DOI 10.1002/pmic.201100496 *Proteomics* 2012, *12*, 1860–1869

RESEARCH ARTICLE

Proteomic analysis of *Taenia solium* metacestode excretion–secretion proteins

Bjorn Victor[1], Kirezi Kanobana[2], Sarah Gabriël[1], Katja Polman[2], Nynke Deckers[3], Pierre Dorny[1], André M. Deelder[4] and Magnus Palmblad[4]

[1] Veterinary Helminthology Unit, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium
[2] Medical Helminthology Unit, Department of Biomedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium
[3] Eurovet Animal Health BV, Bladel, The Netherlands
[4] Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands

The metacestode larval stage of *Taenia solium* is the causal agent of a zoonotic disease called cysticercosis. The disease has an important impact on pork trade (due to porcine cysticercosis) and public health (due to human neurocysticercosis). In order to improve the current diagnostic tools and to get a better understanding of the interaction between *T. solium* metacestodes and their host, there is a need for more information about the proteins that are released by the parasite. In this study, we used protein sequences from different helminths, 1DE, reversed-phase LC, and MS/MS to analyze the excretion–secretion proteins produced by *T. solium* metacestodes from infected pigs. This is the first report of the *T. solium* metacestode excretion–secretion proteome. We report 76 proteins including 27 already described *T. solium* proteins, 17 host proteins and 32 proteins likely to be of *T. solium* origin, but identified using sequences from other helminths.

Received: September 23, 2011
Revised: January 10, 2012
Accepted: March 5, 2012

Keywords:
Animal proteomics / Excretion–secretion / Mass spectrometry / *Taenia solium*

1 Introduction

The *Taenia solium* taeniasis/cysticercosis complex is a neglected zoonotic disease with a serious public health and economic impact in large parts of Africa, Asia, and Latin America [1,3]. The adult stage (tapeworm) of *T. solium* develops in the small intestine of humans (taeniasis) and produces eggs that are shed with the stool. The eggs contain the oncosphere larval stage and may be ingested by both humans and pigs due to poor hygienic conditions and coprophagy, respectively. Oncospheres are liberated in the stomach and penetrate the intestinal mucosa. Once inside the blood flow, they migrate through the body, invade the subcutaneous and muscle tissues and develop into metacestode larval stages (cysticerci or cysts), causing porcine or human cysticercosis. Additionally, oncospheres can migrate to the central nervous system, causing neurocysticercosis (NCC) and potentially provoke epilepsy and other neurological symptoms [4].

Taenia solium metacestodes (and many other helminths/helminth stages) are able to establish and maintain an infection in the host, while being attacked by the host immune system. Helminths have therefore developed diverse mechanisms to protect themselves in this hostile environment. These mechanisms often depend on the production of excretion–secretion proteins (ESPs) [5]. Parasite defense mechanisms and ESPs have been

Correspondence: Dr. Bjorn Victor, Institute of Tropical Medicine, Department of Biomedical Sciences, Nationalestraat 155, 2000 Antwerp, Belgium
E-mail: bvictor@itg.be
Fax: +32-3-2476268

Abbreviations: **ABC**, ammonium bicarbonate; **EITB**, enzyme-linked immunoelectrotransfer blot; **ESPs**, excretion–secretion proteins; **FDR**, false discovery rate; **ICC**, ion charge control; **NCBInr**, National Center for Biotechnology Information nonredundant; **NCC**, neurocysticercosis; **NN**, neural networks; **spp.**, species (plural)

studied and described for other helminths of public health importance such as *Echinococcus* [6,7] and *Schistosoma* [8,9], but the excretion–secretion proteome of *T. solium* has not been well characterized so far. A recent proteomic study investigated *T. solium* oncospheres, as this is the stage that determines the success or failure of the infection [10]. From a diagnostic point of view, ESPs are of specific interest. Furthermore, given the importance of the ESPs for the survival of the parasite, they might also be good vaccine candidates or anthelmintic drug targets.

In the case of *T. solium*, a few specific ESPs (e.g., diagnostic antigen GP50 [11]) have been extensively studied and successfully used in antibody detecting enzyme-linked immunoelectrotransfer blots (EITBs) and ELISAs. The drawback is that antibody detection does not allow the differentiation between an active infection and an aborted infection or contact with oncospheres [12]. Other methods, such as monoclonal antibody-based antigen detection ELISAs demonstrate the presence of viable *T. solium* metacestodes by detecting circulating ESPs in the blood, cerebrospinal fluid, and urine [13], but current tests for porcine cysticercosis are not able to differentiate between different *Taenia* spp. A more thorough knowledge of *T. solium* ESPs would be an important step in improving current diagnostic techniques.

A complete proteomic analysis of many parasite ESPs is still hampered by incomplete protein databases due to the lack of genomic information. Although a *T. solium* genome project has been initiated, no data are yet in the public domain. To get around this limitation, we relied on homology to other helminth species and supplemented the *T. solium* database with protein sequences from other helminths.

Furthermore, our study takes full advantage of the recent advancements in reversed-phase LC-MS/MS that have enabled high throughput analysis of large sample numbers.

In this study, we present the first proteomic analysis of the *T. solium* metacestode ESPs using 1DE and LC-MS/MS.

2 Materials and methods

2.1 Parasite material and ESPs

Taenia solium metacestodes were obtained from naturally infected pigs in Zambia and Peru. In each country, five pigs were selected based on positive tongue palpation. The animals were humanely euthanized and at least 200 cysts were dissected from the muscle tissues of each pig. Cysts were washed three times in sterile PBS, two times in sterile PBS containing penicillin (100 units/mL), streptomycin (100 μg/mL), and amphotericin B (0.25 μg/mL) (Invitrogen, Carlsbad, CA, USA) and incubated in culture dishes (BD, Franklin Lakes, NJ, USA) at a concentration of 50 cysts/20 mL of culture medium (RPMI-1640 with L-glutamine and 25 mM HEPES (Invitrogen) supplemented with penicillin (100 units/mL), streptomycin (100 μg/mL), and amphotericin B (0.25 μg/mL)) for 6 h at 37°C in an atmosphere of 5% CO_2. Because it was likely to contain high amounts of host proteins, the culture medium was discarded after 6 h, replaced with fresh medium and incubated for another 18 h. Culture media from 200 cysts and from the same animal were pooled, centrifuged for 15 min at 3000 × *g*, supplemented with Complete Protease Inhibitor Cocktail Tablets (Roche, Indianapolis, IN, USA) and stored at −80°C pending further processing (ESP 24 h). After removal of the 24 h culture medium, cysts were washed three times in culture medium and incubated for another 24 h at 50 cysts/20 mL. The collection of the 48 h fraction was done as described above (ESP 48 h). Each fraction (per pig, country, and time point) was concentrated to a final volume of 2 mL with an Amicon Stirr Cell (Millipore, Billerica, MA, USA) using a 1 kDa ultrafiltration membrane (Millipore) and Macrosep 1K UF Spin Filters (Pall, Port Washington, NY, USA). Concentrated fractions were aliquoted (100 μL) and stored at −20°C. Cyst viability was assessed by incubating 20 cysts/pig in 10% porcine bile in culture medium and subsequent evaluation of cyst evagination.

