











## Article

# Excretory/Secretory Proteome of Females and Males of the Hookworm *Ancylostoma ceylanicum*

Samuel C. Uzoечи <sup>1</sup>, Bruce A. Rosa <sup>1</sup> , Kumar Sachin Singh <sup>1</sup>, Young-Jun Choi <sup>1</sup>, Bethany K. Bracken <sup>2</sup>, Paul J. Brindley <sup>3</sup> , R. Reid Townsend <sup>4</sup>, Robert Sprung <sup>4</sup> , Bin Zhan <sup>5</sup> , Maria-Elena Bottazzi <sup>5</sup> , John M. Hawdon <sup>3</sup> , Yide Wong <sup>6</sup> , Alex Loukas <sup>6</sup> , Sergej Djuranovic <sup>7</sup>  and Makedonka Mitreva <sup>1,\*</sup> 

<sup>1</sup> Division of Infectious Diseases, Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA

<sup>2</sup> Charles River Analytics, Inc., Cambridge, MA 02138, USA

<sup>3</sup> Department of Microbiology, Immunology & Tropical Medicine, Research Center for Neglected Diseases of Poverty, School of Medicine and Health Sciences, George Washington University, Washington, DC 20037, USA

<sup>4</sup> Division of Endocrinology, Metabolism and Lipid Research, Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA

<sup>5</sup> Department of Pediatric Tropical Medicine, National School of Tropical Medicine, Baylor College of Medicine, Houston, TX 77030, USA

<sup>6</sup> Centre for Molecular Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns 4878, Australia

<sup>7</sup> Department of Cell Biology and Physiology, Internal Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA

\* Correspondence: mmitreva@wustl.edu

**Abstract:** The dynamic host-parasite mechanisms underlying hookworm infection establishment and maintenance in mammalian hosts remain poorly understood but are primarily mediated by hookworm's excretory/secretory products (ESPs), which have a wide spectrum of biological functions. We used ultra-high performance mass spectrometry to comprehensively profile and compare female and male ESPs from the zoonotic human hookworm *Ancylostoma ceylanicum*, which is a natural parasite of dogs, cats, and humans. We improved the genome annotation, decreasing the number of protein-coding genes by 49% while improving completeness from 92 to 96%. Compared to the previous genome annotation, we detected 11% and 10% more spectra in female and male ESPs, respectively, using this improved version, identifying a total of 795 ESPs (70% in both sexes, with the remaining sex-specific). Using functional databases (KEGG, GO and Interpro), common and sex-specific enriched functions were identified. Comparisons with the exclusively human-infective hookworm *Necator americanus* identified species-specific and conserved ESPs. This is the first study identifying ESPs from female and male *A. ceylanicum*. The findings provide a deeper understanding of hookworm protein functions that assure long-term host survival and facilitate future engineering of transgenic hookworms and analysis of regulatory elements mediating the high-level expression of ESPs. Furthermore, the findings expand the list of potential vaccine and diagnostic targets and identify biologics that can be explored for anti-inflammatory potential.

**Keywords:** hookworm; *Ancylostoma ceylanicum*; excretory/secretory; ESP; proteomics



**Citation:** Uzoечи, S.C.; Rosa, B.A.; Singh, K.S.; Choi, Y.-J.; Bracken, B.K.; Brindley, P.J.; Townsend, R.R.; Sprung, R.; Zhan, B.; Bottazzi, M.-E.; et al. Excretory/Secretory Proteome of Females and Males of the Hookworm *Ancylostoma ceylanicum*. *Pathogens* **2023**, *12*, 95. <https://doi.org/10.3390/pathogens12010095>

Academic Editor: Marcello Otake Sato

Received: 21 November 2022

Revised: 20 December 2022

Accepted: 4 January 2023

Published: 6 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Parasitic infections caused by blood-feeding hookworms are an ongoing global health threat. Hookworms are nematode roundworms of the Ancylostomatidae family, a part of the Strongylida order [1]. Humans are infected with hookworms through ingestion and/or skin penetration of the infectious third-stage larvae (iL3). Worldwide, *Ancylostoma duodenale* and *Necator americanus* were believed to be the most prevalent hookworm species, but more recently, *Ancylostoma ceylanicum* appeared as an important human hookworm parasite in some countries [2]. More than 500 million people including children are affected [3],

with the highest prevalence occurring in tropical regions of Asia followed by Sub-Saharan Africa and North and South America [4–7]. According to CDC, light hookworm infections may be asymptomatic and relatively benign in most cases, although individuals with chronic hookworm infection may experience fatigue, abdominal pain, diarrhea, loss of appetite, weight loss, and iron deficiency anemia [7]. In prevalent regions, children between 2–11 years old are more exposed to hookworm infection which retards physical and cognitive development and leads to anemia and malnutrition [8,9]. Hookworm infection is prevalent in resource-limited areas in most endemic regions with limited access to sanitation, and sewage treatment [10]. There is evidence suggesting that diseases concomitant with prolonged hookworm infection may be a predisposing risk factor to other tropical diseases including tuberculosis, HIV/AIDS, and malaria [11,12]. Total productivity losses as a result of global hookworm infections are estimated to be between \$11 billion and \$138 billion annually (depending on the calculation approach used) [13].

The life cycle of *A. ceylanicum* is straightforward, involving just a single mammalian host, including dogs, cats, and humans [14,15], with no intermediate host required. The adult female hookworms in the small intestine lay eggs that are passed out of the host body in the feces, and under favorable conditions, the eggs embryonate in soil and hatch to release the first-stage larva (L1). The larvae undergo a series of molts to reach the infective third stage (iL3) that can be ingested or penetrate the skin of the host [3]. Upon infection, the iL3 parasite moves to the lungs through the circulatory system. Once in the pulmonary vasculature, the iL3 break into the alveolar spaces and migrate up the trachea where they trigger mild irritation that causes the host to cough. Coughing moves the larvae to the mouth, and they are eventually swallowed, whereupon they enter the gastrointestinal tract, ultimately arriving in the duodenum, where they attach to the small intestine, and develop into adult blood-feeding male and female worms [3].

The adult stage of hookworms excretes and secretes a wide range of structurally and functionally distinct macromolecules called excretory/secretory products (ESPs), mostly comprised of arrays of proteins that are believed to interact with host tissues and stimulate parasite survival [16,17]. In addition to these proteins, ESPs include carbohydrates and lipids, and extracellular vesicles containing nucleic acids are constantly pumped from the parasite's outer surfaces. Based on the wide spectrum of biological functions for ESPs, they are thought to be involved in critical host-parasite interactions that mediate parasite survival in the small intestine and contribute to hookworm infection [6]. Because of this, these macromolecules are strong potential vaccines and drug targets for eradicating hookworm infection, and drug targets for many inflammatory diseases [18]. Moreover, hookworm ESPs contain molecules with therapeutic potentials for treating a whole range of immune-mediated inflammatory diseases [19–21]. Published ESP studies are mostly focused on other hookworm species (*N. americanus* [6] and *A. caninum* [22]), with no emphasis on *A. ceylanicum*. By contrast, little is known about the roles of *A. ceylanicum* ESPs in host-parasite interactions and immunoregulation compared to various parasites that serve as model organisms for human infections. Here, we performed a mass spectrometry analysis of the secretome of cultured supernatant of female and male *A. ceylanicum*. About 790 total proteins were detected in both female and male ESPs. The aim was to profile the ES proteins in female and male *A. ceylanicum* to identify common patterns and sex-specific proteins. These findings provide the first comprehensive proteomic characterization of ESPs of female and male *A. ceylanicum* and allowed us to refine the draft genome.

## 2. Materials and Methods

### 2.1. Ethics

The animal study was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of Washington University in St. Louis (Protocol # 20-0323).

## 2.2. Hamster Gavage Procedure

Hamsters were gavaged using the appropriate technique and were carried out in accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health (NIH). Freshly isolated infective *A. ceylanicum* L3 larvae were counted and distributed in individual Eppendorf tubes, and 80–100 L3 were resuspended per tube in 200  $\mu$ L of PBS. A curved 18G needle was used to gavage the hamsters. Hamsters were manually restrained by grasping the loose skin at the scruff of the neck with the thumb and forefinger to immobilize the head and torso. Using the needle held next to the animal, the distance from the tip of the nose to the last rib on the left side was measured. Holding the animal with the nose pointing up, the gavage needle was entered into the mouth over the tongue and directed toward the esophagus on the left side of the throat, gently pressing the needle on the back of the mouth and allowing the animal to swallow the needle. The needle was slowly advanced to the measured distance into the stomach. Once the needle was properly positioned, the plunger was gently depressed to dispense ~200  $\mu$ L of resuspended L3 larvae to the hamster. Once the entire inoculum was dispensed, the gavage needle was slowly removed, and the animal was returned to its cage. Gavaged hamsters were observed for any difficulty breathing or for bleeding from the mouth or nose for at least 5 min before returning them to the animal holding room.

## 2.3. Isolation of Adult *Ancylostoma Ceylanicum*

Twenty-one days post-infection (dpi) with 80 infective *A. ceylanicum* iL3, three Golden/Syrian hamsters were euthanized using the appropriate technique, equipment, and agents. Compressed CO<sub>2</sub> gas in cylinders was used as a source of carbon dioxide, allowing for the controlled inflow of gas to the induction chamber. Hamsters to be euthanized were placed in a clean transparent chamber 100% CO<sub>2</sub> was introduced at a fill rate of 70% of the chamber volume per minute, to achieve a balanced gas mixture with the existing air in the chamber to fulfill the objective of rapid unconsciousness with minimal distress to the animals. The expected time to unconsciousness was usually within 3–5 min. Each hamster was monitored for lack of respiration and faded eye color. CO<sub>2</sub> was maintained for a minimum of 1 min after observing these signs to avoid unintended recovery upon exposure to natural CO<sub>2</sub> concentrations. Upon completion of the procedure, death was confirmed by ascertaining cardiac and respiratory arrest and noting an animal's fixed and dilated pupils. Following this euthanization procedure, the small intestines were removed and longitudinally cut open. The freshly recovered adult *A. ceylanicum* were gently washed three times in 1 $\times$  PBS containing 5 $\times$  Antibiotics/Antimycotics (Gibco, Waltham, MA, USA). Isolation of adult hookworms from hamsters was performed as previously described [23].

