Isolation, cloning, and characterization of a cuticle collagen gene, *Mi-col-5*, in *Meloidogyne incognita*

Sagar Banerjee^{1,2} · Sarvajeet Singh Gill² · Pradeep Kumar Jain³ · Anil Sirohi¹

Received: 13 January 2017 / Accepted: 15 February 2017 / Published online: 27 April 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract Cuticle collagens form a major part of the nematode cuticle and are responsible for maintaining the overall shape of the animal and its protection from the external environment. Although substantial research on cuticle collagen genes has been carried out in *Caenorhabditis elegans*, their isolation and characterization in plant parasitic nematodes have been limited to a few genes only. In this study, a cuticle collagen gene, *Mi-col-5*, was isolated from root-knot nematode, *Meloidogyne incognita*. A partial segment of 402 bp was first cloned and analyzed on Gbrowse followed by subsequent cloning of the 1047 bp long full cDNA specifying the open reading frame. The deduced amino acid sequence showed 92% sequence identity with that of *Mj-col-5*. However, a transmembrane helix was predicted in *Mi-col-5* which was not present in *Mj-col-5*. The conserved pattern of cysteine residues in *Mi-col-5* suggested that it belonged to group 2 of nematode cuticle collagens but with a longer carboxy terminal region as was the case with *Mj-col-5*. Domain prediction revealed the presence of a nematode cuticle collagen N terminal domain and a pfam collagen domain along with collagen triple helix repeats. A phylogenetic tree based on the amino acid sequences showed evolutionary relationship of *Mi-col-5* with cuticle collagens genes of other nematodes. 3D models for *Mi-col-5* were

predicted with the best confidence score of -2.78 . Expression of *Mi-col-5* transcript was found to be maximum in egg masses followed by adult females and J2s suggesting its role in the early stages of the development of the nematode during its life cycle.

Keywords Cuticle collagens · *Mi-col-5* · *Meloidogyne incognita* · Egg masses

Introduction

Structural proteins, collagens are involved in the formation of cuticle in nematodes which maintains the shape of the animal and protects it from the external environment. The collagens constitute around 80% of total protein content of the nematode cuticle (Kingston 1991). Cuticle collagens possess characteristic Gly-X-Y repeats with frequent appearance of proline and hydroxyproline in place of X and Y, respectively (Page and Johnstone 2007). They also share conserved patterns of cysteine residues which forms the basis of their classification into different groups (Johnstone 2000). The collagen genes encode procollagens with molecular masses of approximately 30 kDa, which undergo post-translational modifications and trimerisation in the endoplasmic reticulum. The triple helix structure is brought about by tyrosine-tyrosine bonds while disulphide bonds are involved in cross linking between the triple helices (Koltai et al. 1997; Kramer 1994). In the free living nematode, *C. elegans*, more than 150 collagen genes are found to be involved in the formation of the cuticle. Individual collagen genes are expressed at different stages of the life cycle of the nematode and exoskeletal defects may result from mutation in these individual collagen genes leading to defects in the shape of the animal which

✉ Anil Sirohi
anilsirohi@yahoo.com

¹ Division of Nematology, ICAR-Indian Agricultural Research Institute, Pusa, New Delhi, India

² Centre for Biotechnology, Maharshi Dayanand University, Rohtak, India

³ ICAR-National Research Centre on Plant Biotechnology, Pusa, New Delhi, India

corresponds to lethality (Page et al. 2014; Page and Winter 2003). In plant parasitic nematodes (PPNs) like root-knot nematodes (RKNs) and cyst nematodes (CNs), the cuticle is also involved in their interaction with soil environment as well as with the host (Davies and Curtis 2011). However, not much is known about the genomic organization of the cuticle collagens of PPNs.

Root-knot nematode, *Meloidogyne incognita*, is amongst the most devastating and economically important PPNs. These PPNs have a wide host range and are biotrophic in nature completing most of their life cycle inside the host. The nematode molts multiple times inside the host bringing about spectacular changes in its morphology during development of J2s to adult female. Hence, the cuticle collagen genes and their developmentally regulated expression may play an important role in the establishment of the nematode inside the host. The expression of cuticle collagen genes and their role in synthesis and maintenance of the cuticle has been well studied in *C. elegans*; but in PPNs, it is still an understudied domain. However, the availability of the *M. incognita* whole genome sequence has facilitated ways for identification and cloning of crucial genes like cuticle collagens (Abad et al. 2008). In this study, we have isolated, cloned and characterized a cuticle collagen gene, *Mi-col-5* from *M. incognita*.

