RESEARCH **Open Access**

Identification, isolation, and cloning of the full-length sequence of a gene encoding trypsin inhibitor-like protein (TIL) secreted by the intestinal parasitic nematode *Strongyloides ratti*

Abuelhassan Elshazly Younis[1,2]* and Norbert W. Brattig[2]

Abstract

Background: Intestinal nematodes may interact with the intestinal mucosal host cells through molecules they produced actively in the host environment. Many excretory/secretory products (ESP) of the intestinal nematodes are considered as interesting candidates to be examined in vitro on their tolerance-inducing activities. Nematode-derived protease inhibitors are believed to have multiple, specific abilities for the host immunomodulation, playing on various immune effector mechanisms. Among the previously identified 500 ESP from *Strongyloides ratti*, trypsin inhibitor-like (Sra-TIL) peptides were detected in high concentration in the ESP of the parasitic female (pF). The aim is to identify interesting tolerance-inducing proteins belonging to the longtime adapted human biome which may be potentially applied to moderate inflammatory immune reactions.

Results: In the present study, the full-length cDNA (from the parasitic female) and gDNA of the ~ 1605 bp *Sra-til* is identified and isolated by polymerase chain reaction (PCR), subsequent to cluster/EST analysis via alignment and BLAST search. gDNA and RNA were extracted from infective third stage larvae (iL3) and pF respectively, and the cDNA was synthesized.Then, the *Sra-til* was cloned and sequenced. The genomic organization analysis revealed that introns were absent in the *Sra-til* gene structure. Sequence analyses of both nucleotide and predicted protein sequences were investigated using different bioinformatic tools. The predicted architecture of the translated amino acid sequence showed the presence of a signal peptide for secretion. The architecture of *Sra*-TIL was *Sra*-TIL = TIL, EGF (epidermal growth factor) × 6). Phylogenetic analysis of the deduced amino acid sequence showed how related this novel *Strongyloides* protein to TIL orthologs from other nematodes.

Conclusion: The full-length *Sra-til* gene which encodes an interesting ESP candidate, secreted from pF, was identified, isolated, cloned, and then sequenced. Both nucleotide and amino acid sequences were analyzed by various bioinformatic tools. The predicted innovative structural advantages of *Sra*-TIL polypeptides secreted from a parasitic intestinal stage of *S. ratti* nematode is probably reflective of the abundance of different biological functions, including the protection against host proteases, host immunomodulation via proliferation induction of the damaged intestinal cells as well as wound healing properties.

Keywords: Strongyloides, Excretory/secretory, Trypsin inhibitor-like protein

---

* Correspondence: abssan2000@aswu.edu.eg
[1]Zoology Department, Faculty of Science, Aswan University, Aswan 81528, Egypt
[2]Department of Molecular Medicine, Bernhard Nocht Institute for Tropical Medicine, Bernhard-Nocht-Strasse 74, 20359 Hamburg, Germany

Background

Parasitic worm infections, including nematode infections, represent one of the most prevalent problems in human and veterinary medicine (Newton & Munn, 1999). Soil-transmitted helminths, commonly known as intestinal worms, are the most common infections worldwide, affecting the most deprived communities where infected people generally cannot afford treatment. More than 2 billion humans are infected with gastrointestinal or tissue nematodes and 3.5 billion are exposed to them, which results in tremendous health and economic problems (Chan, 1997; Hotez, 2008). Since infections have tendencies to be chronic, they are destructive to severely infected children, giving rise to anemia, growth retardation, impaired cognitive function, productivity of individuals, and lowered educational accomplishment (Cooper & Bundy, 1988; Guyatt, 2000; Nokes et al., 1992).

Every infection represents a competition between the parasite and the host (Playfair & Bancroft, 2008). The important difference between a free-living animal and a parasite of vertebrates is that the parasite must survive and reproduce against the host defense mechanisms (Wakelin, 1996; Wakelin & Walliker, 1996). Parasites have produced a variety of molecules that intervene with the host's defense system aim to eliminate the undesirable lodger (Nagaraj, Gasser, & Ranganathan, 2008). This alteration is most likely caused by the liberation of soluble substances which bind, degenerate, or otherwise interact with host immune cells (Hewitson, Grainger, & Maizels, 2009; Lightowlers & Rickard, 1988). Molecules generated and released by nematodes that may interfere with the host immune responses include proteases, protease inhibitors, antioxidants, and homologies of host mediators and their receptors (Bungiro & Cappello, 2004).

The genus *Strongyloides* include about some 50 species of intestinal parasitic worms of vertebrates like mammals, birds, reptiles, and amphibians (Grove, 1989; Viney & Lok, 2007). *S. stercoralis*—the major human pathogen species—is an intestinal nematode that has the ability to tolerate host immune attack and survive within the human small intestine for decades and infects at least 100 million people (Concha, Harrington Jr, & Rogers, 2005; Liu & Weller, 1993). This prevalence is likely underestimated since the diagnostic tests are insensitive, and the development of accurate and sensitive methods is needed (Kramme et al., 2011; Montes, Sawhney, & Barros, 2010).

S. ratti is a parasite of rats, used commonly as a model nematode. The excretory/secretory proteins (ESP) from infective larvae (iL3), parasitic females (pF), and free-living stages (flS) of *S. ratti* were identified by LC-MS/MS (Soblik et al., 2011). Proteomic analysis of ESP revealed 79 proteins specific for parasitic females, including a prolyl oligopeptidase, small heat shock proteins, and a trypsin inhibitor-like protein (TIL). *Sra*-TIL

peptides were detected in high concentration in the ESP of the parasitic female. Quantitative RT-PCR was done to confirm that *Sra-til* is parasitic female-specific gene (Soblik et al., 2011). Previously, we characterized some interesting *S. ratti* ESP including heat shock proteins (*Sra*-HSP10 and *Sra*-HSP17s), a homolog of the human macrophage migration inhibitory factor (*Sra*-MIF) which interacts with host immunity in varying ways (Tazir et al., 2009; Younis et al., 2011; Younis et al., 2012).

Sra-TIL is selected in this study, for identification, separation, and cloning of the full-length sequence of the gene and prediction of its structure, function, and phylogeny because it was present among the 25 highest scoring proteins found only in pF ESP (Soblik et al., 2011). It is one of the most interesting ESP candidates, which may be involved directly in the modulation of the host immune responses.

