Transcriptome Profiles of the Protoscoleces of *Echinococcus granulosus* Reveal that Excretory-Secretory Products Are Essential to Metabolic Adaptation

Wei Pan1,2,3, Yujuan Shen1,2,3,*, Xiuming Han4, Ying Wang1,2,3, Hua Liu1,2,3, Yanyan Jianga1,2,3, Yumei Zhang1,2,3, Yanjuan Wang1,2,3, Yuxin Xua1,2,3, Jianping Cao1,2,3

1 National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention, Shanghai, People’s Republic of China, 2 Key Laboratory of Parasite and Vector Biology, Ministry of Health, Shanghai, People’s Republic of China, 3 WHO Collaborating Center for Malaria, Schistosomiasis and Filarialis, Shanghai, People’s Republic of China, 4 Department of Parasitic Diseases, Qinghai Institute for Endemic Disease Prevention and Control, Zong Zhai, Xining, Qinghai, People’s Republic of China

Abstract

**Background:** Cystic hydatid disease (CHD) is caused by the larval stages of the cestode and affects humans and domestic animals worldwide. Protoscoleces (PSCs) are one component of the larval stages that can interact with both definitive and intermediate hosts. Previous genomic and transcriptomic data have provided an overall snapshot of the genomics of the growth and development of this parasite. However, our understanding of how PSCs subvert the immune response of hosts and maintains metabolic adaptation remains unclear. In this study, we used Roche 454 sequencing technology and in silico secretome analysis to explore the transcriptome profiles of the PSCs from *E. granulosus* and elucidate the potential functions of the excretory-secretory proteins (ESPs) released by the parasite.

**Methodology/Principal Findings:** A large number of nonredundant sequences as unigenes were generated (26,514), of which 22,910 (86.4%) were mapped to the newly published *E. granulosus* genome and 17,705 (66.8%) were distributed within the coding sequence (CDS) regions. Of the 2,280 ESPs predicted from the transcriptome, 138 ESPs were inferred to be involved in the metabolism of carbohydrates, while 124 ESPs were inferred to be involved in the metabolism of protein. Eleven ESPs were identified as intracellular enzymes that regulate glycolysis/gluconeogenesis (GL/GN) pathways, while a further 44 antigenic proteins, 25 molecular chaperones and four proteases were highly represented. Many proteins were also found to be significantly enriched in development-related signaling pathways, such as the TGF-β receptor pathways and insulin pathways.

**Conclusions/Significance:** This study provides valuable information on the metabolic adaptation of parasites to their hosts that can be used to aid the development of novel intervention targets for hydatid treatment and control.

Introduction

Cystic hydatid disease (CHD) is a serious parasitic zoonosis that is caused by the larval stages of *Echinococcus granulosus*, a cestode that poses a threat to public health as well as significant economic losses [1,2,3]. At present, more than 3 million people are infected with this parasite [4,5], and the prevalence reaches 10% in some areas [6,7]. The disease is difficult to control because appropriate diagnostic procedures are lacking and the available drugs are inefficient [8].

*E. granulosus* has a complex developmental cycle, involving eggs, oncospheres, protoscoleces (PSCs), and adult stages. Adult parasites live in the small intestine of dogs. After sexual maturation, numerous eggs are produced by the adult parasites and are then excreted with the dog feces. Infections occur in an intermediate host, when eggs containing larvae are ingested. Hydatid cysts (the larval stage or metacestode) develop in the internal organs (primarily in liver and lungs) of intermediate hosts. The larval stages of *E. granulosus* are comprised of two layers of cyst wall: cyst fluid and PSCs [9].

As the only infectious form of the larval stages, PSCs can interact with both definitive and intermediate hosts. They mature into adult parasites when the hydatid cysts are ingested by the definitive host. They can also differentiate into new cysts when released into the body cavity of intermediate hosts upon cyst rupture [10]. Mouse models of CHD are often established via the
intraperitoneal inoculation with PSCs, a method that has been widely applied to drug screening and vaccine development [11,12]. Overall, the PSC is an important infectious reagent that contributes to the transmission of CHD and also an excellent model system in which many aspects of the host-parasite interaction can be studied.

Understanding the elaborate immune evasion strategies and mechanisms of physiological adaptation of the PSCs is critical to ascertain effective intervention targets to control the prevalence of the parasite. In this study, we focus on the role of excretory-secretory products (ESPs) that are released by parasites, as these compounds are exposed directly to the immune system of the hosts and are engaged at the host-parasite interface [13]. The mechanism by which PSCs can subvert the immune environment via ESPs is the key to successful infection. Recently, we found that ESPs from adult *E. granulosus* could downregulate host immune responses by preventing dendritic cells (DC) from maturing, by impairing DC function and by inducing the generation of CD4+ CD25+ FoxP3+ T cells (unpublished data). Previous studies have shown that cyst fluids produced in the intermediate hosts can modulate DC differentiation and cytokine secretion [14], while antigen B released by the germinal cells of *E. granulosus* can direct immature DCs towards the maturation of a Th2 cell response [15]. Moreover, the ESPs from *E. multilocularis* larvae have been found to induce apoptosis and tolerogenic properties in DC *in vitro* [16]. To date, studies have focused primarily on the immune regulation of ESPs by the host, with little work undertaken to investigate the influence of ESPs on the physiological adaptation of parasites to their hosts. Interestingly, several intracellular proteins that were not previously thought to be exposed to the immune system of hosts have recently been identified in the ESPs of PSCs [9,17]. This finding suggests that parasite-derived ESPs are incorporated in the metabolites of the host [18,19].