2.2 1DE

Aliquots were precipitated and washed in ice-cold acetone and resuspended in half the original volume with 10 mM DTT (Sigma-Aldrich Corp., St. Louis, MO, USA) and 3% sodium dodecyl sulfate. When fully resuspended, samples were mixed with NuPAGE LDS Sample Buffer (4x) and separated on a 4–12% NuPAGE Novex Bis-Tris gel (Invitrogen). Gels were stained overnight with the Colloidal Blue Staining Kit (Invitrogen).

2.3 In-gel trypsin digest

Each gel lane was cut in 48 identical slices using a custom designed gel cutter (Gel Company Inc., San Francisco, CA, USA) and all slices were transferred to a 96-well plate (Greiner Bio-One, Frickenhausen, Germany; 1 slice/well). Gel slices were washed three times: first with 100 μL 25 mM ammonium bicarbonate (ABC; Sigma-Aldrich Corp.), then with 100 μL 50% ACN (Biosolve, Valkenswaard, The Netherlands) in ABC and finally with 100 μL 100% ACN. Cystines were reduced by adding 75 μL 10 mM DTT in ABC for 30 min at 56°C. Gel slices were shrunk with 100 μL of 100% ACN and cysteines were alkylated with the addition of 75 μL 50 mM iodoacetamide (Sigma-Aldrich Corp.) in ABC for 20 min (in the dark) at ambient temperature. Gel slices were washed and shrunk again with the addition of 100 μL 25 mM ABC followed by 100 μL 100% ACN. Proteins were digested with sequencing grade porcine trypsin (0.125 μg in 25 μL 25 mM ABC; Promega, Southampton, UK; 25 μL/well) at 37°C for 6 h. From each well, 20 μL of supernatant was extracted and 15 μL of Milli-Q water (Millipore) was added for another

incubation of 30 min at 37°C. A second supernatant sample of 15 μL was extracted and the two supernatants were pooled. TFA (Sigma-Aldrich Corp.) was added to a final concentrationF of 1% to each well to quench the reaction. The total volume per well was reduced to around 15 μL with a vacuum centrifuge and plates were stored at −20°C pending further analysis.

2.4 LC-MS/MS

Peptides were analyzed on an amaZon ETD ion trap mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with an Apollo II ESI source and coupled online with a NanoLC-Ultra 2D plus LC (Eksigent, Dublin, CA, USA). A total volume of 10 μL was first desalted on a C18 PepMap300 trap column (5 μm, 300 μm × 5 mm; Dionex, Sunnyvale, CA, USA) and then separated on a 15 cm C18 ChromXP analytical column (3 μm, 150 × 0.3 mm; Eksigent) by a 45-min linear gradient from 4% to 35% ACN, 0.05% formic acid (Sigma-Aldrich Corp.) with a flow rate of 4 μL/min. Mass spectra were acquired in m/z 300–1300 with the ion charge control (ICC) set to 200 000 and the maximum accumulation time set to 200 ms. The ten most abundant precursors were selected for CID MS/MS in m/z 100–2000 with the ICC and maximum accumulation time set to 200 000 and 50 ms, respectively, but were excluded for 0.5 min after having been selected once. Singly charged ions were excluded for MS/MS. The LC system was controlled by HyStar 3.2 and the ion trap by trapControl 7.0.

2.5 Data analysis

Raw LC-MS/MS data were converted to line spectra mzXML files with the Bruker compassXport tool version 3.0.3. All further data processing was done with the Trans-Proteomic Pipeline 4.4 rev 1 [14]. Database searching was performed with X!Tandem (2009.10.01.1) [15] using the following parameters: precursor mass tolerance window was set between −2.0 and 4.0 Da, while fragment tolerance was set to 0.4 Da, modifications were set for carbamidomethylation of cysteine and oxidation of methionine, k-scoring was enabled, maximum missed cleavages was set to 2 and scoring was done for *b* and *y* ions. The X!Tandem output files were converted to the pepXML file format with tandem2xml (no cutoff was set on X!Tandem hyperscore or E-value). X!Tandem pepXML files from different pigs but corresponding to the same time point and country were merged into PeptideProphet [16] and analyzed with the decoy option enabled [17]. The resulting posterior probabilities for the peptide spectrum matches were further refined by iProphet [18]. Finally, ProteinProphet [19] was used to compute a probability that each protein was present in the sample and to estimate a global false discovery rate (FDR). ProteinProphet output was filtered using the probability threshold that corresponded to an FDR < 1% and proteins with an individual probability of zero were discarded. Filtered protein lists were combined and proteins were organized based on homology using a rough tree (MAFFT; v6.857b) [20]. Further classification was based on Gene Ontology biological process and molecular function information available from the UniProtKB database (http://www.uniprot.org/). SignalP 3.0 server [21] and SecretomeP 2.0 Server [22] were used to predict classical (signal peptide triggered) and nonclassical protein secretions, respectively. The number of spectrum IDs calculated by ProteinProphet was used to estimate the relative protein abundance.