Materials and methods

Nematode culture

For the maintenance of pure culture of *M. incognita* chitwood race 1, young tomato plants (*Solanum lycopersicum* L. cv. Pusa Ruby) were inoculated with fresh second-stage juveniles (J2s). The roots of the infected tomato plants were uprooted 30 days post inoculation and washed with double distilled water, and the egg masses were handpicked and kept in a cavity block. These egg masses were treated with 0.1% HgCl₂ for 1 min for surface sterilization and then washed thrice with double distilled water to remove the surface sterilizing agent. The egg masses were then allowed to hatch at 26–28 °C through a wire gauze covered with double-layered tissue paper into a petri plate filled with double distilled water (Hooper 1986). Freshly hatched J2s were used for further experiments. Adult females were also isolated from the roots of the infected tomato plants 30 days post inoculation under the microscope using a needle.

Isolation of total RNA from different stages of *M. incognita*

Total RNA was isolated from egg masses, adult females, and freshly hatched J2s of *M. incognita* using TRIzol

(ThermoFisher). 1 mL of TRIzol was added to egg masses, J2s and adult females per 100 mg of tissue sample and frozen in liquid nitrogen. The samples were crushed in 1.5-mL centrifuge tubes using a tissue crusher. Finely crushed samples in TRIzol were incubated at room temperature for 5 min followed by addition of 0.2 mL of chloroform/mL of TRIzol reagent. The tubes were vigorously shaken by hand for 15 s and incubated for 5 min at room temperature. The samples were then centrifuged for 15 min at 12000×g at 4 °C for phase separation. The aqueous phases of the samples were taken in new 1.5-mL microcentrifuge tubes and 0.5 mL of 100% isopropanol was added per mL of the TRIzol used. The samples were then incubated at –20 °C for 2 h for precipitation of RNA followed by their centrifugation at 12000×g for 10 min at 4 °C, and the supernatants were removed. The pellets were washed with 1 mL of 75% ethanol/mL of TRIzol used by centrifugation at 7500×g for 5 min at 4 °C. The RNA pellets were air dried for 20 min and then dissolved in 50 µL of nuclease free water per sample. DNase treatment was given to the RNAs and quantification was done using a nanodrop spectrophotometer (Thermo Scientific).

cDNA synthesis, amplification, and cloning of partial and full *Mi-col-5* gene from *M. incognita*

First strand cDNA was synthesized from 300 ng of total RNA of egg masses, J2s, and adult females using Verso first strand cDNA synthesis kit (Thermo Scientific). A 402 bp segment of *Mi-col-5* gene was amplified from the first strand cDNA using primers Col-5-F and Col-5-R (Table 1), designed from already available sequence of *Mj-col-5* from *M. javanica* (Accession No. AF289026.1). The volume of each PCR amplification reaction was 25 µL containing 100 ng of first strand cDNA, 1 × Taq buffer, 10 mmol/L dNTP, 20 µmol/L of each primer, 3.5 mmol/L MgCl₂, and 1.5 U Taq DNA polymerase (Fermentas). The PCR product was purified using geneJET gel extraction kit (Thermo scientific) and cloned into pGEMT easy vector (Promega) using manufacturer's protocol. The recombinant

Table 1 List of primers used and their sequences

S. n.	Primer name	Primer sequence
1	Col-5-F	CAAATCGGAAAAGGACACGTAG
2	Col-5-R	TGGTCCTTTAGCACCAGCAG
3	Col-5-full-F	ATGGAACCTAAAGAGCAGTT
4	Col-5-full-R	TTAATGATATCCACCACTTT
5	Col-5-qrt-F	ACCGAGTTAAACGTGGTTGG
6	Col-5-qrt-R	GGGCCTTGAGATATTGCTGA
7	MI-18 s rRNA-F	TCAACGTGCTTGCTCTACCCTGAA
8	MI-18 s rRNA-R	TGTGTACAAAGGGCAGGGACGTAA

plasmids were transformed into freshly prepared competent cells of *E. coli* DH5 α . The positive clones were selected by blue-white screening using ampicillin (50 mg/L), IPTG (0.5 mM), and X-gal (80 μ g/mL) and sequenced using ABI solid sequencing platform. The partial sequence of *Mi-col-5* thus obtained was submitted to Genbank (Accession No. KF411439.1).

