

RESEARCH

Open Access



Novel insights into the somatic proteome of *Strongyloides stercoralis* infective third-stage larvae

Klevia Dishnica¹, Chiara Piubelli², Marcello Manfredi³, Ravi Teja Kondaveeti², Silvia Stefania Longoni², Monica Degani², Dora Buonfrate², Alejandro Giorgetti¹ and Natalia Tiberti^{2*}

Background Strongyloidiasis is a neglected tropical disease affecting an estimated 600 million people, particularly in resource-limited settings. The infection can persist lifelong due to unusual auto-infective cycle of *Strongyloides stercoralis*. The lack of a diagnostic gold standard and limited knowledge of the mechanisms underpinning this chronic infection are key issues in disease management. To date, only a few proteomics studies have been conducted to elucidate the molecular mechanisms associated with *Strongyloides* parasitism or to highlight novel immunological markers, with the result that our knowledge of *S. stercoralis* proteome remains limited. This study aims at expanding the characterization of *S. stercoralis* infective larvae (iL3) in order to further explore the mechanisms of parasitism and to highlight possible novel targets for serodiagnosis.

Methods iL3 obtained from an infected subject were analysed by high-throughput tandem mass spectrometry. To achieve a more comprehensive characterization of the iL3 proteome we analysed the experimental dataset using an automatic search strategy combined with manual annotation, which included gene ontology (GO) analysis, InterPro annotation, assessment of the homology with *Homo sapiens* and other pathogens of clinical importance and B-cell epitope prediction.

Results Our pipeline identified 430 *S. stercoralis* proteins, 187 (43%) of which were uncharacterized. Oxidoreductases and peptidases were amongst the most represented protein categories, as highlighted by molecular function GO analyses, while membrane and mitochondrial proteins were the most represented cellular component GO categories. A high proportion of proteins bearing the CAP, SCP or thioredoxin domain or belonging to cysteine-rich secretory, transthyretin-like or peptidase protein families were also identified. Additionally, we highlighted nine proteins displaying low homology with *H. sapiens* or other related pathogens and bearing amino acid sequences with immunogenic properties.

Conclusions Our comprehensive description and annotation of the *S. stercoralis* iL3 proteome contribute to expanding the 'omics characterization of this parasite and provide experimental evidence on the most represented proteins associated with *S. stercoralis* parasitism, as inferred from genomic and transcriptomic data. Moreover, novel candidate immunogenic proteins to be evaluated as novel serological diagnostic markers are highlighted.

Keywords *Strongyloides stercoralis*, iL3, Proteomics, Annotation, B-cell epitope prediction, Serodiagnosis

*Correspondence:

Natalia Tiberti

natalia.tiberti@sacrocuore.it

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Human strongyloidiasis caused by *Strongyloides stercoralis* is a soil-transmitted helminthiasis that has recently been listed by the WHO among the tropical neglected diseases requiring control actions in endemic areas [1]. Strongyloidiasis is estimated to affect about 600 million people worldwide [2], mostly in tropical and subtropical regions. However, foci of autochthonous strongyloidiasis have also been reported in temperate areas, including Italy, Spain, Japan, Australia and USA [3]. *Strongyloides stercoralis* belongs to the phylum Nematoda, clade IV [4]. Its life-cycle is complex, alternating between cycles of free-living and parasitic stages. Humans acquire the infection through the penetration of the intact skin by infective filariform larvae (iL3) present in contaminated soil which, once in the host, migrate through different organs. During migration, the larvae moult until they become adult worms, which ultimately settle in the small intestine. Once there, the parthenogenetic females deposit eggs that hatch in rhabditiform larvae (L1), which are then excreted in stools and initiate the free-living cycle. However, some L1 undergo an auto-infective cycle, i.e. mature into invasive filariform larvae, in the large intestine and penetrate the intestinal mucosa or the perianal skin to continue the parasitic life-cycle. This peculiar life-cycle allows *S. stercoralis* to perpetuate the infection, in the absence of treatment, potentially indefinitely [5].

In immunocompetent subjects, the infection mostly leads to a chronic indolent condition; however, changes in the host immune status can cause a dramatic increase in parasite burden, known as hyper-infection or dissemination, which can be life-threatening [5].

The diagnosis of strongyloidiasis is challenging, with most available methods presenting variable sensitivity [5, 6]. The most sensitive diagnostic tools are serological immunoassays [7, 8]. Most commercial assays are based on crude larval antigens, which reduce their specificity and result in a high batch-to-batch variability. The development of assays based on recombinant antigens represents a very promising strategy to avoid the need of constant supply of parasites, to overcome the variability of the antigenic source and to reduce cross reactions with other helminths, factors that affect the performance of current serological tests [5]. Indeed, a novel commercial enzyme-linked immunosorbent assay (ELISA) based on the detection of two recombinant antigens, Ss-NIE and Ss-IR, has recently been developed and evaluated on cryopreserved samples. The test has shown variable accuracy in two different studies, probably due to the lack of a diagnostic gold standard, and has yet to be tested prospectively. Nonetheless, at present it is among the most

sensitive and specific serological tests for strongyloidiasis, further highlighting the potential of recombinant antigens for serodiagnosis [9, 10].

To date, only a few proteomics studies have been conducted to elucidate the molecular mechanisms associated with *Strongyloides* parasitism or to highlight novel immunological markers [11–17]. Consequently, our knowledge of *S. stercoralis* proteome is still limited. In order to highlight novel targets to improve current serodiagnosis and treatment, it is fundamental to expand the molecular understanding derived from ‘omics studies beyond the current state of the art. Indeed, an in depth characterization of the *S. stercoralis* infective larvae proteome might reveal on one hand novel players in the mechanisms of host–pathogen interaction and on the other hand potentially immunogenic proteins to be used for the development of novel diagnostic serological tests, as well as target proteins for new therapeutics.

The aim of this study was to expand the characterization of *S. stercoralis* iL3 proteome as established by high-throughput proteomics, combining automatic search strategies and manual annotation.

Methods

Larvae isolation, protein extraction and digestion

Strongyloides stercoralis larvae were obtained from a human subject. Fresh stools mixed with charcoal and saline were cultivated using the agar plate culture method. iL3 larvae were harvested after 3 days of culture [18], concentrated by centrifugation and incubated with phosphate buffered saline (PBS) supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin and 0.625 µg/ml amphotericin B (all from Gibco, Thermo Fisher Scientific, Waltham, MA, USA) for 2 h at 4 °C. Larvae were then washed twice with cold PBS, counted under the microscope and stored at – 80 °C for future use.

A pellet of 10,000 iL3 was re-suspended in 0.1% RapiGest SF (Waters Corporation, Milford, MA, USA) in 0.1 M triethylammonium bicarbonate buffer (TEAB) pH 8.0, sonicated with breaks on ice and incubated for 10 min at 80 °C, following a protocol reported in [19]. The sample was then centrifuged at 14,000 g for 10 min at 4 °C and the supernatant recovered. Protein concentration was determined by the Qubit protein assay (Life Technologies, Thermo Fisher Scientific). A 75-µg sample of proteins was reduced with 50 mM Tris-(2-carboxyethyl)phosphine hydrochloride (TCEP), alkylated with 15 mM iodoacetamide and digested with 0.25 µg/µl sequencing grade-modified trypsin (Roche, Basel, Switzerland; 1:25 protease to protein ratio). The sample was incubated with 1% trifluoroacetic acid for 45 min at 37 °C to cleave the RapiGest SF surfactant, cleaned with C18

spin columns (Pierce™, Thermo Fisher Scientific) and dried under vacuum prior to liquid chromatography-tandem mass spectrometry (LC–MS/MS) analyses.

