



OPEN ACCESS

EDITED BY
Sasisekhar Bennuru,
NIH, United States

REVIEWED BY
Bruce A. Rosa,
Washington University in St. Louis,
United States
Dong Xia,
Royal Veterinary College (RVC),
United Kingdom

*CORRESPONDENCE
Katerina Artavanis-Tsakonas
✉ ka447@cam.ac.uk

[†]These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to
Parasitology Omics,
a section of the journal
Frontiers in Parasitology

RECEIVED 24 October 2022
ACCEPTED 04 January 2023
PUBLISHED 25 January 2023

CITATION
Nash B, Gregory WF, White RR,
Protasio AV, Gygi SP, Selkirk ME,
Weekes MP and Artavanis-Tsakonas K
(2023) Large-scale proteomic analysis of *T. spiralis*
muscle-stage ESPs identifies a novel
upstream motif for *in silico* prediction
of secreted products.
Front. Parasitol. 2:1078443.
doi: 10.3389/fpara.2023.1078443

COPYRIGHT
© 2023 Nash, Gregory, White, Protasio, Gygi,
Selkirk, Weekes and Artavanis-Tsakonas. This
is an open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Large-scale proteomic analysis of *T. spiralis* muscle-stage ESPs identifies a novel upstream motif for *in silico* prediction of secreted products

Bradley Nash^{1†}, William F. Gregory^{1†}, Rhiannon R. White², Anna V. Protasio¹, Steve P. Gygi³, Murray E. Selkirk², Michael P. Weekes⁴ and Katerina Artavanis-Tsakonas^{1*}

¹Department of Pathology, University of Cambridge, Cambridge, United Kingdom, ²Department of Life Sciences, Imperial College London, London, United Kingdom, ³Department of Cell Biology, Harvard Medical School, Boston, MA, United States, ⁴Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom

The *Trichinella* genus contains parasitic nematodes capable of infecting a wide range of hosts including mammals, birds and reptiles. Like other helminths, *T. spiralis* secretes a complex mixture of bioactive molecules capable of modulating its immediate surroundings and creating a hospitable environment for growth, survival and ultimately transmission. The constitution of these excretory-secretory products (ESPs) changes depending on the tissue niche and the specific stage of parasite development. Unique to *T. spiralis* is a true intracellular stage wherein larvae develop inside striated myotubes. Remarkably, the parasite larvae do not destroy the host cell but rather reprogram it to support their presence and growth. This transformation is largely mediated through stage-specific secretions released into the host cell cytoplasm. In this study, we apply state of the art proteomics and computational approaches to elucidate the composition and functions of muscle-stage *T. spiralis* ESPs. Moreover, we define a recurring, upstream motif associated with the stichosome, the main secretory organ of this worm, and can be used to predict secreted proteins across experimentally less tractable *T. spiralis* life cycle stages.

KEYWORDS

Trichinella, secretions, ES, proteomics, helminth

Introduction

Trichinella spiralis is an obligate parasitic nematode and the leading cause of Trichinellosis (also called Trichinosis). During infection with *T. spiralis*, the new-born larva (NBL) of the parasite invades a terminally differentiated myofiber and resides intracellularly. The growth and development of the larva coincides with the

reprogramming of the host muscle cell into a nurse cell for long-term inhabitation. The invading NBL and the host muscle cells undergo significant change during the process of parasite niche remodelling. The host cell re-enters the cell cycle, loses muscular specialization, and gains a thick collagen capsule. The NBL undergoes significant growth, its cuticle thickens, and it develops the secretory organ, the stichosome, that matures by day 7 post-invasion (Khan, 1966; Lee and Wright, 1984). The proteins secreted from the stichosome are hypothesized to mediate the transformation of a terminally differentiated muscle cell into a nurse cell complex, capable of supporting *T. spiralis* development (Despommier, 1998).

Prior to the sequencing of the *T. spiralis* genome, the identification of *T. spiralis* excretory/secretory products (ESPs) using mass spectrometry was initially achieved using 2-D Electrophoresis for individual spot-based protein identification. The muscle-stage larva (MSL) ESPs were found to be highly tylosylated, isoform rich, and increasing in complexity during developmental progression (Robinson and Connolly, 2005; Robinson et al., 2005). Due to the complex composition of the ESPs, it is expected that the ESPs of the early MSL (eMSL) differ from the ESPs of the mature MSL (mMSL). However, isolation of the eMSL is hindered by their fragility and sensitivity to acidified pepsin up to Day 11 post-invasion (Khan, 1966; Kozek, 1971). The collection of ESPs from stages other than mMSL still requires the development of experimental procedures to effectively overcome these limitations.

Protein identification was greatly improved by the 2011 release of the *T. spiralis* draft genome (PRJNA12603) (Mitreva et al., 2011). The proteomic annotations combined 3,262 genes identified from expressed sequence tag screens on Adult, MSL and NBL stages (Mitreva et al., 2004) with *ab initio* prediction using the software FgenesH and EAnnot. This laid the groundwork for the *T. spiralis* proteome containing a predicted 15,808 genes at the time of publication (Mitreva et al., 2011) and a later expansion to 16,380 genes. This advancement allowed for interpretation of the protein functions of MSL ESPs, identifying proteases, nucleases, glucose metabolizing enzymes, and heat shock proteins from a single screen of only immunogenic bands from 1D-Electrophoresis (Bien et al., 2015).

Previously our lab identified the E2 Ubiquitin-conjugating protein TsUBE2L3 (T01_4545) from the ESPs of the mMSL and demonstrated its localization to the stichosome of the muscle stage larvae (White et al., 2016). Subsequent to this work, a second draft genome (PRJNA257433) has been constructed alongside transcriptomic profiling of pepsin-released MSL of *T. spiralis* and 15 other *Trichinella* species. The proteome annotations represent a dramatic improvement, combining transcriptomic data with homology-directed gene predictions, which were used to train further prediction of non-homologous genes (Korhonen et al., 2016). In this study we sought to define the major activities of the ESPs to predict the host cell processes targeted by *T. spiralis* during muscle stage infection. We identified enriched pathways and abundant protein superfamilies that likely play a critical role in regulating the worm's unique intracellular niche. We also sought to find a way to identify secreted, host-targeted genes bioinformatically. To this end we identified the PouStich motif which we propose to be a stichosome-specific regulatory sequence.

Results and discussion

Collection and annotation of ESPs

Mature L1 larvae were released from rat infected muscle by acidified pepsin digestion and cultured *in vitro*. The ESPs secreted into the culture medium were collected for a limited duration of 24 hours to minimize the risk of obtaining intracellular contaminants from dying worms. Secretions were concentrated and subjected to LC/MS/MS analysis as described in White et al. (2016). We identified 398 ESPs from either of 2 biological replicates and 208 proteins from both replicates (Supplementary Table 1). To understand how the worm modulates the host environment, we proceeded to analyze bioinformatically the proteomic data to predict function, activity, localization, and regulation of these secreted proteins.

KAAS KEGG analysis reveals enrichment of glycolysis pathway components

KEGG Ontology (KO) terms were assigned to ESPs using the KEGG Automatic Annotation Server (KAAS) (Moriya et al., 2007). This analysis assigned KO terms to 48% of the ESPs (190/398) and only 26% of the *T. spiralis* proteome (3661/14,258) (Figure 1). Hypergeometric enrichment was used to determine whether members were present in the ESPs to a greater extent than would be expected based on their presence in the proteome for each KEGG pathway, which would suggest targeting of host cell pathways by *T. spiralis* ESPs (Table 1).

We observed enrichment of proteins involved in the KEGG pathway of hypertrophic cardiomyopathy, which was initially surprising. However, this condition stems from mitochondrial

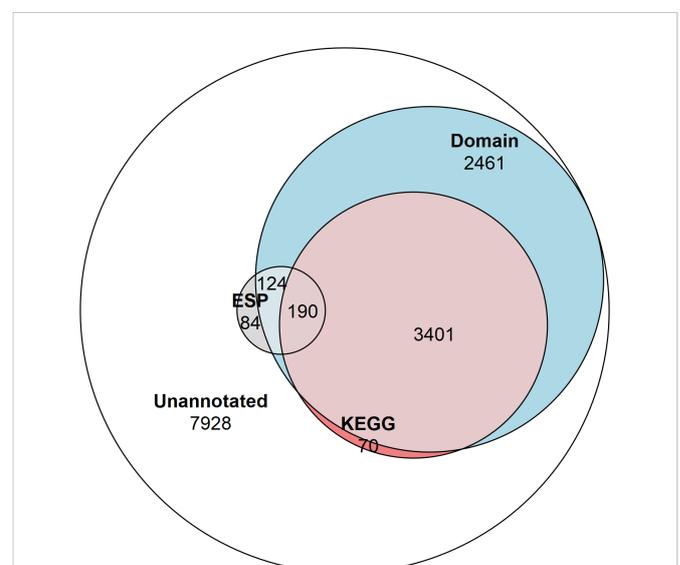


FIGURE 1
Venn diagram showing the proportion of the proteome and ESPs annotated. The number of NCBI Conserved Domain Search annotated proteins (6,176 total/314 ESPs), and KEGG Ontology annotated proteins (3,661 total/190 ESPs), and proteins not annotated by these approaches (7,928 total/84 ESPs).

TABLE 1 Top 10 enriched KEGG Pathways by excretory/secretory proteins (ESPs).