Full *Mi-col-5* sequence was retrieved from G-browse available at *Meloidogyne incognita* resources (http://www6.inra.fr/meloidogyne_incognita). Primers Col-5-full-F and Col-5-full-R (Table 1) were designed manually, and full gene was amplified from first strand cDNA of the adult females. Advantage 2 PCR kit was used for amplification of the full *Mi-col-5* gene using manufacturer's protocol. A 1047 bp long PCR product was purified and cloned into pGEMT easy vector and sequenced as described above. The full gene sequence of *Mi-col-5* was submitted to NCBI database (Accession No. KX372291).

In silico analysis of *Mi-col-5*

The amino acid sequence of *Mi-col-5* was deduced from the nucleotide sequence using expasy translate tool (<http://web.expasy.org/translate/>). The amino acid and nucleotide sequences of *Mi-col-5* were analyzed using NCBI blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Physical and chemical parameters of the predicted amino acid sequence were computed using ProtParam (Gasteiger et al. 2005). Clustal Omega was used for alignment between amino acid sequences of *Mi-col-5* and *Mj-col-5* (Sievers et al. 2011). Domain architecture analysis was done using SMART and MOTIF search (Letunic et al. 2015). SOPMA was used for the analysis of secondary structure of the predicted amino acid sequence of *Mi-col-5* (Combet et al. 2000). Multiple sequence alignment of the deduced amino acid sequence with cuticle collagen proteins identified in *M. incognita* and other species was done using Clustal Omega (Roy et al. 2010). A phylogenetic tree was constructed by maximum likelihood method using MEGA6 (Tamura et al. 2013). Intrinsic folding and unfoldability of *Mi-col-5* was predicted using Fold Index (Prilusky et al. 2005). Prediction of transmembrane domain for *Mi-col-5* and other cuticle collagen proteins was performed using TMHMM server V. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>), and signal peptide prediction was performed through SignalP 4.1 server (Petersen et al. 2011). Subcellular localisation predictor CELLO v2.5 server (Yu et al. 2006) was used for prediction of subcellular localization of *Mi-col-5*. Three-dimensional structure of the protein *Mi-col-5* was predicted by using I-TASSER server (Roy et al. 2010; Yang et al. 2015). The predicted model was evaluated using PROCHECK server (Laskowski et al. 1996).

Differential expression analysis of *Mi-col-5* through q-PCR

Expression of *Mi-col-5* at different stages of the life cycle of *M. incognita* was quantified and analyzed through real-time q-PCR. Total RNA was isolated from egg masses, J2s, and adult females and first strand cDNA was synthesized as described above. qPCR was performed with gene-specific primers for *Mi-col-5* (Table 1) in a CFX96 real-time system (Biorad) using 2X brilliant III SYBR Green q-PCR master mix (Agilent). 18 s rRNA was used as a reference gene for normalization of gene expression levels (Table 1). The amplification reactions were run using the following program: Initial denaturation of 95 °C for 5 min, followed by 35 cycles of 95 °C for 30 s and 57 °C for 1 min. After 35 cycles, a melt curve analysis or dissociation program (95 °C for 15 s, 57 °C for 15 s followed by a slow ramp from 57 to 95 °C) was acquired to ensure the specificity of amplification. Two biological and three technical replicates were used for each of the samples. After obtaining the Ct values, $2^{-\Delta\Delta C_T}$ method was used to quantify the relative fold change in gene expression and Student's *t* test ($p < 0.05$) was performed (Livak and Schmittgen 2001).