Methods

Parasite culture, infection, and harvesting of pF

The *S. ratti* life cycle has been held at the Bernhard Nocht Institute for Tropical Medicine (BNITM, Hamburg) since January 2006. The genetically trusted iL3 for the initial infection were kindly supplied by Prof. Dr. Gerd Pluschke from the Swiss Tropical Institute (STI), Department of Medical Parasitology and Infection Biology, Basel, Switzerland. The life cycle is running in agreement with the German animal protection law. The experimental protocols have been reviewed and confirmed by the responsible federal health authorities of the State of Hamburg (Hamburg, Behoerde fuer Gesundheit und Verbraucherschutz: permission number 89/09). Four to 6-week-old Wistar rats (*Rattus norvegicus*) from Charles River were used to maintain the cycle by serial passage as described (Keiser, Thiemann, Endriss, & Utzinger, 2008; Viney & Lok, 2007). For the collection of different stages, standard methods were used which were established in our laboratory (Soblik et al., 2011; Younis et al., 2011).

Computer-based cluster/EST analyses

Peptides related to Cluster SR02054 and its EST (expressed sequence tag) were identified and searched on www.nematode.net (Martin et al., 2009). The ESTs were screened for more sequence information to get the longest nucleotide sequence which representing partial gene structure. The ESTs constituting this cluster were aligned (using the ClustalW program) and were assembled (manually and using Cap3 program) to get the possible longest DNA sequence and then translated (http://web.expasy.org/translate/) to get the deduced amino acid (aa) sequence.

RNA isolation and reverse transcription

Total RNA was isolated from the freshly prepared pF after 7 days of infection by standard phenol/chloroform method as described before (Tazir et al., 2009). Subsequently, RNA was quantified spectrophotometrically. The quality and integrity of the total RNA were confirmed after staining of 5 μg RNA by loading buffer (Roti®-RNA, ROTH), following the manufacturer's protocol, on an ethidium bromide-stained gel containing formaldehyde. RNA samples were then treated with RNase-free DNase I (Qiagen, Hilden, Germany) and purified with RNeasy MinElute spin column (Qiagen) following the manufacturer's protocol.

A total of ~ 5 μg of purified parasite RNA was used to synthesize the first-strand cDNA using SuperScript® III RT and GeneRacer (Invitrogen). The manufacturer's instructions were followed except for the antisense primer, oligo dT-T7I (Table 1), which was used at a final concentration of 10 μM.

Polymerase chain reaction (PCR)

The partial longest EST obtained from the EST alignment and assembly was used to design specific forward and reverse primers (Table 1). The primers qualities were checked using Oligo Calc software (http://biotools.-nubic.northwestern.edu/OligoCalc.html).

To get 3′-cDNA ends, specific forward primer (tilf1) was used for amplification of *Sra*-TIL, and the oligo dT-T7II was used as the reverse primer. The 5′-ends were obtained using the gene-specific reverse primer (tilr1) and using the 5′ SL-1 as forward primer, related to the nematode spliced leader sequence (Hunter et al., 1997).

The PCR was done as described before (Sambrook, Fritsch, & Maniatis, 1989). In brief, the PCR conditions were 1× thermo buffer (NEB), forward primer (10–20 pmol), reverse primer (10–20 pmol), dNTP mix (Invitrogen; 10 mM), Taq polymerase (NEB; 0.3 μl), DNA 0.1–0.5 μg, and HPLC-H$_2$O to 50 μl. The denaturation step was set to 4 min at 95 °C, and the elongation was performed at 72 °C for 45 s. The annealing temperatures (52–58 °C) are related to the primers' properties and their melting temperatures (Tm), which were calculated using the following formula: $Tm = (A + T) \times 2 + (G + C) \times 4$. The annealing time was set to 45 s, and 30 cycles were repeated.

Agarose gel electrophoresis

The results of the amplifications were tested on agarose gel. Agarose gel electrophoresis was performed as described (Sambrook et al., 1989) using TAE buffer and agarose concentration 1% (w/v). DNA was mixed with the DNA-loading buffer (Fermentas), and DNA ladder was used to monitor fragment size. Five microliters of ethidium bromide per 100 ml agarose solution was added before pouring the gel. The running buffer was 1X TAE and the voltage was set to 100 V. For RNA analysis, the method previously described (Farrell, 1993) was used, using 0.8% (w/v) agarose prepared in 1× MOPS buffer containing 0.7 M as a final concentration of formaldehyde. RNA was mixed with Roti®-RNA Loading buffer (Carl Roth, Karlsruhe, Germany); RNA ladder was used to monitor fragment size. Gels were run overnight in 1× MOPS as running buffer under 25 V. RNA or DNA bands were visualized using a UV transilluminator and immediately photographed.

Cloning and sequencing

Various PCR products were purified by QIAquick® PCR purification Kit (Qiagen, Hilden) and then cloned into pGEM-T Easy vector (Promega Corp., Madison, USA) according to the manufacturer's protocols. The cloned fragments were subsequently transformed into competent *Escherichia coli* TOP10 cells (Invitrogen), and the recombinant plasmids were purified from the bacterial cultures by QIAprep® Miniprep Kit (Qiagen, Hilden), according to the manufacturer's protocols. The resulted products were sequenced via LGC Genomics GmbH, Germany.

Alignment of the various cloned fragments resulted in the full length sequence. Subsequently, primers including the 3′- and the 5′-ends were designed (see Table 1), which included the codon for the initiating methionine and a stop codon. The full-length open reading frame

Table 1 Primers list used for *Sra-til* gene amplification. All primers were ordered from Eurofins MWG/Operon
https://www.eurofinsgenomics.eu/en/home.aspx

Primer name	Sequence (5'-3')	Target
oligodT-T7I	GAGAGAGGATCCAAGTACTAATACGACTCACTATAGGGAGATT24	cDNA synthesis
oligodT-T7II	GAGAGAGGATCCAAGTACTAATACGACTCACTATAGG	3'end
SL-1 (spliced leader)	GGTTTAATTACCCAAGTTTGAG	5'start
tilf1	CAAAAAAATCCACAAAAATGTGGTG	3'end with OdT-T7II (partial)
tilr1	ATCACCTGTTTTAGTGTTTCTTAC	5'start with SL-1 (partial)
tilf0	ATGAGACTATTTATTAGTTTAC	full length
tilr0	TTATGGTCTACTGCCTCTTGGACAT	full length

(ORF) of the *Sra-til* cDNA was captured by PCR using specific forward primer tilf0, and reverse primer tilr0.