KEGG Pathway		No. Hits		Hypergeometric Enrichment
ID	Name	Proteome	ESP	P-value
ko01200	Carbon metabolism	57	16	4.2×10^{-10}
ko00010	Glycolysis/Gluconeogenesis	25	11	2.2×10^{-9}
ko04142	Lysosome	46	13	1.5×10^{-8}
ko00710	Carbon fixation in photosynthetic organisms	13	7	9.4×10^{-7}
ko01230	Biosynthesis of amino acids	31	9	4.5×10^{-6}
ko04141	Protein processing in endoplasmic reticulum	68	12	1.1×10^{-5}
ko00500	Starch and sucrose metabolism	19	7	1.4×10^{-5}
ko05410	Hypertrophic cardiomyopathy	20	7	1.8×10^{-5}
ko00250	Alanine, aspartate, and glutamate metabolism	10	5	8.4×10^{-5}
ko00230	Purine metabolism	38	8	1.6×10^{-4}

Enrichment of KEGG Pathway members within the ESPs at a number of hits greater than expected based on the number of hits in the Proteome was determined using a hypergeometric test for enrichment (Phyfer). The KEGG Pathways enriched by ESPs are ordered based on P value as an indication of the strength and reliability of the enrichment.

dysfunction (Meyers et al., 2013), which is a hallmark of *T. spiralis*-infected muscle cells. Furthermore, specific ESPs were assigned KO terms K05849 (SLC8A, NCX) and K07200 (PRKAG), corresponding to proteins that can cause hypertrophic cardiomyopathy when dysregulated (Khananshvilii, 2013; Banankhah et al., 2018). The secretion of such proteins into the nurse cell could affect both Ca²⁺ concentrations required for normal muscle function and AMP responses which can encourage glycogen accumulation. ESPs were also enriched for carbohydrate metabolism pathways. Amylase, maltase and glucoamylase were identified in ESPs, which most likely serve a digestive function for larvae in the intestinal tract to convert starch and maltose to glucose, which could then be taken up by parasites. More interestingly, the entire glycolytic pathway was represented in ESPs with only 2 omissions, with entry into the pathway provided by glycogen phosphorylase (GP) rather than hexokinase (Figure 2). This suggests that *T. spiralis* mobilizes glycogen stores, and as most animals store glycogen in skeletal muscle and the liver, this could be an important factor in growth and survival of parasites in muscle cells. Glycogen phosphorylase, phosphoglucomutase (PGM) and glucose-6-phosphate isomerase (GPI) are all represented in ESPs, which would drive conversion of glycogen to fructose-6-phosphate. However phosphofructokinase (PFK), a critical regulatory step for flux through glycolysis (Sola-Penna et al., 2010) was missing, either because it was below our level of detection, or suggesting that this step is performed by the host. Enzymes required for conversion of fructose-1,6-bisphosphate (F16BP) to phosphoenolpyruvate (PEP) were all present, with the notable exception of glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

In anaerobic helminths, PEP can be converted to oxaloacetate by phosphoenolpyruvate carboxykinase (PEPCK) and then to pyruvate via malate by the process known as malate dismutation, rather than conversion of PEP to pyruvate by pyruvate kinase (PK) (Komuniecki and Tielens, 2003) (Ward et al., 1969). It was notable that all enzymes necessary for malate dismutation were present in ESPs, and that PK was missing. Enzymes that form the tricarboxylic acid (TCA) cycle

were much less prevalent. Only 3 enzymes, malate dehydrogenase (MDH), citrate synthase (CS) and aconitase were found, which collectively could convert malate through to isocitrate. Further progression through the TCA cycle would therefore presumably be dependent on host enzymes.

It is possible that mobilisation of glycogen stores and secretion of glycolytic enzymes by *T. spiralis* alters host cell metabolism. Blood glucose levels are reduced from day 8 to 13 post infection, which can result in hypothermia (Nishina and Suzuki, 2002; Nishina et al., 2004). Hypoglycaemia during the larval growth stage has been attributed to upregulation of insulin response genes, causing the glucose channel GLUT4 to be trafficked to the membrane, increasing glucose uptake. In this phase, nurse cells also build up glycogen stores (Wu et al., 2009), which may be necessary to sustain the exponential growth of the parasite (Despommier, 1975).

Mature nurse cells exhibit a higher capacity for glucose uptake than L6 rat myoblasts during *in vitro* culture, which is dependent on both GLUT1/4 channels and H⁺/glucose import (Montgomery et al., 2003). Nevertheless, we did not identify any *T. spiralis* secreted proteins which could mediate glucose uptake or drive glycogen synthesis. This suggests that both of these processes are regulated by host cells, whereas catabolism of glycogen may be driven largely by parasite secreted enzymes. Although some parasite glycolytic enzymes have been demonstrated to possess secondary functions (Gómez-Arreaza et al., 2014) (Jiang et al., 2019), the near completeness of the glycolytic pathway represented by *T. spiralis* secreted proteins is strongly suggestive that they function primarily to regulate carbohydrate metabolism.

Domain analysis reveals ESPs are particularly rich in proteases

Though beneficial for the confident identification of orthologs, KO terms can only be assigned to proteins that have a well-studied homolog, with a determined function, within a well-characterised

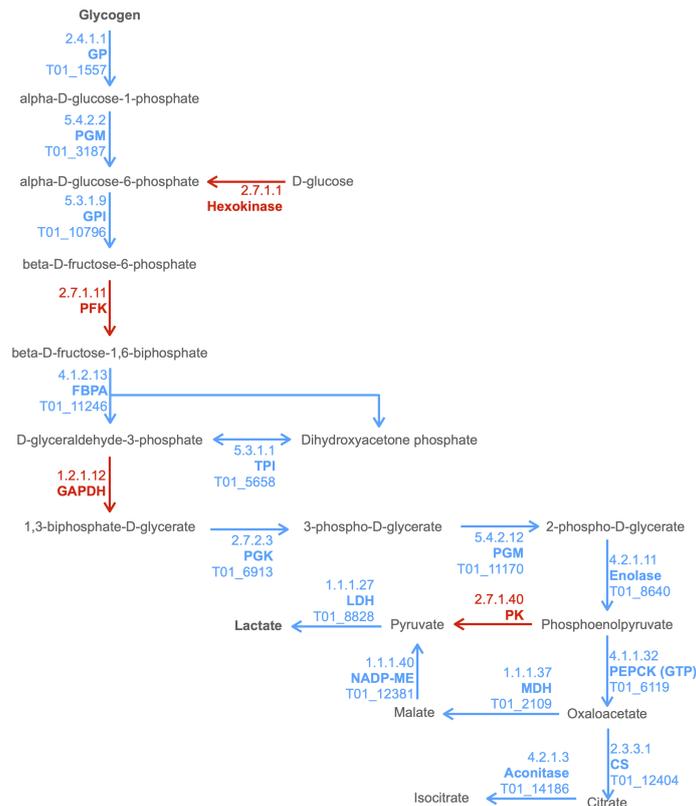


FIGURE 2

Combined excerpt of KEGG Pathways to illustrate ESPs that may be involved in catabolism of polysaccharide storages to lactate and TCA intermediates. The diagram illustrates the presence (blue with *T. spiralis* Gene_Stable_ID) and absence (red) of proteins (with KEGG Enzyme ID and Name) in the *T. spiralis* excretory/secretory proteins (ESPs). Metabolites are displayed in black, with arrows indicating anticipated flux direction. Proteins pertinent to the key stages Polysaccharide metabolism, Glycolysis, Malate dismutation and the TCA cycle have been indicated by differential background colours.

pathway. These strict prerequisites resulted in less than half of the total ESPs being assigned a KEGG term by the KAAS. The NCBI Conserved Domain Search (NCDS) was used to expand the breadth of functional annotation to proteins containing domain signatures without requiring associated KEGG pathway involvement. The NCDS uses the Reverse-Position-Specific iterative Basic Local Alignment Search Tool (RPS-BLAST) approach to query a protein sequence against a collated database (Conserved Domain Database – v3.19-58235 Position-Specific Scoring Matrices) of domains. The NCDS identifies domains which appear specific to sections of an input sequence, rather than the homology of a protein as a whole (Marchler-Bauer et al., 2009; Marchler-Bauer et al., 2011; Lu et al., 2020). The application of the NCDS expanded the functional annotation of 190 (48%) ESPs to 314 (79%). All KEGG annotated ESPs also received domain annotation from the NCDS.

Multiple proteins were found to contain domains from the same superfamily, therefore the frequency of each domain superfamily was counted once per ESP presence and enrichment was scored to identify the ESP domains present at a greater frequency than expected from the proteome (Table 2). In concordance to the finding that carbon metabolism is a broadly enriched KEGG pathway, the AmyAc_family superfamily is identified as a highly enriched conserved domain. The AmyAc_family proteins are glycoside hydrolases whose substrates include the polysaccharides starch and glycogen which could act as an input for glycolysis.

The proteolytic Tryp_SpC superfamily was the most prominent domain superfamily in ESPs. This family also contains the ESP with the 5th highest average normalized spectral abundance factor (NSAF), serine protease 30 (T01_15237). The activity of proteases within *T. spiralis* L1 larvae ESPs has been long established (Criado-Fornelio et al., 1992), with the proteases collected from MSL ESPs demonstrating different substrate specificity to the proteases collected from adult worm ESPs. Indicative of the collagen rich nurse cell complex, MSL ESPs more efficiently degrade gelatin and collagen, while adult ESP proteases are able to digest fibrinogen, hemoglobin and IgG (Ros-Moreno et al., 2000). The Tryp_SpC superfamily is the largest *T. spiralis* protease superfamily and is reported by the International Helminths Genome Consortium to have undergone a large family expansion, specifically within the parasites of Clade I (of which *Trichinella* is a member) (International Helminth Genomes Consortium, 2019). We confirm the presence of reported MSL ESPs, including TS-15-1 (T01_13727) (Park et al., 2018), TspSP-1 (T01_15237) (Romaris et al., 2002), TsSP (T01_10159) (Sun et al., 2018), TsEla (T01_2909) (Hu et al., 2020) and TsChy (T01_6497) (Xu et al., 2020). In a recent study, the Tryp_SpC domain from TS-15-1/T01_13727, found as an ESP, was able to elevate local collagen-I levels in mice after a 14-week treatment (Park et al., 2018). The study uncovered a mechanism dependent on protease activated cell surface receptor F2R-like trypsin receptor 1 (F2RL1) by which an ESP protease could alter host gene expression in construction of the

TABLE 2 Top 10 most enriched domain superfamilies in the excretory/secretory proteins (ESPs).