2.6 Databases

The target database used in this study was constructed from three separate databases. The first database (48 245 protein sequences) contained protein sequences from all *Taenia* species as well as protein sequences from the more studied *Echinococcus* and *Schistosoma* helminths. *Trichinella* spp. protein sequences were also included because it is a common parasite in pigs. This database was extracted from the National Center for Biotechnology Information nonredundant (NCBInr) database on June 15, 2011 with a Perl script that (i) looks up the species identification number for all *Taenia*, *Echinococcus*, *Schistosoma*, and *Trichinella* species in the names.dmp file (ftp://ftp.ncbi.nih.gov/pub/taxonomy/taxdump.tar.gz), (ii) finds the matching GenInfo Identifiers in the gi_taxid_prot.dmp file (ftp://ftp.ncbi.nih.gov/pub/taxonomy/gi_taxid_prot.dmp.gz), and (iii) extracts the corresponding protein sequences from the NCBInr database (ftp://ftp.ncbi.nih.gov/blast/db/FASTA/nr.gz). The second database was a *Sus scrofa* database downloaded from the UniProt website (http://www.uniprot.org/) on June 15, 2011 that contained 1388 SwissProt protein sequences. The third database, the common Repository of Adventitious Proteins database (112 protein sequences; ftp://ftp.thegpm.org/fasta/cRAP/crap.fasta) was included to detect common and/or accidental contaminations in the protein samples. A decoy database with 49 745 reversed sequences was created using decoyfasta and target and decoy databases were concatenated into one final database.

3 Results and discussion

3.1 *Taenia solium* metacestodes and in vitro ESP production

On carcass inspection, Zambian pigs were found to be more massively infected than Peruvian pigs. Cyst viability was 72% and 100% for Peru and Zambia, respectively. After each incubation step, plates were checked. Dead cysts were observed and removed in the Peruvian cultures only. The 24 h fractions from two Zambian pigs were lost due to a failed ammonium

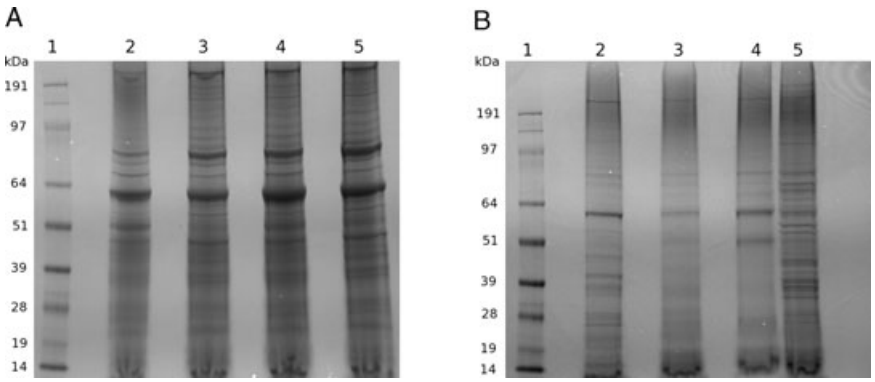

Figure 1. Two examples of SDS-PAGE illustrating (A) the similarities between different ESP fractions from cysts obtained from Peruvian pigs and (B) the differences in protein profiles observed between ESP fractions obtained from cysts from Zambian pigs.

sulphate precipitation. This method was abandoned and all remaining samples were concentrated with filters.

3.2 1DE

Figure 1A shows a protein profile of ESP 24 h and ESP 48 h obtained from cysts coming from two Peruvian pigs. ESP fractions from all Peruvian pigs showed almost identical profiles, irrespective of the incubation time. Figure 1 B illustrates the variation in protein profiles which was observed between ESP 24 h and ESP 48 h fractions obtained from the Zambian pigs.

The discrepancies between profiles from Zambia and Peru could be attributed to differences in ESP production, as well as to proteins leaking from dead/damaged metacestodes or differences in concentration. The aim of the study was to be as complete as possible, not to draw a direct comparison between fractions or countries. Therefore, the differences are not interpreted as erroneous or problematic, but are considered beneficial to the completeness of the study.

3.3 Data analysis

X!Tandem pepXML files from five pigs/ESP 24 h Peru, five pigs/ESP 48 h Peru, five pigs/ESP 48 h Zambia, and three pigs/ESP 24 h Zambia were merged into PeptideProphet, refined with iProphet and further analyzed with ProteinProphet. The ProteinProphet output with FDR < 1% is available in Tables S3– S6 in the Supporting Information. Table 1 summarizes the probability thresholds used to filter the output, the estimated FDRs and the estimated number of correct protein groups, as reported by ProteinProphet. The mixture model in PeptideProphet failed for charge (1+) and (for ESP 24 h Zambia) also for charge (3+).

The combined, filtered and organized output resulted in 76 protein groups containing one or more (homologous) proteins from *T. solium* and/or other organisms (Supporting Information Table S7). Keratin contaminations were omitted from the list. The inclusions of multiple homologous proteins in a protein group is a consequence of the way the database was constructed. However, it does not imply independent evidence for multiple isoforms of the proteins, although, in the case of the 8 kDa protein family, it is highly likely that multiple proteins from this family were present in the samples. SignalP and SecretomeP analyses were done on all proteins. The former reports yes (Y) or no (N) for both the neural networks (NN) and hidden Markov models, while the latter interprets a NN-score above 0.5 as possible secretion (Y).

Table 1. The ProteinProphet estimations for the four merged fractions

Merged fraction	Probability threshold	Estimated FDR (%)	Estimated number of correct protein groups
ESP 24 h Peru	0.90	0.5	82
ESP 48 h Peru	0.90	0.4	85
ESP 24 h Zambia	0.80	0.8	30
ESP 48 h Zambia	0.90	0.5	49

3.4 Proteins identified in *T. solium* ESPs

Table 2 lists the 76 protein groups that were found in the *T. solium* ESPs, organized by Gene Ontology annotations for biological process and molecular function. For simplicity, the protein groups are represented by one protein. The objective of this study was not to compare different time points and/or geographical locations, but to provide a broader coverage and identify ubiquitous components of the ES proteome. Therefore, time points and geographical locations of the proteins were added to Table 2 for completeness only. The results of SignalP and SecretomeP show a number of proteins that are negative for both analyses. These proteins are mostly found in Peruvian ESPs and may be a result of leakage due to cyst damage/death, since dead cysts were observed in Peruvian cultures.

