Molecular & Biochemical Parasitology 167 (2009) 32–40

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Expressed sequence tags of the peanut pod nematode *Ditylenchus africanus*: The first transcriptome analysis of an Anguinid nematode☆

Annelies Haegeman[a], Joachim Jacob[a], Bartel Vanholme[a,1], Tina Kyndt[a], Makedonka Mitreva[b], Godelieve Gheysen[a,*]

[a] *Faculty of Bioscience Engineering (FBE), Department of Molecular Biotechnology, Ghent University, Coupure links 653, B-9000 Ghent, Belgium*
[b] *The Genome Center, Washington University School of Medicine, 4444 Forest Park Boulevard, St. Louis, MO, USA*

ARTICLE INFO

Article history:
Received 28 October 2008
Received in revised form 7 April 2009
Accepted 12 April 2009
Available online 19 April 2009

Keywords:
Transcripts
Parasitism genes
Gene ontology
Anhydrobiosis
RNAi

ABSTRACT

In this study, 4847 expressed sequenced tags (ESTs) from mixed stages of the migratory plant-parasitic nematode *Ditylenchus africanus* (peanut pod nematode) were investigated. It is the first molecular survey of a nematode which belongs to the family of the Anguinidae (order Rhabditida, superfamily Sphaerularioidea). The sequences were clustered into 2596 unigenes, of which 43% did not show any homology to known protein, nucleotide, nematode EST or plant-parasitic nematode genome sequences. Gene ontology mapping revealed that most putative proteins are involved in developmental and reproductive processes. In addition unigenes involved in oxidative stress as well as in anhydrobiosis, such as LEA (late embryogenesis abundant protein) and trehalose-6-phosphate synthase were identified. Other tags showed homology to genes previously described as being involved in parasitism (expansin, SEC-2, calreticulin, 14-3-3b and various allergen proteins). *In situ* hybridization revealed that the expression of a putative expansin and a venom allergen protein was restricted to the gland cell area of the nematode, being in agreement with their presumed role in parasitism. Furthermore, seven putative novel candidate parasitism genes were identified based on the prediction of a signal peptide in the corresponding protein sequence and homologous ESTs exclusively in parasitic nematodes. These genes are interesting for further research and functional characterization. Finally, 34 unigenes were retained as good target candidates for future RNAi experiments, because of their nematode specific nature and observed lethal phenotypes of *Caenorhabditis elegans* homologs.

1. Introduction

Expressed sequence tag (EST) analysis is a relatively cheap and rapid method to obtain a first molecular impression of a species. The technique consists of a random selection of clones from a cDNA library and sequencing of their inserts. Although it is mainly used for gene discovery [1], it can also be used for other goals, e.g. estimation of gene expression level [2], detection of single nucleotide polymorphisms [3], or improving genome annotation [4]. In nematology, the technique is widely used and to date over 1 million ESTs from over 60 species are available (dbEST, NCBI GenBank). Some of these ESTs are derived from cDNA libraries generated from specific life stages or from specific tissues of the nematode. In the case of plant-parasitic nematodes, the transcriptional activity of the pharyngeal glands is of particular interest since gland proteins are injected in the plant tissue during the nematode–host interaction. EST analyses have led to the discovery of many of these parasitism genes that code for hydrolytic enzymes, such as pectate lyase [5] and xylanase [6]. They are involved in maceration of the plant cell walls during migration of the nematode in the plant tissue.

The peanut pod nematode *Ditylenchus africanus* (Wendt, Swart, Vrain and Webster, 1995) was first found in hulls and seeds of groundnut (*Arachis hypogea* L.) in South Africa [7]. It was initially identified as *Ditylenchus destructor*, the potato rot nematode, but experiments showed that it caused no damage to different potato varieties [8]. Therefore the South African population was eventually considered a new race, but later molecular data revealed that this organism rather had to be considered as a new species [9]. Within the nematode order of the Rhabditida, it belongs to the family of the Anguinidae and superfamily Sphaerularioidea [10]. *D. africanus* is a migratory endoparasite with peanut as the main host, causing black discoloration of the seeds and pods. Besides being a parasite of plants, it can also feed and reproduce on the hyphae of com-

☆ *Note:* EST sequence data reported in this paper are available in the GenBank™, EMBL and DDBJ databases under the accession numbers from FE920352 to FE925198.

* Corresponding author. Tel.: +32 92645888; fax: +32 92646219.
E-mail address: godelieve.gheysen@ugent.be (G. Gheysen).
[1] Present address: Department of Plant Systems Biology, Flanders Interuniversitary Institute of Biotechnology, Technologiepark 927, B-9052 Zwijnaarde, Belgium.

0166-6851/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2009.04.004

mon plant pathogenic fungi such as *Aspergillus parasiticus*, *Fusarium oxysporum*, *Botrytis cinerea* and *Rhizoctonia solani*. It has a short life cycle of only 6–7 days at its optimal temperature of 28 °C and can survive long periods of drought by anhydrobiosis [8]. Of several economically important genera in the Anguinidae family, *Ditylenchus* spp. have the widest impact on agriculture [11]. *Ditylenchus dipsaci*, a species complex with an extremely wide host range, is one of the most devastating plant-parasitic nematodes, especially in temperate regions (www.eppo.org). *D. destructor* is an important pest of potato tubers in Europe and North America, while *D. africanus* is a huge problem in the cultivation of groundnuts in South Africa. It was shown that 73% of the seeds collected from different growers was infected [7], and that this mainly has a qualitative effect on the groundnuts leading to a significantly lower income for the farmers [12].