The full-length cDNAs were cloned, after purification, into the pGEM-T Easy vector and transformed into TOP10 cells. The results confirmed by restriction digestion by EcoRI (NEB) according to the supplier's protocol then observed on ethidium bromide-stained gel. For further sequence confirmation, the fragments were sequenced using the M13 forward and gene-specific primers.

Extraction of gDNA and genomic organization

Genomic DNA was isolated from 100,000 iL3 using standard phenol/chloroform method. In brief, the parasite material digested for 3–4 h at 56 °C with proteinase K in ALT buffer (Dneasy kit, Qiagen) under a constant agitation. gDNA was treated with RNase A and was then precipitated with 5.2 M ammonium acetate and dissolved in 20–30 μl HPLC water. Concentrations were determined by spectrophotometry. PCRs were done for both cDNA and gDNA as mentioned above using the specific forward (tilf0) and reverse (tilr0) primers. The resulted PCR products were compared on gel, cloned, and sequenced as mentioned above to investigate the gene genomic organization.

Sequence and phylogenetic analyses

Homology search on the nucleotide and protein database was carried out using the BLAST program with default settings at NCBI (http://www.ncbi.nlm.nih.gov/). Subsequently, to draw an analogy, for pattern and profile deduction, for function and structure prediction and for further bioinformatic analysis, sets of software available at the Expert Protein Analysis System (ExPASy) proteomics server (http://expasy.org/tools/) of the Swiss Institute of Bioinformatics were used, including Simple Modular Architecture Research Tool (SMART). SignalP program (http://www.cbs.dtu.dk/services/SignalP/) which was used for signal peptide prediction. Thereafter, the full length of the *Sra-til* sequence was aligned with the most homologous sequences in the database using CLUSTALW multiple sequence alignment program. The phylogenetic trees were built using online tool software Phylogeny.fr program (Dereeper et al., 2008).

Results

Identification of *Sra-til* full-length cDNA and gene structure

High-quality RNA from pF (Fig. 1) was used in the cDNA synthesis. Initially, the cluster SR02054 was identified by database searches which were performed using the Applied Biosystems ProteinPilotTM search engine (Soblik et al., 2011) as a putative homolog of a scavenger receptor cysteine-rich protein from *Culex pipiens quinquefasciatus* with unknown function. Subsequent to searches on nematode. net, SR02054 was found in
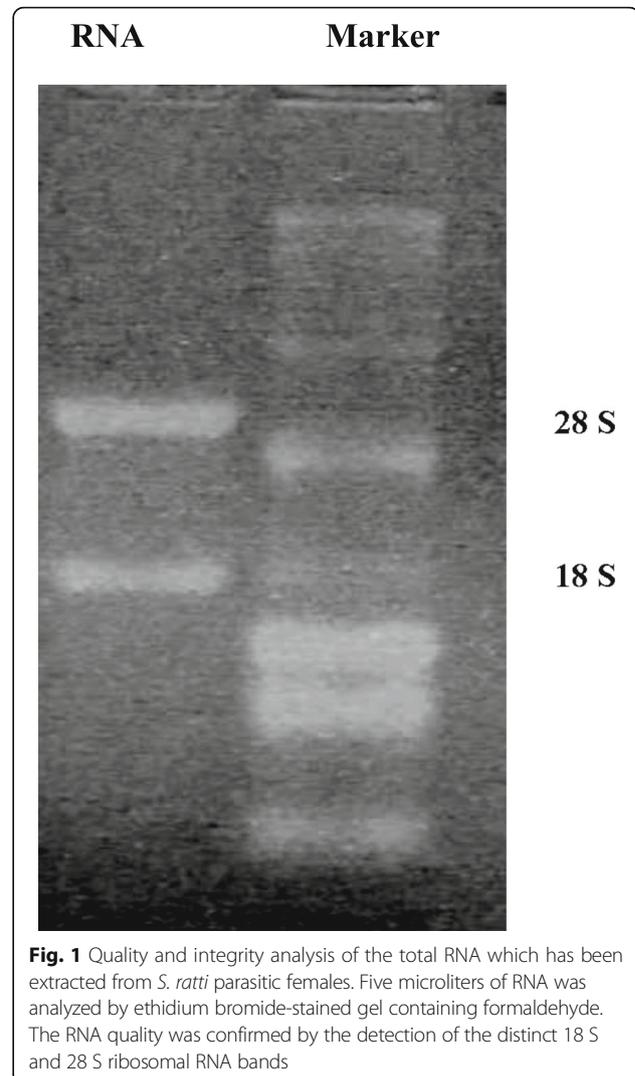
Fig. 1 Quality and integrity analysis of the total RNA which has been extracted from *S. ratti* parasitic females. Five microliters of RNA was analyzed by ethidium bromide-stained gel containing formaldehyde. The RNA quality was confirmed by the detection of the distinct 18 S and 28 S ribosomal RNA bands

Contig 703 and included four EST sequences (kw08g03.y1- kw02a07.y1- ku89b08.y1- kt91a02.y1).

The partial sequence resulted from alignment and assembly of the ESTs of this cluster was used to obtain the full length of the *Sra-til*. After many PCRs with different primers (Table 1), the full length of *Sra-til* was amplified from both pF cDNA and gDNA and resulted in the same size bands ~ 1605 bp (Fig. 2), indicating the absence of introns in the *Sra-til* gene structure. For confirmation, all sequences were cloned, sequenced. and aligned. Interestingly, the sequences from both the pF cDNA and gDNA were 100% identical. For verification, the final *Sra-til* sequence has been submitted to the GenBank database under the accession no. KY441615.