Superfamily		Frequency		Hypergeometric Enrichment
ID	Name	Proteome	ESP	P-value
cl21584	Tryp SPc superfamily	59	21	1.6×10^{-15}
cl00388	Thioredoxin like superfamily	51	15	1.5×10^{-9}
cl00447	Nudix Hydrolase superfamily	24	9	1.7×10^{-6}
cl03084	TTR-52 superfamily	13	7	2.8×10^{-6}
cl11960	Ig superfamily	55	12	3.4×10^{-6}
cl00133	CAP superfamily	14	7	3.7×10^{-6}
cl38930	AmyAc family superfamily	15	7	5.8×10^{-6}
cl36508	PTZ00102 superfamily	11	6	1.4×10^{-5}
cl21504	EGF CA superfamily	67	12	1.9×10^{-5}
cl36828	ER PDI fam superfamily	7	5	2×10^{-5}

The occurrence of a domain associated with a superfamily term was detected by NCBI CDS and each occurrence of a superfamily domain was counted once per protein. The total incidence of proteins containing the domain within the excretory/secretory proteins has been summated and reported as Frequency. Superfamily names with their associated cluster ID are listed in order of P value, calculated using a hypergeometric test for enrichment (Phyper) to quantify the likelihood of the superfamily being enriched within the ESPs relative to their expected frequency based on the abundance of the superfamily within the proteome.

nurse cell complex (Park et al., 2016). One could imagine that the collagen capsule surrounding the nurse cell is likely subject to frequent modification necessitating the function of enzymes involved in both its turnover and its incorporation. We noticed a striking level of diversity in the proteases present outside of the Tryp_SPc superfamily. We identified cysteine, serine and metalloproteases. We also identified an aspartate protease Ts-Asp/T01_3960 that has been detected in all life-stages except the NBL (Park et al., 2012). When not inhibited by pepstatin A, the protease preferentially interacts with the surface of intestinal epithelial cells rather than muscle cells (Xu et al., 2021). To obtain MSL ESPs, the proteins were obtained by briefly culturing the MSL following release from the nurse cell by acidified pepsin. As a consequence, the ESPs collected during this time may constitute a mixture of proteins produced to function in the nurse cell and proteins produced in preparation for the next life-stage of intestinal colonization. This finding would explain the presence of both proteases that have functions demonstrated to be relevant to both nurse cell and intestinal infection stages.

ESPs are enriched in NUDT9-like nudix hydrolases

The nudix enzyme family encompasses over 80,000 proteins that hydrolyze a wide variety of substrates containing a nucleoside diphosphate linked to some other moiety, X (Srouji et al., 2017). They contain a characteristic consensus motif, generally G-x(5)-E-x(5)-[UA]-x-R-E-x(2)-E-E-x-G- (where U is a bulky hydrophobic residue). Members of the nudix family commonly have a wide range of physiological substrates ranging from (d)NTPs, oxidized (d)NTPs, non- nucleoside polyphosphates, and capped mRNAs, however the NUDT9 subfamily has only one known substrate, ADP-ribose (ADPR), which it converts to adenosine 5'-

monophosphate (AMP) and ribose 5'-phosphate (Carreras-Puigvert et al., 2017). Thus, the NUDT9 family is also referred to as ADP-ribose hydrolases.

We identified a large extended family of nine NUDT9-type nudix hydrolases in *T. spiralis* muscle stage ESPs and a further four by searching the *T. spiralis* predicted proteome. In humans NUDT9 homologous sequences are found only in two proteins, the NUDT9 protein itself and in a highly homologous domain forming part of the calcium-permeable melastatin-related transient receptor potential cation channel (TRPM2 NUDT9H). *T. spiralis* nudix from ESP ranged in length from 279-517 amino acids and represent a set of evolutionarily diverging proteins compared to the canonical mammalian NUDT9, with extensions to both N and C termini containing predicted transmembrane domains which may determine protein localization. Although the diversity in both sequence and length across the family of NUDT9 proteins was extensive, a phylogenetic tree could be constructed using an alignment of NUDT9 between the homology blocks (Figure 3A).

Figure 3B shows an alignment of the *T. spiralis* nudix proteins with human NUDT9 and TRPM2 NUDT9H. Strong identity between these proteins and their mammalian counterparts is confined to two homology blocks (boxed in blue and red) either side of the nudix motif and thought to be involved in substrate binding (Perraud et al., 2003). A striking feature of the *T. spiralis* proteins is that all but one (A0A0V1BTP9) lack residues conserved within the consensus motif known to be essential for catalytic activity (Shen et al., 2003). It is therefore likely that the *T. spiralis* proteins are catalytically inactive but still able to bind ADPR.

Human TRPM2 NUDT9H itself diverges from the nudix consensus motif and is unable to hydrolyze ADPR (Iordanov et al., 2016), however it retains the ability to bind ADPR, gating the channel which leads to Ca²⁺ influx into the cell (Iordanov et al., 2016). The TRPM2 channel is activated under conditions of cellular stress as ADPR levels rise as a result of ADPR polymerization and degradation

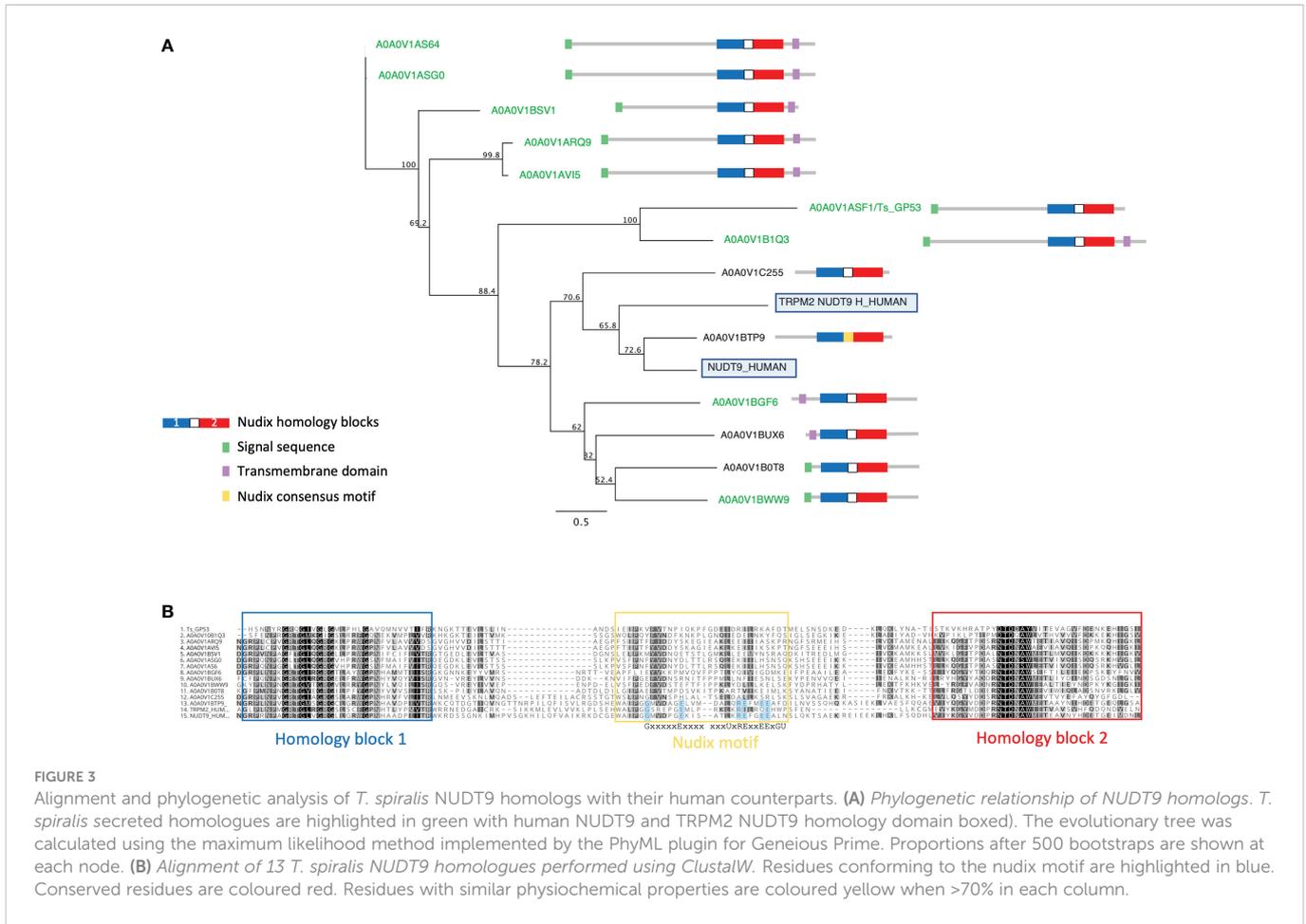


FIGURE 3 Alignment and phylogenetic analysis of *T. spiralis* NUDT9 homologs with their human counterparts. **(A)** Phylogenetic relationship of NUDT9 homologs. *T. spiralis* secreted homologues are highlighted in green with human NUDT9 and TRPM2 NUDT9 homology domain boxed). The evolutionary tree was calculated using the maximum likelihood method implemented by the PhyML plugin for Geneious Prime. Proportions after 500 bootstraps are shown at each node. **(B)** Alignment of 13 *T. spiralis* NUDT9 homologues performed using ClustalW. Residues conforming to the nudix motif are highlighted in blue. Conserved residues are coloured red. Residues with similar physicochemical properties are coloured yellow when >70% in each column.

(Yu et al., 2017). This leads to the release of inflammatory cytokines (Yamamoto et al., 2008; Knowles et al., 2013) and cell death if uncontrolled. This raises the possibility that *T. spiralis* nudix proteins bind free ADPR reducing or preventing Ca²⁺ influx via TRPM2.

It is also possible that *T. spiralis* nudix modulation of available ADPR may have broader effects within the cell such as a reduction in substrate availability for poly-ADPR polymerases (PARP). PARP conjugates ADPR to proteins playing important roles in gene transcription, chromatin organization and stress response (Qi et al., 2019). PARP-deficient mice are resistant to acute sepsis (Soriano et al., 2002) and PARP inhibition prevents addition of mono and poly ADPR to NFκB, an important transcription factor regulating proinflammatory genes (Andreone et al., 2003). PARP also regulates other transcription factors such as AP-1, involved in the transcriptional response to oxidative stress (Bohio et al., 2019). Additionally, PARP has been shown to be involved in the promotion of the M1 macrophage phenotype associated with pro-inflammatory cytokines along with inhibition of M2 macrophages. (Sobczak et al., 2020).