## 2.4. Excretory-Secretory Products (ESPs) Collection

Briefly, adult *A. ceylanicum* were collected in a 10 mm Petri Dish and placed on the inverted microscope stage with a 20 $\times$  objective lens (Nikon Eclipse T<sub>5</sub>2). The freshly recovered adult *A. ceylanicum* were sorted with a nail brush under the microscope to group into females and males. Each group was cultured in three biological replicates at a density of 25 worms/well in 24 well plates. The worms were first cultured for 4 h in a humidified incubator at 37 °C, 5% CO<sub>2</sub>, Dulbecco's Modification of Eagle's Medium (DMEM) (Corning, Corning, NY, USA) supplemented with 2% Antibiotic-Antimycotic (Gibco) and 10% Fetal Bovine Serum (Sigma, St. Louis, MI, USA) [24]. The supernatant was discarded, and the worms were gently rinsed twice in 1 $\times$  PBS containing 5 $\times$  Antibiotics/Antimycotics. Afterward, the worms were subcultured in 2 mL serum-free DMEM supplemented with 2% Antibiotic-Antimycotic at 37 °C, and 5% CO<sub>2</sub> [24]. Every 24 h, over 3 days, at least 1.8 mL of culture supernatant was aspirated from each well with minimal disturbance to the hookworms. Each well was replaced with fresh serum-free medium, and plates were transferred to the incubator, as described [6]. The drawn culture supernatant was centrifuged at 4000 $\times$  g at 4 °C for 30 min and transferred (all but the last 1.6 mL) to a new centrifuge tube and stored at –80 °C. Using this method for further media changes

prevented voided eggs and remaining hamster or dead worm tissue from contaminating the ESPs [6]. After 72 h, no significant alterations in motility or morphological integrity of the worms were detected, suggesting that the drawn media was comprised of ESPs from the normal physiological activity of healthy adult *A. ceylanicum*. However, it should be noted that these ESPs are from worms collected from a hamster host and cultured in media to facilitate the collection, but differences may exist in ESPs actually produced in a human intestine.

At the end of 72 h of media collection, the culture supernatant containing ESPs was gradually thawed in ice. To maximize the peptide yield, the three biological replicates were pooled into a 15 mL tube. Approximately, 14 mL of culture supernatant was recovered from 24 to 72 h of media collection per group. The crude ESPs were further cleaned up by centrifugation at  $7000 \times g$  at  $4\text{ }^{\circ}\text{C}$  for 30 min before concentration and transferred (all but the last 13 mL) to a new tube. The culture media containing ESPs was concentrated using 20 mL Pall Macrosep Advance Centrifugal Devices with a defined 3000 Da molecular weight cutoff membrane. Briefly, 13 mL of culture supernatant was placed into the concentrator sample chamber and capped. The concentrator was then placed into the bucket rotor that contains a 20 mL tube. Concentration by 3K Pall membrane was performed at  $5000 \times g$  at  $4\text{ }^{\circ}\text{C}$  until the retention volume reaches 0.5 mL.

### 2.5. Quantification of Protein Concentration

For total proteomic profiling of adult female and male *A. ceylanicum* proteins, daily ESPs collections of each sex were pooled and quantified. ESPs concentrations were quantified using a Pierce Coomassie Plus Bradford Protein Assay (Thermo Scientific, Waltham, MA, USA). A BSA standard at 1 mg/mL was dissolved in DMEM (Corning) supplemented with 2% Antibiotic-Antimycotic (Gibco). Separately, 5  $\mu\text{L}$  of each standard and ESPs sample were then added to 395  $\mu\text{L}$  of  $1 \times$  Bradford dye in a 96-well plate. Following 5 min of incubation in the dark at room temperature, the optical density (OD) was measured at 595 nm using Eon Microplate Spectrophotometer (Biotek Instruments, Winooski, VT, USA). A standard curve was generated by plotting the graph of OD (y-axis, nm) versus protein concentration (x-axis,  $\mu\text{g}/\text{mL}$ ). The protein concentrations of both female and male ESPs were then estimated based on linear extrapolation from the standard curve.

### 2.6. SDS-PAGE

Approximately 3  $\mu\text{g}$  of *A. ceylanicum* adult ESPs was mixed with  $2 \times$  concentrated Laemmli-SDS buffer (Bio-Rad) with 5% 2-Mercaptoethanol (Sigma-Aldrich). The sample was heated at  $95\text{ }^{\circ}\text{C}$  for 5 min and 50  $\mu\text{L}$  of protein sample was loaded into the wells of 12% Mini-PROTEAN TGX gel (Bio-Rad). The precision protein standard (Bio-Rad) was loaded into the first well at 10  $\mu\text{L}$  without dilution with an SDS sample buffer. Electrophoresis was run for 50 min at 150 V using  $1 \times$  Tris/Glycine buffer (Bio-Rad), [6]. The gel was fixed in (40% ethanol and 10% acetic acid), washed with distilled water, and stained with QC Colloidal Coomassie Stain (Bio-Rad). Gels were imaged in Azure Imaging Systems (C600, Azure Biosystems, Dublin, CA, USA) under LED visible fluorescence detection mode (data not shown).

### 2.7. Sample Preparation for Mass Spectrometry

Concentrated media samples (300  $\mu\text{L}$  of control media, 150  $\mu\text{L}$  of male/female worm media) were precipitated in  $-20\text{ }^{\circ}\text{C}$  acetone (1 part of the sample to 4 parts cold acetone (vol/vol)) overnight followed by centrifugation at 20,000 rcf at for 10 min. Supernatants were removed and media protein pellets were solubilized in 20  $\mu\text{L}$  lysis buffer (8 M urea (Sigma, Cat. No. U4884-500g), 100 mM Tris-HCl buffer (Fisher, Cat. No. BP-152-1), pH 8.0). Samples were reduced with 5 mM DTT (Pierce, Cat. No. A39255) for 1 h at  $37\text{ }^{\circ}\text{C}$ . The reduced Cys residues were alkylated with 10 mM iodoacetamide (Pierce, Cat. No. A39271) for 45 min at room temperature in the dark. Samples were diluted 1:4 with 50 mM Tris-HCl, pH 8.0, and proteins were digested with LysC (Wako Chemicals, Cat. No. 129-02541) (1  $\mu\text{g}$ )

for 2 h at 25 °C in a ThermoMixer (Eppendorf) with shaking at 1000 rpm. Trypsin (Promega, Cat. No. V511C) (1 µg) was added and digestion was performed in the ThermoMixer (100 rpm) overnight at 25 °C. The digestion was quenched by the addition of concentrated formic acid (Fluka) to a final concentration of 1% (vol/vol). Peptides were desalted using stage tips [25] and eluted in 50% (vol/vol) MeCN containing 0.1% (vol/vol) FA. An aliquot (1%) was removed for quantification using the Pierce Quantitative Fluorometric Peptide Assay kit (ThermoFisher, Cat. No. 23290). The remaining peptides were transferred to autosampler vials (Sun-Sri, Cat. No. 200046), dried, and stored at −80 °C.

## 2.8. UPLC-MS

The samples were analyzed using ultra-high-performance mass spectrometry [26] using a hybrid quadrupole Orbitrap LC-MS System, Q-Exactive™ PLUS interfaced to an EASY-nanoLC 1000). A 75 µm i.d. × 50 cm Acclaim PepMap 100 C18 RSLC column (Thermo Scientific™) was equilibrated with 100% solvent A (1%FA) on the nano-LC for a total of 11 µL at 700 bar pressure. Samples in 1% FA (vol/vol) were loaded at a constant pressure of 700 bar. Peptide chromatography was initiated with mobile phase A (1% FA) containing 2% solvent B (100%ACN, 1% FA) for 5 min, then increased to 20% B over 100 min, to 32% B over 20 min, to 95% B over 1 min and held at 95% B for 19 min, with a flow rate of 250 nl/min. Data were acquired in data-dependent mode. Full-scan mass spectra were acquired with the Orbitrap mass analyzer using a scan range of  $m/z = 325$  to 1800 and a mass resolving power set to 70,000. Ten data-dependent high-energy collisional dissociations were performed with a mass resolving power at 17,500, a fixed lower value of  $m/z$  110, an isolation width of 2 Da, and a normalized collision energy setting of 27. The maximum injection time was 60 ms for parent-ion analysis and production analysis. Ions that were selected for MS/MS were dynamically excluded for 30 sec. The automatic gain control was set at a target value of  $1 \times 10^{-6}$  ions for full MS scans and  $1 \times 10^{-5}$  ions for MS2. Peptide ions with charge states of one were excluded from MS2 acquisition. The MS2 spectra from peptides with +2, +3, and +4 charge states were analyzed using Mascot software [27] (Matrix Science, London, UK; version 2.5.1). The mascot was searched against the improved *A. ceylanicum* proteome (22,946 entries, described above), using trypsin enzyme specificity, with a maximum of 2 missed cleavages allowed/peptide. The searches were performed with a fragment ion mass tolerance of 20 ppm and a parent ion tolerance of 20 ppm. Carbamidomethylation of cysteine was specified as a fixed modification. Deamidation of asparagine, deamidation of glutamine, acetylation of protein N-terminus, and oxidation of methionine were entered as variable modifications. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>, accessed on 1 November 2022) via the iProX partner repository [28] with the dataset identifier PXD037410, under project ID IPX0005214000. The accessions for each sample are IPX0005214001 (Adult\_Female\_ESPs), IPX0005214002 (Adult\_Male\_ESPs), and IPX0005214003 (Media-only control). The dataset on iProX includes the .sf3-format Scaffold output file (as described below).