Sequence and phylogenetic analyses

BLAST search of the complete *Sra-til* nucleotide sequence showed 99% identity with a part of a sequence in GenBank

Fig. 2 The PCR products analysis by ethidium bromide-stained gel of the amplified products (slightly above 1500 bp) using *S. ratti* pF cDNA and gDNA
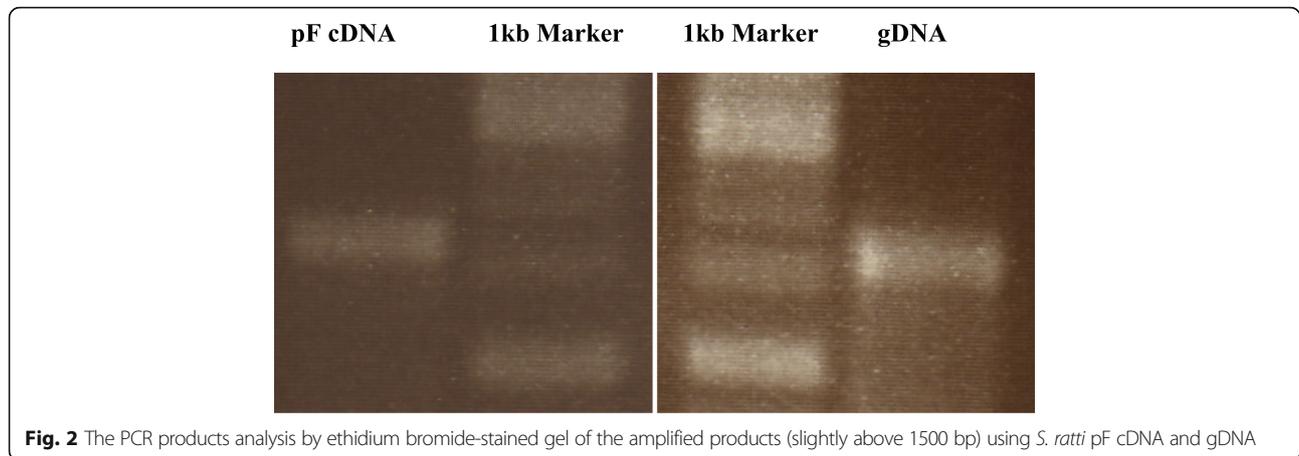
under accession no. LN609529 (*S. ratti* genome assembly) at position 13120855–13122459 and shared 71% identity with parts of the *S. stercoralis* and *S. papillosus* genome assembly sequences (accession no. LL999048 and LM525578, respectively). In addition, *Sra-til* nucleotide sequence showed 69% similarity with parts of the genome assembly sequences of both *S. venezuelensis* (accession no. LM524970) and *Parastrongyloides trichosuri* (accession no. LM523174).

However, BLAST analysis of the translated *Sra*-TIL complete aa sequence showed low similarities (31–43%) to homologous protein sequences from other creatures including nematodes in the GenBank.

Using the full-length cDNA sequence of *Sra-til*, a predicted amino acid (aa) sequence of 534 residues was deduced with a theoretical molecular weight of ~ 57.8 kD and a calculated pI of 8.18. The Simple Modular Architecture Research Tool (SMART) analysis predicted six repeats of epidermal growth factor (EGF)-like domain, in the translated aa sequence of *Sra*-TIL (accession no. ARQ31643), at aa positions 43–77, 229–266, 289–330, 357–394, 422–459, and 482–523 (Fig. 3). A signal peptide, for secretion, (MRLFISLLLILTITLSVVTS) was detected by the SignalP v4.0 program at position 1–20. Only one low compositional complexity was detected by the SEG program at position 268–279.

In addition, protein families database (Pfam) pattern showed a hit of TIL domain at position aa 154–208. TIL domain is found in proteinase inhibitors and typically comprise ten cysteine residues that form five disulfide bonds. Comparison of the *Sra*-TIL domain sequence (55 aa) with other related TIL domains from other parasitic nematodes and insects (Fig. 4), showed 12 conserved residues with complete (no gaps) reactive site regions (inhibitory loops) allaying between the fifth and sixth cysteines (Schechter & Berger, 1967) of the TIL domains analyzed in this figure.

Sra-TIL domains differed from the other nematode and arthropod TIL domains with low to average sequence identities (31–46%).

According to the Pfam entry of the TIL domain (PF01826: http://pfam.xfam.org/family/PF01826), the architectures of serine protease inhibitors, which were identified from some nematodes, were represented only by pure TIL domains (TIL × 1 – TIL × 8) or in combination with other domains like WAP domain (whey acidic protein) in *Trichuris muris* and *Trichinella spiralis*, chaperone DnaJ (Hsp40) in *T. murrelli* and DOMON (dopamine beta-monooxygenase N-terminal) in *Caenorhabditis briggsae*. Like one of *T. muris* TILs, the predicted *Sra*-TIL has a single TIL domain but was found in combination with six EGF domains with predicted architecture (*Sra*-TIL = TIL, EGF × 6). Interestingly, there was more than one TIL protein in the entry PF01826 of different architectures of most nematodes including *S. ratti*, mostly identified as uncharacterized proteins and with different sizes.