Table 2. Proteins ($n = 76$) identified in *T. solium* metacestode ESPs organized by Gene Ontology annotations for biological process and molecular function

Gene ontology classification	Closest organism	gi/sp code	Peru[a)]		Zambia[a)]		SigP/SecP[b)]
Protein			24 h	48 h	24 h	48 h	
1) No Gene Ontology information							
Hypothetical protein	*T. solium*	21726976	–	–	–	+	N/Y
Antigen EM/EG13	*Echinococcus* spp.	Multiple	+	–	–	–	N/N
8 kDa protein family[c)]	*T. solium*	Multiple	–/+	–/+	+	–/+	–/–
Antigen	*T. solium*	29893185	+	+	–	–	N/Y
P27	*T. solium*	333101790	+	+	–	–	N/Y
Apolipoprotein AI binding protein	*E. multilocularis*	223640019	–	+	–	–	N/Y
Immunogenic protein	*T. solium*	Multiple	+	+	–	–	Y/Y
T24	*T. solium*	37786712	+	+	+	+	–/N
Hypothetical protein	*T. solium*	21912540	+	+	–	–	N/N
Diagnostic antigen GP50	*T. solium*	Multiple	+	+	+	+	–/–
Small heat-shock protein	*T. solium*	21665905	+	+	–	–	N/Y
2) Microtubule-based movement/motor activity/tegument-cytoskeleton							
Paramyosin	*T. solium*	Multiple	–	–	–	+	N/–
H17g protein	*T. solium*	34368418	+	+	+	–	N/Y
Actin	Multiple genera[c)]	Multiple	–/+	–/+	–/+	–/+	–/–
Dynein light chain	Multiple genera	Multiple	–/+	+	–	–/+	N/–
Tubulin	Multiple genera	Multiple	–/+	–/+	–/+	–/+	–/–
3) Response to stress/chaperone							
Heat shock protein 70	Multiple genera	Multiple	+	+	–	–/+	N/N
Heat shock protein 90 alpha	*T. asiatica*	124783119	+	+	–	+	N/N
Heat shock protein 90	*T. asiatica*	124783236	+	+	–	–	N/Y
4) Metabolism/biosynthetic processes							
Phosphoglycerate kinase	Multiple genera	Multiple	+	–/+	–	+	N/–
Fructose-bisphosphate aldolase	Multiple genera	Multiple	–/+	+	–/+	–/+	N/N
Enolase	Multiple genera	Multiple	–/+	–/+	–/+	–/+	N/–
Cytosolic malate dehydrogenase	*T. solium*	323361126	+	+	+	+	N/N
Glucose phosphate isomerase	*Echinococcus* spp.	Multiple	+	–/+	–	+	N/Y
Glycogen phosphorylase	*T. spiralis*	316972087	–	+	–	–	N/N
GAPDH[d)]	*T. solium*	149364041	+	+	–	+	N/Y
Triosephosphate isomerase	*T. solium*	38258647	+	+	+	+	N/N
Phosphoenolpyruvate carboxykinase	*T. solium*	283466482	+	+	+	+	N/N
Glucosidase	Multiple genera	Multiple	+	+	–	–	Y/–
Adenosylhomocysteinase	*S. scrofa*	Q710C4	–	+	–	–	N/N
Cell polarity protein	*S. mansoni*	256075333	–	+	–	–	N/N
Carbonic anhydrase 3	*S. scrofa*	Q5S1S4	+	+	–	–	N/N
Creatine kinase M-type	*S. scrofa*	Q5XLD3	+	+	–	–	N/N
Na^+/K^+-transporting ATPase subunit alpha	*T. solium*	74794482	+	+	–	+	N/Y
Nucleoside diphosphate kinase	*Schistosoma* spp.	Multiple	+	+	–	–	N/N
Elongation factor 1 alpha	*T. solium*	Multiple	–	+	–	–	N/N
5) Signal transduction							
GTPase	Multiple genera	Multiple	–/+	–/+	–	–	N/Y
ADP-ribosylation factor	Multiple genera	Multiple	+	–/+	–	–	N/–
Multiple pdz domain protein	*S. mansoni*	256082156	–	–	–	+	N/N
CAMP-dependent protein kinase regulatory subunit	*S. japonicum*	Multiple	+	+	–	–	N/Y
14-3-3 protein	*Echinococcus* spp.	Multiple	+	+	–/+	+	N/–
6) Protein folding							
Cyclophilin	Multiple genera	Multiple	+	+	–/+	+	N/N
Calreticulin	*T. solium*	14029538	+	–	–	–	Y/N
7) Transport							
Glucose transporter TGTP2	*T. solium*	1480799	+	–	–	–	Y/N
Apolipoprotein A-I	*S. scrofa*	P18648	+	–	–	–	Y/Y
Transthyretin	*S. scrofa*	P50390	+	–	–	–	Y/Y

Table 2. Continued.