## 2.9. Genome Annotation Improvement and Peptide Identification

A previously published *A. ceylanicum* genome assembly [29] was re-annotated using the MAKER annotation pipeline v3.1.3 [30] to improve the quality of protein-coding gene models. Repetitive elements were softmasked with RepeatMasker v4.1.2 using a species-specific repeat library created by RepeatModeler v2.0.2 [31]. PacBio cDNA sequencing [32] and Illumina RNA-seq (PRJNA72583) data were assembled and used as transcript evidence. Specifically, Clontech SMARTer cDNA synthesis primer sequences were removed from the PacBio CCS reads using lima v2.2.0 and cutadapt v3.4. After poly-A trimming, these reads were aligned to the reference genome using minimap2 [33] following a two-pass approach [34], and subsequently assembled using StringTie2 v2.1.6 (long-read mode) [35]. Illumina RNA-seq data were adapter/quality trimmed using Trimmomatic v0.39 [36], and reference-aligned using HISAT2 v2.2.0 [37] with the dta option and assembled using

StringTie2 v2.1.6. Protein sequences from WormBase ParaSite Version 15 (October 2020) [38] (*Ancylostoma caninum* PRJNA72585, *Caenorhabditis elegans* PRJNA13758, *Haemonchus contortus* PRJEB506, *Mesorhabditis belari* PRJEB30104, *Micoletzkyia japonica* PRJEB27334, *Necator americanus* PRJNA72135, *Nippostrongylus brasiliensis* PRJEB511, *Oscheius tipulae* PRJEB15512, *Pristionchus pacificus* PRJNA12644) were provided to MAKER as protein homology evidence. Gene predictions generated using BRAKER2 v2.1.6 [39] based on the Illumina RNA-seq alignments, together with the above-mentioned evidence data, were synthesized into gene annotations within the MAKER pipeline. Gene predictions without supporting evidence were excluded in the final annotation build, with the exception of those encoding Pfam domains, as detected by InterProScan v5.50 [40] in order to improve the annotation accuracy by balancing sensitivity and specificity [30,41]. The completeness of annotated gene set was assessed using BUSCO v4.1.4 with Nematoda-specific single-copy orthologs (OrthoDB v10) [42]. The newly annotated, improved gene models were used in subsequent proteomic analysis of ESPs, and are publicly available on Nematode.net [43,44]. Functional annotations, re-mapped life cycle RNA-seq read counts and relative gene expression levels (mRNA levels) [45], and proteomics results (described below) for each of the newly annotated genes are provided in Supplementary Table S1.

Scaffold (version Scaffold\_5.1.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 7.0% probability to achieve an FDR less than 1.0% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [46]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Peptide-spectral match (PSM) counts for each peptide sequence are provided in Supplementary Table S2, and the raw scaffold output file with editable threshold values is provided as part of the iProX dataset (PXD037410).

#### 2.10. Bioinformatic Analysis of Proteomic Sequence Data

Peptide and peptide-spectral-match (PSM) counts were quantified per protein according to the identifications from the Mascot/Scaffold pipeline described above. Proteins in each sample were intersected to identify proteins only detected in the adult female E/S, the adult male E/S, or both adult female and male E/S. The newly annotated *A. ceylanicum* proteins were assigned functional annotations using results from InterProScan v5.42 [40] to identify gene ontology [47] classifications and InterPro functional domains [48], and GhostKOALA v2.2 [49] to assign KEGG [50] annotations. Potentially secreted proteins were identified using both SignalP v5.0 [51] for signal peptides and transmembrane domains, and SecretomeP v2.0 [52] to identify proteins with non-classical secretion sequences (where any proteins with 2 or more transmembrane domains were not classified as secreted). Additional protein naming was performed using PANNZER [53] and Sma3s [54]. Significant functional enrichment for GO terms was performed using GOSTATS v2.50 [55] and for Interpro domains and KEGG pathways using WebGestalt v2019 [56] (FDR-adjusted  $p \leq 0.05$ , minimum 3 genes differentially expressed, for all enrichment testing).

Protein conservation data across nematodes and hosts were quantified using BLAST [57] and OrthoFinder [58]. Protein sequence data were downloaded from WormBase Parasite [38] (WBPS16) for *A. duodenale* (PRJNA72581) and *N. brasiliensis* (PRJEB511). The *N. americanus* proteins were based on the improved annotation of the genome (PRJNA72135) presented in Logan et al., 2020 [6], and available for download on Nematode.net [43,44]. The *A. caninum* protein annotation was based on [59]. In addition, *Homo sapiens* (Human; GRCh38.103) was downloaded from Ensembl [60]. For BLAST comparisons, the top BLAST [57] hit for each *A. ceylanicum* protein was identified, including the E value, alignment length, % identity, and whether the top hit was reciprocal (NCBI blastp v2.7.1+, default settings). Orthofinder [58] (v2.4.1) was used to compare protein sequences across all

species and define orthologous protein families (OPFs), using default values. E/S proteins identified in *N. americanus* were retrieved from Logan et al., 2020 [6] and aligned to all *A. ceylanicum* proteins using the BLAST results.

### 3. Results and Discussion

#### 3.1. Improved Genome Annotation Identified ~10% More Peptides in Male and Female Hookworms

As a first step in the analysis, genome annotation improvement was performed to ensure the best possible and most complete characterization of the ES proteome. We re-annotated the genome of *A. ceylanicum* [29] using an improved gene calling pipeline utilizing RNA-seq data [39] and including more extensive up-to-date homologous protein databases (Table 1). The total number of predicted protein-coding genes decreased by 17,904 from 36,687 to 18,783 with substantial increases in both the mean exon and intron lengths. The mean coding sequence (CDS) length increased by 393 bp with the mean gene length (including introns and UTRs) more than doubling from 3.8 kb to 10 kb. The percentage of BUSCOs [42] detected among the predicted genes increased from 92.5% to 96% (with a 72% reduction in the number of fragmented BUSCOs).

**Table 1.** A comparison of genome annotation statistics, before and after re-annotation.

Statistic	Original Annotation (Schwarz et al., 2015) [29]	Re-Annotation
Number of genes	36,687	18,783
Number of mRNAs	65,583	22,928
Overlapping genes	3714	2170
Contained genes	1421	440
Mean gene length (bp)	3819	10,069
Mean exon length (bp)	138	246
Mean intron length (bp)	647	809
Mean CDS length (bp)	923	1316
Mean mRNAs per gene	1.8	1.2
Mean exons per mRNA	6.7	11.5
Complete BUSCOs	86.9%	94.5%
Fragmented BUSCOs	5.5%	1.5%
Missing BUSCOs	7.5%	4.0%

Across both samples, 700 of the peptide sequences (7.94% of the 8818 total peptides) detected using the improved annotation were not part of the previous genome/proteome annotation for *A. ceylanicum* [29]. These peptides represented 10.8% and 9.7% of the total detected spectra in the adult female and male ESPs samples, respectively.

#### 3.2. Female and Male Hookworm Secrete Similar Number of Proteins in Culture

Applying ultra-high-performance mass spectrometry methods using a hybrid quadrupole Orbitrap LC-MS system, a total of 795 unique proteins were detected (from 20.79 µg of female and 8.41 µg proteins in the male ESPs), with no peptides identified in the media-only control sample. Peptides and spectral counts from *A. ceylanicum* ESPs proteomics samples were identified using a Mascot/Scaffold pipeline (1% FDR peptide threshold, 95% protein confidence threshold, minimum 2 peptides per protein). Female and male hookworm secrete a similar number of ESPs in culture (670 vs. 684, respectively). Some 29.7% ( $n = 236$ ) of all detected proteins were uniquely identified in female and male ESPs (Figure 1A). The relative abundance of all detected proteins in the adult male and adult female is shown for the two samples in Figure 1B, with most proteins having relatively similar abundance in both male and female ESPs. Figure 1C shows the adult-stage normalized mRNA levels (RNA-seq FPKM values) in female and male *A. ceylanicum* based on results from our previous study [45]. The female-specific ES proteins (orange) also have substantially higher mRNA levels in the female vs. male (below the diagonal). Likewise, many of the male-specific ES proteins (blue) also have substantially higher mRNA levels in the males vs. females (above the diagonal). This analysis shows that the sex-specific

ESPs also show sex-specific mRNA levels, which supports their identification using an independent dataset.



**Figure 1.** Distribution of ESPs proteins among female and male *A. ceylanicum*. **(A)** A Venn diagram showing the number of detected proteins in the ESPs of female and male *A. ceylanicum*. **(B)** A comparison of the relative abundance of proteins in the female and male samples, quantified by total number of spectra. Sex-specific ESPs are positioned along the X and Y axes. Brief functional annotations are provided for proteins with high sex specificity, and high overall abundance in both samples. **(C)** The corresponding adult-stage mRNA level for each of the ESPs in female and male *A. ceylanicum* [45].

### 3.3. Conserved ESPs in Female and Male *A. Ceylanicum*

A total of 559 proteins were detected in both female and male hookworm (Supplementary Table S1). Table 2 describes the 20 most abundant proteins in both female and male *A. ceylanicum*. As expected for ESPs, the proteins detected in both females and males were significantly enriched for signal peptides for secretion (63.9% of detected ESPs, vs. 18.6% for all proteins;  $p < 10^{-10}$ , binomial distribution test, MS Excel; the threshold for significance  $p \leq 0.01$ ). They were also enriched for CAP domains (22.2% of proteins vs. 1.9% for all proteins;  $p < 10^{-10}$ ), which belong to sperm-coating protein (SCP)-like proteins, also known as SCP/Tpx-1/Ag5/PR-1/Sc7 domain-containing proteins (SCP/TAPS). This SCP-like family of proteins is highly expressed by parasitic nematodes and trematodes [61]. The major



roles of SCP-like proteins are still unclear, but in hookworms, they have been implicated in larval skin penetration, the transition from the free-living to infectious stages [23,62,63], and host immune response modulation [64,65]. There was significant overlap in detected ESPs with orthologs of the ESPs detected in *N. americanus* in the previous study [6] (9.3% of detected ESPs, vs. 0.46% of all proteins;  $p < 10^{-10}$ , binomial distribution test, MS Excel). Signal peptides, CAP domains, and ESPs are indicated for each protein in Table 2 and Supplementary Tables S1 and S2.

Of the sex-agnostic ESPs, ACEY\_07898-1 (Acey\_s0073.g740) is the most abundant by far, having more than 1.8 times as many total spectra assigned as the second-most abundant ESPs. Acey\_s0073.g740 is a very large protein (3373 amino acids) with conservation only in other hookworm species, so there are no *C. elegans* or other annotated orthologs that can be used for comparison. The reciprocal best BLAST hit in *Necator americanus* (NAME\_09181) was also by far the most abundant ESPs in the Logan et al. study [6], which provides confidence in its annotation and detection in the ESPs in this study and makes it an interesting candidate for future study. Among the most abundant proteins on the top list are also several implicated in immunomodulation, including ACEY\_17598-1: Nematode polyprotein allergen Npa-1 [66], which was more than two-fold more abundant in the male than the female ESPs, and ACEY\_09343-1: SCP-like protein [67], and which had very high mRNA levels in the RNA-seq dataset (Table 2). Npa-1, more commonly known as ABA-1, is a common nematode polyprotein allergen/antigen with a fatty acid binding property [68–70]. Ordinarily, nematodes do not have the potential to synthesize fatty acids and must obtain them from the host environment, and Npa-1 is likely used to meet this demand [71]. In the disease model, Npa-1 has been widely used as a serological marker for nematode infection, especially in diseases caused by *Ascaris lumbricoides* and *Ascaris suum* [72,73]. Npa-1 is abundant in the somatic tissues of the parasitic nematodes and is also discharged into the tissues of the host by the infective stages of nematode parasites [74].