Protein identification by LC–MS/MS

Trypsin-digested protein samples were analysed with a micro-LC system (Eksigent Technologies, Dublin, CA, USA) coupled with the TripleTOF 5600+ system (Sciex, Concord, ON, Canada) equipped with a DuoSpray ion source (Sciex). The stationary phase was a Halo C18 column (0.5 × 100 mm, 2.7 μm; Eksigent Technologies). The mobile phase was a mixture of 0.1% (v/v) formic acid in water (phase A) and 0.1% (v/v) formic acid in acetonitrile (phase B), eluting at a flow rate of 15.0 μl/min at an increasing concentration of solvent B from 2% to 40% in 30 min. Samples were also analysed with nano liquid chromatography using an Acclaim PepMap C18 column 2 μm, 75 μm × 150 mm (Thermo Fisher Scientific) and injection volume of 2 μl. The flow rate was 300 nl/min, phase A was 0.1% formic acid/water and phase B was 80% acetonitrile/0.1% formic acid/20% water. A 2-h gradient was used (3–45%). Identification was performed using a data-dependent acquisition (DDA) method: the MS analysis was carried out using a mass range of 100–1500 Da (time-of-flight scan with an accumulation time of 0.25 s), followed by a MS/MS product ion scan from 200 to 1250 Da (accumulation time of 5.0 ms) with the abundance threshold set at 30 cps (35 candidate ions can be monitored during each cycle) [20]. The MS data were acquired with Analyst TF 1.7 (Sciex). The DDA files were searched using Protein Pilot software v. 4.2 (Sciex) and Mascot v. 2.4 (Matrix Science Inc., Boston, MA, USA) using trypsin as the enzyme, with two missed cleavages, a search tolerance of 50 ppm for the peptide mass tolerance and 0.1 Da for the MS/MS tolerance [21]. Searches were performed using the UniProt Swiss-Prot database for *Strongyloides stercoralis* (version 01/02/2020, taxon: 6248, proteome ID: UP000035681, protein count: 12,978), with a false discovery rate (FDR) fixed at 1%.

The MS proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE [22] partner repository with the dataset identifier PXD037243.

Bioinformatics analyses

The bioinformatics analyses were carried out by automating a standard protocol for the annotation. The protocol included: (i) a BLASTp [23] search using default parameters against Uniprot 2022 [24]; (ii) each protein identifier was then connected with its associated GO terms [25] via UniProt and GO terms were then organized using QuickGO tool [26]; and (iii) proteins were then classified into families or domains, and important sites were predicted, retrieving this information by InterPro

[27]. In order to identify possible candidates for immunogenic epitope prediction, the annotated dataset of proteins was then investigated using BLASTp for homology with the human protein database and with a list encompassing 29 clinically relevant pathogens (24 helminths and 5 *Plasmodium* spp.) that might co-infect individuals with strongyloidiasis (Fig. 1). The threshold for considering an *S. stercoralis* protein as having low homology with proteins of human or with those of other pathogens' origin was empirically established based on the BLASTp e-value obtained for the L3NieAg.01 (AC: Q9UA16), which is known to have a good specificity when used in serodiagnosis [9]. Thus, a BLASTp e-value threshold of 4E-25 and 2E-30 was applied for the comparison with *H. sapiens* or with other pathogens, respectively.

Linear B-cell epitopes were predicted from protein sequences using the different web-based tools available via the Immune Epitope Database Analysis Resource (IEDB; available at <http://tools.iedb.org/main/>). The following physicochemical properties of individual residues were explored and scored: beta-turn, surface accessibility, antigenicity and hydrophilicity, as already reported in the literature [28]. All residues having an individual score equal or higher than the average protein score were highlighted. In parallel, prediction was also performed using BepiPred-2.0, which combines a hidden Markov model (HMM) with an amino acid propensity scale [29]. Proteins of potential interest for bearing B-cell epitopes were then manually analysed and selected based on the following parameters: (i) sequences of at least 8 amino acids; (ii) a BepiPred-2.0 score > 0.5 (range 0–1); and (iii) at least three physicochemical properties above their thresholds (calculated as the mean of the scores of all individual residues). The specific sequences of interest were highlighted and visualized in the proteins three-dimensional model using Pymol v2.4.1 [30]. Due to the lack of structural characterization of *S. stercoralis* proteins on Protein Data Bank (PDB) [31], selected proteins of interest containing predicted epitopes were structurally predicted using AlphaFold [32].

Results and discussion

In the present study, we analysed the proteome of *S. stercoralis* infective larvae by LC-MS/MS and performed a semi-automated annotation of the dataset to achieve a more in depth characterization of larval proteome and to predict potential immunogenic proteins of interest for the development of new sero-diagnostic tools. The study flowchart is reported in Fig. 1. Our high-throughput MS analysis identified 430 proteins (2 unique peptides, 1% FDR), which to the best of our knowledge is the largest experimental proteome of *S. stercoralis* iL3 reported to date (Additional file 1: Table S1). Indeed, only one study