Results and discussion

Isolation and cloning of *Mi-col-5* from *M. incognita*

Among root knot nematodes, only three cuticle collagen genes viz. *Lemmi-5*, *mi-col-1*, and *mi-col-2* have been identified in *M. incognita*, while *Mjcol-3* and *Mj-col-5* have been identified in *M. javanica* (Koltai et al. 1997; Liu et al. 2001; Ray and Hussey 1995; Van Der Eycken et al. 1994; Wang et al. 1998). In potato cyst nematode *Globodera pallida*, cuticle collagen genes *gp-col-1*, *gp-col-2* and *gp-col-8* have been isolated and characterized (Gray et al. 2001; Jones et al. 1996). In this study, we have isolated and characterized a putative cuticle collagen gene, *Mi-col-5*, in *M. incognita*. Initially a partial segment of 402 bp of the gene *Mi-col-5* was amplified from adult female cDNA of *M. incognita* using the primers designed based on the already available sequence of *Mj-col-5* (Liu et al. 2001). After cloning and sequencing of this partial segment (Fig. 1), its sequence was submitted to NCBI Genbank which showed 99% sequence identity with *Mj-col-5*. The partial sequence was further analyzed on G-browse platform of 'Resources for *Meloidogyne incognita*' to predict the full sequence of the gene based on whole genome sequence data of *M. incognita*. A 1593 bp long DNA segment was predicted to code for *Mi-col-5* out of which the exonic sequence consisted of 1047 bp specifying the open reading frame. The full cDNA sequence was

Fig. 1 Restriction digestion of pGEMT clones with EcoRI releasing a 402 bp partial fragment of *Mi-col-5*. Lane M 1 Kb ladder. lanes 1–4 positive pGEMT clones. lanes 5–7 negative plasmids

Fig. 2 Amplification of full *Mi-col-5* gene from adult female cDNA. Lane M, 1 Kb plus ladder. lane 1 adult female cDNA

amplified (Fig. 2), cloned in pGEMT easy vector, sequenced and submitted to NCBI. At the nucleotide level, this sequence showed 98% sequence identity with *Meloidogyne enterolobii* collagen mRNA (accession no. KU350654.1) and 96% sequence identity with *M. javanica* *Mj-col-5* gene.

In silico analysis of *Mi-col-5*

A predicted amino acid sequence of 348 residues was deduced using the cDNA sequence of *Mi-col-5* with a calculated molecular mass of 35.12 kD and a theoretical pI of 5.79. Primary structural properties of *Mi-col-5* are listed in Table 2. The predicted protein sequence showed 92% identity with *Mj-col-5* from *M. javanica* (accession no. AAK83075.1), 67% identity with CRE-ROL-8 protein from *Caenorhabditis remanei* (accession no. XP_003113796.1) and 66% identity with ROL-8 protein

Table 2 Primary structural properties of *Mi-col-5*

S. n.	Parameter	Theoretical prediction
1	Molecular weight (kDa)	35.12
2	Isoelectric point	5.79
3	Total no. of negatively charged residues (Asp + Glu)	33
4	Total no. of positively charged residues (Arg + Lys)	27
5	Extinction coefficient ($M^{-1} \text{ cm}^{-1}$, at 280 nm)	27,430
6	Instability index	49.76
7	Aliphatic index	49.08
8	Grand average of hydropathicity (GRAVY)	−0.625

from *C. elegans* (accession no. NP_495582.1). Predicted amino acid sequence of *Mi-col-5* suggests that it belongs to group 2 of the cuticle collagen genes based on the pattern of conserved cysteine residues according to the classification proposed by Johnstone 2000 (Fig. 3). However, Like *Mj-col-5*, *Mi-col-5* also has a 12 amino acid longer carboxy terminal than other members of the group 2 with an additional tyrosine residue. Since the tyrosine residues in the carboxy terminus of the nematode cuticle genes are supposedly involved in the collagen cross linking, the presence of an extra tyrosine residue in *Mi-col-5* suggests a different pattern of its cross linking compared to other members of the group 2 (Cox 1992).