KEGG pathway, gene ontology (GO) and Interpro domain enrichment among the proteins in both female and male (Table 3) indicated very strong enrichment for peptidases across all three enrichment tests, especially cysteine peptidases, as indicated by Interpro and GO. Cysteine peptidases are known to be particularly important for maintaining parasitic helminth survival in the host, including roles in the invasion of host tissues and in molting and embryogenesis [75]. Very strongly enriched is also the “exosome” KEGG term ( $p < 10^{-10}$ ) in hookworms; secreted extracellular vesicles are known to interact with host cells and induce immunomodulatory effects, through the action of both proteins and miRNAs [76].

Among the abundant proteins in both female and male *A. ceylanicum* ESPs, the most differentially abundant protein in female (216 spectra vs. 62 in males) is ACEY\_08810-1: Vitellogenin (VIT) (Figure 1B). Vitellogenins are large glycolipoproteins representing the most conserved family of yolk proteins in oviparous species, where the proteins are naturally synthesized in extraovarian organs, transported to the circulatory system, and absorbed by the ovary [77,78]. Studies looking at the control of germline development and the role of soma-germline interactions in *C. elegans* revealed that vitellogenins, which serves as a nutrient source for the developing embryo, are made in the intestine, secreted into the somatic tissue, and taken up by oocytes through receptor-mediated endocytosis [79,80]. Vitellogenin, in addition to fatty acid and retinol-binding (FAR) proteins, are the main two lipid transport and storage proteins engaged in cross-membrane transport [81]. According to Wei et al. [81] more than three vitellogenins are fully expressed in the hookworm gut. Vitellogenin comprises four main domains including phosvitin, lipovitellin-1, lipovitellin-2, and a von Willebrand factor type D domain (vWFD) [82,83]. The three-dimensional structure of vitellogenin shows that lipovitellin and phosvitin domains form amphipathic structures of a lipid-binding pocket to accept their lipid cargo [84].

**Table 2.** The top twenty most abundant proteins in both female and male *A. ceylanicum*. Detected in Logan et al., 2020 [6]. Y = the top *N. americanus* BLAST hit is an ES product in that dataset; “-” = the top *N. americanus* BLAST hit is not an ES product in that dataset; “(N/A)” = there is no significant ( $E < 10^{-5}$ ) BLAST hit in the *N. americanus* protein set. Normalized mRNA levels (Fragments per kilobase per million reads, FPKM) based on the RNA-seq dataset from Bernot et al., 2020 [45].

Protein ID	Best Annotation	Conservation	Detected in Logan et al., 2020 [6]	Sig. pep.	CAP Domain	Adult E/S Proteomics				Avg. mRNA Level (FPKM)	
						Peptide Count		Spectral Count		Ad F	Ad M
						F	M	F	M		
ACEY_07898-1	Acey_s0073.g740	Clade V-specific	Y	-	-	241	261	788	874	638.1	486.6
ACEY_17598-1	5H561: Npa-1-Nematode polyprotein allergen	Clade V conserved (no human)	Y	Y	-	94	128	271	637	302.0	217.0
ACEY_09343-1	ANCCEY_02418: SCP-like protein	<i>Ancylostoma</i> -specific	Y	Y	Y	40	44	374	432	1622.5	3254.1
ACEY_06577-1	ASP2: SCP-like protein	Conserved in humans	Y	Y	Y	48	47	302	328	1876.3	2123.7
ACEY_15233-3	<i>A. ceylanicum</i> -specific	<i>A. ceylanicum</i> -specific	(N/A)	Y	-	19	19	343	277	126.8	140.3
ACEY_15233-2	<i>A. ceylanicum</i> -specific	<i>A. ceylanicum</i> -specific	(N/A)	Y	-	20	17	299	250	126.8	140.3
ACEY_02017-1	Acey_s0009.g591	<i>Ancylostoma</i> -specific	(N/A)	Y	-	18	13	200	206	4718.3	4053.5
ACEY_18401-1	IPR001283: Cysteine-rich secretory protein	<i>Ancylostoma</i> -specific	-	Y	Y	33	39	169	232	857.2	955.1
ACEY_07544-1	asp6: SCP-like protein	Clade V-specific	Y	Y	Y	27	28	184	210	1552.1	1634.6
ACEY_02018-1	Acey_s0009.g591	<i>Ancylostoma</i> -specific	(N/A)	-	-	18	15	183	186	2183.4	2253.6
ACEY_17268-1	SCP domain-containing protein	<i>Ancylostoma</i> -specific	(N/A)	Y	Y	18	22	170	158	16,104	18,856
ACEY_09344-1	asp6: SCP-like protein	<i>Ancylostoma</i> -specific	Y	Y	Y	34	38	159	161	1072.0	1697.4
ACEY_07899-1	Acey_s0073.g746	Clade V-specific	Y	-	-	48	57	130	160	352.1	276.2
ACEY_07543-1	asp6: Secreted protein 6	Clade V-specific	Y	Y	Y	31	35	123	166	449.4	657.6
ACEY_08810-1	VIT: Vitellogenin	Clade V conserved (no human)	Y	-	-	121	53	216	62	3155.6	0.3
ACEY_01671-3	asp5: SCP-like protein	Clade V conserved (no human)	Y	Y	Y	35	36	132	144	1053.7	1217.8
ACEY_13362-1	mtp: Metalloendopeptidase	<i>Ancylostoma</i> -specific	-	Y	-	27	23	131	130	213.3	380.1
ACEY_07900-1	Acey_s0073.g749	Clade V-specific	Y	-	-	63	62	140	117	136.1	162.5
ACEY_09375-1	IPR035109: Ancylostoma-associated secreted protein	<i>Ancylostoma</i> -specific	(N/A)	-	-	8	9	108	113	2634.1	2577.3
ACEY_02245-1	IPR035109: Ancylostoma-associated secreted protein	<i>Ancylostoma</i> -specific	(N/A)	Y	-	23	25	104	98	3281.0	3512.8

**Table 3.** Significantly enriched KEGG pathways, Gene Ontology (molecular function) terms, and Interpro domains among the 559 proteins detected in both male and female ESPs. The top 10 enriched terms are shown for each enrichment category.

Term	Description	Total Term Size	Number of Sig. Genes	FDR-Adjusted <i>p</i> Value
KEGG				
4147	Exosome	578	59	0
1002	Peptidases and inhibitors	406	53	0
3110	Chaperones and folding catalysts	206	32	0
4210	Apoptosis	108	26	0
4612	Antigen processing and presentation	73	23	0
4621	NOD-like receptor signaling pathway	86	21	$6.3 \times 10^{-14}$
536	Glycosaminoglycan binding proteins	126	24	$1.9 \times 10^{-13}$
4142	Lysosome	224	30	$1.1 \times 10^{-12}$
4140	Autophagy-animal	137	24	$1.1 \times 10^{-12}$
10	Glycolysis/Gluconeogenesis	62	10	$8.9 \times 10^{-5}$
GO molecular function				
GO: 0008233	peptidase activity	531	68	$2.1 \times 10^{-24}$
GO: 0016787	hydrolase activity	1238	92	$4.5 \times 10^{-17}$
GO: 0043169	cation binding	733	61	$1.5 \times 10^{-12}$
GO: 0046872	metal ion binding	727	60	$2.9 \times 10^{-12}$
GO: 0004175	endopeptidase activity	358	40	$5.2 \times 10^{-12}$
GO: 0004222	metalloendopeptidase activity	204	30	$5.2 \times 10^{-12}$
GO: 0008234	cysteine-type peptidase activity	105	22	$5.3 \times 10^{-12}$
GO: 0140096	catalytic activity, acting on a protein	1136	76	$1.9 \times 10^{-11}$
GO: 0008237	metallopeptidase activity	248	32	$2.1 \times 10^{-11}$
GO: 0046914	transition metal ion binding	468	40	$1.3 \times 10^{-8}$
Interpro domains				
IPR035940	CAP superfamily	465	136	0
IPR014044	CAP domain	394	124	0
IPR001283	Cysteine-rich secretory protein-related	306	94	0
IPR000668	Peptidase C1A, papain C-terminal	61	23	$4.7 \times 10^{-14}$
IPR025660	Cysteine peptidase, histidine active site	47	20	$2.6 \times 10^{-13}$
IPR000169	Cysteine peptidase, cysteine active site	52	20	$2.2 \times 10^{-12}$
IPR035109	Ancylostoma-associated secreted protein related	136	27	$3.0 \times 10^{-9}$
IPR025661	Cysteine peptidase, asparagine active site	39	15	$4.1 \times 10^{-9}$
IPR038765	Papain-like cysteine peptidase superfamily	110	23	$2.6 \times 10^{-8}$
IPR001506	Peptidase M12A	123	24	$4.1 \times 10^{-8}$

### 3.4. Proteins Detected Only in Female ESPs

Similar to the ESPs detected in both sexes, the 111 female-specific ESPs (14.6% of all female detected proteins) were enriched for signal peptides (49.5% of female-specific ESPs, vs. 18.6% for all proteins;  $p < 10^{-10}$ , binomial distribution test, MS Excel) and CAP domains (9.0% vs. 1.9%;  $p = 8.7 \times 10^{-6}$ ). This included SCP domain-containing protein (ACEY\_06131-2) and Heat shock protein 90 (ACEY\_06074-1) (Table 4). In previous research, many SCP-like proteins that were abundant in the ESPs of the iL3 stage of the canine hookworm *A. caninum* were considered “*Ancylostoma*-secreted proteins” [23,63] (ASPs), and many ASP proteins are identified as being only in the female ESPs here, including the second-most abundant protein, asp4. The most abundant female-specific ESPs also include heat shock proteins (Hsp), which act as molecular chaperones essential for regulating the folding and stability of different proteins during periods of increased temperature and other stress [85]. In *A. ceylanicum* female ES, only two GO terms were enriched among the 111 proteins detected. (Table 5). The represented GO terms related to molecular functions were related to “binding”, including pyridoxal phosphate binding (GO: 0030170) and vitamin B6 binding (GO: 0070279) (Table 5). Hookworm infection can induce vitamin B6 deficiency [86], and another parasitic helminth, *Schistosoma mansoni*, is known to acquire and metabolize host vitamin B6 during some life cycle stages [87], so these results are suggestive of a similar mechanism in *A. ceylanicum* adult females.