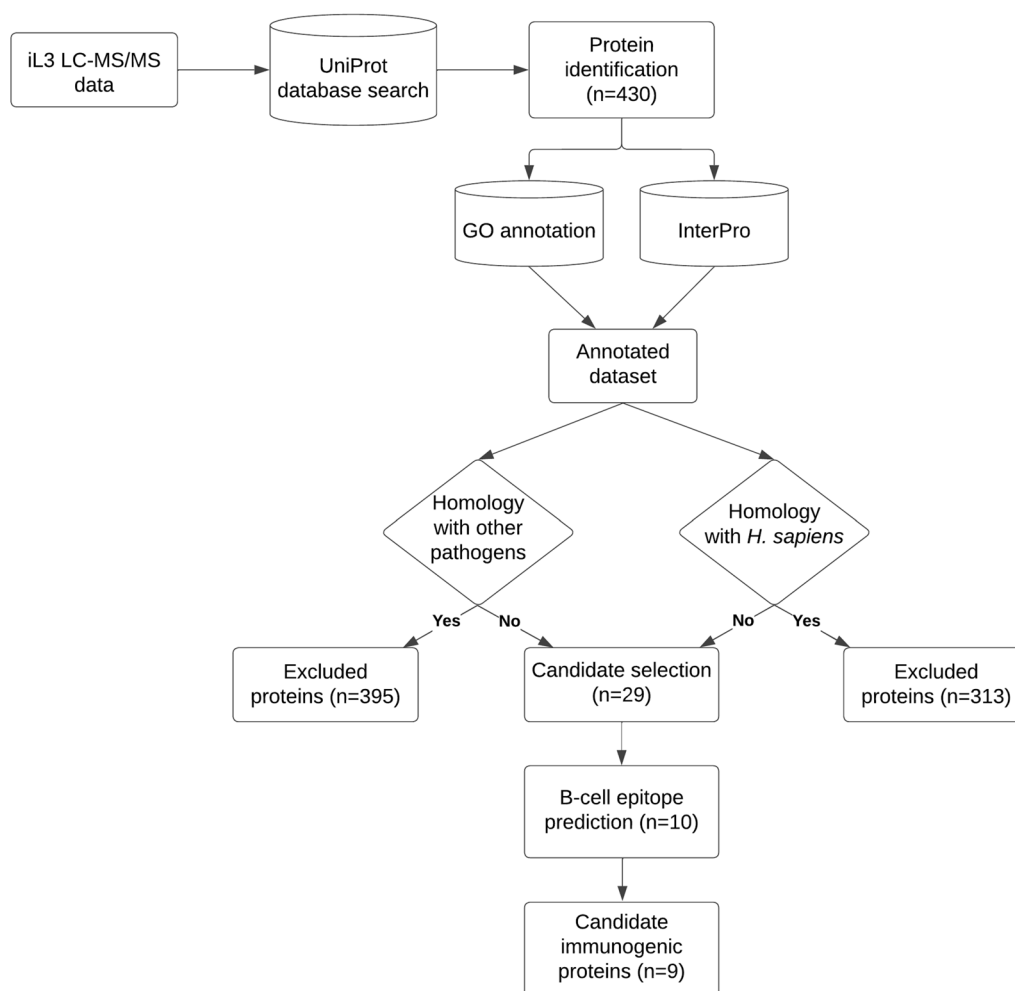


Fig. 1 Study flowchart. Pipeline followed in the present study. GO, Gene ontology; iL3, infective filariform larvae; LC-MS/MS, Liquid chromatography-tandem mass spectrometry. Other pathogens include: *Ancylostoma duodenale*; *Ancylostoma ceylanicum*; *Necator americanus*; *Ascaris lumbricoides*; *Trichuris trichiura*; *Toxocara canis*; *Loa loa*; *Mansonella perstans*; *Mansonella ozzardi*; *Wuchereria bancrofti*; *Onchocerca volvulus*; *Brugia malayi*; *Brugia timori*; *Dirofilaria immitis*; *Dirofilaria repens*; *Trichinella spiralis*; *Taenia saginata*; *Taenia solium*; *Echinococcus granulosus*; *Hymenolepis nana*; *Schistosoma mansoni*; *Schistosoma haematobium*; *Schistosoma japonicum*; *Fasciola hepatica*; *Plasmodium falciparum*; *Plasmodium vivax*; *Plasmodium ovale*; *Plasmodium malariae*; *Plasmodium knowlesi*

had previously employed untargeted proteomics to investigate the *S. stercoralis* iL3 proteome; however, due to the lack of a reference genome at that time, only 26 proteins were identified [12].

A high-quality draft genome of *S. stercoralis* was assembled in 2016 (42.6 Mb) and predicted to contain 13,098 protein-coding genes [11], facilitating the annotation of ‘omics data, although a reference genome has yet to be assembled. Hunt et al. performed an in depth investigation of the genomic bases of parasitism in the *Strongyloides* clade by comparing distinct life stages of different *Strongyloides* species, the closely related *Parastrogyloides trichosuri* and the free-living *Rhabditophanes* at the genome, transcriptome and proteome level [11].

This comparison has allowed researchers to propose protein categories with a putative role in parasitism that are expanded in the *S. stercoralis* genome or abundantly transcribed in iL3. These include proteinases (astacins—metallopeptidases, aspartic proteases, prolyl oligopeptidase), protease inhibitors, SCP/TAPS proteins, transthyretin-like proteins and acetylcholinesterases [11, 33]. Interestingly, the same protein families were also identified in *Strongyloides venezuelensis* iL3 [15]. Similarly, next generation RNA sequencing was employed to evaluate the association between larval development in an *S. stercoralis* laboratory strain (i.e. PV001) and the expression of specific genes homologous of *Caenorhabditis elegans*, in

which they were reported to be involved in dauer arrest or activation [34].

In our dataset, 43% of the identified protein sequences (i.e. 187 protein matches) corresponded to uncharacterized proteins according to UniProt database 2022 for *Strongyloides stercoralis*. In order to achieve a better characterization of the dataset we performed a semi-automated annotation through GO and InterPro functional analyses (Additional file 1: Table S1). The cellular component (CC) GO analysis highlighted a prevalence of membrane and mitochondrial proteins, which together accounted for > 40% of the annotated terms (Fig. 2a). Almost half (47%) of the molecular function (MF) GO terms had binding activities, with nucleic acid and nucleotide binding being the most represented sub-categories, while 40% were associated with enzymatic activities (Fig. 2b). Intriguingly, within this latter group, the most represented term corresponded to oxidoreductase activity, accounting for 33% of all GO terms associated with catalytic activities. It could be speculated that infective larvae might need to counteract the oxidative stress either derived from their particularly active cellular metabolism or as a defence mechanism against the host immune response [35, 36]. Interestingly, we identified three of the four major antioxidant enzyme families involved in the response against reactive oxygen species (ROS), namely glutathione peroxidase, superoxide dismutase and peroxiredoxin/thioredoxin. Antioxidant enzymes are known to be important in nematodes, and an evolutionary analysis has been recently published [35]. Antioxidative enzymes were also identified in the excretory-secretory products (ESPs) from the different life stages of *Strongyloides ratti* [14] and *S. venezuelensis* iL3 [17]; it would be interesting to evaluate whether the expression of these proteins is modulated during parasite development. The biological process (BP) GO analysis also highlighted that *S. stercoralis* iL3 larvae express a high number of proteins involved in metabolic (42%) and cellular processes (44%) (Fig. 2c). Notably, the most represented cellular processes were translation and cellular respiration, which is in agreement with the high number of nucleic acid/nucleotide binding proteins and mitochondrial proteins and further supports the observation of a highly active metabolic state of infective larvae. The InterPro analysis allowed a further classification of the identified proteins, including those uncharacterized, either into families or on the basis of the presence of specific domains within their amino acid sequence. The most frequent InterPro domain and family entries are reported in Fig. 3, while the entire annotation is available as Additional file 1: Table S1. The CAP domain, SCP domain and thioredoxin domain were the most commonly represented protein domains, while several proteins were

annotated as belonging to cysteine-rich secretory, transthyretin-like or peptidase protein families, making them the most represented protein families in the iL3 proteome. Overall, our functional analysis provides experimental evidence that confirms previous data on the most represented proteins associated with *S. stercoralis* parasitism, as inferred from genomic and transcriptomic data [11, 33, 37, 38], as well as with proteomics analyses of *S. ratti* and *S. venezuelensis* iL3 ESP [14, 17]. In particular, in our dataset we identified a high proportion of proteins with peptidase activity; such proteins have already been highlighted as potentially involved in parasitism as they are upregulated in the adult parasitic female stage of *S. ratti* and *S. stercoralis* [11]. Indeed, these proteins, including metalloproteases and metallopeptidases (such as astacin-like proteins), are involved in tissue degradation. This is a fundamental process in the initial phases of the infection for the penetration of host tissues and in parasite migration through the host body—even though peptidases could also contribute to immune evasion [39]. Other protein categories known to be associated with *S. stercoralis* parasitism and identified in our study include: (i) galectins, involved in pathogen adhesion to the host cells and activation of host innate and adaptive immunity [40]; (ii) transthyretin-like proteins; and (iii) SCP/TAPS-/CAP-domain containing proteins, with putative immunomodulatory properties in parasitic nematodes [41]. The expansion of SCP/TAPS coding genes in *Strongyloides* and *Parasitstrongyloides* compared to *Rabditophanes* suggests that their gene products might be associated with human parasitism [11]. In our study, we identified 11 proteins either as SCP domain-containing proteins ($n=7$) or as uncharacterized proteins containing the SCP-domain according to the InterPro analysis (IPR034113). Most of the protein categories that had already been proposed as associated with iL3 parasitism were thus experimentally confirmed in our proteomics study of *S. stercoralis* iL3 with 43 protein matches (Additional file 2: Table S2). It is worth noting that almost 50% of these proteins were uncharacterized and were assigned to those categories only following the GO and InterPro semi-automated annotation. The importance of focusing 'omics studies not only on known and characterized genes and proteins, but especially on those "novel" or uncharacterized ones was already highlighted more than 10 years ago. Such an approach can achieve a more in depth knowledge of the molecular mechanisms associated with pathology but also with pathogen development [42], especially for organisms whose genome and proteome are not fully annotated, such as *S. stercoralis*.