Further investigations into the mechanisms of physiological adaptation of ESPs released by PSC have been hampered due to the paucity of information regarding ESPs. Although studies have utilized proteomics to identify the constituents of ESPs [9,20–22], very few have been identified. This is largely because of interference from host proteins [20–21] and because of technical limitations of the methodologies used. In recent years, however, the combination of transcriptomics and proteomics has enabled the identification of an increasing number of parasitic proteins [23,24].

In this study, we used Roche 454 sequencing technology and *in silico* secretome analysis to explore the transcriptome profiles of *E. granulosus* PSCs and to elucidate the potential functions of the ESPs released by the parasite.

**Materials and Methods**

**Ethics statement**

This study was performed in strict accordance with the recommendations provided in the Guide for the Care and Use of Laboratory Animals of the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention. The protocol was approved by the Laboratory Animal Welfare & Ethics Committee (LAWEC), National Institute of Parasitic Diseases, Chinese Center for Diseases Control and Prevention (Permit Number: IPD 2011-006).

**Sample collection**

Hydatid cysts were collected from the livers of a naturally infected sheep in a slaughterhouse in Qinghai, China. Cyst fluids containing PSCs were sucked out of the cysts using a sterile syringe. After natural sedimentation for 10 min, PSCs were carefully collected from the sediment of cyst fluids and washed 10 times with saline solution. We then added 2 mL of Trizol reagent (Invitrogen, USA) to the well-washed PSCs. After continuous mixing with a pipette, the PSCs were stored at −80°C prior to use.

**Genotyping the PSCs**

Genomic DNA from the PSCs was extracted using the DNeasy tissue kit (Qiagen, Hilden, Germany) and used as a template for a polymerase chain reaction (PCR) [25]. The two primer pairs were used to amplify the mitochondrial genes of *Echinococcus* species: cytochrome oxidase subunit 1 (*cox1*) gene (F: 5′-TTGATTTTGCCAGTTTGATGC-3′; and R: 5′-GAACCTAACCACATAACATACTG-3′); and cytochrome b (*cytb*) gene (F: 5′-GTGAGATGTCTTTGCGCTG-3′; and R: 5′-CTTGGTGACACCACCTAATA-3′). Each 25-μL reaction mixture contained 1 μL of template DNA, 12.5 μL Premix Taq™ mix (TaKaRa Biomedicals, Tokyo, Japan), 1 μL of 10 μM of each primer, and 9.5 μL nuclease-free water. The procedure of PCR amplification consisted of 94°C for 1 min, 30 cycles of 94°C for 30 s, 56°C for 30 s, and 72°C for 1 min, followed by 72°C for 10 min, with a final holding step at 4°C. The PCR products were directly sequenced with a Dye Terminator Cycle Sequencing Kit (Amersham Biosciences, Tokyo, Japan) and ABI 3730 DNA Analyzer (Applied Biosystems, Foster City, USA).

**cDNA library preparation, Roche 454 sequencing and sequence assembly**

The total RNA was extracted from the PSCs in TRIzol reagent, and RNA quality was performed by gel electrophoresis with a 2100 Bioanalyzer (Agilent Technology, Santa Clara, USA). The sequencing protocol followed that described in Liao et al. [26], and was carried out at the Shanghai OE Biotech Company. cDNA was synthesized using 2 μg of total RNA with the SMART cDNA synthesis kit (Clontech Laboratories, Mountain View, USA) according to the manufacturer’s instructions. The cDNA library
was constructed using a GS-FLX Titanium General Library Preparation Kit (Roche, Branford, USA) without normalization [27], and then sequenced using a half run on the Roche 454 GS-FLX Titanium platform. The modules built-in Newbler 2.5.3 (a de novo sequence assembly software, Roche, USA) was used to remove low quality sequences and assemble the remaining sequences. Briefly, the quality score trimming filter trims back from the 3' end of reads and was based on estimated quality scores (not the final quality scores) derived from an internal calibrated signal histogram. The error rate in a sliding window (default size of 40 bp) was calculated from the estimated quality scores and multiplied by an empirical scaling factor (default of 1.1). The window was moved leftwards until the estimated error rate in the window was <1.0% (by default). If the resulting read was less than 40 bp (default), the read was discarded and not counted (numTrimmedTooShortQuality metric). After removing low quality sequences and sequencing adaptors, the remaining sequencing reads were assembled using the Newbler 2.5.3 with the 'extend low depth overlaps' parameter. All of the ESTs from the Roche 454 were used to run the final assembly. The resulting isotig consensus sequences and singletons were referred as 'unigenes' in the following study.

Bioinformatic analyses of transcriptomic sequence data

The software SOAP2 was used to map the raw sequence reads to the nonredundant sequence data [28]. Briefly, raw reads were aligned to the assembled, nonredundant transcriptomic data, to ensure that each read was mapped to a unique transcript. Reads mapped to more than one transcript were randomly assigned to one unique transcript, to ensure that they were recorded only once. Reads per kilobase per million reads (RPKM), the evaluation index of relative assessment of transcript abundance, was calculated using the standard formula [29].

Unigene sequences were compared (using BLASTn with a cutoff E-value of 1e-5) to public sequences available in NCBI nonreducent (Nr) and STRING databases, and to five entire genome sequences (E. multilocularis [30], E. granulosus [31], Schistosoma hematobium [32], S. japonicum [33], S. mansoni [34]).

