis* which are each predicted to have a single transmembrane domain close to the N-terminus of the proteins.

RT-PCR showed that the mRNA derived from the *Mt-pgsA* gene is present in all nematode life cycle stages examined (Fig. 3). The identity of the PCR product obtained was confirmed by excising it from the gel, cloning and sequencing. These experiments also confirmed that *cis*-splicing occurred at the sites predicted from the genomic sequence. In nematodes, the 5'-termini of many mRNAs are added post-transcriptionally by a *trans*-splicing mechanism using one of two specific sequences, SL1 and SL2. RT-PCR experiments using the SL1- and SL2-derived primers produced no bands in any life cycle stages tested, suggesting that the *Mt-pgsA* mRNA is not *trans*-spliced in vivo. This feature has also been observed for other nematode genes horizontally acquired from bacteria including the cellulase and pectate lyases of *G. rostochiensis* [6,7].

Experiments using in situ hybridization showed that the *Mt-pgsA* mRNA is present solely in the nematode intestine (Fig. 4). The transcript was present throughout the entire length of the intestine. Control reactions, using probes generated from the sense strand, gave no staining (not shown).

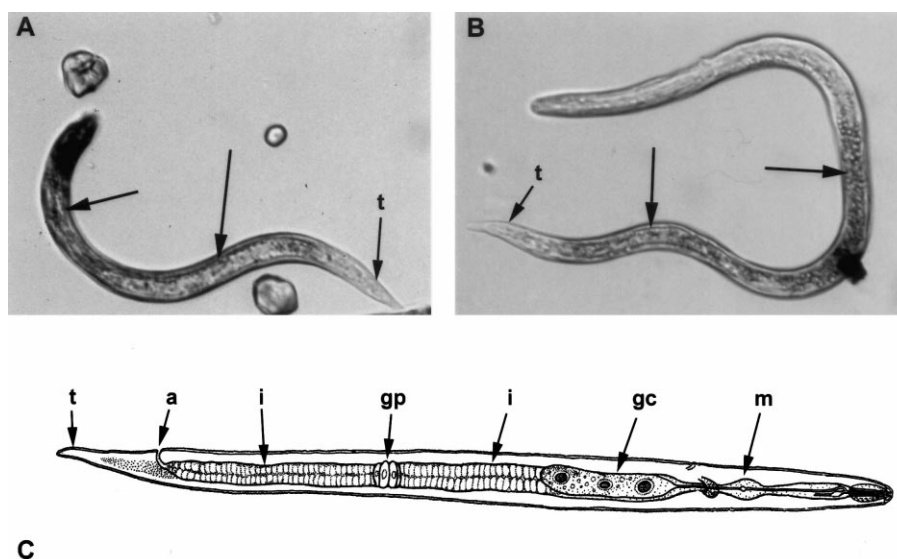


Fig. 4. A,B: In situ hybridization reaction showing that the *Mt-pgsA* gene is expressed in the intestine (arrows) of *Mt*. C: Generalized diagram showing location of nematode intestine in relation to other body structures (copied with permission of Comparative Parasitology and the Helminthological Society of Washington [32]. t: Tail; a: anus; i: intestine; gp: genital primordial; gc: oesophageal gland cells; m: metacarpus.

4. Discussion

We have identified a gene from a plant parasitic nematode which, while showing striking similarity to bacterial genes, has no apparent homologs in other eukaryotes. Several lines of evidence indicate that the *Mt* gene is a genuine nematode gene rather than a gene from a bacterial symbiont or contaminant. First, the genomic fragment from which the *Mt-pgsA* gene derives contains a chitin synthase gene, which has been shown to be a nematode gene [17]. Second, the genomic organization of the *Mt-pgsA* gene is typical of that of other nematode genes. Introns, similar in size to those found in nematodes, are present, and we have shown these to be processed in the expected manner. The mRNA derived from the gene contains many of the features expected for eukaryotic genes, including a polyadenylation signal and a polyA tail, and is present specifically in one nematode tissue. We conclude, therefore, that this gene has been acquired by *Mt* as a result of horizontal gene transfer (xenology). An alternative possibility is that the gene has been retained in *Mt* but has been lost in other species. However, several lines of evidence suggest this is not the case. First, studies on nematodes have been greatly facilitated by both the *C. elegans* genome project and the many EST projects that have been undertaken on a wide range of parasitic nematodes from across the phylum. None of the sequences in any of these databases show similarity to the *Mt-pgsA* sequence described here. Secondly, no sequences with similarity to the *Mt-pgsA* sequence has been described from any other phyletic groups. Therefore, for the presence of the *Mt-pgsA* gene in *Mt* to have arisen by gene loss alone would require loss of orthologous genes from other phyletic groups as well as many nematode lineages. We conclude that horizontal transfer to a nematode from a bacterium is the most realistic explanation for the findings described here.