The objective of our study was not only to improve our knowledge of the iL3 proteome with novel experimental evidence, but also to predict—among those identified

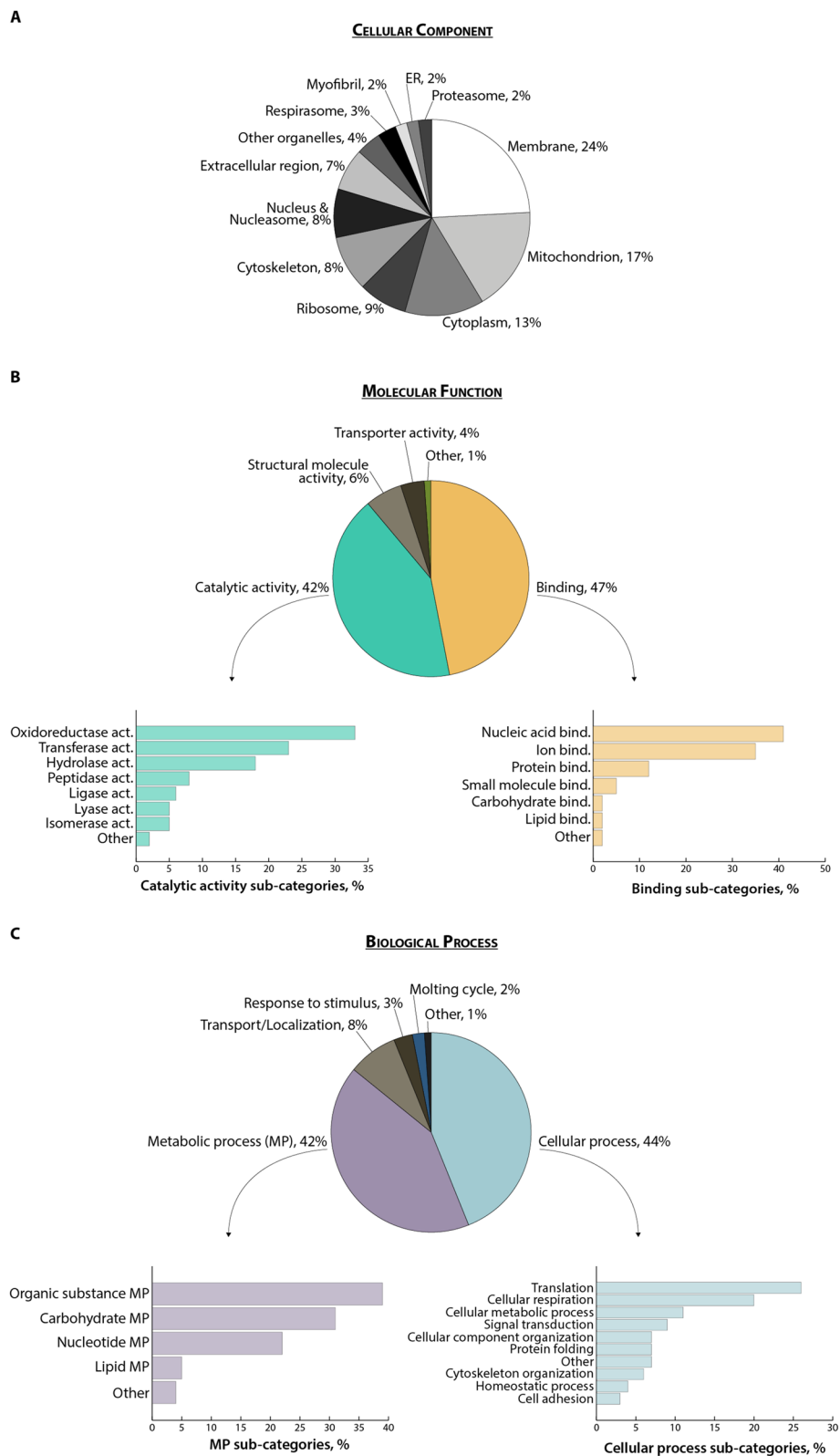


Fig. 2 Gene ontology results. Frequency of the GO terms for the three categories cellular component (A), molecular function (B) and biological process (C) across identified proteins. ER, Endoplasmic reticulum

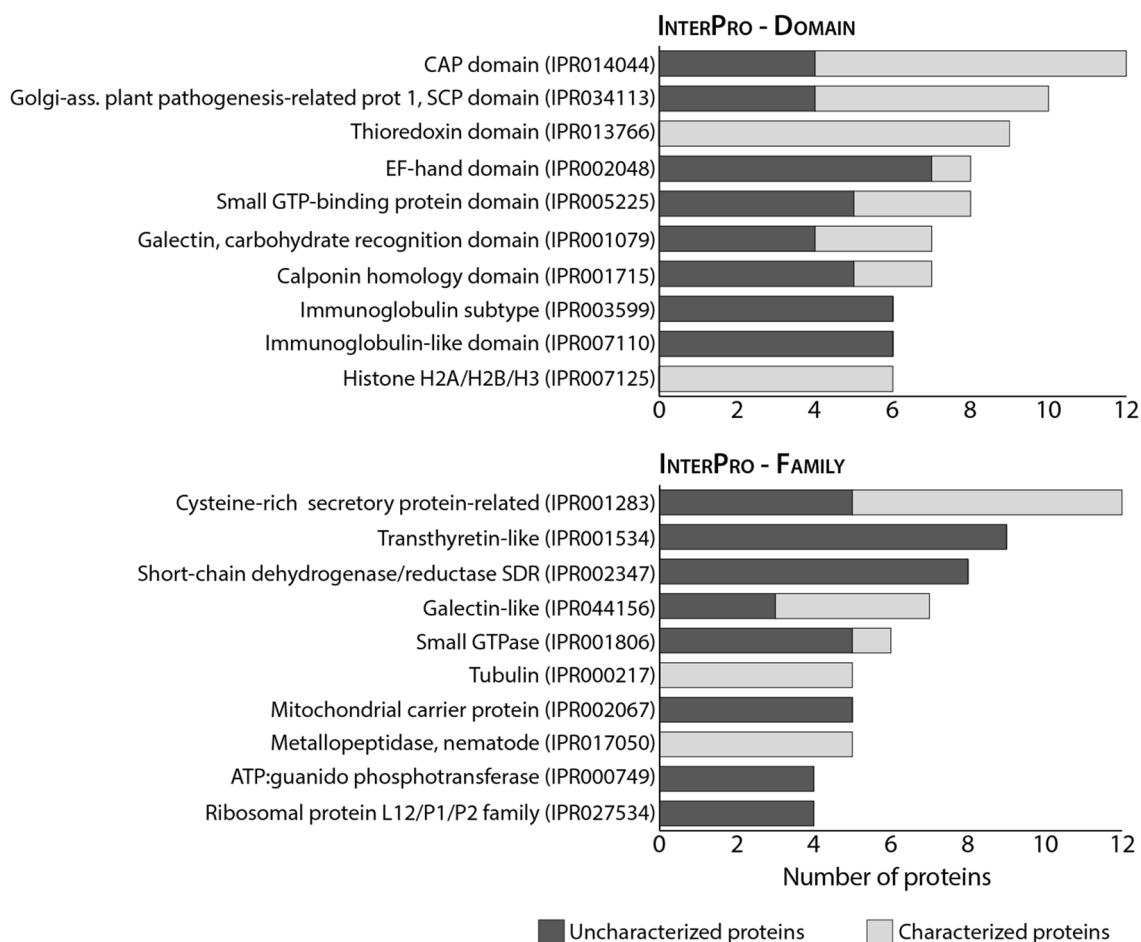


Fig. 3 InterPro annotation results. The top 10 most frequent InterPro terms for the categories domain and family represented among all identified proteins ($n=430$), as established by InterPro annotation. For each term, the number of uncharacterized and characterized proteins is represented in different color shades. The complete annotation is reported in Additional file 1: Table S1

from a clinical isolate—potential immunogenic proteins that could be useful for the development of novel serological tests for the accurate diagnosis of human strongyloidiasis or as vaccine candidates, as potentially recognized by antibodies present in patients’ serum.