After conceptual translation from the predicted coding domains of individual transcriptomic sequences, the functions of the potential proteins were predicted using InterProScan [35], employing the default parameters. According to their homology with conserved domains and with protein families, proteins inferred for E. granulosus PSC (EgPSC) were assigned to three gene ontology (GO) categories, including molecular function, cellular component and biological process [36]. The pathway analysis of inferred proteins was carried out using the KEGG (Kyoto Encyclopedia of Genes and Genomes) database [37].

In silico secretome analysis

Excretory-secretory proteins (ESPs) were predicted according to the methods described by Garg and Ranganathan [38,39]. Briefly, the secretory proteins were predicted utilizing the following five tools: ESTScan 3.0.3 [40] to translate the unigenes into putative proteins; SecretomeP 1.0 [41] for non-classical secreted proteins; SignalP 4.1 [42] for classical secreted proteins; TargetP 1.1 [43] for trimming mitochondrial proteins; and TMHMM 2.0 [44] for trimming transmembrane proteins. The predicted proteins with no transmembrane helices were thought to be ESPs.

In addition to traditional computational approaches for ESPs prediction, we also predicted E. granulosus ESPs (EgESPs) using BLASTP [45]. Based on their homology, a list of ESP sequences that included 478 nucleotides and 1,126 proteins was obtained to extract ESPs from the proteins that were predicted to be non-secretory by SecretomeP. Those ESPs had been identified in experiments in other species (S. mansoni, S. japonicum, Brugia malayi, Ankylostoma caninum, Teladorsagia circumcincta, Fasciola hepatica and Clinorchis sinensis) [46–59]. In this approach, a correct match for protein (Query) to protein (Subject) was designated when the query ratio was >80% of their length and identity ≥60, while a correct match for protein (Query) to nucleic acids (Subject) was designated when the query ratio was >80% of their length and identity ≥90.

All potential ESPs were blasted with known ESP sequences from E. granulosus (including nucleotide and protein sequences [9,7,20–22] and our unpublished data) to validate the in silico secretome analysis. They were then annotated against GO, KEGG, Reactome (http://www.reactome.org/ReactomeGWT/entrypoint.html) and Panther (http://www.pantherdb.org/) databases to identify functional groups and pathway annotations. Enrichment of KEGG pathways for genes with significant expression was calculated utilizing a classical hypergeometric distribution statistical comparison of the query gene list against all predicted E. granulosus genes. Caenorhabditis elegans pathways were used as a reference. Calculated P-values were subjected to FDR correction, with p<0.05 taken as the threshold for significance.

Accession number

The transcriptome data is stored in Sequence Read Archive (SRA, No. SRP040541, http://www.ncbi.nlm.nih.gov/sra/?term=SRP040541).

Results/Discussion

Genotyping of E. granulosus PSCs

The genotype of E. granulosus PSCs used in this study was sheep G1, as the PCR fragment amplified from cytb gene showed the highest identity (99%) to the E. granulosus G1 genotype referenced in GenBank (accession AF297617, S1 Figure). This was consistent with the fact that sheep G1 strain is the most common strain worldwide [60].

Roche 454 transcriptome sequencing and reads assembly

A total of 330,188 raw reads (mean length = 411.8 bp) were generated. The data is stored in Sequence Read Archive (SRA, No. SRP040541). After trimming to remove adaptors, low quality reads and polyN tail sequences, 329,927 clean reads remained (mean length = 400.3 bp; Table 1). Clean reads were assembled and produced about 26,514 unigenes ranging in size form 150–3,357 bp (mean = 501.5 bp). These included 4,175 isoforms ranging in size from 154 to 3,357 bp and 22,339 singletons of 150 to 1,710 bp. Approximately 84% of the isoforms were >500 bp, while most singletons (85.97%) were between 300 and 800 bp in size (Table 1, S2 Figure). The numbers of EgPSCs unigenes matching known sequences are listed in Table 1. In summary, 26,514 unigenes were inferred from our transcriptome. The large majority of these (17,861, 67.4%) exhibited the highest level of homology to proteins in E. multilocularis, followed by proteins from E. granulosus (17,732; 66.9%), Caenorhabditis elegans (8,946; 33.7%) and S. mansoni (2,159; 17.5%). Moreover, 22,910 (86.4%) contigs were mapped to the E. granulosus genome and 17,705 (66.8%) of these were distributed within the coding sequence (CDS) region, which suggested that our results were reliable.

Annotation of the transcriptome

Proteins predicted from EgPSCs transcriptome were categorized using Blast2Go [61]. A total of 5,846 were assigned at least
Potential secretome database

PSCs are an important, infectious component of the larval stages of *E. granulosus* that can interact with both definitive and intermediate hosts [10]. The adaptive mechanisms that facilitate this interaction between host and parasite is of great interest to our understanding of the transmission of this widespread disease. Preliminary investigations suggest that parasites secrete certain molecules to assist in host tissue colonization [13]. We therefore focused on the components of ESPs released by PSCs and their potential roles in the physiological adaptation to their hosts and/or themselves.

Of the 26,514 unigenes identified, 19,576 were translated into proteins by ESTScan, 437 proteins were predicted to be classical secreted proteins using SignalP, while 592 were predicted to be non-classical secreted proteins according to SecretomeP. The classical and non-classical proteins were then analyzed using TargetP software for mitochondrial proteins, which resulted in the removal of 25 proteins. A further 123 transmembrane proteins were removed from the secretory protein dataset by TMHMM. In total, we obtained 881 ESPs using the four tools. A further 1,399 proteins that showed a high degree of similarity to experimentally identified secreted proteins were added by the Blast program. Thus, a total of 2,280 proteins were finally predicted as secretory proteins (Table 2).