These effects have several features in common with those reported for one of the most abundant secreted *T. spiralis* NUDT9 proteins A0A0V1ASF1 (also known as GP53/Ts53). It has been shown to suppress inflammatory responses in mouse models of sepsis (Yang et al., 2016) and colitis (Du et al., 2011) and prevent death in mice from systemic administration of LPS (Chen et al., 2016).

Administration of recombinant GP53 results in the activation of M2 macrophages, perhaps via the modulation of PARP activity by GP53. It is therefore possible that *T. spiralis* nudix proteins regulate PARP activity by modulating ADPR availability interfering with the addition of ADPR to proteins and the synthesis of poly-ADPR.

ESPs contain components of the extracellular matrix and enzymes associated with ECM assembly

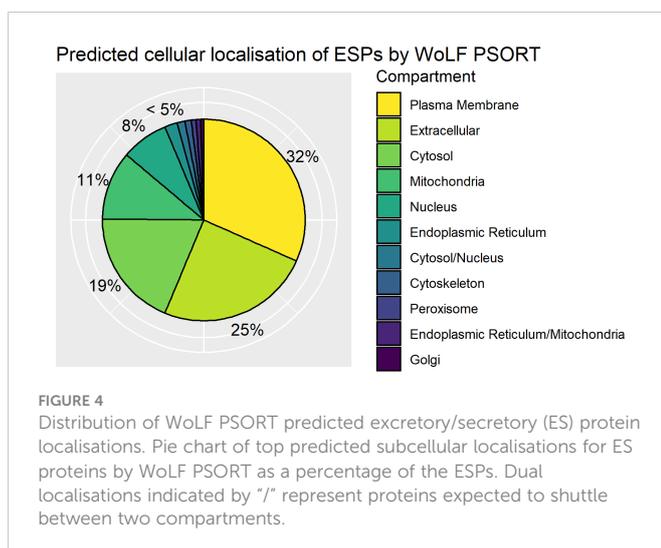
After invasion of a muscle cell *T. spiralis* manipulates its environment outside the cell culminating in the formation of the collagenous capsule of the nurse cell which persists for the life-span of the host. Amongst the ESP proteins we find evidence that this manipulation continues after the nurse cell is formed. Of note is the secretion of two members of the Quiescin sulphhydryl oxidase (QSOX) family, proteins that catalyze disulphide bond formation. They are secreted into the extracellular environment by fibroblasts and are essential for the incorporation of laminin into the extracellular matrix (Ilani et al., 2013). Importantly, we also see secretion of laminin by *T. spiralis*.

In addition, we see *T. spiralis* secreting a wide range of proteins that make up, or are associated with the ECM including six collagen family members, teneurin, and a homologue of the *C. elegans* protein, MUP-4. Teneurin is involved in the assembly of basement membrane

collagen in *C.elegans*, interacts with dystroglycan, integrin and laminin and, when disrupted, leads to detachment of epidermis from muscle cells (Topf and Chiquet-Ehrismann, 2011). MUP-4 is essential for correct embryonic epithelial morphogenesis and maintenance of muscle position. It is also required for attachment of the apical epithelial surface and the cuticular matrix (Hong et al., 2001). Another homologue of a cytoskeletal component, vinculin, is found within cell-cell and cell-ECM junctions and released by *T. spiralis*. We therefore hypothesize that ESP could contribute to the maintenance of the nurse cell within the muscle tissue niche during its long-term and stable infection.

Abundance of ESPs predicted to localize to the host-parasite interface

ESPs are released directly into the host cell and have been demonstrated to localize to different organelles of the nurse cell (Lee et al., 1991). To complement functional annotation of the ESPs we also used WoLF PSORT to predict subcellular localization of ESPs within the infected host cell based on inherent protein sequence properties. WoLF PSORT assigns localization of a protein by similarity to the features of the nearest neighbor class with a known subcellular localization (Horton et al., 2007). WoLF PSORT predicted plasma membrane (32%) and extracellular (25%) localization based on protein features for greater than half the identified ESPs. The analysis indicates that a significant proportion of the ESP is expected to be trafficked from *T. spiralis* through the nurse cell, to presumably mediate interaction with the host and the nurse cell complex (Figure 4). The importance of the ESPs that function at the nurse cell complex interface is supported by enrichment of classically extracellularly functioning conserved domains. We identify enrichment in the ESPs for the Ig superfamily and EGF_CA superfamilies that typically occur on the cell surface. The finding supports that despite the parasite's intracellular confinement, many ESPs may be trafficked to the plasma membrane and contribute to the extracellular environment.



Only a minority of ESPs contain classical secretion signal peptides

Since our identified ESPs are secreted, one would expect them to contain signal peptides. Surprisingly, we found that only 36% (143 proteins) contained a recognizable N-terminal secretion signal using SignalP6.0 (Teufel et al., 2022). The low incidence of secretion signals hints at the complexity of the ESPs collected from mMSL culture supernatant and suggests they have classical secretion dependent and independent origins. Our results may under-represent the number of ESPs containing secretion signals, since their recognition is entirely dependent on the N-terminal sequence of the protein. Many of the proteomes' gene models were constructed from transcriptomics data and protein sequences have not been confirmed experimentally. Therefore, up or down-stream mis-predictions of translation start sites may hinder the detection of secretion signals.

There is a hypergeometric enrichment (P value = 1.4×10^{-83}) of canonical secretion signal containing proteins in the ESPs (143 proteins) compared to the total gene pool (743 proteins). Furthermore, we found that the majority of ESPs anticipated to localize extracellularly also contain a canonical secretion signal, again detected by SignalP6.0 (Supplementary Table 1). However, 25% of predicted extracellular proteins and 66% of predicted plasma membrane localizing proteins did not. Despite all the ESPs being collected externally from the nematode, the variable requirement for a secretion signal is still poorly understood. In *T. spiralis*, the stichosome is a dedicated secretory organ composed of α and β -stichocytes that contain respective α and β -granules. ES antigens are detected within the granules of the stichosome and predominantly exit from the mouth of the organism (Despommier and Müller, 1976). However, non-canonically secreted proteins may also be transported to the host cell as components of extracellular vesicles (EVs) that have been widely shown to have immunomodulatory effects such as upregulating anti-inflammatory cytokines (IL-10, IL-4, IL-13 and TGF- β) and Th2-associated transcription factors Foxp3 and GATA3 (Yang et al., 2020). Given their immunomodulatory effects, MSL EVs must be capable of transporting proteins from the worm to the cell surface where they can interface with the immune system. A recent study that proteomically profiled EVs derived from MSL-stage *T. spiralis* identified 753 proteins (Gao et al., 2022). Of these, 175 were also present in our dataset (Supplementary Table 1). The overlap is perhaps not surprising given the fact that our ESP preparation would be expected to capture EVs. In the interest of understanding *Trichinella* biology, future studies to determine the subset of ESPs that are derived from the stichocytes and the subset of ESPs that originate from EVs of the cuticle would be greatly beneficial.

Identification of a novel ESP-associated motif

The development of the dedicated secretory organ of *T. spiralis* does not progress until after the NBL establishes itself in the host muscle cell suggestive of an important role in nurse cell development. Since a large proportion of the ESPs are expected to originate from the stichosome, we hypothesized that the ESPs could be used to identify a

stichosome-specific regulatory motif. A regulatory motif enriched upstream of genes identified as ESPs could, in turn, be used to predict stichosome-derived, secreted proteins across the life-cycle. Since ESPs are impossible to collect in sufficient quantities and properly study during the earliest stages of MSL development, such a predictive tool would be valuable in understanding how *T. spiralis* initiates muscle reprogramming.

We searched for a motif enriched in the region of 1000 bp upstream from the Transcription Start Site (TSS) of the 398 identified ESPs. The equivalent 1000 bp region was recovered for each of the 14,258 genes in the PRJNA257433 *Trichinella spiralis* genome (Korhonen et al., 2016) and used as background control sequences. The Homer2 software (Heinz et al., 2010) was used to construct Hidden Markov Models (HMMs) for motifs that appear more frequently in the “regulated” target dataset (ESP) compared to the background of “unregulated” sequences (Reference Proteome). Motif-direction outside of the core promoter region does not significantly influence the transcription regulating effect of a transcription factor on a gene (Lis and Walther, 2016), hence the motif search was performed on both strands.