The origin of the *Mt* gene is unknown. Similarity searches show that of the genes in the databases, it is most similar to a gene from *M. tuberculosis*. The nematode protein also shares several structural features with the *M. tuberculosis* protein, most notably the absence of a transmembrane domain which

is present in the homologous genes from other bacteria including *B. anthracis* and *B. subtilis*. Other bacterial homologs also lack this transmembrane domain. The observation that this gene is absent in the *C. elegans* genome suggests that the acquisition of this gene is an event that has occurred after the separation of *C. elegans* and *Meloidogyne* lineages from a common ancestor. The absence of similar genes from any of the parasitic nematode EST datasets coupled with the very high similarity between the *Mt* and bacterial genes also argues for a relatively recent acquisition. Intriguingly, the *cap* genes of at least some bacteria (e.g. *B. anthracis*) are encoded on plasmids. This is also the case for at least some of the other genes acquired by nematodes via horizontal gene transfer such as cellulases and pectate lyases, and suggests that plasmid-encoded genes may be more easily acquired by nematodes.

Mt-pgsA contains six introns. These introns have many features typical of nematode introns, but this is to be expected since they will be removed by a nematode-splicing apparatus. The origin of these introns is not known. There are currently two schools of thought regarding the evolution of introns in eukaryotic cells. One (intron-late) suggests that eukaryotic genes acquired introns and that prokaryotic genes have never contained introns. The other (intron-early) suggests that early organisms all possessed introns and that prokaryotic lineages subsequently lost these sequences, whereas eukaryotic organisms refined the systems present for their removal and exploited the presence of introns and splicing to allow generation of new genes through events such as alternative splicing. Some evidence exists to support this hypothesis. It is, therefore, feasible that *Mt-pgsA* was acquired with introns, which were subsequently lost from the bacterial lineage. However, it seems likely (see above) that the transfer of the *Mt-pgsA* gene to nematodes has occurred more recently than the events leading to loss of introns by prokaryotes (if this occurred). It seems more likely that *Mt-pgsA* acquired introns after entry into the nematode genome. Some evidence to suggest that this is a reasonable assumption comes from studies on nematode cellulases which were also thought to have been horizontally acquired by nematodes from bacteria. These genes also con-

tain introns and the position of these introns is not correlated with the functional domains of the protein sequence [11]. This evidence supports the intron-late hypothesis above and also suggests, therefore, that the introns were acquired after entry into the nematode genome. Similar analysis of the protein encoded by *Mt-pgsA* is problematic as little is known about the details of the biochemical function of this protein or its homologs in bacteria.

Although it is not known whether homologs of the other two genes involved in synthesis of the polyglutamate capsule are present in *Mt* genome, the similarity of the *Mt-pgsA* with the corresponding genes from *M. tuberculosis*, *B. subtilis* and *B. anthracis* suggests that the product of this gene is involved in synthesis of polyglutamate. Polyglutamate plays an important physiological role in bacterial cells, forming part of the bacterial cell envelope. This structure is one of the most complex and physiologically important components of the bacterial cell. The presence of a polyglutamate capsule, which represents the outermost component of the cell, is known to be a virulence factor, conferring strong resistance to phagocytosis upon entry into the bacterial host [30]. The roles of polyglutamate in *Mt* remain to be determined. The nematode intestine (the site of expression of the *Mt-pgsA* gene) is formed by an internal tube of epithelial cells bearing microvilli which project into the lumen. Proteins synthesized by the intestine may have a structural role in this tissue or may be involved in digestion or transport of absorbed substances to the rest of the nematode body. Polyglutamate itself may have a role in the process of feeding, but the difficulties of using *Mt* as an experimental organism mean that it has not been possible to determine the precise role of the *Mt-PGSA* protein.

In bacteria, horizontal gene transfer is recognized as a mechanism which can lead to the gain of resistance against adverse conditions in the environment, the acquisition of a novel biosynthetic or degradative capacity, or the acquisition of a new protein whose function or structure can be of some utility in a new organismal niche [31]. In the case of nematodes, the acquisition of a novel gene may have benefits in terms of parasitic capacity, as can be seen in the case of cellulases or other plant cell wall degrading enzymes or in terms of more adaptive phenotype.

Future studies will be aimed at uncovering the precise function of the protein encoded by *Mt-pgsA* and determining the extent of its distribution within *Meloidogyne* and related genera.

Acknowledgements: This work was funded by Ministero della Università e della Ricerca Scientifica e Tecnologica (C.D.G.) and by Scottish Executive Rural Affairs Division project Number SCR/514/97 (J.T.J.). P.V. received a fellowship from the Consorzio Interuniversitario Biotecnologie. The authors thank Alison Prior for assistance with *in situ* hybridization.