To validate the *in silico* secretome analysis, we compiled a list of all experimentally identified ESP sequences of *E. granulosus* from the NCBI database and from previous studies (47 nucleotides and 77 proteins) [9,17,20–22], and then blasted the putative ESP sequences with the known ESP sequences (see Table S2). Ninety-one proteins were successfully mapped to the known ES proteins, of which 18 shared 100% identity and 33 shared 95%–99% identity. In addition, most known ESPs from other parasites [62] were matched successfully to those identified in our study. More importantly, domains in ESPs of *Teladorsagia circumcincta* (including metridin-like ShK toxin, lectin, protease inhibitor I29, and allergen V5/Tpx-1) were also found in the ESPs of *EgPSC*, which strengthens the concept that parasites employ universal ESPs to mediate parasite-host interplay [53]. Overall, these data suggest that the ESPs of *EgPSCs* identified in this study were reliable.

To date, there have been five proteomic studies regarding *E. granulosus* that have identified just 157 ESPs among them [9,17,20–22]. In this study, approximately 500 ESP domains were found, including known proteins (Table S3), a result that significantly expands the known ES components of *EgPSCs*. For example, WD40 repeats [63,64], G-protein-coupled receptor (GPCR) [65] and Cadherin [66] all presented novel ESPs that were involved in parasite development-related processes. Recent studies using genome-wide and transcriptome data provide comprehensive information about the growth and development of *E. granulosus* [51,67]. The results of this study extend this

---

Table 1. Summary of the nucleotide sequence data for EgPSCs prior to and following assembly, with detailed bioinformatic annotation and analyses.

| Raw reads | 330188 |
|-----------|--------|
| Unigenes (average length; min-max length) | 26514 (510.5; 150–3357) |
| Containing an open reading frame (%) | 19576 (73.8) |
| With homologues in *E. granulosus* (%) | 17732 (66.9) |
| *E. multilocularis* | 17861 (67.4) |
| *Caenorhabditis elegans* | 8946 (33.7) |
| *Clonorchis sinensis* | 2540 (20.6) |
| *Schistosoma mansoni* | 2159 (17.5) |
| *Schistosoma japonicum* | 1485 (12.1) |
| *Escherichia coli* | 159 (1.3) |
| Returning STRING results (%) | 3188 (12.0) |
| Returning NCBI NR results (%) | 12408 (46.8) |
| Gene Ontology (%) | 5846 (22.0) |
| Number of biological process terms (level 2) | 24 |
| Cellular component | 20 |
| Molecular function | 14 |
| Returning a KOBAS result (%) | 5657 (21.3) |
| Number of predicted biological pathways | 306 |

DOI:10.1371/journal.pntd.0003392.t001
Table 2. Prediction of secretory-excretory proteins (ESP) from the transcriptome of EgPSCs.

| Classification                                      | No. of predicted proteins | Prediction tools |
|-----------------------------------------------------|---------------------------|-----------------|
| Unigene                                             | 26514                     | Newbler         |
| Protein                                             | 19576                     | ESTScan-3.0.3   |
| Classic secreted proteins                           | 437                       | SignaIP 4.1, Web|
| Non-classical secretory proteins                    | 592                       | SecretomeP 1.0  |
| Mitochondrial proteins                              | 25                        | TargetP 1.1, Web|
| Transmembrane proteins                              | 123                       | TMHMM 2.0, Web  |
| Homologues of experimentally verified proteins      | 1399                      | Blast-2.2.27    |
| Total secreted proteins predicted                    | 2280                      |                 |

doi:10.1371/journal.pntd.0003392.t002

information and pave the way to a greater understanding of how PSCs utilize ESPs to survive in hosts.

ES proteins annotation

The putative ESPs were allocated to functional categories based on InterPro domains and GO categories. Of the 2,280 proteins predicted from EgPSCs, the largest functional group represented ‘binding’ (n = 201, GO: 000488), followed by ‘catalytic activity’ (n = 196, GO: 0003824) for ‘molecular function’, ‘metabolic process’ (n = 190, GO: 0008152) and ‘cellular process’ (n = 181, GO: 0009987) for ‘biological process’, and ‘cell part’ (n = 200, GO: 0044464) and ‘cell’ (n = 200, GO: 0005623) for ‘cellular component’ (S5 Figure).

The pathway enrichment analysis for identified ESPs was performed using KOBAS v2.0 software and more than 400 pathways were identified, of which 33 were statistically significant (Table 3). The term for ‘Huntington disease’ represented the most significant group (39, corrected p < 0.0001), followed by Phagosome (37, p < 0.0001), Protein folding (22, p < 0.0001) and Chaperonin-mediated protein folding (16, p < 0.0001).

Of the 2,280 putative ESPs, only 1,406 were mapped to known functions (Table S3). These proteins included not only many common and abundant ‘house-keeping proteins’ (e.g., ribosome proteins, cytochrome subunits, and enzymes involved in carbohydrate and protein metabolism), but also some rare but interesting proteins (e.g., putative receptor and antigenic proteins). This highlights the important roles of ESPs in parasite survival and development within hostile host environments. Below, we characterize these potential ESPs in greater detail.

Metabolism of carbohydrates for parasite energy and nutrition

The interaction of pathogens with mammalian hosts leads to a variety of physiology responses that drive the adaptation of the interacting partners to their new environments and conditions [19]. The ESPs released by parasites might be important actors in this process of adaptation, because they are involved in the metabolism of carbohydrates [68]. We identified a total of 122 domains (summarized in Table S4), of which, 32 proteins were identified to have a higher level of expression in the parasite (Table 4).