D. africanus is an interesting nematode to subject to EST analysis for several reasons. First, no molecular knowledge is available for this species or any other member of the Anguinidae. Second, it has a different taxonomic classification in contrast to other EST studies examining plant-parasitic nematodes, which focused mainly on species from the superfamily Tylenchoidea (grouping cyst and root-knot nematodes as well as migratory nematodes such as *Radopholus* and *Pratylenchus*). Third, it is a plant-parasitic nematode facultatively feeding on fungi. Only one nematode with a similar feeding habit has been investigated using a comparable approach (*Bursaphelenchus xylophilus*; superfamily Aphelenchoidea) [13]. Finally, it adds to the EST dataset of migratory nematodes, therefore empowering comparative studies. The focus for small-scale EST projects is gradually shifting from sedentary nematodes towards migratory nematodes. In the last few years, EST data have become available for several migratory nematodes such as *Pratylenchus penetrans* [14], *Radopholus similis* [15], *B. xylophilus* [13] and *Xiphinema index* [16].

2. Materials and methods

2.1. Nematode culture, cDNA library construction and EST generation

D. africanus was cultured at 25 °C on carrot discs in small Petri dishes (Ø 35 mm) under sterile conditions. Carrot discs were infected with approximately 100 individual nematodes. Six weeks after inoculation, nematodes were collected by rinsing the carrot discs with sterile water. RNA was extracted from approximately 10,000 individuals with TRIzol (Invitrogen, Carlsbad, CA, USA) as described by Jacob et al. [15]. A cDNA library was constructed using the SMART cDNA Library Construction Kit (Clontech, Palo Alto, CA, USA) according to the manufacturer's instructions, starting from 1 μg RNA. The resulting *D. africanus* mixed stage library contained over 8×10^5 primary transformants. Random colonies were sequenced using the M13 forward or reverse primer at the Genome Center (Washington University, St.-Louis, MO, USA). Resulting sequences were submitted to the EST division of GenBank (dbEST).

2.2. Cleaning and clustering of the EST sequences

The sequences were cleaned using Seqclean (www.tigr.org) with a locally downloaded vector database and default parameter settings. The cleaned dataset was clustered using the TIGR Gene Indices Clustering Tool (TGICL) [17] and sequences were assembled by CAP3 [18] using default settings. The fragmentation (i.e. the percentage of unigenes which are redundant) was estimated with ESTstat [19]. The obtained unigenes served as a basis for following analyses.

2.3. Homology searches

Basic Local Alignment Search Tool (blast) analyses [20] were done both locally and by using netblast. Blastx and blastn searches were conducted with all unigenes against the NCBI protein and nucleotide database. Additionally, blastn and tblastx searches were done against the genomes of *Meloidogyne incognita* (http://meloidogyne.toulouse.inra.fr/) [21] and *Meloidogyne hapla* (http://www.hapla.org/) [22]. Since most data available for nematodes are in the EST database, a tblastx search against all nematode ESTs was done. In-house perl scripts parsed the resulting hits for species names and unigenes were subsequently classified into different categories (nematodes, invertebrates, plants, animals, fungi, prokaryotes) according to the species names derived from the blast hits. In parallel, a blastx search was conducted against proteins of the model organism *Caenorhabditis elegans*. Resulting top hits were searched for RNAi phenotypes using WormBase [23]. For all *C. elegans* homologs with an RNAi lethal phenotype, gene ontology (GO) terms were retrieved and visualized with WEGO [24]. Blastn against all mitochondrial nematode genes revealed putative mitochondrial unigenes.

To annotate GO terms, BLAST2GO [25] was used on all unigenes using default parameters. Blastx value cut-off was chosen at $E < 1e{-}4$. GOSlim view was used and GO graphs were generated with a node scoring filter of 25 for "molecular function", 50 for "biological process" and 12 for "cellular component".

2.4. Translation into putative proteins

All unigenes were translated with OrfPredictor [26]. The blastx output was used to select the correct reading frame for translation. The minimum amino acid number for predicted protein sequences was set to 40. Nucleotide sequences were trimmed to their coding parts and the overall GC content as well as the average GC content of the first (GC1), second (GC2) and third nucleotide (GC3) of the codons was calculated. Unigenes without predicted open reading frame were considered to be non-coding, and for these sequences only the GC content was calculated. Signal peptides for secretion in the predicted proteins were predicted by SignalP 3.0 [27]. A signal peptide was only assigned to a sequence when both the Hidden Markov Model (HMM) and the neural network predicted its presence. Moreover, a transmembrane domain search (http://protfun.net/services/TMHMM/) on the mature putative proteins revealed whether the protein was retained in the membrane.

2.5. Dot blot analysis

Forward (F) and reverse (R) primers were developed to amplify fragments of CL1, CL7, CL270, CL371, CL406, FE921742 and FE922861 (Table 1) by polymerase chain reaction (PCR). The reaction mixture contained 300 ng purified plasmids of the cDNA library as template, 0.5 μM of each primer, 4 mM dNTPs, 1.5 mM $MgCl_2$, 20 mM Tris–HCl (pH 8.3), 50 mM KCl and 1 U of *Taq* DNA polymerase (Invitrogen, Carlsbad, CA, USA). PCR conditions were as follows: 35 cycles of 30 s at 94 °C, 30 s at 54 °C and 40 s at 72 °C. PCR products were loaded on a 0.5× TAE 1.5% agarose gel and the resulting fragments were ligated into pGEM-T (Promega, Madison, WI, USA) and the ligation mixture was used for transformation into *Escherichia coli* DH5α cells. Transformed cells were selected on LB agar plates supplemented with 100 μg/ml carbenicillin. Plasmids of positive colonies were extracted using the Nucleobond AX Kit (Macherey-Nagel, Düren, Germany) and the inserts were sequenced at AGOWA (Berlin, Germany). Probes were generated by a PCR under conditions as described above with purified plasmid with the correct insert as template. Resulting PCR products