Fig. 3 SMART analysis of the *Sra*-TIL protein sequence predicted six EGF domains and a TIL domain. A signal peptide, for secretion, was detected by the SignalP v4.0 program at position 1–20. Only one low compositional complexity was detected by the SEG program at position 268–279
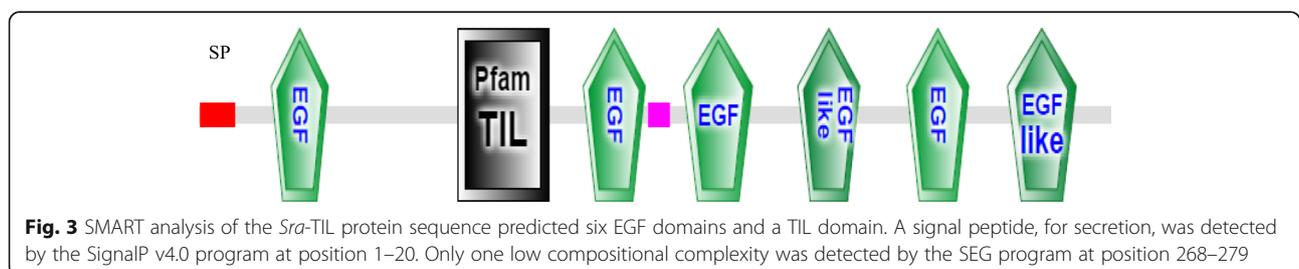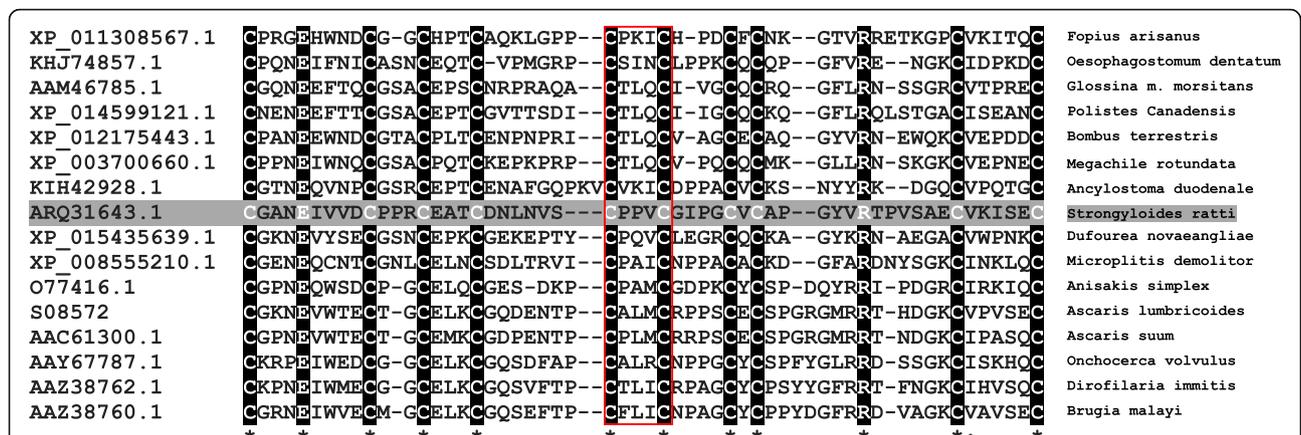
Fig. 4 CLUSTAL multiple sequence alignment by MUSCLE (3.8) of selected 16 aa sequences of TIL domains in some nematodes (including the *Sra*-TIL domain which is highlighted in gray) and arthropods. Samples are labeled according to the GenBank accession numbers (left) and name of species (right). Symbols: identical (asterisks; black highlighted) and no aa (dashes). The active inhibitory site is labeled with red boundary

Based on DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST), phylogenetic analysis of the predicted full-length *Sra*-TIL aa sequence with the most related nematode serine protease inhibitors (Fig. 5), showed two main clusters. The present studied *Sra*-TIL is belonging to a sub-cluster with *Trichinella sp.* and *C. remanei* TILs. It is very closely related to a deduced TIL sequence from *S. ratti* which previously submitted to the GenBank under accession number CEF69530.

Discussion

Nematode-derived protease inhibitors have been suggested to play multiple roles in the survival of the parasite by suppression of exogenous host proteases (Morris & Sakanari, 1994; Peanasky et al., 1984a, b) and to rule endogenous proteases implicated in its development and reproduction (Ford et al., 2005;

Lustigman, Brotman, Huima, Prince, & McKerrow, 1992). They are also believed to have multiple, specific abilities for the host immunomodulation, playing on various immune effector mechanisms (Hartmann & Lucius, 2003).

In our previous study (Soblik et al., 2011), *Sra*-TIL peptides, which is a serine protease inhibitor, has been found abundantly in the Excretory/Secretory products of the parasitic female only but not in the ESP of the infective larvae (iL3) or any other free-living stages. In the same study, qRT-PCR experiment also showed that *Sra-til* is a specific gene for pF.

This study agrees with the genomic study (Hunt et al., 2016) which reported that eight different gene families likely to be important in nematode parasitism in *S. ratti*, including that encoding trypsin inhibitor-like proteins.
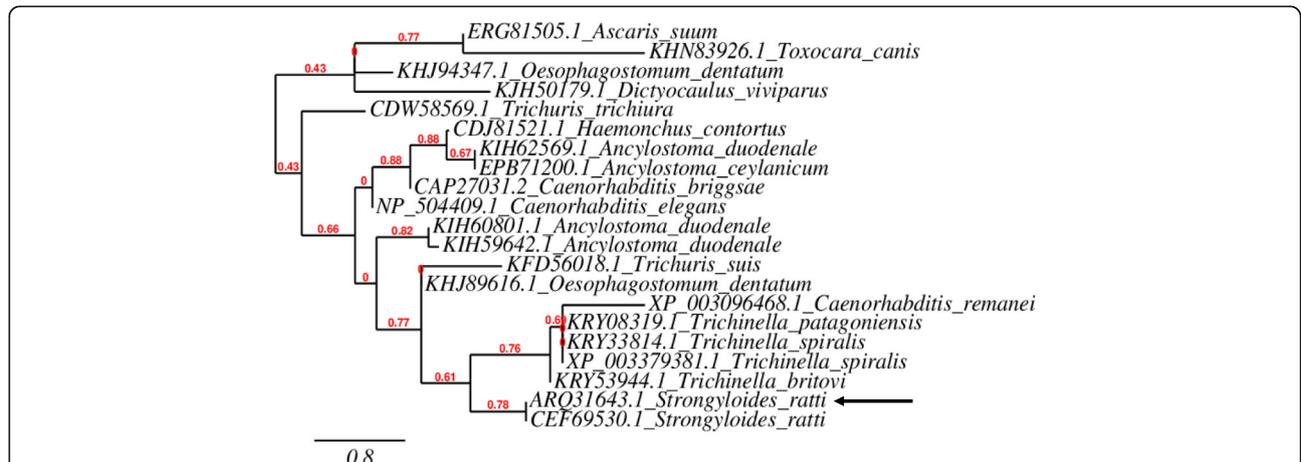
Fig. 5 Phylogenetic tree of 21 TIL orthologs from nematodes resulted from DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST) of the deduced *Sra*-TIL protein sequence, using Phylogeny.fr program (http://www.phylogeny.fr/). Bootstrap support values are shown on the branches. The compared proteins are represented by GenBank accession number and the name of species

In the present study, the full-length cDNA encoding trypsin inhibitor-like protein from the intestinal nematode S. ratti was identified, cloned, and analyzed.

The full-length of *Sra-til* was amplified from both pF cDNA and gDNA and resulted in the same size bands ~ 1605 bp. Unlike most filarial nematodes, TILs (Ford et al., 2005) and introns were absent in the *Sra-til* gene structure.