E. granulosus has evolved an optimal strategy to gain energy and nutrition from its host using ESPs (Fig. 1). Firstly, the parasite can regulate glycolysis (GL). We identified nine enzymes associated with GL, including the rate-limiting enzymes PFK1 and pyruvate kinase. Through GL, non-essential amino acids (e.g., glutamine, aspartic acid, arginine, proline, histidine, alanine, tyrosine and cysteine), fatty acids, adenine and hypoxanthine nucleotides, as well as pyrimidine, could be synthesized to support parasite development and growth. Alternatively, glucose and other carbohydrates could be synthesized via gluconeogenesis (GN) when alternative carbon sources (e.g., glucogenic amino acids, lactate, and glyceral) were available. In addition to the reversible enzymatic GL steps, several reactions are essential in the GN pathway from pyruvate via oxaloacetate to glucose: the reactions catalyzed by pyruvate carboxylase, phosphoenolpyruvate carboxykinase (PEPCK), fructose-1, 6-biphosphatase, and glucose-6-phosphatase leading to oxaloacetate, phosphoenolpyruvate (PEP), fructose-6-phosphate, and glucose. Finally, tricarboxylic acid (TCA) enzymes, such as aconitate hydratase, succinate dehydrogenase complex, malate dehydrogenase, were identified in the TCA cycle. Other enzymes involved in carbohydrate metabolism are shown in Table 4.

Certain enzymes have been recognized to play key roles in the development of parasites. Phosphoglucone isomerase (PGI), one of glycolytic enzymes, has been found to stimulate parasite growth and the formation of novel blood vessels nearby the developing metacestode [69]. Vaccinating mice with recombinant PGI increases their resistance towards a secondary infection challenge [69]. Similarly, PEPCK is a novel egg antigen of S. mansoni [70] and an abundant protein in adult parasites that is related to numerous metabolic pathways (e.g., endocrine function, excretion and carbohydrate metabolism [22]).

To date, only five ESPs have been identified to participate in this metabolic process [17]. The results of this study support the role of these proteins in metabolic adaptation to their hosts and, more importantly, demonstrate that many more ESPs may be used by E. granulosus to regulate carbohydrate metabolism. Further work is required to identify these additional ESPs and establish their functions.

Control of parasite homeostasis

Following infection with E. granulosus, the intermediate host produces a significant immune response that affects the growth and development of parasites [71,72], while the parasites initiate effective evasion mechanisms to counteract adverse host environments. In this study, we found that 36 ESP domains were molecular chaperones (Table S5), and identified a further 25 proteins that were present with high levels of abundance (Table 4), including several novel molecules (heat shock proteins, HSP90 and HSP40, universal stress protein [Usp], calreticulin, calcineurin B, GrpE in the HSP60 family and Grp96). HSP90 was the most strongly expressed of all the molecular chaperons (Fig. 2), suggesting it is
Table 3. Pathway enrichment analysis of 1406 ESPs in the EgPSCs transcriptomes.

| Category                   | Term                        | Pathway database | Pathway Id | Sample number | Background number | P-Value   | Corrected P-value |
|----------------------------|-----------------------------|------------------|------------|---------------|------------------|-----------|------------------|
| Carbohydrate metabolism    | Pentose phosphate pathway   | KEGG             | ce00030    | 13            | 18               | 9.41E-10  | 4.21E-08         |
|                            | Glycolysis/Gluconeogenesis  | KEGG             | ce00010    | 19            | 40               | 2.62E-08  | 8.80E-07         |
|                            | Gluconeogenesis             | Reactome         | —          | 12            | 19               | 4.87E-08  | 1.51E-06         |
|                            | Glycolysis                  | Reactome         | —          | 7             | 8                | 9.66E-08  | 2.60E-06         |
|                            | Starch and sucrose metabolism | KEGG         | ce00550    | 10            | 26               | 0.000253  | 0.04432          |
|                            | Fructose and mannose metabolism | KEGG         | ce00051    | 8             | 23               | 0.001801  | 0.026766         |
|                            | Amino sugar and nucleotide sugar metabolism | KEGG | ce00520 | 10 | 32 | 0.001974 | 0.0241093 |
|                            | Glucose metabolism          | Reactome         | —          | 18            | 30               | 2.63E-10  | 1.32E-08         |
|                            | Carbon metabolism           | KEGG             | ce101200   | 24            | 78               | 1.50E-05  | 0.0003552        |
|                            | Biosynthesis of amino acids | KEGG             | ce101230   | 17            | 65               | 0.001585  | 0.0212919        |
| Signal transduction        | Heterotrimeric G-protein signaling pathway | PANTHER | P00026    | 7             | 8                | 9.66E-08  | 2.60E-06         |
|                            | Calcium signaling pathway   | KEGG             | ce04020    | 13            | 37               | 0.000157  | 0.0030182        |
|                            | IFN-alpha/beta pathways     | Reactome         | —          | 3             | 5                | 0.001394  | 0.0193684        |
|                            | TGF-beta receptor signaling | Reactome         | —          | 3             | 6                | 0.003739  | 0.0367554        |
|                            | Apoptosis signaling pathway | PANTHER         | P00006     | 7             | 18               | 0.00122   | 0.0182041        |
| Proteins metabolism        | Protein folding             | Reactome         | —          | 22            | 23               | 6.32E-21  | 8.49E-19         |
|                            | Metabolism of proteins      | Reactome         | —          | 64            | 293              | 1.75E-05  | 0.0003918        |
|                            | Mitochondrial protein import | Reactome       | —          | 9             | 22               | 0.000238  | 0.0043566        |
|                            | Chaperonin-mediated protein folding | Reactome | — | 16 | 19 | 1.57E-13 | 1.59E-11 |
|                            | Post-chaperonin tubulin folding pathway | Reactome | — | 8 | 9 | 1.28E-08 | 5.16E-07 |
|                            | Activation of chaperones by ATF6-alpha | Reactome | — | 5 | 7 | 3.41E-05 | 0.000723 |
|                            | Calnexin/calreticulin cycle | Reactome         | —          | 5             | 12               | 0.002474  | 0.0293186        |
| Gene expression             | MicroRNA (miRNA) biogenesis | Reactome         | —          | 6             | 16               | 0.002776  | 0.0310737        |
| Genetic information processing | Spliceosome              | KEGG             | ce03040    | 25            | 103              | 0.000786  | 0.0121906        |
| Transport and catabolism    | Phagosome                  | KEGG             | ce04145    | 37            | 55               | 2.53E-21  | 5.10E-19         |
| Disease pathway             | Huntington disease         | PANTHER         | P00029     | 39            | 41               | 2.31E-34  | 9.29E-32         |
|                            | Parkinson disease          | PANTHER         | P00049     | 15            | 31               | 3.76E-07  | 9.46E-06         |