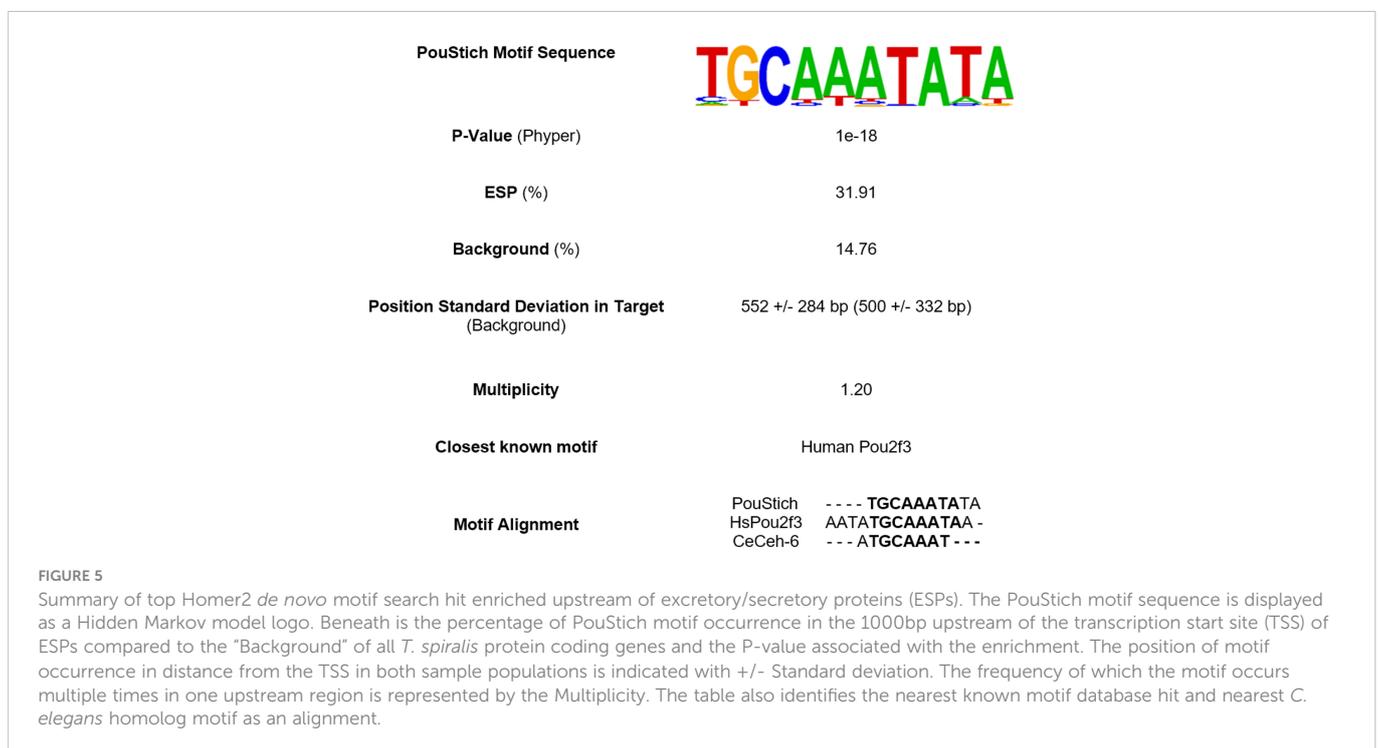
The Homer2 motif search using known transcription factor binding sites, including those from the nematode model organism *C. elegans*, failed to identify any significantly enriched motifs. Given the absence of a stichosome in *C. elegans*, this result was not discouraging and highlights the divergence of regulatory processes in the genetically distinct nematode Clades. To ensure the exclusion of false positive hits, we interrogated ESPs with 2 or more UPSMs. Only 1 *de novo* motif was significantly enriched upstream of the TSS of these ESP genes. This motif has the highest similarity to the known human POU Class 2 Homeobox 3 (HsPou2f3) binding site (MA0627.2/Jaspar(0.761) and is highly enriched in what we hypothesize to be stichosome-derived proteins, so we named it PouStich (Figure 5).

Prediction of PouStich motif-binding proteins

BLASTp was used to identify proteins from *T. spiralis* that are similar in amino acid sequence to HsPou2f3 and therefore, likely to bind the PouStich motif. We identified Pou domain, class 3, transcription factor 2-A protein (TsPou3f2-a) as the closest *T. spiralis* hit to HsPou2f3. However, the human and the *T. spiralis* Pou domain proteins fall into different classes and appear evolutionarily distant. To explore a more similar background, BLASTp was used to interrogate the nematode *C. elegans* and identified the class 3 Pou domain containing transcription factor ceh-6 (CeCeh-6) (Figures 6A, B).

CeCeh-6 is the only class 3 Pou domain-containing protein identified in *C. elegans* and has a demonstrated role in proper excretory cell differentiation. Class 3 Pou domains are conserved across many phyla as an excretory/secretory cell regulator and are also implicated in the nervous system (Bürglin and Ruvkun, 2001). *C. elegans* has 3 Pou domain containing homeobox genes, with CeCeh-6 being the sole POU-III family member. Using a lacZ reporter, CeCeh-6 expression has been localized to discrete neural cell populations and the excretory cells. An effective null mutant for CeCeh-6 caused embryonic and larval lethality with larvae developing bodily vacuoles. The phenotype could be mimicked by laser ablation of the excretory cells, concluding that larval lethality was secondary to impaired excretory cell function (Bürglin and Ruvkun, 2001). The involvement of CeCeh-6 in excretory cell function in *C. elegans* is encouraging, as it supports the idea that the PouStich motif may mediate ESPs production.

Pou domain-containing proteins typically interact with DNA sequences through the combination of the Pou domain and the Homeobox (Hox) domain (Figure 6A). The TsPou3f2-a sequence has 81% identity with CeCeh-6 across the Pou domain (red) but only



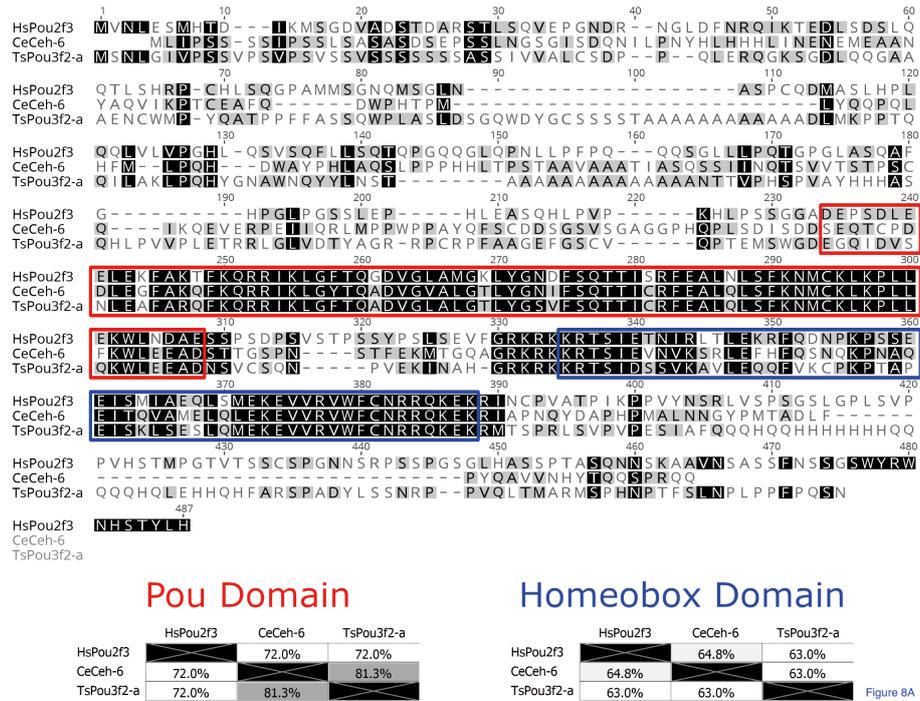


Figure 8A

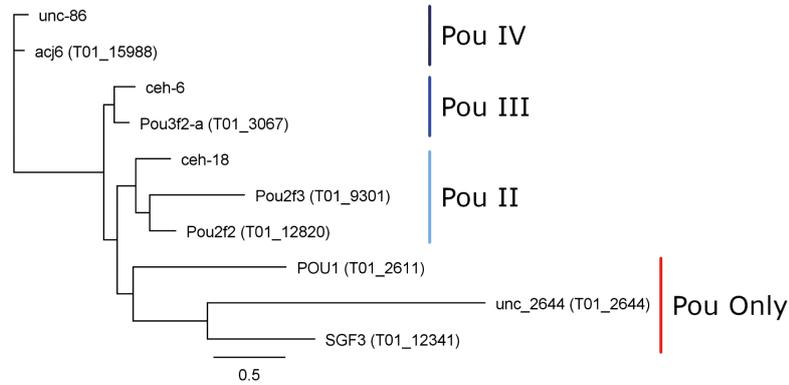


FIGURE 6

Pou domain containing octamer motif binding proteins. (A) MUSCLE alignment of the Human Pou2f3 with the closest (C) elegans and T. spiralis BLASTp hits. Amino acid shading highlights similarity. In the red box is the Pou domain and in the blue box is the homeobox domain, the two essential conserved regions of Pou-domain containing octamer motif binding proteins. (B) Phylogenetic tree from alignment of the T. spiralis and (C) elegans Pou domains. The tree was constructed with Phylml from aligned Pou domains across all sequences shown above using the Whelan and Goldman substitution model. The Pou domain type is annotated adjacent to the tree groupings. T. spiralis genes are included in brackets next to the gene name, while only (C) elegans gene names are provided.

72% identity with HsPou2f3. Across the Hox domain (blue) TsPou3f2-a shares 63% identity with both HsPou2f3 and CeCeh-6. Taken together the results indicate that the Pou domain of TsPou3f2-a would bind a sequence more similar to the CeCeh-6 Pou interaction motif than that of HsPou2f3.

Using the NCDS, we identified a further 6 Pou domain-containing proteins from the Trichinella spiralis reference proteome and found that each C. elegans Pou domain containing protein has an identifiable homolog in Trichinella spiralis. This could suggest that nematodes require a minimum of three Pou domain-containing proteins of Class IV, Class III and Class II. In Trichinella spiralis however, a presumed duplication has resulted in the gain of a second

Class II Pou domain containing protein. One Class II protein may function as a required nematode factor similarly to C. elegans CeCeh-18, while the other Class II protein might regulate the stichosome, an organ absent in C. elegans.

Using the PouStich motif to predict novel ESPs

Hypothesizing that the PouStich motif is predictive of stichosome origin, we used Homer2 to search for the PouStich motif in the region upstream of all reference proteome genes. We identified 2,157 genes

containing the PouStich motif within 1000bp from the TSS as a direct match or with 1 tolerable mismatch (Motif Score = 9.216498), including 127 ESP genes from this study (Supplementary Table 2). From the identified ESPs the protein with the highest average NSAF, *mcd-1* contained an upstream PouStich motif as did secreted from muscle stage larvae 3, *sml-3* (T01_15139). The motif was also found upstream of proteins that were anticipated as ESPs but unconfirmed by peptides, such as *sml-5* (T01_15780) and the glycolytic enzyme glyceraldehyde 3-phosphate dehydrogenase, *gpd-1* (T01_7457). The proteins *mcd-1*, *sml-3*, and *sml-5* have been demonstrated to localize to the MSL stichosome (Guiliano et al., 2009; Srey et al., 2020).