**Table 4.** The top 20 most abundant proteins that are only detected in the female ESPs samples. Detected in Logan et al., 2020 [6]. Y = the top *N. americanus* BLAST hit is an ES product in that dataset; “-” = the top *N. americanus* BLAST hit is not an ES product in that dataset; “(N/A)” = there is no significant ( $E < 10^{-5}$ ) BLAST hit in the *N. americanus* protein set. Normalized mRNA levels (Fragments per kilobase per million reads, FPKM) based on RNA-seq dataset from Bernot et al., 2020 [45].

Protein ID	Best Annotation	Conservation	Detected in Logan et al., 2020 [6]	Sig. pep.	CAP Domain	Adult E/S Proteomics				Avg. mRNA Level (FPKM)	
						Peptide Count		Spectral Count		Ad F	Ad M
						F	M	F	M		
ACEY_06131-2	SCP domain-containing protein	<i>Ancylostoma</i> -specific	Y	Y	Y	14	0	33	0	15.7	41.7
ACEY_11619-1	asp4: Secreted protein 4	<i>Ancylostoma</i> -specific	-	Y	Y	7	0	16	0	197.3	0.10
ACEY_18091-1	ANCCEY_04696: SCP-like protein	<i>Ancylostoma</i> -specific	Y	-	Y	6	0	15	0	0.03	0.06
ACEY_06074-1	hspD: Heat shock protein 90	Conserved in humans	-	-	-	11	0	12	0	1143.4	544.7
ACEY_00553-1	nsbp: Putative nucleosome binding protein	Conserved in humans	-	-	-	9	0	9	0	1037.1	701.5
ACEY_03322-1	IPR014044:CAP domain	<i>A. ceylanicum</i> -specific	Y	Y	Y	8	0	9	0	0.16	36.2
ACEY_06771-1	HGS: Metalloendopeptidase	<i>A. ceylanicum</i> -specific	-	Y	-	8	0	9	0	1852.0	0.31
ACEY_17074-1	IPR035109: <i>Ancylostoma</i> -associated secreted protein	<i>A. ceylanicum</i> -specific	(N/A)	-	-	8	0	9	0	49.0	42.3
ACEY_06945-1	ttn-1: Immunoglobulin I-set domain protein	Conserved in humans	-	-	-	8	0	8	0	0.29	0.37
ACEY_03325-1	IPR002172:Low-density lipoprotein (LDL) receptor	Clade V-specific	(N/A)	Y	-	7	0	8	0	0.26	2796.7
ACEY_16559-1	r2: Ribonucleoside-diphosphate reductase	Conserved in humans	-	-	-	6	0	7	0	315	48
ACEY_16507-1	CBG16650: Deoxyribonuclease II	Clade V conserved (no human)	-	-	-	5	0	7	0	44.5	70.7
ACEY_02929-1	CBG06778: Trypsin Inhibitor like cysteine rich domain	Clade V conserved (no human)	-	Y	-	6	0	6	0	129.9	1.2
ACEY_08051-1	Acey_s0076.g1061	Clade V-specific	-	Y	-	6	0	6	0	146.0	0.14
ACEY_14605-1	hspa4: Heat shock protein 105 kDa	Conserved in humans	-	-	-	6	0	6	0	23.8	15.3
ACEY_04573-1	-	Clade V conserved (no human)	Y	Y	-	5	0	6	0	0.12	632.5
ACEY_15954-1	IPR035109: <i>Ancylostoma</i> -associated secreted protein	Clade V-specific	-	Y	-	5	0	6	0	32.9	54.5
ACEY_00191-1	GOT2: Aspartate aminotransferase, mitochondrial	Conserved in humans	-	-	-	5	0	6	0	169.2	56.4
ACEY_03760-1	lgmn: Legumain	Conserved in humans	-	Y	-	5	0	6	0	279.9	256.6
ACEY_10061-1	spna2: Spectrin alpha chain	Conserved in humans	-	-	-	5	0	6	0	31.0	24.2

**Table 5.** Gene Ontology (GO) term analysis, and molecular function in adult female ESPs of *A. ceylanicum*.

Term	Description	Total Term Size	Number of Sig. Genes	FDR-Adjusted <i>p</i> Value
GO molecular function				
GO: 0030170	pyridoxal phosphate binding	35	3	0.037
GO: 0070279	vitamin B6 binding	35	3	0.037

### 3.5. Proteins Detected Only in Male Hookworm

Similarly to the ESPs detected in both sexes and the female-specific ESPs, the 125 male-specific ESPs in *A. ceylanicum* (18.3% of all male-detected proteins) were significantly enriched for signal peptides (53.6% of male-specific ESPs, vs. 18.6% for all proteins;  $p < 10^{-10}$ , binomial distribution test, MS Excel) and CAP domains (16.0% vs. 1.9%;  $p < 10^{-10}$ ). The most abundant male-specific ESPs was ACEY\_02196-1, annotated as ttl1: precursor transthyretin-like protein 1 (Table 6). In a previous study, across 20 nematode secretomes, transthyretin proteins were the only component common across all species in the adult stage [88], highlighting an important yet unknown molecular function for these proteins. In mammals, transthyretin transports the hormones thyroxine and retinol into the circulation [89].

To further understand the functions of proteins identified in the ESPs of adult male *A. ceylanicum*, GO, KEGG, and Interpro domain functional enrichment was performed. In *A. ceylanicum* male ES, only 1 KEGG term was enriched among the 125 proteins detected, whereas 7 GO terms and 11 Interpro IDs were enriched (Table 7). Interestingly, the only male-specific KEGG-enriched term, lectin (4091) has been previously characterized for its importance in the reproductive physiology of *A. ceylanicum* [90]. Lectins are a single or multi-domain large family of calcium-dependent receptors with a wide variety of carbohydrate-recognition domains [91–93]. Our Interpro functional analysis predicted three C-type lectin domains-containing proteins including C-type lectin-like/link domain superfamily (IPR016186), C-type lectin fold (IPR016187), and C-type lectin-like (IPR001304) (Table 7). This lectin domain comprises about 110 to 130 residues. There is increasing evidence that these C-type lectins play an essential role in modulating helminth-mediated immune responses [91]. While these represent a set of male-specific C-type lectins, four others were detected in both males and females, and one was detected only in females. Similarly to the ESPs in both the females and males, the seven enriched GO terms among the male-specific ESPs were all related to peptidase activity (Table 7).

**Table 6.** The top twenty most abundant proteins that are only detected in the male samples. Detected in Logan et al., 2020 [6]. Y = the top *N. americanus* BLAST hit is an ES product in that dataset; “-” = the top *N. americanus* BLAST hit is not an ES product in that dataset; “(N/A)” = there is no significant ( $E < 10^{-5}$ ) BLAST hit in the *N. americanus* protein set. Normalized mRNA levels (Fragments per kilobase per million reads, FPKM) based on the RNA-seq dataset from Bernot et al., 2020 [45].

Protein ID	Best Annotation	Conservation	Detected in Logan et al., 2020 [6]	Sig. pep.	CAP Domain	Adult E/S Proteomics				Avg. mRNA Level (FPKM)	
						Peptide Count		Spectral Count		Ad F	Ad M
						F	M	F	M		
ACEY_02196-1	ttl1: Precursor transthyretin like protein 1	Clade V conserved (no human)	-	Y	-	0	7	0	35	354.7	490.3
ACEY_13694-1	Venom allergen/ <i>Ancylostoma</i> secreted protein-like 18	<i>Ancylostoma</i> -specific	-	Y	Y	0	6	0	31	223.7	184.5
ACEY_06112-1	NECAME_07497: SCP-like protein	<i>Ancylostoma</i> -specific	Y	Y	Y	0	8	0	27	0.21	52.4
ACEY_07390-1	Acey_s0065.g3604	Clade V-specific	-	Y	-	0	17	0	22	0.94	12.7
ACEY_15398-1	inx: MFP2b	Clade V conserved (no human)	-	-	-	0	13	0	17	0.14	177.2
ACEY_06326-1	Acey_s0048.g1585: SCP-like protein	Clade V-specific	-	Y	Y	0	10	0	17	9.95	17.6
ACEY_01267-1	Nematode fatty acid retinoid binding protein	<i>Ancylostoma</i> -specific	Y	-	-	0	14	0	14	0.80	1623.7
ACEY_15396-1	inx: MFP2b	Clade V conserved (no human)	-	-	-	0	12	0	14	0.080	95.5
ACEY_08562-1	Acey_s0087.g2034: SCP-like protein	<i>Ancylostoma</i> -specific	(N/A)	Y	Y	0	9	0	14	0.04	852.7
ACEY_16228-1	SCP domain-containing protein	<i>A. ceylanicum</i> -specific	Y	-	Y	0	4	0	14	2.58	3.0
ACEY_13234-1	von Willebrand factor domain containing protein	Clade V conserved (no human)	-	-	-	0	9	0	11	1.2	12
ACEY_01957-1	asp6: SCP-like protein	<i>Ancylostoma</i> -specific	Y	-	Y	0	8	0	11	0.12	1.0
ACEY_09376-1	ARS: Arylsulfatase	Conserved in humans	-	Y	-	0	9	0	10	2.1	74.7
ACEY_00325-1	asp5: SCP-like protein	<i>Ancylostoma</i> -specific	Y	Y	Y	0	7	0	10	1.0	158.7
ACEY_02166-1	Acey_s0010.g910	Clade V-specific	-	Y	-	0	8	0	9	0.10	498.8
ACEY_15765-1	<i>Ancylostoma</i> -specific	<i>Ancylostoma</i> -specific	(N/A)	Y	-	0	7	0	9	0.43	119.5
ACEY_02233-1	ANCCEY_04287: SCP-like protein	Clade V-specific	Y	Y	Y	0	7	0	8	0.51	241.0
ACEY_15766-1	Acey_s0439.g1485	<i>Ancylostoma</i> -specific	(N/A)	Y	-	0	6	0	8	11.7	141.1
ACEY_02212-1	asp6: SCP-like protein	<i>Ancylostoma</i> -specific	Y	Y	Y	0	6	0	8	0.032	0.29
ACEY_07632-1	PPN: Kunitz/Bovine pancreatic trypsin inhibitor	<i>Ancylostoma</i> -specific	-	Y	-	0	6	0	8	0.038	252.7

**Table 7.** KEGG, GO, and Interpro analysis of molecular function and domain functional annotations in adult male ESPs of *A. ceylanicum*.