Others
### Table 3. Cont.

| Category Term | Pathway databasea | Pathway Idc | Sample numberd | Background numbe | P-Valuef | Corrected P-valuef |
|---------------|------------------|-------------|----------------|-----------------|----------|--------------------|
| Cytoskeletal regulation by Rho GTPase | PANTHER | P00016 | 15 | 18 | 1.07E-12 | 7.16E-11 |
| CCT/TriC | Reactome | — | 15 | 18 | 1.07E-12 | 7.16E-11 |
| mRNA splicing - minor pathway | Reactome | — | 13 | 39 | 0.0003 | 0.0050358 |
| N-glycan trimming in ER and CNX/CRT | Reactome | — | 6 | 13 | 0.000599 | 0.0096585 |
| Adenine and hypoxanthine salvage pathway | PANTHER | P02723 | 3 | 6 | 0.003739 | 0.0367554 |

KEGG enrichment analysis was performed by KOBAS 2.0 (http://kobas.cbi.pku.edu.cn/home.do). Caenorhabditis elegans pathways were used as a reference. The ESP corresponding to each pathway can be found in Table S9. Pathway databases mapped by KOBAS including KEGG pathway: http://www.genome.jp/kegg/pathway.html; reactome: http://www.reactome.org/ReactomeGWT/entrypoint.html; PANTHER: http://www.pantherdb.org/. 

aKEGG enrichment analysis was performed by KOBAS 2.0 (http://kobas.cbi.pku.edu.cn/home.do). Caenorhabditis elegans pathways were used as a reference. The ESP corresponding to each pathway can be found in Table S9. bPathway databases mapped by KOBAS including KEGG pathway: http://www.genome.jp/kegg/pathway.html; reactome: http://www.reactome.org/ReactomeGWT/entrypoint.html; PANTHER: http://www.pantherdb.org/. cPathway identified in specific database. “-” means not given. dThe number of input proteins mapped to the particular pathway. eThe number of identified proteins mapped to the particular pathway. fOnly significant results (p < 0.05) were shown. The statistical method was a hypergeometric test, whereas the FDR correction method was from Benjamini and Hochberg (1995). doi:10.1371/journal.pntd.0003392.t003

Transcriptome Profiles of the Protoscoleces of Echinococcus granulosus

Under normal conditions, the tegument of Echinococcus granulosus is highly immunogenic for the host. It is a complex multilayered membrane that extends from the parasite surface and may secrete proteases or inhibitors to digest host proteins, or to protect itself from digestion by endogenous or host-derived proteinases. In this study, 39 proteases, including serine, aspartic, metallo- and cysteine proteinases, and five inhibitors, were inferred among the set of ESPs (see Table S6). Several of these (serine, cysteine, and the proteinase inhibitors) are likely to be important targets for parasite intervention and control [77–79]. However, only three proteases and two protease inhibitors were strongly expressed in the set of ESPs (Table 4). More sensitive technologies will therefore be required to identify other proteinases that were expressed at lower levels of abundance.

In contrast, the action of antioxidant enzymes is a key component of parasite survival during infection. In this study, seven ESPs were identified as antioxidant enzymes, including glutathione transferase, peroxiredoxin, thioredoxin, Ca2+/Zn2+ superoxide dismutase, and neuronal nicotinic synapse protein. These molecules might be utilized by the parasite to detoxify the reactive oxygen species produced by the host environments [80].

### Direct regulation of host immunological responses

In previous experiments we demonstrated that following infection with EgPSCs the microenvironment of the murine peripheral immune system undergoes several changes. These included T cell activation and the accumulation of immunosuppressive cells, such as myeloid-derived suppressor cells (MDSC) and CD4+CD25+FoxP3+ T cells (Treg) [71]. Such alterations might occur via the action of ESPs as many ESPs have been found to redirect host immune responses [13,17]. In this study, we found several ESPs that contribute to immune regulation following infection (Table 4). Tegument protein (Teg) is known to induce a biased Th2 cell immune response related to chronic infection [81], while 14-3-3 proteins are associated with resistance to the immune responses mediated by local cells [82]. In addition, the antigen B (AgB) family are important in immune evasion because the antigen is secreted at variable amounts [83], and have also been demonstrated to direct immature DC maturation towards a preferential Th2 immune response [15].