Different serine protease inhibitors have been identified from various parasitic nematodes, including *Ascaris suum*, *A. lumbricoides* (Gronenborn, Nilges, Peanasky, & Clore, 1990; Peanasky et al., 1984, b), *Anisakis simplex* (Morris & Sakanari, 1994) *Ancylostoma caninum*, *A. ceylanicum*, *A. duodenale* (Mieszczanek, Harrison, & Cappello, 2004; Rios-Steiner, Murakami, Tulinsky, & Arni, 2007; Jin et al., 2011), *Trichuris suis* (Rhoads, Fetterer, & Hill, 2000), and *Onchocerca volvulus* (Ford et al., 2005; Lustigman et al., 1992) where they were believed to have a variety of functions to protect the nematode from intestinal serine proteases or to be involved in molting and development of the nematodes.

The inhibitory function of these proteins is corresponding to the TIL domains (trypsin inhibitor-like) as identified by pFam. The predicted *Sra*-TIL has a single TIL domain but was found in combination with six EGF domains with predicted architecture (*Sra*-TIL = TIL, EGF × 6). The full deduced protein sequence of *Sra*-TIL showed a hit for TIL domain of 55 aa and typically comprise ten cysteine residues that form five disulfide bonds. Reactive site of TIL domain determines the targeted enzyme, which may be trypsin, elastase, chymotrypsin, or cathepsin G. *Sra*-TIL domain has a complete reactive site (inhibitory loop) found between the fifth and sixth cysteines (Schechter & Berger, 1967) like some TILs investigated in *O. volvulus* (Ford et al., 2005).

Nineteen different genes containing TIL domains are found in the *Caenorhabditis elegans* genome (*C. elegans* Sequencing Consortium, 1998), while 37 proteins of *C. elegans* matched the TIL domain were found in the interpro entry IPR002919, which were identified mostly in databases as uncharacterized or theoretical proteins with unknown function(s). Proteins with TIL domains have been detected in arthropods (Cierpicki, Bania, & Otlewski, 2000; Lung et al., 2002; Parkinson et al., 2004; Zhu & Li, 2002), where their functions have not yet been fully explained, although they were stated to be implicated in defining the lifespan or to be constituents of venom. TIL family proteins have also been identified from amphibians where they are suggested to have antimicrobial effects (Ali et al., 2002). In mammals, TILs have been characterized in a wide range of proteins like von Willebrand factor or zonadhesin, which have been reported to play a part in sperm gamete recognition, adhesion, or blood clotting (Bonthron et al., 1986; Gao & Garbers, 1998).

EGFs belong to a common group of proteins that share a repeat pattern, including a number of conserved Cys residues. Growth factors are implicated in cell recognition and proliferation. The EGF motif is found frequently in nature, particularly in extracellular proteins (Campbell et al., 1990). EGF and EGF receptor (EGFR) exerts a fundamental role in wound healing through stimulation, proliferation, and migration of cells (Bodnar, 2013; Wenczak, Lynch, & Nanney, 1992). EGFR also identified in some parasites, and it was believed that EGF signaling could be conserved in helminths and might be directly involved in host-parasite molecular interplay (Dissous, Khayath, Vicogne, & Capron, 2006; Vicogne et al., 2004).

The innovative structural advantages of *Sra*-TIL polypeptides secreted from parasitic nematodes is probably reflective of an abundance of different biological functions, including the protection against host proteases, host immunomodulation via proliferation induction of the damaged intestinal cells as well as wound healing properties.

Molecular expression and functional assays are under planning for independent study, in the future, to examine the biological functions of the recombinantly expressed *Sra*-TIL, including antigenic properties, cell binding, cytokine release, and enzymatic biochemical experiments.

Conclusions

The full-length *Sra-til* gene, which encodes an interesting ESP candidate, secreted from pF of S. ratti, was identified, isolated, cloned and then sequenced. Both nucleotide and amino acid sequences were subjected to analyses by various bioinformatic tools like alignment, BLAST search, structure prediction, and phylogenetic analysis. The predicted innovative structural advantages of *Sra*-TIL polypeptides secreted (signal peptide was predicted) from a parasitic intestinal stage of the S. ratti nematode, showing predicted host protease inhibition (TIL × 1) and cell proliferation/wound healing (EGF × 6) putative features, are probably reflective of the abundance of different biological functions, including the protection against host proteases, host immunomodulation via proliferation induction of the damaged intestinal cells as well as wound healing properties.

Abbreviations
aa: Amino acid; BLAST: Basic local alignment search tool; cDNA: Complementary DNA; DELTA-BLAST: Domain Enhanced Lookup Time Accelerated BLAST; DnaJ: Hsp40; DOMON: Dopamine beta-monooxygenase N-terminal; EGF: Epidermal growth factor; ESP: Excretory/secretory products; EST: Expressed sequence tag; flS: Free-living stages; gDNA: Genomic DNA; iL3: Infective third-stage larvae; ORF: Open reading frame; PCR: Polymerase chain reaction; pF: Parasitic female; Pfam: Protein families database; SMART: Simple Modular Architecture Research Tool; *Sra-til*: Trypsin inhibitor-like gene from *Strongyloides ratti*; *Sra*-TIL: Trypsin inhibitor-like protein from *S. ratti*; TIL: Trypsin inhibitor-like; WAP: Whey acidic protein

Availability of data and materials

The authors confirm the availability of all supporting data.

Authors' contributions

The practical part of this work was done by AY in the BNITM, Hamburg, during his post-doctoral fellowship in 2014. Then the data analyses, subsequent interpretations, and writing were done in Aswan University. NB is the supervisor who prepared the research plan and contributed in the writing of this manuscript. Both authors read and approved the final manuscript.

Ethics approval

Animal experimentation was conducted at the animal facility of the Bernhard Nocht Institute for Tropical Medicine in agreement with the German animal protection law under the supervision of a veterinarian. The experimental protocols have been reviewed and approved by the responsible federal health authorities of the State of Hamburg (Hamburg, Behoerde fuer Gesundheit und Verbraucherschutz: permission number 89/09). The proposal of the study was accepted by the BNITM and Aswan University.

Consent for publication

Aswan University and the BNITM agreed to publish these data.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Ali, M. F., Lips, K. R., Knoop, F. C., Fritzsch, B., Miller, C., & Conlon, J. M. (2002). Antimicrobial peptides and protease inhibitors in the skin secretions of the crawfish frog, *Rana areolata*. *Biochimica et Biophysica Acta*, *1601*(1), 55–63.