Not all genes expressed within the stichosome would be secreted as ESPs, some would likely be required for maintenance and differentiation of the secretory organ itself. It is likely that additional factors are required to identify the true ESPs amongst the motif containing genes. Of the 2,157 PouStich motif-regulated genes we detected co-occurrence of a Secretion Signal in only 175, 61 of these were also identified in our proteomic screen (Figure 7 and Supplementary Table 1). It remains to be seen whether the other 114 proteins are secreted, supporting our hypothesis that the PouStich motif is an important regulator of the secretory proteins of *T. spiralis*.

Conclusions and perspectives

Recent advancements in the construction of the *T. spiralis* proteome has enabled more extensive analysis of the secretome across the lifecycle. Combining these advances with modern protein annotation software allowed us to gain insight into the functions of the ESPs involved in this unique host-pathogen interface. We observed that the mMSL secretome is enriched for proteins predicted to affect metabolism. Furthermore, we identify superfamilies within the ESPs that predict multiple mechanisms

through which the MSL can regulate the host immune system and remodel the extracellular environment during infection. We use the genomic context of ESPs identified in this study to highlight a novel motif, the PouStich motif, enriched upstream of their TSS. Though more experimental validation is required, the PouStich motif represents an exciting possibility that a portion of the ESPs are regulated by a common motif. Therefore, the PouStich motif in combination with other ESP associated patterns could allow the identification of ESPs from other currently inaccessible developmental stages.

Materials and methods

ESP preparation and LC/MS/MS analysis

ESP were collected and processed as described in White et al. (2016) (White et al., 2016). Briefly, in two separate biological replicates, mature L1 larvae were released from rat infected muscle by acidified pepsin digestion and maintained in a culture flask in serum-free medium. The ESPs secreted into the culture medium were collected for 24 hours, concentrated using a 10 kDa protein filter, digested into peptides using trypsin, enriched and subjected to LC/MS/MS analysis. To avoid contamination with proteins released from dead or dying parasites, uncoiled and floating *T. spiralis* larvae were removed prior to culturing. In addition, to ensure a low level of parasite death, larvae were cultured for 24 h only, before supernatant was collected for SP purification. Here, for replicate 1, an LTQ Orbitrap XL2 (Thermo) equipped with an Agilent 1200 Bin Pump and a Famos microautosampler (LC Packings) was used. Peptides were separated with a gradient of 6 to 35% CH₃CN in 0.125% formic acid over 90 min. MS data was acquired between 300 and 1500 m/z at 60,000 fwhm with lockmass enabled (371.101233 m/z). CID spectra were acquired in the LTQ with MSMS switching operating in a top 10 DDA fashion. For replicate 2, an Orbitrap QExactive (Thermo) equipped with an Accela 600 quaternary pump and a Famos microautosampler (LC Packings) was used. Peptides were separated with a gradient of 6 to 25% CH₃CN in 0.125% formic acid over 120 min. MS data was acquired between 300 and 1500 m/z at 70,000 fwhm. CID spectra were acquired in the LTQ with MSMS switching operating in a top 20 DDA fashion.

Mass spectra were processed using a Sequest-based software pipeline for quantitative proteomics developed by Professor Steven Gygi's laboratory at Harvard Medical School. After converting RAW files to mxml format, MS data were searched against a database of protein sequences. A combined database was constructed from (a) pig trypsin, human and the *Rattus Norvegicus* UniProt proteome (download: 26/04/2022), (b) the UniProt *Trichinella spiralis* 'one protein per gene' reference proteome (14,258 genes, UP000054776); (c) common contaminants such as porcine trypsin and endoproteinase LysC. The combined database was concatenated with a reverse database composed of all protein sequences in reversed order. Sequest searches were performed using a 20 ppm precursor ion tolerance. Product ion tolerance was set to 1 Da. Oxidation of methionine residues (15.99492 Da) was set as a variable modification.

To control the fraction of erroneous protein identifications, a target-decoy strategy was employed (Huttlin et al., 2010). Peptide

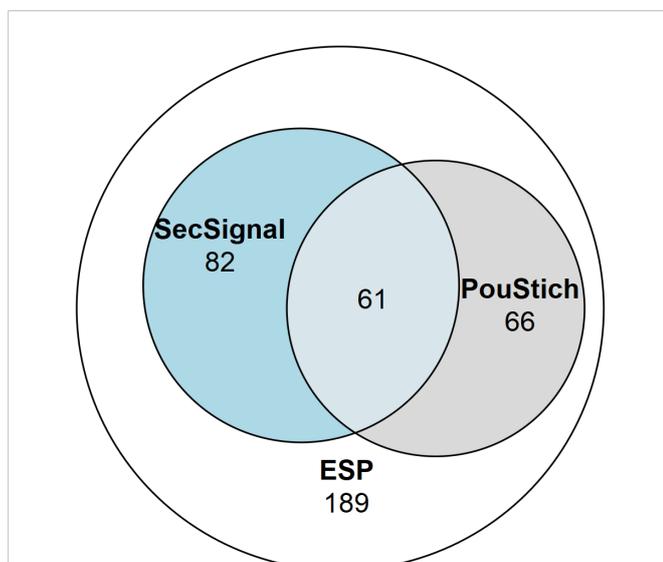


FIGURE 7

Venn diagram showing the proportion of the ESPs annotated with motifs. 398 ESPs were identified in this study, of which 143 contained a SignalP6.0 detected Secretion Signal, and 127 ESPs contained a PouStich motif upstream of the transcription start site. The 61 ESPs that contain a Secretion Signal and PouStich motif are a subset of high confidence predicted ESPs.

spectral matches (PSMs) were filtered to a peptide-level false discovery rate (FDR) of 2% with subsequent filtering to attain a final protein-level FDR of 0.9%. PSM filtering was performed using a linear discriminant analysis, as described previously (Huttlin et al., 2010). This distinguishes correct from incorrect peptide IDs in a manner analogous to the widely used Percolator algorithm (<https://noble.gs.washington.edu/proj/percolator/>), though employing a distinct machine learning algorithm. The following parameters were considered: XCorr, ΔC_n , missed cleavages, peptide length, charge state, and precursor mass accuracy. Peptides shorter than seven amino acids in length or with XCorr less than 1.0 were excluded prior to LDA filtering. Two missed cleavages were allowed. Protein assembly was guided by principles of parsimony to produce the smallest set of proteins necessary to account for all observed peptides [algorithm described previously (Huttlin et al., 2010)].

The mass spectrometry raw files and associated peptide and protein identification data have been deposited in the ProteomeXchange Consortium *via* the iProX partner repository with identifier IPX0005645000.

Bioinformatic analysis

Each identified ESP was queried against the Conserved Domain Database (v3.19 -58235 Position-Specific Scoring Matrices) using the NCBI Conserved Domain Search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)? with a significance threshold E-value of $< 1 \times 10^{-2}$ (Lu et al., 2020). Gene ontology analysis was performed with gProfiler2 using the organism *trspirprjna257433* (Kolberg et al., 2020). Kyoto Encyclopedia of Genes and Genomes (KEGG) ontology (KO) terms were assigned to proteins using the KEGG Automatic Annotation Server (KAAS) (https://www.genome.jp/kaas-bin/kaas_main) using the Bi-directional Best Hit Assignment Method from Blast results against the default eukaryotic representative gene set with the additional inclusion of *T. spiralis* as a reference organism (Moriya et al., 2007). Canonical Eukaryotic secretion signals were detected at the N-termini of peptides using the SignalP-6.0 server (<https://services.healthtech.dtu.dk/service.php?SignalP>) (Teufel et al., 2022). WoLF PSORT (<https://wolfsort.hgc.jp/>) cellular localizations were determined by reporting the highest scoring localization for each protein using the Animal organism type (Horton et al., 2007).

Statistical analysis

The R function `phyper` was used to perform hypergeometric testing for enrichment. The P values were calculated using the script as follows: `phyper(q - 1, m, n - m, k, lower.tail = FALSE)` where `q` and `m` are constant and equal to `No.` Hits in the Subject and the Background respectively and while `n` and `k` are the total members of the Background and the Subject respectively. The arguments `q - 1` and `lower.tail = FALSE` tests the hypothesis that enrichment is greater than or equal to the test population. The p-value is adjusted using the Benjamini-Hochberg method to account for the FDR associated with multiple comparisons testing. Therefore, the P-value represents the probability that the hits within the subject population are greater than

expected based on the hits in the background population as a result of chance, whereby $P \leq 0.05$ suggests enrichment.

Sequence analysis

Sequence analysis was performed using Geneious 11.1.5. Sequences were aligned using Muscle 3.8.425. A phylogenetic tree was produced from aligned sequences using the PhyML 3.3.20180621 plugin with the Whelan and Goldman substitution model (Guindon et al., 2010).

Identification of enriched motifs

The 1000bp region upstream of the Transcription Start Site (TSS) was retrieved for each Reference proteome protein using BioMart (<https://parasite.wormbase.org/biomart/martview/>) hosted on WormBase Parasite (Korhonen et al., 2016; Howe et al., 2017). For Homer2 (<http://homer.ucsd.edu/homer>) (Heinz et al., 2010) input sequences, all the recovered sequences were used as the background, while the subject was the subset of sequences from identified ESP genes. The motif search was performed at a higher stringency level of proteins with 2 or more unique peptide spectrum matches (UPSs) to minimize the inclusion of false positive hits. The `findMotifs.pl` function was used to search for significantly enriched (P value $< 1e-11$) *de novo* motifs across all stringency levels, using `-fasta` to specify the custom background file and using `fasta` as the placeholder organism. Motif occurrence in the genome was determined using the `-find` option with the output `.motif` file, to search the background sequences for the significantly enriched PouStich motif.

Data availability statement

The data presented in the study are deposited in the ProteomeXchange Consortium *via* the iProX partner repository with the accession number IPX0005645000.

Ethics statement

This study used material derived from Sprague-Dawley rats obtained from Harlan LTD UK. The Animal Welfare Ethical Review Board at Imperial College London did not require the study to be reviewed or approved by an ethics committee because these materials were a gift from experiments licensed by and performed under the UK Home Office Animals (Scientific Procedures) Act Personal Project Licence number PFA8EC7B7: 'Immunity and immunomodulation in helminth infection.