Gene Ontology classification / Protein	Closest organism	gi/sp code	Peru[a)] 24 h	Peru[a)] 48 h	Zambia[a)] 24 h	Zambia[a)] 48 h	SigP/SecP[b)]
Hemopexin	*S. scrofa*	P50828	+	+	–	–	Y/Y
Serum albumin	*S. scrofa*	P08835	+	+	+	+	Y/Y
Transferrin	*S. scrofa*	Multiple	+	+	–/+	–/+	–/–
8) Cell redox homeostasis/detoxification/oxidoreductase activity							
Thioredoxin	*E. granulosus*	29337032	+	+	–	+	N/N
2-Cys peroxiredoxin	*T. solium*	Multiple	–/+	–/+	+	–/+	N/Y
Cytosolic Cu/Zn-superoxide dismutase	*T. solium*	18252397	+	+	–	+	N/N
Glutathione S-transferase	*T. solium*	Multiple	+	+	–/+	–/+	N/–
Ferritin	*T. saginata*	1297064	+	+	–	+	N/N
Aldo-keto reductase	*Schistosoma* spp.	Multiple	+	+	–	–	N/–
9) Proteolysis/endopeptidase activity							
Trypsin-like protein	*T. solium*	311335041	+	+	+	+	Y/Y
Cathepsin L-like cysteine proteinase	*T. solium*	Multiple	+	+	–	–	Y/Y
Proteasome subunit alpha 2 (T01 family)	*S. mansoni*	256083548	–	+	–	+	N/N
Haptoglobin	*S. scrofa*	Q8SPS7	+	+	+	–	Y/Y
Trypsin	*S. scrofa*	P00761	+	+	+	+	N/Y
10) Endopeptidase inhibitor activity							
Immunogenic protein Ts11	*T. solium*	7339849	+	+	–	–	Y/Y
Alpha-1-antitrypsin	*S. scrofa*	P50447	+	+	–	–	Y/Y
Leukocyte elastase inhibitor	*S. scrofa*	P80229	+	–	–	–	N/Y
Interalpha-trypsin inhibitor heavy chain	*S. scrofa*	P79263	+	+	–	–	Y/Y
Alpha-2-HS-glycoprotein (Fragment)	*S. scrofa*	P29700	+	+	–	–	Y/Y
11) Binding (miscellaneous)							
Cytosolic fatty acid binding protein	*T. solium*	82412213	+	+	+	+	N/N
Calcium-binding protein	*Schistosoma* spp.	Multiple	–	+	–	–	N/Y
Annexin	*T. solium*	Multiple	+	–/+	–/+	+	N/–
Actin-binding protein	Multiple genera	Multiple	+	+	–	–	–/–
Tuftelin interacting protein	*S. mansoni*	256076696	+	–	–	–	N/N
Hemoglobin subunit beta	*S. scrofa*	P02067	+	+	–	–	N/N
Ig lambda chain C region	*S. scrofa*	P01846	+	+	+	+	N/Y
12) Translation							
Ubiquitin	Multiple genera	Multiple	+	+	–	–/+	N/Y
13) Cell division							
Cell division control protein	*S. mansoni*	Multiple	–	+	–	–	N/N
14) Defense response to bacterium							
Protegrin (and related)	*S. scrofa*	Multiple	–	–	+	–	–/–
15) Catalytic activity							
Transketolase	*E. multilocularis*	27526313	+	+	–	–	N/Y

a) Proteins present in Peru/Zambia in the 24 h/48 h fraction are described as + while – indicates absence. –/+ indicates there were multiple entries for this protein with variable presence in the fractions.
b) SignalP and SecretomeP (SigP/SecP) results are described as yes (Y) or no (N). SignalP is Y when both the neural network and the hidden Markov models agree. (–) is reported when not all individual proteins in that groups have the same result.
c) For simplicity, the 8 kDa protein family is presented as one protein although it is highly likely that multiple proteins from this family were present in the samples.
d) GAPDH: glyceraldehyde-3-phosphate dehydrogenase.

Gene Ontology group 1 includes proteins with little information on molecular function or biological processes in the UniProtKB. However, this group includes proteins that are widely used in immunodiagnostic assays. Group 2 contains proteins related to muscle (paramyosin), microtubule-based processes (dynein light chain), protein polymerization (tubulin), and tegument/cytoskeleton (H17g protein and actin). Group 3 are the stress response/chaperone proteins and group 4 lists proteins involved in various metabolic activities, e.g., glycolysis (phosphoglycerate kinase, fructose-bisphosphate aldolase, enolase, glucose phosphate isomerase, GADPH) and gluconeogenesis (phosphoenolpyruvate carboxykinase) as well as biosynthetic processes/protein synthesis (Na^+/K^+-transporting ATPase,

nucleoside diphosphate kinase, elongation factor 1 alpha). Group 5 entails signal transduction, group 6 presents protein folding, and group 7 contains transport proteins. Group 8 contains the enzymatic antioxidant system of Taeniidae (cytosolic Cu/Zn-superoxide dismutase, glutathione S-transferase, 2-cys peroxiredoxin) and proteins with oxidoreductase activity (ferritin, aldo-keto reductase). Group 9 holds proteolysis related proteins with serine-type (trypsin-like protein, trypsin), cysteine-type (cathepsin L-like cysteine proteinase), and threonine-type (proteasome subunit alpha) endopeptidase activity. Trypsin was introduced during the protein digestion step. Group 10 contains proteins with cysteine-type (alpha-2-HS-glycoprotein [Fragment], immunogenic protein Ts11) and serine-type endopeptidase inhibitor activity (alpha-1-antitrypsin, leukocyte elastase inhibitor, interalpha-trypsin inhibitor). Group 11 are binding proteins, e.g., fatty acid (fatty acid binding protein), calcium (annexin, calcium-binding protein), actin (actin-binding protein), DNA (tuftelin interacting protein), oxygen (hemoglobin), and antigen-binding proteins (Ig lambda chain C region). Groups 12, 13, 14, and 15 contain proteins involved in translation, cell division, defense response, and catalytic activity, respectively.

3.4.1 Relative protein abundance

Protein abundance showed similar patterns in Peru ESP 24 h and 48 h: porcine albumin and phosphoenolpyruvate carboxykinase were most abundant followed by members of the 8 kDa protein family, immunogenic protein, trypsin-like protein, enolase, 2-cys peroxiredoxin and 14-3-3 protein. Host proteins (serotransferrin, Ig) were present in both ESPs although more abundant in the 24 h fraction. Zambia ESP 24 h showed high abundance for members of the 8 kDa protein family and porcine albumin, followed by immunogenic protein, trypsin-like protein, cyclophilin, porcine Ig, phosphoenolpyruvate carboxykinase and porcine serotransferrin. Zambia ESP 48 h showed highest abundance for members of the 8 kDa protein family and phosphoenolpyruvate carboxykinase, followed by immunogenic protein, porcine albumin, glutathione S-transferase, enolase, and 2-cys peroxiredoxin.

The abundance of albumin in the ESPs can be explained by its abundance in blood (despite multiple washings) and possibly because it was secreted by the cysts. Secretion of host albumin has been described in *Taenia crassiceps* metacestodes where it may be involved in osmoregulation [23].