Term	Description	Total Term Size	Number of Sig. Genes	FDR-Adjusted <i>p</i> Value
KEGG				
4091	Lectins	43	4	$3.7 \times 10^{-2}$
GO molecular function				
GO: 0061135	endopeptidase regulator activity	146	13	$1.8 \times 10^{-10}$
GO: 0004866	endopeptidase inhibitor activity	146	13	$1.8 \times 10^{-10}$
GO: 0030414	peptidase inhibitor activity	152	13	$2.0 \times 10^{-10}$
GO: 0061134	peptidase regulator activity	156	13	$2.1 \times 10^{-10}$
GO: 0004867	serine-type endopeptidase inhibitor activity	130	12	$3.4 \times 10^{-10}$
GO: 0030234	enzyme regulator activity	273	14	$1.4 \times 10^{-8}$
GO: 0098772	molecular function regulator	298	14	$3.7 \times 10^{-8}$
Interpro domains				
IPR035940	CAP superfamily	465	22	$3.2 \times 10^{-7}$
IPR014044	CAP domain	394	20	$3.5 \times 10^{-7}$
IPR021010	Cytosolic motility protein	7	5	$7.2 \times 10^{-7}$
IPR036880	Pancreatic trypsin inhibitor Kunitz domain superfamily	112	11	$2.6 \times 10^{-6}$
IPR002223	Pancreatic trypsin inhibitor Kunitz domain	107	10	$1.7 \times 10^{-5}$
IPR001283	Cysteine-rich secretory protein-related	306	15	$3.6 \times 10^{-5}$
IPR020901	Proteinase inhibitor I2, Kunitz, conserved site	76	8	$1.1 \times 10^{-4}$
IPR016186	C-type lectin-like/link domain superfamily	111	9	$1.7 \times 10^{-4}$
IPR016187	C-type lectin fold	117	9	$2.4 \times 10^{-4}$
IPR001304	C-type lectin-like	102	8	$7.5 \times 10^{-4}$
IPR008962	PapD-like superfamily	56	5	0.025

#### 4. Conclusions

Here we report the exclusive sensitivity of an efficient mass spectrometry strategy combined with improved genome annotation of *A. ceylanicum* to identify secreted proteins in the ESPs of the adult female and male parasites. This is the first comprehensive profiling of the ESPs proteomes of female and male *A. ceylanicum*. The protein composition of profiled ESPs in adult *A. ceylanicum* leads to the conclusion that adult females and males secrete similar numbers of proteins into culture media, many of which are shared between the sexes, and a subset being differentially represented among the sexes, and related to sex-specific biology. This study increases our understanding of *A. ceylanicum* biology. Our results offer new leads for the development of vaccines and drugs to control the parasites, as well as the identification of new biologics that can be explored for their anti-inflammatory properties. The identified sex-specific secretomes of *A. ceylanicum* will facilitate other analyses, such as the enrichment of regulatory elements and motifs that mediate the expression of matching genes, information that should facilitate the engineering of transgenic *A. ceylanicum* to understand functions of specific genes, or to over-express different proteins of interest.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/pathogens12010095/s1>, Table S1: Gene database including functional annotations (Sma3s, Pannzer2, KEGG, Interpro, Gene Ontology), conservation data, remapped relative mRNA levels (FPKM) for RNA-seq samples spanning the *A. ceylanicum* life cycle, and peptide and summed peptide spectral match (PSM) counts for all detected proteins in the ESPs samples. Table S2: Peptide spectral match (PSM) counts per detected peptide, according to output from Scaffold.

**Author Contributions:** Conceptualization, M.M., S.D., P.J.B., B.K.B. and A.L.; methodology and experiments, S.C.U., R.R.T., K.S.S., R.S. and Y.W.; formal analysis, S.C.U., B.A.R. and Y.-J.C.; resources, J.M.H., S.C.U., M.M., K.S.S., B.Z. and M.-E.B.; writing—original draft preparation, S.C.U., B.A.R., Y.-J.C. and M.M.; writing—review and editing, S.C.U., B.A.R., Y.-J.C., B.K.B., P.J.B., R.R.T., A.L., J.M.H. and M.M.; visualization, S.C.U., B.A.R.; supervision, M.M., S.D.; funding acquisition, M.M., S.D., P.J.B., B.K.B. and A.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The funding was obtained from work partially supported by the Defense Advanced Research Projects Agency (DARPA) and Naval Information Warfare Center Pacific (NIWC Pacific), under Contract No. N66001-21-C-4013 (Approved for Public Release, Distribution Unlimited). The WU-PSR is supported in part by the Washington University in St. Louis Institute of Clinical and Translational Sciences (NCATS UL1 TR000448), the Washington University in St. Louis Mass Spectrometry Research Resource (NIGMS P41 GM103422), and the Siteman Comprehensive Cancer Center Support Grant (NCI P30 CA091842).

**Institutional Review Board Statement:** The animal study was reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of Washington University in St. Louis (Protocol # 20-0323).

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>, accessed on 1 November 2022) via the iProX partner repository with the dataset identifier PXD037410, under project ID IPX0005214000. The accessions for each sample are IPX0005214001 (Adult\_Female\_ESPs), IPX0005214002 (Adult\_Male\_ESPs) and IPX0005214003 (Media-only control).

**Acknowledgments:** The expert technical assistance of Petra Erdmann-Gilmore, Yiling Mi, Alan Davis, and Rose Connors from Washington University Proteomics Shared Resource (WU-PSR) is gratefully acknowledged. We thank John Martin from Mitreva laboratory for data submission.

**Conflicts of Interest:** B.B. is employed by Charles River Analytics, Inc. The remaining authors declare no conflict of interest. The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense of the U.S. Government. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Hodda, M. Phylum Nematoda: A classification, catalogue and index of valid genera, with a census of valid species. *Zootaxa* **2022**, *5114*, 1–289. [[CrossRef](#)] [[PubMed](#)]
2. Inpankaew, T.; Schär, F.; Dalsgaard, A.; Khieu, V.; Chimnoi, W.; Chhoun, C.; Sok, D.; Marti, H.; Muth, S.; Odermatt, P.; et al. High prevalence of *Ancylostoma ceylanicum* hookworm infections in humans, Cambodia, 2012. *Emerg. Infect. Dis.* **2014**, *20*, 976–982. [[CrossRef](#)] [[PubMed](#)]
3. Loukas, A.; Hotez, P.J.; Diemert, D.; Yazdanbakhsh, M.; McCarthy, J.S.; Correa-Oliveira, R.; Croese, J.; Bethony, J.M. Hookworm infection. *Nat. Rev. Dis. Prim.* **2016**, *2*, 16088. [[CrossRef](#)] [[PubMed](#)]
4. De Silva, N.R.; Brooker, S.; Hotez, P.J.; Montresor, A.; Engels, D.; Savioli, L. Soil-transmitted helminth infections: Updating the global picture. *Trends Parasitol.* **2003**, *19*, 547–551. [[CrossRef](#)]
5. Ngui, R.; Lim, Y.A.; Traub, R.; Mahmud, R.; Mistam, M.S. Epidemiological and genetic data supporting the transmission of *Ancylostoma ceylanicum* among human and domestic animals. *PLoS Negl. Trop. Dis.* **2012**, *6*, e1522. [[CrossRef](#)]
6. Logan, J.; Pearson, M.S.; Manda, S.S.; Choi, Y.J.; Field, M.; Eichenberger, R.M.; Mulvenna, J.; Nagaraj, S.H.; Fujiwara, R.T.; Gazzinelli-Guimaraes, P.; et al. Comprehensive analysis of the secreted proteome of adult *Necator americanus* hookworms. *PLoS Negl. Trop. Dis.* **2020**, *14*, e0008237. [[CrossRef](#)]
7. Hotez, P.J.; Brooker, S.; Bethony, J.M.; Bottazzi, M.E.; Loukas, A.; Xiao, S. Hookworm infection. *N. Engl. J. Med.* **2004**, *351*, 799–807. [[CrossRef](#)]
8. Humphries, D.; Simms, B.T.; Davey, D.; Otchere, J.; Quagraine, J.; Terryah, S.; Newton, S.; Berg, E.; Harrison, L.M.; Boakye, D.; et al. Hookworm infection among school age children in Kintampo north municipality, Ghana: Nutritional risk factors and response to albendazole treatment. *Am. J. Trop. Med. Hyg.* **2013**, *89*, 540–548. [[CrossRef](#)]
9. Stephenson, L.S.; Latham, M.C.; Adams, E.J.; Kinoti, S.N.; Pertet, A. Physical fitness, growth and appetite of Kenyan school boys with hookworm, *Trichuris trichiura* and *Ascaris lumbricoides* infections are improved four months after a single dose of albendazole. *J. Nutr.* **1993**, *123*, 1036–1046. [[CrossRef](#)]
10. Hotez, P. Hookworm and poverty. *Ann. N. Y. Acad. Sci.* **2008**, *1136*, 38–44. [[CrossRef](#)]
11. Pullan, R.L.; Bethony, J.M.; Geiger, S.M.; Cundill, B.; Correa-Oliveira, R.; Quinnell, R.J.; Brooker, S. Human helminth co-infection: Analysis of spatial patterns and risk factors in a Brazilian community. *PLoS Negl. Trop. Dis.* **2008**, *2*, e352. [[CrossRef](#)]
12. Geiger, S.M.; Massara, C.L.; Bethony, J.; Soboslay, P.T.; Corrêa-Oliveira, R. Cellular responses and cytokine production in post-treatment hookworm patients from an endemic area in Brazil. *Clin. Exp. Immunol.* **2004**, *136*, 334–340. [[CrossRef](#)]
13. Bartsch, S.M.; Hotez, P.J.; Asti, L.; Zapf, K.M.; Bottazzi, M.E.; Diemert, D.J.; Lee, B.Y. The Global Economic and Health Burden of Human Hookworm Infection. *PLoS Negl. Trop. Dis.* **2016**, *10*, e0004922. [[CrossRef](#)]