A few studies dating back to the 1990s reported the investigation of the humoral immune response associated with the intensity of infection and the detection of immunoreactive iL3 polypeptides [43–45]. However, only a couple of studies applied immuno-proteomics MS/MS-based approaches to also identify the immunogenic proteins recognized by antibodies from infected subjects [13, 46]. Indeed, Rodpai and colleagues confirmed by immunoblotting the high frequency of some protein bands that had previously been reported as immunoreactive [43–45, 47], but also identified them by tandem MS based on protein homology with *S. ratti* [46]. In particular, they identified a 26-kDa band corresponding to 14–3–3 protein and a 29-kDa band corresponding to ADP/ATP translocase 4.

Additional antigenic proteins were further identified by the same group after they had improved sample separation through two-dimensional gel electrophoresis prior to immunoblotting [13]; the majority of these proteins were also identified in the present study.

In silico approaches can be used as an alternative, or a complement, to the experimental identification of immuno-reactive proteins. The advent of immunoinformatics has actually led to the development of a number of tools that can assist researchers in B-cell epitope prediction. Moreover, it has been shown that using multiple prediction methods results in a more accurate epitope prediction than using individual tools [28, 48]. In agreement with this, in the present study we combined the use of a machine learning-based algorithm (i.e. BepiPred-2.0 [29]) and the evaluation of several physicochemical residue properties to predict linear B-cell epitopes. In order to avoid the selection of proteins highly conserved across helminths or similar to human

Table 1 List of potentially immunogenic proteins and the predicted B-cell epitopes

Protein AC	Protein name	e-value ^a vs <i>Homo sapiens</i>	e-value ^a vs other pathogens ^b	Protein properties relevant for selection	Epitope sequence
A0A0K0E6J0	SCP domain-containing protein	7.00E−24	Minimum 2E−25 Maximum 5.5	Extracellular region [GO:0005576] Integral component of membrane [GO:0016021]	¹⁰⁵ VTQPPRPTARPF SRNPE ¹²⁸ KPAPRPTIPPKTAKPG
A0A0K0DY51	Uncharacterized protein	3.00E−06	Minimum 3E−11 Maximum 7.7	IPR014044: CAP domain IPR035940: CAP superfamily [11] IPR018244: Allergen V5/Tpx-1-related, conserved site [11]	¹⁴⁷ APPNNRIDPMYIPNPDE ¹⁶ ESKNEEVHPT ⁴⁰ AVEPPAETPAE ⁷² PVETTTETP ¹⁴⁰ PVETPAETSV-DAPTENPTEVSADVPSTE ¹⁸⁶ SVPEQSVKEIEEPSVTEVQCP
A0A0K0ECK4	Galectin	1.00E−13	Minimum 1E−22 Maximum 7.5	IPR001079: Galectin, carbohydrate recognition domain	¹⁰ TYDHIPKESYSIQPR ⁹¹ KWQHEERTPKGNPFK
A0A0K0EG68	SCP domain-containing protein	7.00E−10	Minimum 2E−11 Maximum 1.8	IPR001283: Cysteine-rich secretory protein-related IPR014044: CAP domain [11]	⁶⁶ RPTNRPINKKPIKPNKPK ¹⁰⁵ PKPPGPRPKPPG ¹²³ GPRPKPPG ¹³⁷ GPKPKPTTTKPKPKPTTTKPKPKPTTTKPKPTTQPPT ¹²⁵ MSPEKSPRYIYPPE
A0A0K0EMX1	NTR domain-containing protein	3.00E−05	Minimum 7E−15 Maximum 8.0	Extracellular region [GO:0005576] IPR001820: Protease inhibitor I35 (TIMP) [11]	¹⁴⁵ EVKNNLRTN
Q9UA16	L3NieAg.01 (Fragment)	4.00E−25	Minimum 2E−30 Maximum 4.0	Extracellular region [GO:0005576] IPR014044: CAP domain [11]	⁷⁵ YNYDNDKA ¹³¹ LEHDPKNRIE
A0A0K0E2F4	Uncharacterized protein	1.00E + 00	Minimum 1.1E−02 Maximum 4.1	Integral component of membrane [GO:0016021] IPR007863: Peptidase M16, C-terminal [11]	⁴⁰ IDNQPAYV ⁷⁵ HKIPHEPKASAREGVDGDEEDGASDTF
A0A0K0DTP5	SCP domain-containing protein	1.00E−13	Minimum 4E−16 Maximum 5.9	IPR014044: CAP domain [11]	¹⁰⁸ KQHNYDRDT
A0A0K0ELA9	Uncharacterized protein	2.90E−01	Minimum 4E−13 Maximum 8.6	Integral component of membrane [GO:0016021] IPR001534: Transthyretin-like [11]	⁴⁴ FGKKDFSTKDLEPKNLKD
A0A0K0E132	Uncharacterized protein	8.40E−01	Minimum 2E−19 Maximum 7.3	Integral component of membrane [GO:0016021] IPR001534: Transthyretin-like [11]	No B-cell epitope found

^a e-value from BLASTp search

^b The largest and smaller e-value from the homology analysis with other pathogens are reported. Other pathogens include: *Ancylostoma duodenale*; *Ancylostoma ceylanicum*; *Necator americanus*; *Ascaris lumbricoides*; *Trichuris trichiura*; *Toxocara canis*; *Loa loa*; *Mansonella perstans*; *Mansonella ozzardi*; *Wuchereria bancrofti*; *Onchocerca volvulus*; *Brugia malayi*; *Brugia timori*; *Dirofilaria immitis*; *Dirofilaria repens*; *Trichinella spiralis*; *Taenia saginata*; *Taenia solium*; *Echinococcus granulosus*; *Hymenolepis nana*; *Schistosoma mansoni*; *Schistosoma haematobium*; *Schistosoma japonicum*; *Fasciola hepatica*; *Plasmodium falciparum*; *Plasmodium vivax*; *Plasmodium ovale*; *Plasmodium malariae*; *Plasmodium knowlesi*.

ones, we first excluded all those having high homology, as established by our BLASTp analysis (Additional file 3: Table S3). Among the 29 proteins showing limited

homology with *Homo sapiens* or other pathogens of clinical relevance, we selected 10 for use in the prediction of the presence of B-cell epitopes (Table 1). This selection

was based on the following criteria: (i) proteins already highlighted as potentially associated with *S. stercoralis* parasitism [11, 33] or as immunogenic [13]; (ii) extra-cellular or plasma membrane proteins as per CC GO terms; (iii) proteins with peptidase activity according to the MF GO terms; and (iv) proteins associated with relevant InterPro domain, family or homologous super-family (namely transthyretin-like domain, CAP domain, galectin, cysteine-rich, peptidase, protease inhibitors). According to UniProt, 60% of the selected proteins are already characterized, while the remaining 40% are still uncharacterized (Table 1). The six characterized proteins included three SCP domain-containing proteins (ACs: A0A0K0E6J0, A0A0K0EG68, A0A0K0DTP5), galectin (AC: A0A0K0ECK4), NTR domain-containing protein