Bodnar, R. J. (2013). Epidermal growth factor and epidermal growth factor receptor: The yin and yang in the treatment of cutaneous wounds and cancer. *Adv Wound Care (New Rochelle)*, *2*, 24–29.

Bonthron, D., Orr, E. C., Mitsock, L. M., Ginsburg, D., Handin, R. I., & Orkin, S. H. (1986). Nucleotide sequence of pre-pro-von Willebrand factor cDNA. *Nucleic Acids Research*, *14*(17), 7125–7127.

Bungiro, R., & Cappello, M. (2004). Hookworm infection: New developments and prospects for control. *Current Opinion in Infectious Diseases*, *17*, 421–426.

C. elegans Sequencing Consortium (1998). Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science*, *282*(5396), 2012–2018.

Campbell, I. D., Baron, M., Cooke, R. M., Dudgeon, T. J., Fallon, A., Harvey, T. S., & Tappin, M. J. (1990). Structure-function relationships in epidermal growth factor (EGF) and transforming growth factor-alpha (TGF-alpha). *Biochemical Pharmacology*, *40*, 35–40.

Chan, M. S. (1997). The global burden of intestinal nematode infections—fifty years on. *Parasitology Today*, *13*, 438–443.

Cierpicki, T., Bania, J., & Otlewski, J. (2000). NMR solution structure of *Apis mellifera* chymotrypsin/cathepsin G inhibitor-1 (AMCI-1): Structural similarity with *Ascaris* protease inhibitors. *Protein Science*, *9*, 976–984.

Concha, R., Harrington Jr., W., & Rogers, A. I. (2005). Intestinal *strongyloidiasis*: Recognition, management, and determinants of outcome. *Journal of Clinical Gastroenterology*, *39*, 203–211.

Cooper, E. S., & Bundy, D. A. (1988). Trichuris is not trivial. *Parasitology Today*, *4*, 301–306.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: Robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research*, *36*, W465–W469.

Dissous, C., Khayath, N., Vicogne, J., & Capron, M. (2006). Growth factor receptors in helminth parasites: Signalling and host-parasite relationships. *FEBS Letters*, *580*, 2968–2975.

Farrell, R. E. (1993). *RNA methodologies: A laboratory guide for isolation and characterization*, (p. 322). San Diego; London: Academic Press.

Ford, L., Guiliano, D. B., Oksov, Z. Y., Debnath, A. K., Liu, J., Williams, S. A., … Lustigman, S. (2005). Characterization of a novel filarial serine protease inhibitor, *OV*-SPI-1, from *Onchocerca volvulus*, with potential multifunctional roles during development of the parasite. *The Journal of Biological Chemistry*, *280*(49), 40845–40856.

Gao, Z., & Garbers, D. (1998). Species diversity in the structure of zonadhesin, a sperm-specific membrane protein containing multiple cell adhesion molecule-like domains. *The Journal of Biological Chemistry*, *273*, 3415–3421.

Gronenborn, A. M., Nilges, M., Peanasky, R. J., & Clore, G. M. (1990). Sequential resonance assignment and secondary structure determination of the *Ascaris* trypsin inhibitor, a member of a novel class of proteinase inhibitors. *Biochemistry*, *29*, 183–189.

Grove, D. I. (1989). *Strongyloidiasis: A major roundworm infection of man*, (p. 336). London: Taylor & Francis.

Guyatt, H. (2000). Do intestinal nematodes affect productivity in adulthood? *Parasitology today (Personal ed)*, *16*, 153–158.

Hartmann, S., & Lucius, R. (2003). Modulation of host immune responses by nematode cystatins. *International Journal for Parasitology*, *33*(11), 1291–1302.

Hewitson, J. P., Grainger, J. R., & Maizels, R. M. (2009). Helminth immunoregulation: The role of parasite secreted proteins in modulating host immunity. *Molecular and Biochemical Parasitology*, *167*, 1–11.

Hotez, P. J. (2008). Tropical diseases research: Thirty years and counting. *PLoS Neglected Tropical Diseases*, *2*, e329.

Hunt, V. L., Tsai, I. J., Coghlan, A., Reid, A. J., Holroyd, N., Foth, B. J., et al. (2016). The genomic basis of parasitism in the *Strongyloides* clade of nematodes. *Nature Genetics*, *48*, 299–307.

Hunter, T., Bannister, W. H., & Hunter, G. J. (1997). Cloning, expression, and characterization of two manganese superoxide dismutases from Caenorhabditis elegans. *Journal of Biological Chemistry*, *272*, 28652–28659.

Jin, X., Deng, L., Li, H., Zhang, Z., He, Q., et al. (2011). Identification and characterization of a serine protease inhibitor with two trypsin inhibitor-like domains from the human hookworm Ancylostoma duodenale. *Parasitology Research*, *108*, 287–295.

Keiser, J., Thiemann, K., Endriss, Y., & Utzinger, J. (2008). *Strongyloides ratti*: In vitro and in vivo activity of tribendimidine. *PLoS Neglected Tropical Diseases*, *2*, e136.

Kramme, S., Nissen, N., Soblik, H., Erttmann, K., Tannich, E., Fleischer, B., … Brattig, N. (2011). Novel real-time PCR for the universal detection of *Strongyloides spp*. *Journal of Medical Microbiology*, *60*(Pt 4), 454–458.

Lightowlers, M.W., Rickard, M.D. 1988. Excretory-secretory products of helminth parasites: Effects on host immune responses. Parasitology 96: Suppl S123-S166.

Liu, L. X., & Weller, P. F. (1993). Strongyloidiasis and other intestinal nematode infections. *Infectious Disease Clinics of North America*, *7*(3), 655–682.

Lung, O., Tram, U., Finnerty, C. M., Eipper-Mains, M. A., Kalb, J. M., & Wolfner, M. F. (2002). The *Drosophila melanogaster* seminal fluid protein Acp62F is a protease inhibitor that is toxic upon ectopic expression. *Genetics*, *160*, 211–224.

Lustigman, S., Brotman, B., Huima, T., Prince, A. M., & McKerrow, J. H. (1992). Molecular cloning and characterization of onchocystatin, a cysteine proteinase inhibitor of *Onchocerca volvulus*. *The Journal of Biological Chemistry*, *267*(24), 17339–17346.