14. Bungiro, R.D., Jr.; Anderson, B.R.; Cappello, M. Oral transfer of adult *Ancylostoma ceylanicum* hookworms into permissive and nonpermissive host species. *Infect. Immun.* **2003**, *71*, 1880–1886. [[CrossRef](#)] [[PubMed](#)]
15. Smout, F.A.; Thompson, R.C.; Skerratt, L.F. First report of *Ancylostoma ceylanicum* in wild canids. *Int. J. Parasitol. Parasites Wildl.* **2013**, *2*, 173–177. [[CrossRef](#)] [[PubMed](#)]
16. Kucera, K.; Harrison, L.M.; Cappello, M.; Modis, Y. *Ancylostoma ceylanicum* excretory-secretory protein 2 adopts a netrin-like fold and defines a novel family of nematode proteins. *J. Mol. Biol.* **2011**, *408*, 9–17. [[CrossRef](#)] [[PubMed](#)]
17. Maizels, R.M.; Smits, H.H.; McSorley, H.J. Modulation of Host Immunity by Helminths: The Expanding Repertoire of Parasite Effector Molecules. *Immunity* **2018**, *49*, 801–818. [[CrossRef](#)] [[PubMed](#)]
18. Abuzeid, A.M.I.; Zhou, X.; Huang, Y.; Li, G. Twenty-five-year research progress in hookworm excretory/secretory products. *Parasit. Vectors* **2020**, *13*, 136. [[CrossRef](#)] [[PubMed](#)]
19. Blount, D.; Hooi, D.; Feary, J.; Venn, A.; Telford, G.; Brown, A.; Britton, J.; Pritchard, D. Immunologic profiles of persons recruited for a randomized, placebo-controlled clinical trial of hookworm infection. *Am. J. Trop. Med. Hyg.* **2009**, *81*, 911–916. [[CrossRef](#)] [[PubMed](#)]
20. Smallwood, T.B.; Giacomini, P.R.; Loukas, A.; Mulvanna, J.P.; Clark, R.J.; Miles, J.J. Helminth Immunomodulation in Autoimmune Disease. *Front. Immunol.* **2017**, *8*, 453. [[CrossRef](#)]
21. Navarro, S.; Pickering, D.A.; Ferreira, I.B.; Jones, L.; Ryan, S.; Troy, S.; Leech, A.; Hotez, P.J.; Zhan, B.; Laha, T.; et al. Hookworm recombinant protein promotes regulatory T cell responses that suppress experimental asthma. *Sci. Transl. Med.* **2016**, *8*, 362ra143. [[CrossRef](#)] [[PubMed](#)]
22. Morante, T.; Shepherd, C.; Constantinoiu, C.; Loukas, A.; Sotillo, J. Revisiting the *Ancylostoma Caninum* Secretome Provides New Information on Hookworm-Host Interactions. *Proteomics* **2017**, *17*, 1700186. [[CrossRef](#)] [[PubMed](#)]
23. Hawdon, J.M.; Narasimhan, S.; Hotez, P.J. *Ancylostoma* secreted protein 2: Cloning and characterization of a second member of a family of nematode secreted proteins from *Ancylostoma caninum*. *Mol. Biochem. Parasitol.* **1999**, *99*, 149–165. [[CrossRef](#)] [[PubMed](#)]
24. Kenney, E.T.; Mann, V.H.; Ittiprasert, W.; Rosa, B.A.; Mitreva, M.; Bracken, B.K.; Loukas, A.; Brindley, P.J.; Sotillo, J. Differential Excretory/Secretory Proteome of the Adult Female and Male Stages of the Human Blood Fluke, *Schistosoma mansoni*. *Front. Parasitol.* **2022**, *1*, 950774. [[CrossRef](#)]
25. Mertins, P.; Tang, L.C.; Krug, K.; Clark, D.J.; Gritsenko, M.A.; Chen, L.; Clauser, K.R.; Clauss, T.R.; Shah, P.; Gillette, M.A.; et al. Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. *Nat. Protoc.* **2018**, *13*, 1632–1661. [[CrossRef](#)]
26. Contrepois, K.; Ezan, E.; Mann, C.; Fenaille, F. Ultra-high performance liquid chromatography-mass spectrometry for the fast profiling of histone post-translational modifications. *J. Proteome Res.* **2010**, *9*, 5501–5509. [[CrossRef](#)]
27. Perkins, D.N.; Pappin, D.J.; Creasy, D.M.; Cottrell, J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551–3567. [[CrossRef](#)]
28. Chen, T.; Ma, J.; Liu, Y.; Chen, Z.; Xiao, N.; Lu, Y.; Fu, Y.; Yang, C.; Li, M.; Wu, S.; et al. iProX in 2021: Connecting proteomics data sharing with big data. *Nucleic Acids Res.* **2022**, *50*, D1522–D1527. [[CrossRef](#)]
29. Schwarz, E.M.; Hu, Y.; Antoshechkin, I.; Miller, M.M.; Sternberg, P.W.; Aroian, R.V. The genome and transcriptome of the zoonotic hookworm *Ancylostoma ceylanicum* identify infection-specific gene families. *Nat. Genet.* **2015**, *47*, 416–422. [[CrossRef](#)]
30. Holt, C.; Yandell, M. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinform.* **2011**, *12*, 491. [[CrossRef](#)]
31. Flynn, J.M.; Hubley, R.; Goubert, C.; Rosen, J.; Clark, A.G.; Feschotte, C.; Smit, A.F. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 9451–9457. [[CrossRef](#)] [[PubMed](#)]
32. Magrini, V.; Gao, X.; Rosa, B.A.; McGrath, S.; Zhang, X.; Hallsworth-Pepin, K.; Martin, J.; Hawdon, J.; Wilson, R.K.; Mitreva, M. Improving eukaryotic genome annotation using single molecule mRNA sequencing. *BMC Genom.* **2018**, *19*, 172. [[CrossRef](#)] [[PubMed](#)]
33. Li, H. Minimap and miniasm: Fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* **2016**, *32*, 2103–2110. [[CrossRef](#)] [[PubMed](#)]
34. Parker, M.T.; Knop, K.; Barton, G.J.; Simpson, G.G. 2passtools: Two-pass alignment using machine-learning-filtered splice junctions increases the accuracy of intron detection in long-read RNA sequencing. *Genome Biol.* **2021**, *22*, 72. [[CrossRef](#)] [[PubMed](#)]
35. Kovaka, S.; Zimin, A.V.; Pertea, G.M.; Razaghi, R.; Salzberg, S.L.; Pertea, M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **2019**, *20*, 278. [[CrossRef](#)]
36. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
37. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [[CrossRef](#)]
38. Howe, K.L.; Bolt, B.J.; Shafie, M.; Kersey, P.; Berriman, M. WormBase ParaSite—A comprehensive resource for helminth genomics. *Mol. Biochem. Parasitol.* **2017**, *215*, 2–10. [[CrossRef](#)]
39. Hoff, K.J.; Lange, S.; Lomsadze, A.; Borodovsky, M.; Stanke, M. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics* **2016**, *32*, 767–769. [[CrossRef](#)]
40. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]

41. Campbell, M.S.; Law, M.; Holt, C.; Stein, J.C.; Moghe, G.D.; Hufnagel, D.E.; Lei, J.; Achawanantakun, R.; Jiao, D.; Lawrence, C.J.; et al. MAKER-P: A tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiol.* **2014**, *164*, 513–524. [[CrossRef](#)] [[PubMed](#)]
42. Simao, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [[CrossRef](#)] [[PubMed](#)]
43. Martin, J.; Rosa, B.A.; Ozersky, P.; Hallsworth-Pepin, K.; Zhang, X.; Bhonagiri-Palsikar, V.; Tyagi, R.; Wang, Q.; Choi, Y.J.; Gao, X.; et al. Helminth.net: Expansions to Nematode.net and an introduction to Trematode.net. *Nucleic Acids Res.* **2015**, *43*, D698–D706. [[CrossRef](#)] [[PubMed](#)]
44. Martin, J.; Tyagi, R.; Rosa, B.A.; Mitreva, M. A Multi-Omics Database for Parasitic Nematodes and Trematodes. *Methods Mol. Biol.* **2018**, *1757*, 371–397. [[CrossRef](#)] [[PubMed](#)]
45. Bernot, J.P.; Rudy, G.; Erickson, P.T.; Ratnappan, R.; Haile, M.; Rosa, B.A.; Mitreva, M.; O'Halloran, D.M.; Hawdon, J.M. Transcriptomic analysis of hookworm *Ancylostoma ceylanicum* life cycle stages reveals changes in G-protein coupled receptor diversity associated with the onset of parasitism. *Int. J. Parasitol.* **2020**, *50*, 603–610. [[CrossRef](#)] [[PubMed](#)]
46. Nesvizhskii, A.I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75*, 4646–4658. [[CrossRef](#)]
47. Gene Ontology, C. The Gene Ontology resource: Enriching a Gold mine. *Nucleic Acids Res.* **2021**, *49*, D325–D334. [[CrossRef](#)]
48. Blum, M.; Chang, H.Y.; Chuguransky, S.; Grego, T.; Kandasaamy, S.; Mitchell, A.; Nuka, G.; Paysan-Lafosse, T.; Qureshi, M.; Raj, S.; et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* **2021**, *49*, D344–D354. [[CrossRef](#)]
49. Kanehisa, M.; Sato, Y.; Morishima, K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* **2016**, *428*, 726–731. [[CrossRef](#)]
50. Kanehisa, M.; Furumichi, M.; Sato, Y.; Ishiguro-Watanabe, M.; Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **2021**, *49*, D545–D551. [[CrossRef](#)]
51. Almagro Armenteros, J.J.; Tsirigos, K.D.; Sønderby, C.K.; Petersen, T.N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **2019**, *37*, 420–423. [[CrossRef](#)] [[PubMed](#)]
52. Bendtsen, J.D.; Jensen, L.J.; Blom, N.; Von Heijne, G.; Brunak, S. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* **2004**, *17*, 349–356. [[CrossRef](#)] [[PubMed](#)]
53. Toronen, P.; Holm, L. PANNZER-A practical tool for protein function prediction. *Protein Sci.* **2022**, *31*, 118–128. [[CrossRef](#)] [[PubMed](#)]
54. Casimiro-Soriguer, C.S.; Munoz-Merida, A.; Perez-Pulido, A.J. Sma3s: A universal tool for easy functional annotation of proteomes and transcriptomes. *Proteomics* **2017**, *17*, 1700071. [[CrossRef](#)] [[PubMed](#)]
55. Falcon, S.; Gentleman, R. Using GOstats to test gene lists for GO term association. *Bioinformatics* **2007**, *23*, 257–258. [[CrossRef](#)]
56. Liao, Y.; Wang, J.; Jaehnic, E.J.; Shi, Z.; Zhang, B. WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **2019**, *47*, W199–W205. [[CrossRef](#)]
57. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
58. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238. [[CrossRef](#)]
59. Comparative genomics of the major parasitic worms. *Nat. Genet.* **2019**, *51*, 163–174. [[CrossRef](#)]
60. Howe, K.L.; Achuthan, P.; Allen, J.; Allen, J.; Alvarez-Jarreta, J.; Amode, M.R.; Armean, I.M.; Azov, A.G.; Bennett, R.; Bhai, J.; et al. Ensembl 2021. *Nucleic Acids Res.* **2021**, *49*, D884–D891. [[CrossRef](#)]
61. Cantacessi, C.; Gasser, R.B. SCP/TAPS proteins in helminths—Where to from now? *Mol. Cell. Probes* **2012**, *26*, 54–59. [[CrossRef](#)] [[PubMed](#)]
62. Moser, J.M.; Freitas, T.; Arasu, P.; Gibson, G. Gene expression profiles associated with the transition to parasitism in *Ancylostoma caninum* larvae. *Mol. Biochem. Parasitol.* **2005**, *143*, 39–48. [[CrossRef](#)] [[PubMed](#)]
63. Hawdon, J.M.; Jones, B.F.; Hoffman, D.R.; Hotez, P.J. Cloning and characterization of *Ancylostoma*-secreted protein. A novel protein associated with the transition to parasitism by infective hookworm larvae. *J. Biol. Chem.* **1996**, *271*, 6672–6678. [[CrossRef](#)]
64. Chen, J.; Hu, X.; He, S.; Wang, L.; Hu, D.; Wang, X.; Zheng, M.; Yang, Y.; Liang, C.; Xu, J.; et al. Expression and immune response analysis of *Schistosoma japonicum* VAL-1, a homologue of vespid venom allergens. *Parasitol. Res.* **2010**, *106*, 1413–1418. [[CrossRef](#)] [[PubMed](#)]
65. Ryan, S.M.; Ruscher, R.; Johnston, W.A.; Pickering, D.A.; Kennedy, M.W.; Smith, B.O.; Jones, L.; Buitrago, G.; Field, M.A.; Esterman, A.J.; et al. Novel antiinflammatory biologics shaped by parasite-host coevolution. *Proc. Natl. Acad. Sci. USA* **2022**, *119*, e2202795119. [[CrossRef](#)]
66. Kennedy, M.W. The nematode polyprotein allergens/antigens. *Parasitol. Today* **2000**, *16*, 373–380. [[CrossRef](#)]
67. Moyle, M.; Foster, D.L.; McGrath, D.E.; Brown, S.M.; Laroche, Y.; De Meutter, J.; Stanssens, P.; Bogowitz, C.A.; Fried, V.A.; Ely, J.A.; et al. A hookworm glycoprotein that inhibits neutrophil function is a ligand of the integrin CD11b/CD18. *J. Biol. Chem.* **1994**, *269*, 10008–10015. [[CrossRef](#)]