Extinction coefficient (at 280 nm) was calculated to be 27430 ($M^{-1} \text{ cm}^{-1}$) assuming all pairs of cysteine residues form cystines and 27930 ($M^{-1} \text{ cm}^{-1}$) assuming all cysteine residues were reduced. The aliphatic index and Grand Average of Hydropathicity (GRAVY) were found to be 49.08 and −0.625, respectively. 63.5% of the amino acids in the predicted protein are hydrophobic (Ala, Cys, Gly, Ile, Leu, Met, Phe, Pro, Val) while 36.5% of the amino acids were found to be polar (Arg, Asn, Asp, Gln, Glu, His, Lys, Ser, Thr, Trp, Tyr). Secondary structure prediction revealed that 50 amino acid residues (14.37%) were involved in alpha helices, 44 residues (12.64%) in extended strand, 16 residues (4.60%) in β -turn, and 238 residues (68.39%) in random coil. On comparison of the predicted secondary structures of *Mi-col-5* and *Mj-col-5*, most of the difference was observed within the first 50 amino acids with deviation in the distribution of α helices in this region (Fig. 4). TMHMM results indicate the presence of a transmembrane helix in *Mi-col-5* between amino acids positions 21–43 involving approximately 22.5 amino acid residues (Fig. 5). Interestingly, most of the difference between *Mi-col-5* and *Mj-col-5* lies in the amino acid composition of this region with the latter lacking a transmembrane helix (Fig. 6). However, transmembrane helices

Fig. 5 TMHMM predictions for *Mi-col-5* (a), *Mj-col-5* (b). Blue, red, and pink colors represent the portions of the amino acids inside, in the transmembrane helix and outside the cell, respectively

Fig. 6 Clustal Omega analysis of the *Mi-col-5* and *Mj-col-5* amino acid sequences. Area under the black box shows the transmembrane helix region of *Mi-col-5* and its difference with *Mj-col-5*

Mi-col-5	MEPKQFCLKETDEHRQMRRIAFVAVVVSTVAVIASVVTLPMLYSYVQSFQSHLIVEIDH	60
Mj-col-5	MEPKQFCLKETDEHRQMRRIAFIFEKNVFLKDEHRQMRRIAFYSYVQSFQSHLIVEIDH	60

Mi-col-5	CKAKSRDMWLEMTALQIGKGVDRVKRGWLFQQWVPENGYEPAQTGPSNTVQSAISQGPS	120
Mj-col-5	CKAKSRDMWLEMTALQIGKGVDRVKRGWLFQQWVPENGYEPAQTGPSNTVQSAISQGPS	120

Mi-col-5	GATYGQGAAGYQPVVAPPEPAPVCCTCHQGGPPGPIGPEGEPDGEDGPNKGDTSGKDAQ	180
Mj-col-5	GATYGQGAAGYQPVVAPPKPAPVCCTCHQGGPPGPIGPEGEPDGEDGPNKGDTSGKIDAR	180

Mi-col-5	ILPAPLEPPCIICPPGPAGPQGPAGAKGPPGPKGSPGEPKDGVPGEQGMAGQHGPGRP	240
Mj-col-5	ILPAPLEPPCIICPPGPAGPQGPAGAKGPPG--SLGEPKDGVPGEQGMVQHGPPGRP	237

Mi-col-5	GREGRPAGPSGRLIPVPGPQGPAGPPGVGPPGAPGAAGPPGQSFEGPPGPPGEPGRP	300
Mj-col-5	GREGRPAGPSGRLIPVPGPQGPAGPPGVGPPGAPGAAGPPGQSFEGPPGPPGEPGRP	297

Mi-col-5	GREGRPGGPGAPPPGQDGEKGSCEHCPEPTPPGYFAEASAKSGGYH	348
Mj-col-5	GREGRPGGPGAPPPGQDGEKGSCEHCPEPTPPGYFAEASAKSGGYH	345

avored region, while 43.5% of the residues were in the additionally allowed region, 17.1% residues in the generously allowed region, and 6.9% residues in the disallowed region. However, PROCHECK analysis of Model 3 with a C-score of -4.47 showed 52.8% residues in the most favored region, 34.7% residues in the additionally allowed region, 6% residues in the generously allowed region, and 6.5%

residues in the disallowed region. An important aspect of the Ramachandran plots for the predicted models is G factors, which is a combination of different parameters like psi-psi distribution, chi1-chi2 distribution, etc. The overall average G factors value for model 1 is -1.45 which falls into highly unusual category, while that for model 3 is -0.99 falling into unusual category.