(AC: A0A0K0EMX1) and L3NieAg.01 (AC: Q9UA16 also known as Ss-NIE). It is worth noting that galectins have already been reported to be involved in host–pathogen interaction and to display immuno-regulatory properties in *S. ratti* [49]. Also, several commercial and in-house assays already use the recombinant Ss-NIE for *S. stercoralis* serodiagnosis [9, 50–53]. Ss-NIE is also included in a commercial research use only (RUO) serological test, together with Ss-IR, which was not identified in our dataset [9]. The presence of Ss-NIE within our selection further supports the validity of our approach. In our study we also identified most of the proteins already highlighted as potentially immunogenic by Rodpai and colleagues [13, 46]. However, since these proteins displayed high homology with either human or other related

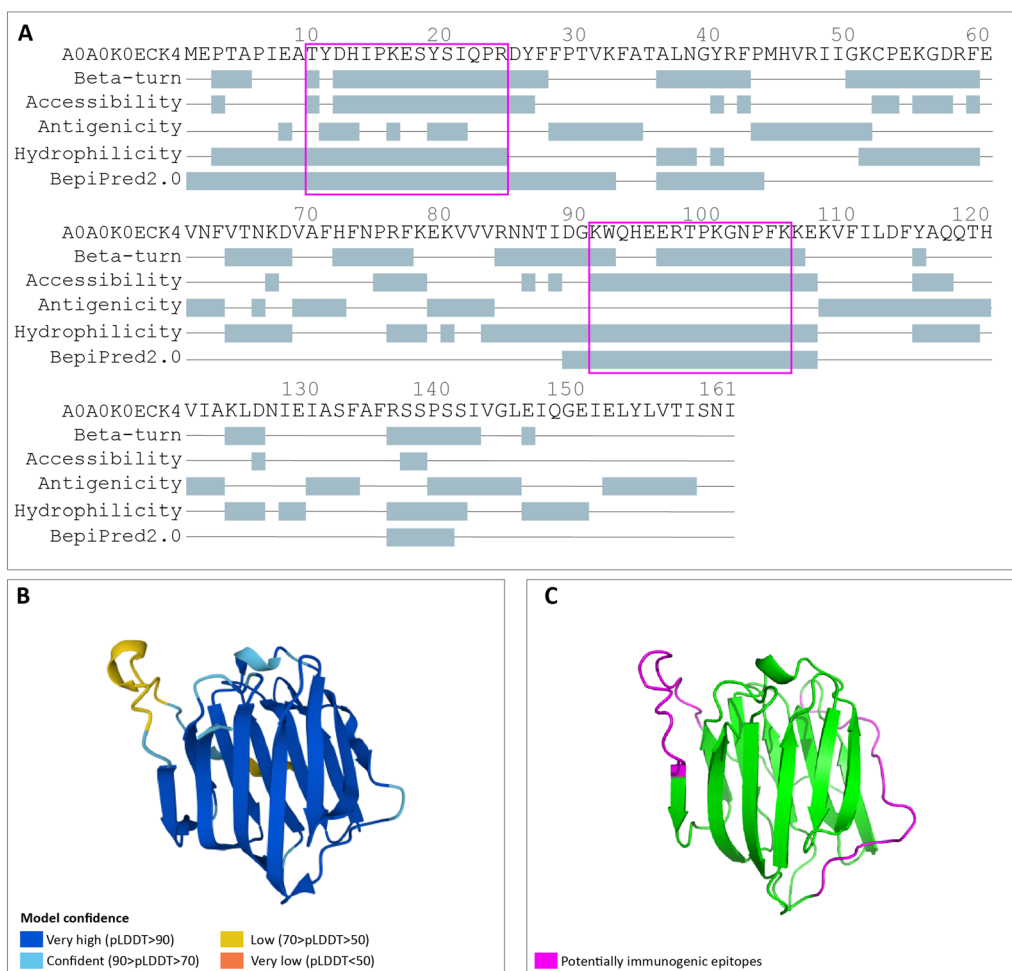


Fig. 4 B-cell epitope prediction results. The results for the protein A0A0K0ECK4—galectin are reported as an example. **a** FASTA sequence showing the results obtained with each tool (Chou & Fasman Beta-Turn Prediction; Emini Surface Accessibility Prediction; Kolaskar & Tongaonkar Antigenicity; Parker Hydrophilicity Prediction, BepiPred2.0; all available via <http://tools.iedb.org/bcell/>). All residues having a score above their threshold are highlighted in grey. The purple squares indicate the sequences highlighted as being potentially immunogenic as reported in the [Methods](#) section. **b** Protein structures as predicted by AlphaFold showing the model confidence. **c** Mapping of the potentially immunogenic epitopes on the protein structure. The same images for all other selected proteins are reported in Additional file 4: Figures S1–S8

pathogen proteins, their analysis for epitope prediction was not pursued (Additional file 2: Table S2, Additional file 3: Table S3).

The B-cell epitope prediction highlighted that nine out of the 10 selected proteins contained epitopes with high consensus across the different tools employed (Table 1; Fig. 4; Additional file 4: Figures S1–S8); these were therefore considered as potentially immunogenic. The remaining protein (AC A0A0K0E132, uncharacterized protein) did not display potentially immunogenic epitopes as per our analysis. The structural models, together with a confidence estimation as per AlphaFold, are reported in Fig. 4 and Additional file 4: Figure S1–S8. In agreement with the results obtained from different web-based prediction tools, all epitopes were exposed to the external environment, thus potentially accessible for antibody binding. However, some epitopes fell within regions of the structure which was modelled with low confidence. This could be explained by the fact that immunogenic epitopes often fall within highly variable regions, and there is a lower confidence in the structure as predicted by AlphaFold.

A recent work employed a reverse in silico approach to predict immunogenic proteins from the *S. stercoralis* proteome available in UniProt [54]. However, none of the proteins proposed as potentially immunogenic was identified in our dataset, probably because the analysis was performed on the entire *S. stercoralis* proteome, without taking into account the parasite developmental stage.

In the present study we did not perform a comparison of protein expression between larval developmental stages, as has been done at the transcriptomic level or for other *Strongyloides* species [11, 14, 33, 34, 38], thus we cannot speculate on the role of iL3 proteins in larval development. However, a comparison using quantitative proteomics of different larval stages might contribute to corroborate these transcriptomics data and might identify novel proteins potentially involved in parasitism and/or in parasite development that could be of interest for the development of novel disease control strategies. Similarly, investigations should be extended to the study of ESPs released from *S. stercoralis* iL3, as has already done for *S. ratti* [14] and *S. venezuelensis* [17], as these could highlight additional candidates for serodiagnosis.

Proteomics data on *S. stercoralis* are still limited, and the reference database and proteome are in continuous evolution. Therefore, some proteins ID here reported might change in the future.

Conclusions

In conclusion, we provide the largest experimental dataset of the *S. stercoralis* iL3 proteome. By presenting, for the first time, an extensive proteomics dataset from the analysis of iL3 isolated from a clinical sample, our study brings knowledge on the *S. stercoralis* proteome to a level comparable to our knowledge on its close relatives *S. ratti* and *S. venezuelensis* [55]. These data may be useful for future studies as they represent a step towards filling the current gap in experimental proteomics data. Indeed, a broader expertise about protein expression in *S. stercoralis* larvae, as well as their modulation during different developmental stages, will be essential for identifying novel therapeutic and vaccine targets.