References

- [1] Sasser, J.N. and Freckman, D.W. (1987) in: *Vistas on Nematology* (Veech, J.A. and Dickson, D.W., Eds.), pp. 7–14, Society of Nematologists, Hyattsville, MD.
- [2] Williamson, V.M. and Hussey, R.S. (1996) *Plant Cell* 8, 1735–1742.
- [3] Niebel, A., de Almeida Engeler, J., Hemerly, A., Ferreira, P., Inze, D., Van Montagu, M. and Gheysen, G. (1996) *Plant J.* 10, 1037–1043.
- [4] Davis, E.L., Hussey, R.S., Baum, T.J., Bakker, J., Schots, A., Rosso, M.N. and Abad, P. (2000) *Annu. Rev. Phytopathol.* 38, 365–396.
- [5] Blaxter, M. (1998) *Science* 282, 2041–2046.
- [6] Smant, G., Stokkermans, J.P.W.G., Yan, Y., de Boer, J.M., Baum, T.J., Wang, X., Hussey, R.S., Gommers, F.J., Henrissat, B., Davis, E.L., Helder, J., Schots, A. and Bakker, J. (1998) *Proc. Natl. Acad. Sci. USA* 95, 4906–4911.
- [7] Popeijus, H., Overmars, H., Jones, J., Blok, V., Govers, A., Helder, J., Schots, A., Bakker, J. and Smant, G. (2000) *Nature* 406, 36–37.
- [8] Rosso, M.N., Favery, B., Piotte, C., Arthaud, L., De Boer, J.M., Hussey, R.S., Bakker, J., Baum, T.J. and Abad, P. (1999) *Mol. Plant-Microbe Interact.* 12, 585–591.
- [9] Lambert, K.N., Allen, K.D. and Sussez, I.M. (1999) *Mol. Plant-Microbe Interact.* 12, 328–336.
- [10] Popeijus, H., Blok, V.C., Cardle, L., Bakker, E., Phillips, M.S., Helder, J., Smant, G. and Jones, J.T. (2000) *Nematology* 2, 567–574.
- [11] Yan, Y., Smant, G., Stokkermans, J., Qin, L., Helder, J., Baum, T., Schots, A. and Davis, E. (1998) *Gene* 220, 61–70.
- [12] Lawrence, J.G. (1999) *Curr. Opin. Microbiol.* 2, 519–523.
- [13] De la Cruz, F. and Davies, J. (2000) *Trends Microbiol.* 8, 12–133.
- [14] Keen, N.T. and Roberts, P.A. (1998) *Proc. Natl. Acad. Sci. USA* 95, 4789–4790.
- [15] Bird, D.M. and Koltai, H. (2000) *J. Plant Growth Regul.* 19, 183–184.
- [16] De Giorgi, C., De Luca, F. and Di Vito, M. (1997) *Mol. Gen. Genet.* 25, 589–598.
- [17] Veronico, P., Gray, L.J., Jones, J.T., Bazzicalupo, P., Arbucci, S., Cortese, M.R., Di Vito, M. and De Giorgi, C. (2001) *Mol. Gen. Genomics* 266, 28–34.
- [18] Fickett, J.W. (1982) *Nucleic Acid Res.* 10, 5303–5318.
- [19] Corpet, F. (1988) *Nucleic Acid Res.* 16, 10881–10890.
- [20] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [21] De Boer, J.M., Yan, Y., Smant, G., Davis, E.L. and Baum, T.J. (1998) *J. Nematol.* 30, 309–312.
- [22] Philipp, W.J., Poulet, S., Eiglmeier, K., Pascopella, L., Balasubramanian, V., Heym, B., Bergh, S., Bloom, B.R., Jacobs Jr., W.R. and Cole, S.T. (1996) *Proc. Natl. Acad. Sci. USA* 93, 3132–3137.
- [23] Makino, S., Uchida, I., Terakado, N., Sasakawa, C. and Yoshikawa, M. (1989) *J. Bacteriol.* 171, 722–730.
- [24] Ashiuchi, M., Soda, K. and Misono, H. (1999) *Biochem. Biophys. Res. Commun.* 263, 6–12.
- [25] Bird, D.M. and Riddle, D.L. (1994) *J. Nematol.* 26, 138–143.
- [26] Blumenthal, T. and Steward, K. (1997) in: *C. elegans II* (Riddle, D.L., Blumenthal, T., Meyer, B.J. and Priess, J.R., Eds.), pp. 117–145, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [27] Kozak, M. (1983) *Microbiol. Rev.* 47, 1–45.
- [28] Pesole, G., Fiormarino, G. and Saccone, C. (1994) *Gene* 140, 219–225.
- [29] Sonhammer, E.L.L., Von Heijne, G. and Krogh, A. (1998) in: *Proceedings of the Sixth International Conference on Intelligent Systems for Molecular Biology* (Glasgow, J., Littlejohn, T., Major, F., Lathrop, R., Sankoff, D. and Sensen, C., Eds.), pp. 175–182, AAAI Press, Menlo Park, CA.
- [30] Troy, F.A. (1973) *J. Biol. Chem.* 248, 305–315.
- [31] Doolittle, W.F. (1999) *Trends Cell Biol.* 9, M5–M8.
- [32] Endo, B.Y. (1988) *Proc. Helminthol. Soc. Wash.* 55, 117–131.