Author contributions

KA-T and MS conceived of and designed the study. RW collected all experimental material and MW and SG performed

the proteomic analyses. BN, WG, RW and AP analyzed and organized the data. KA-T, BN and WG wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by a Wellcome Trust Career Development Award [WT085054MA], a BBSRC Project Grant [BB/R001642/1] both held by KAT and an MRC Studentship awarded to RW. BN was supported by a Department of Pathology PhD Studentship and Rosetrees Trust PhD Consumables [M703] and PhD Plus Grants [PhD2021/100037]. MW is supported by Medical Research Council Project Grant MR/W025647/1 and SG is supported by an NIH RO1 grant [GM067945]. For the purpose of open access the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

References

- Andreone, T. L., O'Connor, M., Denenberg, A., Hake, P. W., and Zingarelli, B. (2003). Poly(ADP-ribose) polymerase-1 regulates activation of activator protein-1 in murine fibroblasts. *J. Immunol.* 170, 2113–2120. doi: 10.4049/jimmunol.170.4.2113
- Banankhah, P., Fishbein, G. A., Dota, A., and Ardehali, R. (2018). Cardiac manifestations of PRKAG2 mutation. *BMC Med. Genet.* 19. doi: 10.1186/s12881-017-0512-6
- Bien, J., Cabaj, W., and Moskwa, B. (2015). Proteomic analysis of potential immunoreactive proteins from muscle larvae and adult worms of trichinella spiralis in experimentally infected pigs. *Folia Parasitol.* 62, 022. doi: 10.14411/fp.2015.022
- Bohio, A. A., Sattout, A., Wang, R., Wang, K., Sah, R. K., Guo, X., et al. (2019). c-Abl-mediated tyrosine phosphorylation of PARP1 is crucial for expression of proinflammatory genes. *J. Immunol.* 203, 1521–1531. doi: 10.4049/jimmunol.1801616
- Bürglin, T. R., and Ruvkun, G. (2001). Regulation of ectodermal and excretory function by the c. elegans POU homeobox gene ceh-6. *Development* 128, 779–790. doi: 10.1242/dev.128.5.779
- Carreras-Puigvert, J., Zitnik, M., Jemth, A.-S., Carter, M., Unterlass, J. E., Hallström, B., et al. (2017). A comprehensive structural, biochemical and biological profiling of the human NUDIX hydrolase family. *Nat. Commun.* 8, 1541. doi: 10.1038/s41467-017-01642-w
- Chen, Z.-B., Tang, H., Liang, Y.-B., Yang, W., Wu, J.-G., Hu, X.-C., et al. (2016). Recombinant trichinella spiralis 53-kDa protein activates M2 macrophages and attenuates the LPS-induced damage of endotoxemia. *Innate Immun.* 22, 419–432. doi: 10.1177/1753425916651984
- Criado-Fornelio, A., de Armas-Serra, C., Giménez-Pardo, C., Casado-Escribano, N., Jiménez-González, A., and Rodríguez-Cabeiro, F. (1992). Proteolytic enzymes from trichinella spiralis larvae. *Vet. Parasitol.* 45, 133–140. doi: 10.1016/0304-4017(92)90035-8
- Despommier, D. (1975). Adaptive changes in muscle fibers infected with trichinella spiralis. *Am. J. Pathol.* 78, 477–496.
- Despommier, D. D. (1998). How does trichinella spiralis make itself at home? *Parasitol. Today* 14, 318–323. doi: 10.1016/s0169-4758(98)01287-3
- Despommier, D. D., and Müller, M. (1976). The stichosome and its secretion granules in the mature muscle larva of trichinella spiralis. *J. Parasitol.* 62, 775–785. doi: 10.2307/3278960
- Du, L., Tang, H., Ma, Z., Xu, J., Gao, W., Chen, J., et al. (2011). The protective effect of the recombinant 53-kDa protein of trichinella spiralis on experimental colitis in mice. *Dig. Dis. Sci.* 56, 2810–2817. doi: 10.1007/s10620-011-1689-8
- Gao, X., Yang, Y., Liu, X., Xu, F., Wang, Y., Liu, L., et al. (2022). Extracellular vesicles from trichinella spiralis: Proteomic analysis and protective immunity. *PLoS Negl. Trop. Dis.* 16, e0010528. doi: 10.1371/journal.pntd.0010528
- Gómez-Arreaza, A., Acosta, H., Quiñones, W., Concepción, J. L., Michels, P. A. M., and Avilán, L. (2014). Extracellular functions of glycolytic enzymes of parasites: Unpredicted

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpara.2023.1078443/full#supplementary-material>

use of ancient proteins. *Mol. Biochem. Parasitol.* 193, 75–81. doi: 10.1016/j.molbiopara.2014.02.005

Guiliano, D. B., Oksov, Y., Lustigman, S., and Gounaris, K. (2009). And Selkirk, m Characterisation of novel protein families secreted by muscle stage larvae of trichinella spiralis. *E. Int. J. Parasitol.* 39, 515–524. doi: 10.1016/j.ijpara.2008.09.012

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., et al. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and b cell identities. *Mol. Cell* 38, 576–589. doi: 10.1016/j.molcel.2010.05.004

Hong, L., Elbl, T., Ward, J., Franzini-Armstrong, C., Rybicka, K. K., Gatewood, B. K., et al. (2001). MUP-4 is a novel transmembrane protein with functions in epithelial cell adhesion in caenorhabditis elegans. *J. Cell Biol.* 154, 403–414. doi: 10.1083/jcb.200007075

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* 35, W585–W587. doi: 10.1093/nar/gkm259

Howe, K. L., Bolt, B. J., Shafie, M., Kersey, P., and Berriman, M. (2017). WormBase ParaSite – a comprehensive resource for helminth genomics. *Mol. Biochem. Parasitol.* 215, 2–10. doi: 10.1016/j.molbiopara.2016.11.005

Hu, C. X., Jiang, P., Yue, X., Zeng, J., Zhang, X. Z., Song, Y. Y., et al. (2020). Molecular characterization of a trichinella spiralis elastase-1 and its potential as a diagnostic antigen for trichinellosis. *Parasitol. Vectors* 13, 97. doi: 10.1186/s13071-020-3981-y

Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., et al. (2010). A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* 143, 1174–1189. doi: 10.1016/j.cell.2010.12.001

Ilani, T., Alon, A., Grossman, I., Horowitz, B., Kartvelishvili, E., Cohen, S. R., et al. (2013). A secreted disulfide catalyst controls extracellular matrix composition and function. *Science* 341, 74–76. doi: 10.1126/science.1238279

International Helminth Genomes Consortium (2019). Comparative genomics of the major parasitic worms. *Nat. Genet.* 51, 163–174. doi: 10.1038/s41588-018-0262-1

Jordanov, I., Mihályi, C., Tóth, B., and Csanády, L. (2016). The proposed channel-enzyme transient receptor potential melastatin 2 does not possess ADP ribose hydrolase activity. *Elife* 5, e17600. doi: 10.7554/eLife.17600

Jiang, P., Zao, Y. J., Yan, S. W., Song, Y. Y., Yang, D. M., Dai, L. Y., et al. (2019). Molecular characterization of a trichinella spiralis enolase and its interaction with the host's plasminogen. *Vet. Res.* 50, 106. doi: 10.1186/s13567-019-0727-y

Khan, Z. A. (1966). The postembryonic development of trichinella spiralis with special reference to ecdysis. *J. Parasitol.* 52, 248–259. doi: 10.2307/3276480