Notably, cysteine proteinases have been reported to inhibit Th1 immune response via the induction of IL-4, which is the main cytokine responsible for Th2 differentiation [84]. HSP70 has been shown to stimulate both of types of response in CHD patients [85]. Also, the intraperitoneal injection of calreticulin (CRT) significantly influences Th1/Th2 balance [86]. Hence, these proteins might be novel immunoregulatory molecules that contribute to immune evasion.

### Signaling pathways

We found that EgPSC possesses many signaling pathways such as PI3K-Akt, mitogen-activated protein kinase (MAPK), Wnt, calcium, HIF-1, insulin, estrogen and chemokine signaling (Table S1). However, in the putative set of ESPs, only G-protein, calcium,
| GI Number       | Description                           | Species*  | Functionb |
|-----------------|---------------------------------------|-----------|-----------|
| 116242320       | Lysosomal pro-X carboxypeptidase       | C. sinensis |           |
| 111036376       | Cathepsin L-like proteinase            | E. multilocularis | A       |
| 498960202       | Lysosome membrane protein 2-like isom. | M. zebra |           |
| 226478810       | Cytochrome c-type heme lyase           | S. japonicum |           |
| 223037336       | Kunitz protein 8                      | E. granulosus |           |
| 124783098       | Ribosomal protein S18                 | T. asiatica |           |
| 56753617        | Ribosomal protein L21                 | S. japonicum |           |
| 226483022       | Putative small subunit ribosomal protein S27Ae | S. japonicum |           |
| 256074063       | 605 ribosomal protein L9              | S. mansoni |           |
| 392495090       | Ribosomal protein S13                 | S. erinaceieuropaei |           |
| 421975923       | 605 ribosomal protein L7              | S. erinaceieuropaei |           |
| 29841212        | Putative ribosomal protein L27A protein | S. japonicum |           |
| 60692924        | Ribosomal protein                     | S. japonicum |           |
| 421975956       | Putative ribosomal protein S25         | S. erinaceieuropaei |           |
| 358340304       | U1 small nuclear ribonucleoprotein A   | C. sinensis |           |
| 358332789       | Ribosomal RNA-processing protein 9     | C. sinensis |           |
| 256078860       | U3 small nucleolar ribonucleoprotein protein imp4 | S. mansoni |           |
| 55976640        | Actin-1/4 actin                       | T. solium |           |
| 207298859       | Beta-actin                            | A. transmontanus |           |
| 543766          | Actin-1                               | E. granulosus |           |
| 133721998       | Actin                                 | G. viridula |           |
| 29337144        | Tubulin beta-2 chain                  | E. multilocularis |           |
| 29337143        | Tubulin beta-3 chain                  | E. multilocularis |           |
| 29337145        | Tubulin beta-1 chain                  | E. multilocularis |           |
| 401087689       | Tubulin alpha-1C chain-like           | T. rubripes |           |
| 311992220       | Tropomyosin 2 high molecular weight isom. | M. corti | A       |
| 29337029        | Tropomyosin                           | E. multilocularis | A       |
| 168071448       | Tropomyosin B                         | E. granulosus | A       |
| 256086965       | Myosin heavy chain                    | S. mansoni | A       |
| 547974          | Paramyosin                            | E. granulosus | A       |
| 432897369       | Dynein light chain 2, cytoplasmic-like | O. latipes | A       |
| 171473974       | Dynein light chain LC6                | S. japonicum | A       |
| 405970739       | Dynein light chain 2, cytoplasmic      | C. gigas | A       |
| 68071557        | Dynein light chain 1                  | P. berghei | A       |
| 29467010        | Dynein light chain                    | E. multilocularis | A       |
| 226487996       | Nucleolar protein 5                   | S. japonicum |           |
| 226487430       | Myophillin                            | S. japonicum | A       |
| 29336625        | Myophillin                            | E. granulosus | A       |
| 256086246       | Histone H3                            | S. mansoni |           |
| 344240017       | Histone H2A type 1                    | C. griseus |           |
| 358338242       | Histone H2AV                           | C. sinensis |           |
| 405975240       | Histone H2A                           | C. gigas |           |
| 358331974       | PHD finger protein 7                  | C. sinensis |           |
| 343887008       | Heat shock protein 90 alpha            | K. marmoratus | A, D   |
| 1661112         | Heat shock 70 kDa protein, partial    | M. corti | A       |
| 29336623        | Heat shock cognate 70 kDa protein      | E. granulosus | A       |
Table 4. Cont.

| GI Number | Description | Species* | Functionb |
|-----------|-------------|----------|-----------|
| 124783198 | Heat shock protein gp96 | T. asiatica | |
| 124783152 | 40, partial | T. asiatica | A |
| 124783287 | Chaperonin | T. asiatica | |
| 256082744 | T-complex protein 1 epsilon subunit | S. mansoni | |
| 421975972 | T-complex protein 1 subunit alpha | S. erinaceieuropaei | |
| 349934375 | T-complex protein 1 subunit zeta | C. sinensis | |
| 358342604 | Molecular chaperone GrpE | C. sinensis | |
| 318064648 | DnaJ-like protein subfamily b member 11 | I. punctatus | |
| 312065499 | Protein disulfide isomerase | L. loa | |
| 256081230 | Ubiquitin-conjugating enzyme E2r | S. mansoni | |
| 29841024 | 265 proteasome regulatory complex subunit p42A | S. japonicum | |
| 226470558 | Proteasome subunit beta type 4 | S. japonicum | |
| 56754539 | 205 proteasome subunit alpha 8 | S. japonicum | |
| 29336773 | Putative growth regulator 14-3-3 | E. granulosus | A, ST |
| 62178030 | Putative 14-3-3 protein | E. granulosus | A, ST |
| 148613837 | Calreticulin | E. granulosus | A |
| 444792465 | Calcineurin B | E. granulosus | A |
| 353530026 | Calcineurin B | E. granulosus | A |