Our semi-automated annotation allowed us to confirm the presence—at the proteome level—of protein categories potentially involved in parasitism that to date were only inferred from genomics and transcriptomics data. Moreover, additional protein groups deserving further investigation, such as oxidoreductases, were also highlighted. Finally, we also propose a number of immunogenic protein candidates that, if experimentally confirmed, might be considered in the future for the development of novel serological diagnostic tests that could make the diagnosis of this neglected tropical disease more reliable and accurate.

Abbreviations

BP	Biological process
CC	Cellular component
ESP	Excretory-secretory product
GO	Gene ontology
iL3	Infective filariform larvae
L1	Rhabditiform larva
LC–MS/MS	Liquid chromatography-tandem mass spectrometry
MF	Molecular function
PBS	Phosphate buffered saline
ROS	Reactive oxygen species

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13071-023-05675-7>.

Additional file 1: Table S1. Annotated dataset. The dataset includes peptide list, protein identification, gene ontology annotation and InterPro annotation.

Additional file 2: Table S2. Proteins identified in the present study and already reported in the literature as: (i) associated with *Strongyloides* parasitism; (ii) part of iL3 proteome; (iii) potentially immunogenic.

Additional file 3: Table S3. Homology with *Homo sapiens* and other pathogens of clinical importance as potentially responsible for co-infections with *S. stercoralis*.

Additional file 4: Figures S1–S8. B-cell epitope prediction results. For each figure: **a** FASTA sequence showing the results obtained with each tool (Chou & Fasman Beta-Turn Prediction; Emini Surface Accessibility Prediction; Kolaskar & Tongaonkar Antigenicity; Parker Hydrophilicity Prediction; BepiPred2.0; all available via <http://tools.iedb.org/bcell/>). All residues having a score above their threshold are highlighted in grey. The purple

squares indicate the sequences highlighted as potentially immunogenic, as reported in the **Methods** section. For each figure: **b** Protein structures as predicted by AlphaFold showing the model confidence; **c** mapping of the potentially immunogenic epitopes on the protein structure.

Acknowledgements

The authors thank Prof. Daniela Cecconi (Department of Biotechnology, University of Verona), for technical support for sample sonication.

Author contributions

CP, SSL and NT conceived the study. MD provided the biological sample. MM and NT performed the experiments. KD, MM, RTK, AG, NT performed the data analysis. SSL drew the graphic abstract. KD, CP, MM, DB, AG, NT wrote the manuscript. All authors read, revised and approved the final manuscript.

Funding

This work was funded by the Italian Ministry of Health “Fondi Ricerca Corrente–L2P4” to “IRCCS Sacro Cuore–Don Calabria Hospital”.

Availability of data and materials

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD037243.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Biotechnology, University of Verona, Verona, Italy. ²Department of Infectious, Tropical Diseases and Microbiology, IRCCS Sacro Cuore Don Calabria Hospital, Negrar Di Valpolicella, Italy. ³Department of Translational Medicine, University of Piemonte Orientale, Novara, Italy.

Received: 7 November 2022 Accepted: 18 January 2023

Published online: 31 January 2023

References

- WHO. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. 2021. <https://www.who.int/publications/i/item/9789240010352>. Accessed on 27 Jan 2023
- Buonfrate D, Bisanzio D, Giorli G, Odermatt P, Furst T, Greenaway C, et al. The global prevalence of *Strongyloides stercoralis* infection. *Pathogens*. 2020;9:468.
- Schar F, Trostorf U, Giardina F, Khieu V, Muth S, Marti H, et al. *Strongyloides stercoralis*: global distribution and risk factors. *PLoS Negl Trop Dis*. 2013;7:e2288.
- International Helminth Genomes Consortium. Comparative genomics of the major parasitic worms. *Nat Genet*. 2019;51:163–74.
- Nutman TB. Human infection with *Strongyloides stercoralis* and other related *Strongyloides* species. *Parasitology*. 2017;144:263–73.
- Tamarozzi F, Longoni SS, Mazzi C, Rizzi E, Noordin R, Buonfrate D. The accuracy of a recombinant antigen immunochromatographic test for the detection of *Strongyloides stercoralis* infection in migrants from sub-Saharan Africa. *Parasit Vectors*. 2022;15:142.
- Buonfrate D, Marrone R, Silva R, Mirisola C, Ragusa A, Mistretta M, et al. Prevalence of strongyloidiasis in a cohort of migrants in Italy and accuracy of a novel ELISA assay for *S. stercoralis* infection, a cross-sectional study. *Microorganisms*. 2021;9:401.
- Anderson NW, Klein DM, Dornink SM, Jespersen DJ, Kubofcik J, Nutman TB, et al. Comparison of three immunoassays for detection of antibodies to *Strongyloides stercoralis*. *Clin Vaccine Immunol*. 2014;21:732–6.
- Sears WJ, Nutman TB. Strongy detect: Preliminary validation of a prototype recombinant Ss-NIE/Ss-IR based ELISA to detect *Strongyloides stercoralis* infection. *PLoS Negl Trop Dis*. 2022;16:e0010126.
- Tamarozzi F, Longoni SS, Mazzi C, Pettene S, Montresor A, Mahanty S, et al. Diagnostic accuracy of a novel enzyme-linked immunoassay for the detection of IgG and IgG4 against *Strongyloides stercoralis* based on the recombinant antigens NIE/SsIR. *Parasit Vectors*. 2021;14:412.
- Hunt VL, Tsai IJ, Coghlan A, Reid AJ, Holroyd N, Foth BJ, et al. The genomic basis of parasitism in the *Strongyloides* clade of nematodes. *Nat Genet*. 2016;48:299–307.
- Marcilla A, Sotillo J, Perez-Garcia A, Igual-Adell R, Valero ML, Sanchez-Pino MM, et al. Proteomic analysis of *Strongyloides stercoralis* L3 larvae. *Parasitology*. 2010;137:1577–83.
- Rodpai R, Intapan PM, Thanchomngang T, Sanpool O, Janwan P, Laummanwai P, et al. Identification of antigenic proteins in *Strongyloides stercoralis* by proteomic analysis. *Parasitol Res*. 2017;116:1687–93.
- Soblik H, Younis AE, Mitreva M, Renard BY, Kirchner M, Geisinger F, et al. Life cycle stage-resolved proteomic analysis of the excretome/secretome from *Strongyloides ratti*—identification of stage-specific proteases. *Mol Cell Proteomics*. 2011;10:010157.
- Fonseca PDM, Corral MA, Cosenza-Contreras M, Meisel D, Melo GB, Antunes MMS, et al. Shotgun proteomics of *Strongyloides venezuelensis* infective third stage larvae: Insights into host-parasite interaction and novel targets for diagnostics. *Mol Biochem Parasitol*. 2020;235:111249.
- Corral MA, Paula FM, Meisel DM, Castilho VL, Goncalves EM, Levy D, et al. Potential immunological markers for diagnosis of human strongyloidiasis using heterologous antigens. *Parasitology*. 2017;144:124–30.
- Maeda Y, Palomares-Rius JE, Hino A, Afrin T, Mondal SI, Nakatake A, et al. Secretome analysis of *Strongyloides venezuelensis* parasitic stages reveals that soluble and insoluble proteins are involved in its parasitism. *Parasit Vectors*. 2019;12:21.
- Lok JB. *Strongyloides stercoralis*: a model for translational research on parasitic nematode biology. *WormBook*. 2007. <https://doi.org/10.1895/wormbook.1.134.1>.
- Dozio V, Sanchez JC. Profiling the proteomic inflammatory state of human astrocytes using DIA mass spectrometry. *J Neuroinflammation*. 2018;15:331.
- Brandi J, Cheri S, Manfredi M, Di Carlo C, Vita Vanella V, Federici F, et al. Exploring the wound healing, anti-inflammatory, anti-pathogenic and proteomic effects of lactic acid bacteria on keratinocytes. *Sci Rep*. 2020;10:11572.
- Martinotti S, Patrone M, Manfredi M, Gosetti F, Pedrazzi M, Marengo E, et al. HMGB1 osteo-modulatory action on osteosarcoma SaOS-2 cell line: an integrated study from biochemical and -omics approaches. *J Cell Biochem*. 2016;117:2559–69.
- Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res*. 2019;47:D442–50.
- McGinnis S, Madden TL. Blast: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res*. 2004;32:W20–25.
- UniProt C. Uniprot: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49:D480–9.
- Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, et al. The gene ontology (GO) database and informatics resource. *Nucleic Acids Res*. 2004;32:D258–261.
- Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. QuickGo: a web-based tool for gene ontology searching. *Bioinformatics*. 2009;25:3045–6.
- Blum M, Chang HY, Chuguransky S, Grego T, Kandasamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res*. 2021;49:D344–54.
- Zobayer N, Hossain AA, Rahman MA. A combined view of B-cell epitope features in antigens. *Bioinformatics*. 2019;15:530–4.
- Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res*. 2017;45:W24–9.