3.4.2 *Taenia solium* proteins

Of the 76 proteins listed in Table 2, 27 have been described in literature specifically for *T. solium*. Many of these proteins have been studied due to their importance as diagnostic antigens in antibody detecting assays [13], e.g., diagnostic antigen GP50 [11], T24 [24], and the 8 kDa protein family [25], [26], which are the three protein groups that represent the seven diagnostic bands in the EITB [27]. Members of the 8 kDa protein family are believed to be hydrophobic ligand binding proteins [25] and a similarity to the related *E. granulosus* AgB might indicate an immunoregulatory role, enabling the parasite to reduce an inflammatory response [28] and manipulate the Th1–Th2 balance [29]. Paramyosin [30], 14-3-3 protein [10], actin [31], P27, small heat-shock protein [32] and phosphoenolpyruvate carboxykinase have been described as immunoreactive proteins recognized by NCC-positive human sera [33]. Paramyosin in particular is also reported to inhibit complement formation [34]. P27 is homologous to P-29 (*E. granulosus*) and Antigen 6 (*E. multilocularis*). Several studies have shown similarities between sequences and epitopes of P-29, Antigen S, Antigen 6, and even Antigen 5, which possesses similar protease activity as the *T. solium* trypsin-like protein [35,38]. The enzymatic antioxidant system of Taeniidae involves Cu/Zn-superoxide dismutase [39], glutathione S-transferase [40,43] and 2-cys peroxiredoxin [44]. These proteins have been considered as targets for drugs and vaccines [45]. Cysteine proteases help the parasite to evade the host immune system in multiple ways: they possess the ability to lyse host immunoglobulins [46] and in vitro tests have proven their capability to deplete CD4+ lymphocytes by inducing apoptosis [47,48]. Furthermore, cathepsin L-like cysteine peptidase has been used for differential diagnosis of infections with *Taenia* species with a loop-mediated isothermal amplification assay [49]. Annexin B2 is known to be associated with inflammatory reaction [50]. Other studies have shown anticoagulant activity [51]. Antigen cC1 [52] (homologous to annexin) was suggested as a candidate for the development of a vaccine against *T. solium* cysticercosis. Calreticulin controls intracellular Ca^{2+} homeostasis, but can also assist in secretion, protein synthesis, and control of protein folding [53]. Cytosolic malate dehydrogenase is involved in the citric acid metabolism and has recently been studied because it might potentially serve as a target for drugs [54]. Glucose transporter TGTP2 is believed to mediate sugar uptake and is predominantly localized on the surface of the cyst [55]. It is however not recognized by the host immune system [56].

3.4.3 Proteins matching sequences from other species

The interpretation of the 49 proteins that are not identified by *T. solium* requires careful consideration to assess if they are likely to be of helminth or host origin.

In most cases, spectra matching proteins from other helminths ($n = 32$) can be assumed to derive from *T. solium* proteins not present in the *T. solium* database. A number of proteomic studies have been carried out on other helminth parasites such as *Schistosoma* spp. [8,9] and *E. granulosus* metacestodes [7]. Many of those proteins are identified in this study as well and it is reasonable to assume that the proposed function is similar. Although a few of these proteins (e.g., enolase, tubulin, and phosphoglycerate kinase) are

identified also by *S. scrofa*, the proteins are unlikely to be of host origin due to their nature or function and presence in other helminths.

A total of 17 proteins were identified only by *S. scrofa* and because most of them are commonly found in blood (e.g., Ig, hemoglobin, endopeptidase inhibitors, and transport proteins), they are almost certainly of host origin. They may be considered as contamination, although importantly, some host proteins (e.g., albumin) are also known to be taken up and secreted by cysts [23].

4 Concluding remarks

This study presents the first report of the *T. solium* metacestode excretion–secretion proteome. Since the genome of *T. solium* is not yet sequenced and only 401 protein entries were found in the NCBInr database, we relied on homology to other helminths to increase the proteome coverage. We report 76 proteins including 27 already described *T. solium* proteins, 17 host proteins, and 32 proteins likely to be of *T. solium* origin but identified using sequences from other helminths, effectively demonstrating the value of this approach. Until a reference genome for *T. solium* becomes available, it is likely that several proteins in the ESPs remain unidentified. The fact that many of the proteins found in this study are involved in parasite survival strategies and have been described in other helminths as well indicates that the excretion–secretion proteome is not all that different between species or even related genera. Therefore, it is reasonable to assume that the proposed protein functions are similar. Albeit the close homology to another protein may limit the specificity of diagnostic tests based on that protein, it also has the potential to widen the usable range of a vaccine or drug.

Future research could include sequencing and annotating the genomes of *Taenia* spp. as this will greatly improve the distinction between host and parasite proteins as well as the identification of new proteins. Furthermore, it will allow for a more detailed study of small differences between homologous proteins.

The authors thank the University of Zambia, School of Veterinary Medicine (Zambia) and the Universidad National Mayor de San Marcos (Lima, Peru) for assistance and use of facilities, and Hannah Scott and Hans Dalebout (LUMC, Leiden, The Netherlands) for technical support. The research leading to these results has received funding from The Research Foundation—Flanders (FWO) (project number: G.0192.10N) and the European Union's Seventh Framework Program (FP7/2007-2013) under grant agreement no. 221948 (ICONZ). The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

The authors have declared no conflict of interest.

5 References

[1] Flisser, A., Sarti, E., Lightowlers, M. W., Schantz, P. M., Neurocysticercosis: regional status, epidemiology, impact and control measures in the Americas. *Acta Trop.* 2003, *87*, 43–51.

[2] Rajshekhar, V., Joshi, D. D., Doanh, N. Q., van De, N. et al., *Taenia solium* taeniosis/cysticercosis in Asia: epidemiology, impact and issues. *Acta Trop.* 2003, *87*, 53–60.

[3] Phiri, I. K., Ngowi, H., Afonso, S., Matenga, E. et al., The emergence of *Taenia solium* cysticercosis in Eastern and Southern Africa as a serious agricultural problem and public health risk. *Acta Trop.* 2003, *87*, 13–23.

[4] White, A. C., Jr, Neurocysticercosis: updates on epidemiology, pathogenesis, diagnosis, and management. *Annu. Rev. Med.* 2000, *51*, 187–206.

[5] Hewitson, J. P., Grainger, J. R. Maizels, R. M., Helminth immunoregulation: the role of parasite secreted proteins in modulating host immunity. *Mol. Biochem. Parasitol.* 2009, *167*, 1–11.

[6] Siracusano, A., Riganò, R., Ortona, E., Profumo, E. et al., Immunomodulatory mechanisms during *Echinococcus granulosus* infection. *Exp. Parasitol.* 2008, *119*, 483–489.