- Khananshvil, D. (2013). The SLC8 gene family of sodium-calcium exchangers (NCX) - structure, function, and regulation in health and disease. *Mol. Aspects Med.* 34, 220–235. doi: 10.1016/j.mam.2012.07.003
- Knowles, H., Li, Y., and Perraud, A.-L. (2013). The TRPM2 ion channel, an oxidative stress and metabolic sensor regulating innate immunity and inflammation. *Immunol. Res.* 55, 241–248. doi: 10.1007/s12026-012-8373-8
- Kolberg, R. (2020). Kuzmin, vilo, and Peterson gprofiler2—an r package for gene list functional enrichment analysis and namespace conversion toolset g: Profiler. *F1000Res.* 9, 709. doi: 10.12688/f1000research.24956.2
- Komuniecki, T., and Tielens, A. G. M. (2003). “Chapter 14—carbohydrate and energy metabolism in parasitic helminths,” in *Molecular medical parasitology* (Academic Press), 339–358. doi: 10.1016/B978-012473346-6/50017-X
- Korhonen, P. K., Pozio, E., La Rosa, G., Chang, B. C. H., Koehler, A. V., Hoberg, E. P., et al. (2016). Phylogenomic and biogeographic reconstruction of the trichinella complex. *Nat. Commun.* 7, 10513. doi: 10.1038/ncomms10513
- Kozek, W. J. (1971). The molting pattern in trichinella spiralis. i. a light microscope study. *J. Parasitol.* 57, 1015–1028. doi: 10.2307/3277858
- Lee, D. L., Ko, R. C., Yi, X. Y., and Yeung, M. H. (1991). Trichinella spiralis: Antigenic epitopes from the stichocytes detected in the hypertrophic nuclei and cytoplasm of the parasitized muscle fibre (nurse cell) of the host. *Parasitology* 102 Pt 1, 117–123. doi: 10.1017/S003118200006042X
- Lee, D. L., and Wright, K. A. (1984). And shivers, r A freeze-fracture study of the surface of the infective-stage larva of the nematode trichinella. *R. Tissue Cell* 16, 819–828. doi: 10.1016/0040-8166(84)90012-0
- Lis, M., and Walther, D. (2016). Erratum to: “The orientation of transcription factor binding site motifs in gene promoter regions: Does it matter?” *BMC Genomics* 17, 310 doi: 10.1186/s12864-016-2630-5
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., et al. (2020). CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* 48, D265–D268. doi: 10.1093/nar/gkz991
- Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., et al. (2009). CDD: specific functional annotation with the conserved domain database. *Nucleic Acids Res.* 37, D205–D210. doi: 10.1093/nar/gkn845
- Marchler-Bauer, A., Lu, S., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., et al. (2011). CDD: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Res.* 39, D225–D229. doi: 10.1093/nar/gkq1189
- Meyers, D. E., Basha, H. I., and Koenig, M. K. (2013). Mitochondrial cardiomyopathy: Pathophysiology, diagnosis, and management. *Tex. Heart Inst. J.* 40, 385–394.
- Mitrev, M., Jasmer, D. P., Appleton, J., Martin, J., Dante, M., Wylie, T., et al. (2004). Gene discovery in the adenophorean nematode trichinella spiralis: An analysis of transcription from three life cycle stages. *Mol. Biochem. Parasitol.* 137, 277–291. doi: 10.1016/j.molbiopara.2004.05.015
- Mitrev, M., Jasmer, D. P., Zarlenga, D. S., Wang, Z., Abubucker, S., Martin, J., et al. (2011). The draft genome of the parasitic nematode trichinella spiralis. *Nat. Genet.* 43, 228–235. doi: 10.1038/ng.769
- Montgomery, J. M., Augostini, P., and Stewart, G. L. (2003). Glucose uptake and metabolism in the trichinella spiralis nurse cell. *Int. J. Parasitol.* 33, 401–412. doi: 10.1016/S0020-7519(03)00013-4
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35, W182–W185. doi: 10.1093/nar/gkm321
- Nishina, M., and Suzuki, M. (2002). Biphasic hypothermia in mice infected with a parasitic nematode, trichinella spiralis. *Exp. Anim.* 51, 207–211. doi: 10.1538/expanim.51.207
- Nishina, M., Suzuki, M., and Matsushita, K. (2004). Trichinella spiralis: Activity of the cerebral pyruvate recycling pathway of the host (mouse) in hypoglycemia induced by the infection. *Exp. Parasitol.* 106, 62–65. doi: 10.1016/j.exppara.2004.01.007
- Park, M. K., Cho, M. K., Kang, S. A., Kim, B. Y., and Yu, H. S. (2016). The induction of the collagen capsule synthesis by trichinella spiralis is closely related to protease-activated receptor 2. *Vet. Parasitol.* 230, 56–61. doi: 10.1016/j.vetpar.2016.11.001
- Park, M. K., Kim, H.-J., Cho, M. K., Kang, S. A., Park, S. Y., Jang, S. B., et al. (2018). Identification of a host collagen inducing factor from the excretory secretory proteins of trichinella spiralis. *PLoS Negl. Trop. Dis.* 12, e0006516. doi: 10.1371/journal.pntd.0006516
- Park, J. N., Park, S. K., Cho, M. K., Park, M.-K., Kang, S. A., Kim, D.-H., et al. (2012). Molecular characterization of 45kDa aspartic protease of trichinella spiralis. *Veterinary Parasitol.* 190, 510–518. doi: 10.1016/j.vetpar.2012.06.029
- Perraud, A.-L., Shen, B., Dunn, C. A., Rippe, K., Smith, M. K., Bessman, M. J., et al. (2003). NUDT9, a member of the nudix hydrolase family, is an evolutionarily conserved mitochondrial ADP-ribose pyrophosphatase. *J. Biol. Chem.* 278, 1794–1801. doi: 10.1074/jbc.M205601200
- Qi, H., Price, B. D., and Day, T. A. (2019). Multiple roles for mono- and Poly(ADP-ribose) in regulating stress responses. *Trends Genet.* 35, 159–172. doi: 10.1016/j.tig.2018.12.002
- Robinson, M. W., and Connolly, B. (2005). Proteomic analysis of the excretory-secretory proteins of the trichinella spiralis L1 larva, a nematode parasite of skeletal muscle. *Proteomics* 5, 4525–4532. doi: 10.1002/pmic.200402057
- Robinson, M. W., Gare, D. C., and Connolly, B. (2005). Profiling excretory/secretory proteins of trichinella spiralis muscle larvae by two-dimensional gel electrophoresis and mass spectrometry. *Vet. Parasitol.* 132, 37–41. doi: 10.1016/j.vetpar.2005.05.019
- Romaris, F., North, S. J., Gagliardo, L. F., Butcher, B. A., Ghosh, K., Beiting, D. P., et al. (2002). A putative serine protease among the excretory–secretory glycoproteins of L1 trichinella spiralis. *Mol. Biochem. Parasitol.* 122, 149–160. doi: 10.1016/S0166-6851(02)00094-4
- Ros-Moreno, R. M., Vázquez-López, C., Giménez-Pardo, C., de Armas-Serra, C., and Rodríguez-Cabeiro, F. (2000). A study of proteases throughout the life cycle of trichinella spiralis. *Folia Parasitol.* 47, 49–54. doi: 10.14411/fp.2000.009
- Shen, B. W., Perraud, A. L., Scharenberg, A., and Stoddard, B. L. (2003). The crystal structure and mutational analysis of human NUDT9. *J. Mol. Biol.* 332, 385–398. doi: 10.1016/S0022-2836(03)00954-9
- Sobczak, M., Zyma, M., and Robaszkiewicz, A. (2020). The role of PARP1 in monocyte and macrophage commitment and specification: Future perspectives and limitations for the treatment of monocyte and macrophage relevant diseases with PARP inhibitors. *Cells* 9, 2040. doi: 10.3390/cells902040
- Sola-Penna, M., Da Silva, D., Coelho, W. S., Marinho-Carvalho, M. M., and Zancan, P. (2010). Regulation of mammalian muscle type 6-phosphofructo-1-kinase and its implication for the control of the metabolism. *IUBMB Life* 62, 791–796. doi: 10.1002/iub.393
- Soriano, F. G., Liaudet, L., Szabó, E., Virág, L., Mabley, J. G., Pacher, P., et al. (2002). Resistance to acute septic peritonitis in poly(ADP-ribose) polymerase-1-deficient mice. *Shock* 17, 286–292. doi: 10.1097/00024382-200204000-00008
- Srey, M. T., Taccogna, A., Oksov, Y., Lustigman, S., Tai, P.-Y., Acord, J., et al. (2020). Vaccination with novel low-molecular weight proteins secreted from trichinella spiralis inhibits establishment of infection. *PLoS Negl. Trop. Dis.* 14, e0008842. doi: 10.1371/journal.pntd.0008842
- Srouji, J. R., Xu, A., Park, A., Kirsch, J. F., and Brenner, S. E. (2017). The evolution of function within the nudix homology clan. *Proteins: Structure Function Bioinf.* 85, 775–811. doi: 10.1002/prot.25223
- Sun, G. G., Ren, H. N., Liu, R. D., Song, Y. Y., Qi, X., Hu, C. X., et al. (2018). Molecular characterization of a putative serine protease from trichinella spiralis and its elicited immune protection. *Vet. Res.* 49, 59. doi: 10.1186/s13567-018-0555-5
- Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gislason, M. H., Pihl, S. I., Tsirigos, K. D., et al. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* 40, 1023–1025. doi: 10.1038/s41587-021-01156-3
- Topf, U., and Chiquet-Ehrismann, R. (2011). Genetic interaction between caenorhabditis elegans tenascin ten-1 and prolyl 4-hydroxylase phy-1 and their function in collagen IV-mediated basement membrane integrity during late elongation of the embryo. *Mol. Biol. Cell* 22, 3331–3343. doi: 10.1091/mbc.e10-10-0853
- Ward, C. W., Castro, G. A., and Fairbairn, D. (1969). Carbon dioxide fixation and phosphoenolpyruvate metabolism in trichinella spiralis larvae. *J. Parasitol.* 55, 67–71. doi: 10.2307/3277348
- White, R. R., Ponsford, A. H., Weekes, M. P., Rodrigues, R. B., Ascher, D. B., Mol, M., et al. (2016). Ubiquitin-dependent modification of skeletal muscle by the parasitic nematode, trichinella spiralis. *PLoS Pathog.* 12, e1005977. doi: 10.1371/journal.ppat.1005977
- Wu, Z., Nagano, I., Kajita, K., Nishina, M., and Takahashi, Y. (2009). Hypoglycaemia induced by trichinella infection is due to the increase of glucose uptake in infected muscle cells. *Int. J. Parasitol.* 39, 427–434. doi: 10.1016/j.ijpara.2008.09.001
- Xu, J., Xu, YXY, Yue, WW, Hao, HN, Liu, RD, Long, SR, et al. (2021). Specific binding of aspartic protease and enterocytes promotes Trichinella spiralis invasion of murine intestinal epithelium cells. *Trop. Agric. Res* 38 (1), 160–171. doi: 10.47665/tb.38.1.028
- Xu, J., Liu, R. D., Long, S. R., Song, Y. Y., Jiang, P., Zhang, X., et al. (2020). Characterization of a chymotrypsin-like enzyme from trichinella spiralis and its facilitation of larva penetration into the host’s enteral epithelial cells. *Res. Vet. Sci.* 128, 1–8. doi: 10.1016/j.rvsc.2019.10.018
- Yamamoto, S., Shimizu, S., Kiyonaka, S., Takahashi, N., Wajima, T., Hara, Y., et al. (2008). TRPM2-mediated Ca²⁺ influx induces chemokine production in monocytes that aggravates inflammatory neutrophil infiltration. *Nat. Med.* 14, 738–747. doi: 10.1038/nm1758
- Yang, C., Liu, H., Li, X., Sui, S., and Liu, Y. (2016). Protective effect of recombinant trichinella 53-kDa protein in sepsis and the effect on macrophages. *Exp. Ther. Med.* 12, 965–968. doi: 10.3892/etm.2016.3354
- Yang, Y., Liu, L., Liu, X., Zhang, Y., Shi, H., Jia, W., et al. (2020). Extracellular vesicles derived from trichinella spiralis muscle larvae ameliorate TNBS-induced colitis in mice. *Front. Immunol.* 11. doi: 10.3389/fimmu.2020.01174
- Yu, P., Xue, X., Zhang, J., Hu, X., Wu, Y., Jiang, L.-H., et al. (2017). Identification of the ADPR binding pocket in the NUDT9 homology domain of TRPM2. *J. Gen. Physiol.* 149, 219–235. doi: 10.1085/jgp.201611675