30. Schrödinger L, DeLano W. Pymol. 2020. <http://www.pymol.org/pymol>. Accessed on 27 Jan 2023
31. Cokelaer T, Pultz D, Harder LM, Serra-Musach J, Saez-Rodriguez J. BioS-ervices: a common Python package to access biological web services programmatically. *Bioinformatics*. 2013;29:3241–2.
32. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583–9.
33. Hunt VL, Tsai IJ, Selkirk ME, Viney M. The genome of *Strongyloides* spp. gives insights into protein families with a putative role in nematode parasitism. *Parasitology*. 2017;144:343–58.
34. Stoltzfus JD, Minot S, Berriman M, Nolan TJ, Lok JB. RNAseq analysis of the parasitic nematode *Strongyloides stercoralis* reveals divergent regulation of canonical dauer pathways. *PLoS Negl Trop Dis*. 2012;6:e1854.
35. Xu L, Yang J, Xu M, Shan D, Wu Z, Yuan D. Speciation and adaptive evolution reshape antioxidant enzymatic system diversity across the phylum Nematoda. *BMC Biol*. 2020;18:181.
36. Henkle-Duhrsen K, Kampkotter A. Antioxidant enzyme families in parasitic nematodes. *Mol Biochem Parasitol*. 2001;114:129–42.
37. Marcilla A, Garg G, Bernal D, Ranganathan S, Forment J, Ortiz J, et al. The transcriptome analysis of *Strongyloides stercoralis* L3i larvae reveals targets for intervention in a neglected disease. *PLoS Negl Trop Dis*. 2012;6:e1513.
38. Pomari E, Malerba G, Veschetti L, Franceschi A, Moron Dalla Tor L, Deiana M, et al. Identification of miRNAs of *Strongyloides stercoralis* L1 and iL3 larvae isolated from human stool. *Sci Rep*. 2022;12:9957.
39. Varatharajulu R, Parandaman V, Ndao M, Andersen JF, Neva FA. *Strongyloides stercoralis* excretory/secretory protein strongylastacin specifically recognized by IgE antibodies in infected human sera. *Microbiol Immunol*. 2011;55:115–22.
40. Donskow-Lysoniewska K, Maruszewska-Cheruiyot M, Stear M. The interaction of host and nematode galectins influences the outcome of gastrointestinal nematode infections. *Parasitology*. 2021;148:648–54.
41. Cantacessi C, Gasser RB. SCP/TAPS proteins in helminths—where to from now? *Mol Cell Probes*. 2012;26:54–9.
42. Pawlowski K. Uncharacterized/hypothetical proteins in biomedical “omics” experiments: is novelty being swept under the carpet? *Brief Funct Genomic Proteomic*. 2008;7:283–90.
43. Sato Y, Inoue F, Matsuyama R, Shiroma Y. Immunoblot analysis of antibodies in human strongyloidiasis. *Trans R Soc Trop Med Hyg*. 1990;84:403–6.
44. Atkins NS, Conway DJ, Lindo JF, Bailey JW, Bundy DA. L3 antigen-specific antibody isotype responses in human strongyloidiasis: correlations with larval output. *Parasite Immunol*. 1999;21:517–26.
45. Conway DJ, Bailey JW, Lindo JF, Robinson RD, Bundy DA, Bianco AE. Serum IgG reactivity with 41-, 31-, and 28-kDa larval proteins of *Strongyloides stercoralis* in individuals with strongyloidiasis. *J Infect Dis*. 1993;168:784–7.
46. Rodpai R, Intapan PM, Thanchomngang T, Sanpool O, Janwan P, Laummaunwai P, et al. *Strongyloides stercoralis* diagnostic polypeptides for human strongyloidiasis and their proteomic analysis. *Parasitol Res*. 2016;115:4007–12.
47. Sudre AP, Siqueira RC, Barreto MG, Peralta RH, Macedo HW, Peralta JM. Identification of a 26-kDa protein fraction as an important antigen for application in the immunodiagnosis of strongyloidiasis. *Parasitol Res*. 2007;101:1117–23.
48. Potocnakova L, Bhide M, Pulzova LB. An introduction to B-cell epitope mapping and in silico epitope prediction. *J Immunol Res*. 2016;2016:6760830.
49. Ditgen D, Anandarajah EM, Reinhardt A, Younis AE, Witt S, Hansmann J, et al. Comparative characterization of two galectins excreted-secreted from intestine-dwelling parasitic versus free-living females of the soil-transmitted nematode *Strongyloides*. *Mol Biochem Parasitol*. 2018;225:73–83.
50. Bisoffi Z, Buonfrate D, Sequi M, Mejia R, Cimino RO, Krolewiecki AJ, et al. Diagnostic accuracy of five serologic tests for *Strongyloides stercoralis* infection. *PLoS Negl Trop Dis*. 2014;8:e2640.
51. Yunus MH, Arifin N, Balachandra D, Anuar NS, Noordin R. Lateral flow dipstick test for serodiagnosis of strongyloidiasis. *Am J Trop Med Hyg*. 2019;101:432–5.
52. Balachandra D, Rahumatullah A, Lim TS, Mustafa FH, Ahmad H, Anuar NS, et al. A new antigen detection ELISA for the diagnosis of strongyloides infection. *Acta Trop*. 2021;221:105986.
53. Rascoe LN, Price C, Shin SH, McAuliffe I, Priest JW, Handali S. Development of Ss-NIE-1 recombinant antigen based assays for immunodiagnosis of strongyloidiasis. *PLoS Negl Trop Dis*. 2015;9:e0003694.
54. Culma MF. *Strongyloides stercoralis* proteome: A reverse approach to the identification of potential immunogenic candidates. *Microb Pathog*. 2021;152:104545.
55. Jaleta TG, Lok JB. Advances in the molecular and cellular biology of *Strongyloides* spp. *Curr Trop Med Rep*. 2019;6:161–78.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