Fig. 7 Folding predictions for *Mi-col-5*. Predicted folded and unfolded regions of *Mi-col-5* (a), amino acid sequence for *Mi-col-5* showing ordered segments in green and disordered segments in red (b)

Expression analysis of *Mi-col-5* by qRT PCR

Stage-specific expression analysis of *Mi-col-5* transcript was investigated in egg masses, J2s, and adult females. Taking expression levels of *Mi-col-5* at J2 stage as reference, 27-fold upregulation was observed in the expression in egg masses, whereas a 12-fold upregulation was observed in the adult females (Fig. 10). In *C. elegans*, increased expression of cuticle collagen genes has been observed during cuticle synthesis just before molting (Johnstone and Barry 1996). Similarly, a higher expression in the egg masses suggests the role of *Mi-col-5* in the earlier stages of nematode cuticle development, while the expression in the adult females may be attributed to the role of *Mi-col-5* in maintenance and thickening of adult female cuticle. These results are in agreement with the expression pattern of *Mj-col-5* (Liu et al. 2001). Wang et al. (1998) reported expression patterns of *Mi-col-1*, *Mi-col-2*, and *Lemmi-5* showing highest level of expression in adult females followed by parasitic J3s or J4s and parasitic J2 s for all the three genes. Expression of none of these genes could be detected in pre-parasitic J2 s. Similarly, a higher expression of *gp-col-1* and *gp-col-2* was detected in virgin and adult females of *G. pallida* (Gray et al. 2001). In *M. javanica*, the expression pattern of *Mj-col-3* showed a

Fig. 8 Phylogenetic tree constructed by maximum likelihood method using MEGA 6 showing evolutionary relationship of *Mi-col-5* with cuticle collagens of other plant parasitic, animal parasitic, and free living nematodes

Fig. 9 Molecular modeling of *Mi-col-5* using I-TASSER server. Predicted 3D models of *Mi-col-5* representing model 1 and model 3, respectively, (a) and (c), Ramachandran plots for predicted model 1 and model 3, respectively, (b) and (d)

higher expression in developing eggs followed by J4 while the transcription level was very low in adult females (Koltai et al. 1997). Interestingly, cuticle collagen genes *gp-col-1*, *gp-col-2*, *Lemmi-5*, *Mi-col-1*, and *Mi-col-2* having similar expression patterns were placed in the same cluster (cluster II) in the phylogenetic tree while *Mi-col-5*, *Mi-col-3*, and *Mj-col-5* representing cluster I exhibited similar expression patterns. This suggests a strong correlation between the phylogenetic relationship and expression patterns of cuticle collagen genes in plant parasitic nematode. The differential expression of different cuticle

collagen genes at different stages in root-knot nematodes suggests the involvement of these genes at different developmental stages and molting of the nematodes.

Conclusion

A full length cuticle collagen gene from *M. incognita* was isolated, characterized with the help of bioinformatic tools and its differential expression was studied at various developmental stages in the life cycle of the nematode. A

Fig. 10 Expression of *Mi-col-5* in different developmental stages of *M. incognita*

comparison of this gene with other characterized cuticle collagen genes from PPNs was also made leading to understanding of its evolutionary relationships and classification. Further studies focused on RNAi-based silencing of *Mi-col-5* can provide a more detailed insight into the role of this gene in the development and maintenance of the cuticle. Development of transgenic plants expressing dsRNAs of this cuticle collagen gene can be another future approach for engineering resistance against *M. incognita* by hindering the structural organization of the nematode cuticle.

Acknowledgements The authors are thankful to Indian Council of Agricultural Research for providing financial support through NFBSFARA project.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interest regarding the publication of this article.