68. Xia, Y.; Spence, H.J.; Moore, J.; Heaney, N.; McDermott, L.; Cooper, A.; Watson, D.G.; Mei, B.; Komuniecki, R.; Kennedy, M.W. The ABA-1 allergen of *Ascaris lumbricoides*: Sequence polymorphism, stage and tissue-specific expression, lipid binding function, and protein biophysical properties. *Parasitology* **2000**, *120 Pt 2*, 211–224. [[CrossRef](#)]
69. Acevedo, N.; Sánchez, J.; Erler, A.; Mercado, D.; Briza, P.; Kennedy, M.; Fernandez, A.; Gutierrez, M.; Chua, K.Y.; Cheong, N.; et al. IgE cross-reactivity between *Ascaris* and domestic mite allergens: The role of tropomyosin and the nematode polyprotein ABA-1. *Allergy* **2009**, *64*, 1635–1643. [[CrossRef](#)]
70. Kennedy, M.W.; Brass, A.; McCrudden, A.B.; Price, N.C.; Kelly, S.M.; Cooper, A. The ABA-1 allergen of the parasitic nematode *Ascaris suum*: Fatty acid and retinoid binding function and structural characterization. *Biochemistry* **1995**, *34*, 6700–6710. [[CrossRef](#)]
71. Fairfax, K.C.; Harrison, L.M.; Cappello, M. Molecular cloning and characterization of a nematode polyprotein antigen/allergen from the human and animal hookworm *Ancylostoma ceylanicum*. *Mol. Biochem. Parasitol.* **2014**, *198*, 37–44. [[CrossRef](#)] [[PubMed](#)]
72. Turner, J.D.; Faulkner, H.; Kamgno, J.; Kennedy, M.W.; Behnke, J.; Boussinesq, M.; Bradley, J.E. Allergen-specific IgE and IgG4 are markers of resistance and susceptibility in a human intestinal nematode infection. *Microbes Infect.* **2005**, *7*, 990–996. [[CrossRef](#)] [[PubMed](#)]
73. McSharry, C.; Xia, Y.; Holland, C.V.; Kennedy, M.W. Natural immunity to *Ascaris lumbricoides* associated with immunoglobulin E antibody to ABA-1 allergen and inflammation indicators in children. *Infect. Immun.* **1999**, *67*, 484–489. [[CrossRef](#)] [[PubMed](#)]
74. Kennedy, M.W.; Qureshi, F. Stage-specific secreted antigens of the parasitic larval stages of the nematode *Ascaris*. *Immunology* **1986**, *58*, 515–522.
75. Grote, A.; Caffrey, C.R.; Rebello, K.M.; Smith, D.; Dalton, J.P.; Lustigman, S. Cysteine proteases during larval migration and development of helminths in their final host. *PLoS Negl. Trop. Dis.* **2018**, *12*, e0005919. [[CrossRef](#)]
76. Eichenberger, R.M.; Ryan, S.; Jones, L.; Buitrago, G.; Polster, R.; Montes de Oca, M.; Zuvelek, J.; Giacomini, P.R.; Dent, L.A.; Engwerda, C.R.; et al. Hookworm Secreted Extracellular Vesicles Interact With Host Cells and Prevent Inducible Colitis in Mice. *Front. Immunol.* **2018**, *9*, 850. [[CrossRef](#)]
77. Erdmann, E.A.; Abraham, O.; Hundley, H.A. *Caenorhabditis elegans* expressing a Vitellogenin::GFP fusion protein show reduced embryo content and brood size. *MicroPubl. Biol.* **2022**. [[CrossRef](#)]
78. Ramos, I.; Machado, E.; Masuda, H.; Gomes, F. Open questions on the functional biology of the yolk granules during embryo development. *Mol. Reprod. Dev.* **2022**, *89*, 86–94. [[CrossRef](#)]
79. Grant, B.; Hirsh, D. Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. *Mol. Biol. Cell* **1999**, *10*, 4311–4326. [[CrossRef](#)]
80. Hall, D.H.; Winfrey, V.P.; Blaeuer, G.; Hoffman, L.H.; Furuta, T.; Rose, K.L.; Hobert, O.; Greenstein, D. Ultrastructural features of the adult hermaphrodite gonad of *Caenorhabditis elegans*: Relations between the germ line and soma. *Dev. Biol.* **1999**, *212*, 101–123. [[CrossRef](#)]
81. Wei, J.; Damania, A.; Gao, X.; Liu, Z.; Mejia, R.; Mitreva, M.; Strych, U.; Bottazzi, M.E.; Hotez, P.J.; Zhan, B. The hookworm *Ancylostoma ceylanicum* intestinal transcriptome provides a platform for selecting drug and vaccine candidates. *Parasit. Vectors* **2016**, *9*, 518. [[CrossRef](#)] [[PubMed](#)]
82. Yamamura, J.; Adachi, T.; Aoki, N.; Nakajima, H.; Nakamura, R.; Matsuda, T. Precursor-product relationship between chicken vitellogenin and the yolk proteins: The 40 kDa yolk plasma glycoprotein is derived from the C-terminal cysteine-rich domain of vitellogenin II. *Biochim. Biophys. Acta* **1995**, *1244*, 384–394. [[CrossRef](#)] [[PubMed](#)]
83. Baker, M.E. Invertebrate vitellogenin is homologous to human von Willebrand factor. *Biochem. J.* **1988**, *256*, 1059–1061. [[CrossRef](#)] [[PubMed](#)]
84. Finn, R.N. Vertebrate yolk complexes and the functional implications of phosvitins and other subdomains in vitellogenins. *Biol. Reprod.* **2007**, *76*, 926–935. [[CrossRef](#)] [[PubMed](#)]
85. Gelmedin, V.; Delaney, A.; Jennelle, L.; Hawdon, J.M. Expression profile of heat shock response factors during hookworm larval activation and parasitic development. *Mol. Biochem. Parasitol.* **2015**, *202*, 1–14. [[CrossRef](#)] [[PubMed](#)]
86. Nontasut, P.; Changbumrung, S.; Muennoo, C.; Hongthong, K.; Vudhivai, N.; Sanguankiat, S.; Yaemput, S. Vitamin B1, B2 and B6 deficiency in primary school children infected with hookworm. *Southeast Asian J. Trop. Med. Public Health* **1996**, *27*, 47–50.
87. Da’dara, A.A.; Elzoheiry, M.; El-Beshbishi, S.N.; Skelly, P.J. Vitamin B6 Acquisition and Metabolism in *Schistosoma mansoni*. *Front. Immunol.* **2020**, *11*, 622162. [[CrossRef](#)]
88. Tritten, L.; Ballesteros, C.; Beech, R.; Geary, T.G.; Moreno, Y. Mining nematode protein secretomes to explain lifestyle and host specificity. *PLoS Negl. Trop. Dis.* **2021**, *15*, e0009828. [[CrossRef](#)]
89. Magalhaes, J.; Eira, J.; Liz, M.A. The role of transthyretin in cell biology: Impact on human pathophysiology. *Cell. Mol. Life Sci.* **2021**, *78*, 6105–6117. [[CrossRef](#)]
90. Brown, A.C.; Harrison, L.M.; Kapulkin, W.; Jones, B.F.; Sinha, A.; Savage, A.; Villalon, N.; Cappello, M. Molecular cloning and characterization of a C-type lectin from *Ancylostoma ceylanicum*: Evidence for a role in hookworm reproductive physiology. *Mol. Biochem. Parasitol.* **2007**, *151*, 141–147. [[CrossRef](#)]
91. van Die, I.; Cummings, R.D. The Mannose Receptor in Regulation of Helminth-Mediated Host Immunity. *Front. Immunol.* **2017**, *8*, 1677. [[CrossRef](#)] [[PubMed](#)]

92. Drickamer, K. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **1988**, *263*, 9557–9560. [[CrossRef](#)] [[PubMed](#)]
93. Drickamer, K. Evolution of Ca<sup>(2+)</sup>-dependent animal lectins. *Prog. Nucleic Acid Res. Mol. Biol.* **1993**, *45*, 207–232. [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.