[7] Monteiro, K. M., de Carvalho, M. O., Zaha, A. Ferreira, H. B., Proteomic analysis of the *Echinococcus granulosus* metacestode during infection of its intermediate host. *Proteomics* 2010, *10*, 1985–1999.

[8] Guillou, F., Roger, E., Moné, Y., Rognon, A. et al., Excretory-secretory proteome of larval *Schistosoma mansoni* and *Echinostoma caproni*, two parasites of *Biomphalaria glabrata*. *Mol. Biochem. Parasitol.* 2007, *155*, 45–56.

[9] Liu, F., Cui, S.-J., Hu, W., Feng, Z. et al., Excretory/secretory proteome of the adult developmental stage of human blood fluke, *Schistosoma japonicum*. *Mol. Cell. Proteomics* 2009, *8*, 1236–1251.

[10] Santivañez, S. J., Hernández-González, A., Chile, N., Oleaga, A. et al., Proteomic study of activated *Taenia solium* oncospheres. *Mol. Biochem. Parasitol.* 2010, *171*, 32–39.

[11] Hancock, K., Pattabhi, S., Greene, R. M., Yushak, M. L. et al., Characterization and cloning of GP50, a *Taenia solium* antigen diagnostic for cysticercosis. *Mol. Biochem. Parasitol.* 2004, *133*, 115–124.

[12] García, H. H., Gonzalez, A. E., Gilman, R. H., Palacios, L. G. et al., Short report: transient antibody response in *Taenia solium* infection in field conditions-a major contributor to high seroprevalence. *Am. J. Trop. Med. Hyg.* 2001, *65*, 31–32.

[13] Deckers, N., Dorny, P., Immunodiagnosis of *Taenia solium* taeniosis/cysticercosis. *Trends Parasitol.* 2010, *26*, 137–144.

[14] Keller, A., Eng, J., Zhang, N., Li, X.-J. Aebersold, R., A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol. Syst. Biol.* 2005, *1*, 1–8.

[15] Craig, R., Beavis, R. C., TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 2004, *20*, 1466–1467.

[16] Keller, A., Nesvizhskii, A. I., Kolker, E. Aebersold, R., Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 2002, *74*, 5383–5392.

[17] Choi, H., Nesvizhskii, A. I., Semisupervised model-based validation of peptide identifications in mass spectrometry-based proteomics. *J. Proteome Res.* 2008, *7*, 254–265.

[18] Shteynberg, D., Deutsch, E. W., Lam, H., Eng, J. K. et al., iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell. Proteomics* 2011, *10*, M111.007690.

[19] Nesvizhskii, A. I., Keller, A., Kolker, E., Aebersold, R., A statistical model for identifying proteins by MS/MS. *Anal. Chem.* 2003, *75*, 4646–4658.

[20] Katoh, K., Toh, H., Recent developments in the MAFFT multiple sequence alignment program. *Briefings Bioinf.* 2008, *9*, 286–298.

[21] Bendtsen, J. D., Nielsen, H., von Heijne, G., Brunak, S., Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 2004, *340*, 783–795.

[22] Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G. et al., Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng., Des. Sel.* 2004, *17*, 349–356.

[23] Aldridge, J. R., Jr, Jennette, M. A., Kuhn, R. E., Uptake and secretion of host proteins by *Taenia crassiceps* metacestodes. *J. Parasitol.* 2006, *92*, 1101–1102.

[24] Hancock, K., Pattabhi, S., Whitfield, F. W., Yushak, M. L. et al., Characterization and cloning of T24, a *Taenia solium* antigen diagnostic for cysticercosis. *Mol. Biochem. Parasitol.* 2006, *147*, 109–117.

[25] Hancock, K., Khan, A., Williams, F. B., Yushak, M. L. et al., Characterization of the 8-kilodalton antigens of *Taenia solium* metacestodes and evaluation of their use in an enzyme-linked immunosorbent assay for serodiagnosis. *J. Clin. Microbiol.* 2003, *41*, 2577–2586.

[26] Ferrer, E., Bonay, P., Foster-Cuevas, M., González, L. et al., Molecular cloning and characterisation of Ts8B1, Ts8B2 and Ts8B3, three new members of the *Taenia solium* metacestode 8 kDa diagnostic antigen family. *Mol. Biochem. Parasitol.* 2007, *152*, 90–100.

[27] Tsang, V. C., Brand, J. A., Boyer, A. E., An enzyme-linked immunoelectrotransfer blot assay and glycoprotein antigens for diagnosing human cysticercosis (*Taenia solium*). *J. Infect. Dis.* 1989, *159*, 50–59.

[28] Shepherd, J. C., Aitken, A., McManus, D. P., A protein secreted *in vivo* by *Echinococcus granulosus* inhibits elastase activity and neutrophil chemotaxis. *Mol. Biochem. Parasitol.* 1991, *44*, 81–90.

[29] Riganò, R., Profumo, E., Bruschi, F., Carulli, G. et al., Modulation of human immune response by *Echinococcus granulosus* antigen B and its possible role in evading host defenses. *Infect. Immun.* 2001, *69*, 288–296.

[30] Vargas-Parada, L., Laclette, J. P., Gene structure of *Taenia solium* paramyosin. *Parasitol. Res.* 2003, *89*, 375–378.

[31] Campos, A., Bernard, P., Fauconnier, A., Landa, A. et al., Cloning and sequencing of two actin genes from *Taenia solium* (Cestoda). *Mol. Biochem. Parasitol.* 1990, *40*, 87–93.

[32] Ferrer, E., González, L. M., Foster-Cuevas, M., Cortéz, M. M. et al., *Taenia solium*: characterization of a small heat shock protein (Tsol-sHSP35.6) and its possible relevance to the diagnosis and pathogenesis of neurocysticercosis. *Exp. Parasitol.* 2005, *110*, 1–11.

[33] Salazar-Anton, F., Lindh, J., *Taenia solium*: a two-dimensional Western blotting method combined with the use of an EST-library for the identification of immunogenic proteins recognized by sera from neurocysticercosis patients. *Exp. Parasitol.* 2011, *128*, 371–376.

[34] Laclette, J. P., Shoemaker, C. B., Richter, D., Arcos, L. et al., Paramyosin inhibits complement C1. *J. Immunol.* 1992, *148*, 124–128.