References

Abad P, Gouzy J, Aury JM, Castagnone-Sereno P, Danchin EG, Deleury E, Perfus-Barbeoch L, Anthouard V, Artiguenave F, Blok VC, Caillaud MC, Coutinho PM, Dasilva C, De Luca F, Deau F, Esquibet M, Flutre T, Goldstone JV, Hamamouch N, Hewezi T, Jaillon O, Jubin C, Leonetti P, Magliano M, Maier TR, Markov GV, McVeigh P, Pesole G, Poulai J, Robinson-Rechavi M, Sallet E, Segurens B, Steinbach D, Tytgat T, Ugarte E, van Ghelder C, Veronico P, Baum TJ, Blaxter M, Blev-Zacheo T, Davis EL, Ewbank JJ, Favery B, Grenier E, Henrissat B, Jones JT, Laudet V, Maule AG, Quesneville H, Rosso MN, Schiex T, Smant G, Weissenbach J, Wincker P (2008) Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat Biotechnol* 26:909–915

Combet C, Blanchet C, Geourjon C, Deleage G (2000) NPS@: network protein sequence analysis. *Trends Biochem Sci* 25:147–150

Cox GN (1992) Molecular and biochemical aspects of nematode collagens. *J Parasitol* 78:1–15

Davies KG, Curtis RSG (2011) Cuticle surface coat of plant-parasitic nematodes. *Ann Rev Phytopathol* 49:135–156

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the expasy server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana, Totowa, pp 571–607

Gray LJ, Curtis RH, Jones JT (2001) Characterisation of a collagen gene subfamily from the potato cyst nematode *Globodera pallida*. *Gene* 263:67–75

Hooper DJ (1986) Extraction of free-living stages from soil. In: Southey JF (ed) *Laboratory methods for work with plant and soil nematodes*. Ministry of agriculture, fisheries and food, London, pp 5–30

Johnstone IL (2000) Cuticle collagen genes expression in *Caenorhabditis elegans*. *Trends Genet* 16:21–27

Johnstone IL, Barry JD (1996) Temporal reiteration of a precise gene expression pattern during nematode development. *EMBO J* 15:3633–3639

Jones JT, Curtis RH, Wightman PJ, Burrows PR (1996) Isolation and characterization of a putative collagen gene from the potato cyst nematode *Globodera pallida*. *Parasitology* 113:581–588

Kingston IB (1991) Nematode collagen genes. *Parasitol Today* 7:11–15

Koltai H, Chejanovsky N, Raccach B, Spiegel Y (1997) The first isolated collagen gene from *Meloidogyne javanica* is developmentally regulated. *Gene* 196:191–199

Kramer JM (1994) Structures and functions of collagens in *Caenorhabditis elegans*. *FASEB J* 8:329–336

Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8:477–486

Letunic I, Doerks T, Bork P (2015) SMART: recent updates, new developments and status in. *Nucleic Acids Res* 43:D257–D260

Liu J, Koltai H, Chejanovsky N, Spiegel Y (2001) Isolation of a novel collagen gene (*MJ-COL-5*) in *Meloidogyne javanica* and analysis of its expression pattern. *J Parasitol* 87(4):801–807

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25:402–408

Page AP, Johnstone IL (2007) The cuticle. In: *The C. elegans research community* (eds.) *WormBook*. doi:10.1895/wormbook.1.138.1

Page AP, Winter AD (2003) Enzymes involved in the biogenesis of the nematode cuticle. *Adv Parasit* 53:85–148

Page AP, Stepek G, Winter AD, Pertab D (2014) Enzymology of the nematode cuticle: a potential drug target? *Int J Parasitol Drugs Drug Resist* 4:133–141

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8(785–786):1–15

Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 21(16):3435–3438

Ray C, Hussey RS (1995) Evidence for proteolytic processing of a cuticle collagen in a plant-parasitic nematode. *Mol. Biochem Parasit* 72:243–246

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5:725–738

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729

- Van Der Eycken W, Dealmeidaengler J, Vanmontagu M, Gheysen G (1994) Identification and analysis of a cuticular collagen coding gene from the plant-parasitic nematode *Meloidogyne incognita*. *Gene* 151:237–242
- Wang T, Deom CM, Hussey RS (1998) Identification of a *Meloidogyne incognita* cuticle collagen gene and characterization of the developmental expression of three collagen genes in parasitic stages. *Mol Biochem Parasit* 93:131–134
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 12:7–8
- Yu CS, Chen YC, Lu CH, Hwang JK (2006) Prediction of protein subcellular localization. *Proteins* 64:643–651