Martin, J., Abubucker, S., Wylie, T., Yin, Y., Wang, Z., & Mitreva, M. (2009). Nematode. net update 2008: improvements enabling more efficient data mining and comparative nematode genomics. *Nucleic Acids Research*, *37*, D571–D578.

Mieszczanek, J., Harrison, L. M., & Cappello, M. (2004). *Ancylostoma ceylanicum* anticoagulant peptide-1: Role of the predicted reactive site amino acid in mediating inhibition of coagulation factors Xa and VIIa. *Molecular and Biochemical Parasitology*, *137*(1), 151–159.

Montes, M., Sawhney, C., & Barros, N. (2010). *Strongyloides stercoralis*: There but not seen. *Current Opinion in Infectious Diseases*, *23*, 500–504.

Morris, S. R., & Sakanari, J. A. (1994). Characterization of the serine protease and serine protease inhibitor from the tissue-penetrating nematode *Anisakis simplex*. *The Journal of Biological Chemistry*, *269*(44), 27650–27656.

Nagaraj, S. H., Gasser, R. B., & Ranganathan, S. (2008). Needles in the EST haystack: Large-scale identification and analysis of excretory-secretory (ES) proteins in parasitic nematodes using expressed sequence tags (ESTs). *PLoS Neglected Tropical Diseases*, *2*, e301.

Newton, S. E., & Munn, E. A. (1999). The development of vaccines against gastrointestinal nematode parasites, particularly Haemonchus contortus. *Parasitology Today*, *15*, 116–122.

Nokes, C., Grantham-McGregor, S. M., Sawyer, A. W., Cooper, E. S., Robinson, B. A., & Bundy, D. A. (1992). Moderate to heavy infections of *Trichuris trichiura* affect cognitive function in Jamaican school children. *Parasitology*, *104*(Pt 3), 539–547.

Parkinson, N. M., Conyers, C., Keen, J., MacNicoll, A., Smith, I., Audsley, N., et al. (2004). Towards a comprehensive view of the primary structure of venom proteins from the parasitoid wasp *Pimpla hypochondriaca*. *Insect Biochemistry and Molecular Biology*, *34*, 565–571.

Peanasky, R. J., Bentz, Y., Homandberg, G. A., Minor, S. T., & Babin, D. R. (1984). The isoinhibitors of chymotrypsin/elastase from *Ascaris lumbricoides*: The reactive site. *Archives of Biochemistry and Biophysics*, *232*, 135–142.

Peanasky, R. J., Bentz, Y., Paulson, B., Graham, D. L., & Babin, D. R. (1984). The isoinhibitors of chymotrypsin/elastase from *Ascaris lumbricoides*: Isolation by affinity chromatography and association with the enzymes. *Archives of Biochemistry and Biophysics*, *232*, 127–134.

Playfair, J. H. L., & Bancroft, G. J. (2008). *Infection and immunity, 3rd ed. edn*, (p. 360). Oxford University Press: Oxford.

Rhoads, M. L., Fetterer, R. H., & Hill, D. E. (2000). *Trichuris suis*: A secretory serine protease inhibitor. *Experimental Parasitology*, *94*(1), 1–7.

Rios-Steiner, J. L., Murakami, M. T., Tulinsky, A., & Arni, R. K. (2007). Active and exo-site inhibition of human factor Xa: Structure of des-Gla factor Xa inhibited by NAP5, a potent nematode anticoagulant protein from *Ancylostoma caninum*. *Journal of Molecular Biology*, *371*, 774–786.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning, 2nd ed. edn*, (p. 1626). Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

Schechter, I., & Berger, A. (1967). On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications*, *27*(2), 157–162.

Soblik, H., Younis, A. E., Mitreva, M., Renard, B. Y., Kirchner, M., Geisinger, F., … Brattig, N. W. (2011). Life cycle stage-resolved proteomic analysis of the excretome/secretome from *Strongyloides ratti*—identification of stage-specific proteases. *Molecular & Cellular Proteomics*, *10*, M111.010157.

Tazir, Y., Steisslinger, V., Soblik, H., Younis, A. E., Beckmann, S., Grevelding, C. G., … Erttmann, K. D. (2009). Molecular and functional characterisation of the heat shock protein 10 of *Strongyloides ratti*. *Molecular and Biochemical Parasitology*, *168*, 149–157.

Vicogne, J., Cailliau, K., Tulasne, D., Browaeys, E., Yan, Y. T., Fafeur, V., … Dissous, C. (2004). Conservation of epidermal growth factor receptor function in the human parasitic helminth *Schistosoma mansoni*. *The Journal of Biological Chemistry*, *279*, 37407–37414.

Viney, M.E., Lok, J.B. 2007. Strongyloides spp. WormBook: the online review of C elegans biology.1–15. Available at website http://www.wormbook.org.

Wakelin, D. (1996). Immunology and genetics of zoonotic infections involving parasites. *Comparative Immunology, Microbiology and Infectious Diseases*, *19*, 255–265.

Wakelin, D., Walliker, D. 1996. Genetics of host and parasite-implications for immunity, epidemiology and evolution. Parasitology 112: Suppl S1-S4.

Wenczak, B. A., Lynch, J. B., & Nanney, L. B. (1992). Epidermal growth factor receptor distribution in burn wounds. Implications for growth factor-mediated repair. *The Journal of Clinical Investigation*, *90*, 2392–2401.

Younis, A. E., Geisinger, F., Ajonina-Ekoti, I., Soblik, H., Steen, H., Mitreva, M., … Brattig, N. W. (2011). Stage-specific excretory-secretory small heat shock proteins from the parasitic nematode *Strongyloides ratti*—putative links to host's intestinal mucosal defense system. *The FEBS Journal*, *278*, 3319–3336.

Younis, A. E., Soblik, H., Ajonina-Ekoti, I., Erttmann, K. D., Luersen, K., Liebau, E., & Brattig, N. W. (2012). Characterization of a secreted macrophage migration inhibitory factor homologue of the parasitic nematode *Strongyloides* acting at the parasite-host cell interface. *Microbes and Infection*, *14*, 279–289.

Zhu, S. Y., & Li, W. X. (2002). Precursors of three unique cysteine–rich peptides from the scorpion Buthus martensii Karsch. *Comp Biochem Physiol. Part B, Biochem Mol Biol*, *131*(4), 749–756.