[35] Lorenzo, C., Salinas, G., Brugnini, A., Wernstedt, C. et al., *Echinococcus granulosus* antigen 5 is closely related to proteases of the trypsin family. *Biochem. J.* 2003, *369*, 191–188.

[36] Siles-Lucas, M., Gottstein, B., Felleisen, R. S., Identification of a differentially expressed *Echinococcus multilocularis* protein Em6 potentially related to antigen 5 of *Echinococcus granulosus*. *Parasite Immunol.* 1998, *20*, 473–481.

[37] González, G., Spinelli, P., Lorenzo, C., Hellman, U. et al., Molecular characterization of P-29, a metacestode-specific component of *Echinococcus granulosus* which is immunologically related to, but distinct from, antigen 5. *Mol. Biochem. Parasitol.* 2000, *105*, 177–184.

[38] Ben Nouir, N., Gianinazzi, C., Gorcii, M., Müller, N. et al., Isolation and molecular characterization of recombinant *Echinococcus granulosus* P29 protein (recP29) and its assessment for the post-surgical serological follow-up of human cystic echinococcosis in young patients. *Trans. R. Soc. Trop. Med. Hyg.* 2009, *103*, 355–364.

[39] Castellanos-González, A., Jiménez, L., Landa, A., Cloning, production and characterisation of a recombinant Cu/Zn superoxide dismutase from *Taenia solium*. *Int. J. Parasitol.* 2002, *32*, 1175–1182.

[40] Vibanco-Pérez, N., Jiménez, L., Mendoza-Hernández, G., Landa, A., Characterization of a recombinant mu-class glutathione S-transferase from *Taenia solium*. *Parasitol. Res.* 2002, *88*, 398–404.

[41] Plancarte, A., Rendon, J. L., Landa, A., Purification, characterization and kinetic properties of the *Taenia solium* glutathione S-transferase isoform 26.5 kDa. *Parasitol. Res.* 2004, *93*, 137–144.

[42] Torres-Rivera, A., Landa, A., Cooperative kinetics of the recombinant glutathione transferase of *Taenia solium* and characterization of the enzyme. *Arch. Biochem. Biophys.* 2008, *477*, 372–378.

[43] Nguyen, H. A., Bae, Y.-A., Lee, E.-G., Kim, S.-H. et al., A novel sigma-like glutathione transferase of *Taenia solium* metacestode. *Int. J. Parasitol.* 2010, *40*, 1097–1106.

[44] Vaca-Paniagua, F., Parra-Unda, R., Landa, A., Characterization of one typical 2-Cys peroxiredoxin gene of *Taenia solium* and *Taenia crassiceps*. *Parasitol. Res.* 2009, *105*, 781–787.

[45] Vaca-Paniagua, F., Torres-Rivera, A., Parra-Unda, R., Landa, A., *Taenia solium*: antioxidant metabolism enzymes as targets for cestocidal drugs and vaccines. *Curr. Top. Med. Chem.* 2008, *8*, 393–399.

[46] White, A. C., Jr, Baig, S., Chappell, C. L., Characterization of a cysteine proteinase from *Taenia crassiceps* cysts. *Mol. Biochem. Parasitol.* 1997, *85*, 243–253.

[47] Molinari, J. L., Mejia, H., White, A. C., Jr, Garrido, E. et al., *Taenia solium*: a cysteine protease secreted by metacestodes depletes human CD4 lymphocytes *in vitro*. *Exp. Parasitol.* 2000, *94*, 133–142.

[48] Tato, P., Fernández, A. M., Solano, S., Borgonio, V. et al., A cysteine protease from *Taenia solium* metacestodes induce apoptosis in human CD4+ T-cells. *Parasitol. Res.* 2004, *92*, 197–204.

[49] Nkouawa, A., Sako, Y., Nakao, M., Nakaya, K. et al., Loop-mediated isothermal amplification method for differentiation and rapid detection of *Taenia* species. *J. Clin. Microbiol.* 2009, *47*, 168–174.

[50] Gao, Y.-J., Yan, H.-L., Ding, F.-X., Lu, Y.-M. et al., Annexin B1 at the host-parasite interface of the *Taenia solium* cysticercus: Secreted and associated with inflammatory reaction. *Acta Trop.* 2007, *101*, 192–199.

[51] Lu, Y.-M., Wang, N., Wang, J.-J., Wang, K.-H. et al., Expression, purification, and characterization of a novel Ca(2+)- and phospholipid-binding protein annexin B2. *Mol. Biol. Rep.* 2010, *37*, 1591–1596.

[52] Guo, Y.-J., Sun, S.-H., Zhang, Y., Chen, Z.-H. et al., Protection of pigs against *Taenia solium* cysticercosis using recombinant antigen or in combination with DNA vaccine. *Vaccine* 2004, *22*, 3841–3847.

[53] Mendlovic, F., Ostoa-Saloma, P., Solís, C. F., Martínez-Ocaña, J. et al., Cloning, characterization, and functional expression of *Taenia solium* calreticulin. *J. Parasitol.* 2004, *90*, 891–893.

[54] Nava, G., Laclette, J. P., Bobes, R., Carrero, J. C. et al., Cloning, sequencing and functional expression of cytosolic malate dehydrogenase from *Taenia solium*: purification and characterization of the recombinant enzyme. *Exp. Parasitol.* 2011, *128*, 217–224.

[55] Rodríguez-Contreras, D., Skelly, P. J., Landa, A., Shoemaker, C. B. et al., Molecular and functional characterization and tissue localization of 2 glucose transporter homologues (TGTP1 and TGTP2) from the tapeworm *Taenia solium*. *Parasitology* 1998, *117*, 579–588.

[56] Rodríguez-Contreras, D., de, I. T. P., Velasco, J., Shoemaker, C. B. et al., The *Taenia solium* glucose transporters TGTP1 and TGTP2 are not immunologically recognized by cysticercotic humans and swine. *Parasitol. Res.* 2002, *88*, 280–282.

[57] Pérez-Sánchez, R., Valero, M. L., Ramajo-Hernández, A., Siles-Lucas, M. et al., A proteomic approach to the identification of tegumental proteins of male and female *Schistosoma bovis* worms. *Mol. Biochem. Parasitol.* 2008, *161*, 112–123.