**Carbohydrate metabolism**

| GI Number | Description | Species* | Functionb |
|-----------|-------------|----------|-----------|
| 167541050 | Phosphoglycerate mutase | C. sinensis | |
| 358333945 | Phosphoglycerate kinase | C. sinensis | |
| 262192839 | Enolase | E. granulosus | A |
| 62178020 | Putative glucose phosphate isomerase | E. granulosus | A |
| 29336626 | 78 kDa glucose-regulated protein, GRP-78 | E. multilocularis | |
| 328789193 | UTP-glucose-1-phosphate uridylyltransferase isoform 1 | A. mellifera | |
| 6016079 | Glyceraldehyde-3-phosphate dehydrogenase | E. multilocularis | A |
| 338827784 | Glucose-6-phosphatase | E. granulosus | |
| 470364276 | UDP-glucose dehydrogenase | C. owczarzaki | |
| 470610058 | Cyclophilin B | T. truncatus | A, S, M |
| 31077167 | Cyclophilin | T. truncatus | A, S, M |
| 358252886 | Dehydrodolichyl diphosphate synthase | C. sinensis | |
| 338827788 | Phosphoenolpyruvate carboxykinase | E. granulosus | A |
| 358334589 | Dolichyl-phosphate beta-glucosyltransferase | C. sinensis | |
| 256090534 | Phosphoglucomutase | S. mansoni | |
| 46406288 | Malate dehydrogenase | E. granulosus | A |
| 29841093 | Citrate synthase | S. japonicum | |
| 29336561 | Fructose-bisphosphate aldolase | E. multilocularis | A |
| 56682906 | Hypoxanthine-guanine phosphoribosyltransferase | S. japonicum | |
| 256082514 | Uridine cytidine kinase I | S. mansoni | |
| 358336324 | Sterol O-acetyltransferase | C. sinensis | |
| 256085769 | Methyltransferase | S. mansoni | |
| 170579277 | Lysyl-tRNA synthetase | B. malayi | |
| 256071828 | Polyadenylate binding protein | S. mansoni | |

**Oxidation/reduction**

| GI Number | Description | Species* | Functionb |
|-----------|-------------|----------|-----------|
| 29337026 | Thioredoxin peroxidase | E. granulosus | A |
| 1004227 | Glutathione transferase | E. multilocularis | A |
| 341616326 | Peroxiredoxin 3 | C. sinensis | A |
| 347948498 | Cu2+/Zn2+ superoxide dismutase (SOD1) | T. solium | A, T |
| 29337032 | Thioredoxin | E. granulosus | A |
IFN-α/β, TGF-β receptor and apoptosis signaling pathways were dominant (Table S7), which indicated their importance in parasite-host interactions and physiological processes.

Notably, we found that G-protein-coupled receptors (GPCRs), TGF-β and insulin signaling pathways might closely associate with the development of EgPSCs. For example, GPCRs can activate the G-proteins located within the cell. They work cooperatively to deliver varied signals, which in turn regulate various physiological processes [87]. However, the exact function of G-protein signaling in parasites remain unclear.

Studies have shown that TGF-β and insulin signaling pathways in C. elegans can trigger an ‘alternative’ developmental pathway, and can regulate and transit the environmental stresses on the first larval stage of the parasite [88,89]. In particular, the disruption of both signaling pathways leads to arrested development in this species [90,91]. Indeed, the TGF-β pathway is speculated to regulate developmental events in parasitic nematodes [92], as molecules involved in the TGF-β pathway have been found in several parasitic nematodes including Brugia pahangi, Brugia malayi and Parastrongyloides trichosuri [93–95]. The role of TGF-signaling in E. granulosus development and growth warrants further investigation. A recent study revealed that host insulin acts as a stimulant for parasite development within the host liver and that E. multilocularis senses the hormones of hosts through an evolutionary-conserved insulin signaling pathway, which demonstrates the importance of insulin signaling for parasite survival [96].

Potential targets for diagnosis and vaccine development

CHD has a global distribution and causes high rates of morbidity and has a high socio-economic burden in several countries [97]. The Eg95 vaccine induces a high antibody titer in sheep and goats, which protects them against CHD [98]. However, due to antigenic variation caused by genotypic diversity [99], the common Eg95 vaccine does not bind the antibodies of all E. granulosus species, which limits its utility. We suggest that the ESPs of EgPSCs are an excellent alternative candidate for a vaccine, as they are easy to prepare and safer for human health.

### Table 4. Cont.

| GI Number       | Description                                         | Species* | Function^{b} |
|-----------------|-----------------------------------------------------|----------|--------------|
| 358340540       | Thioredoxin domain-containing protein 9              | C. sinensis | A            |
| 94556988        | Neuronal nitric oxide synthase protein inhibitor     | T. solium | PI           |
| 256070830       | Peroxidasin                                         | S. mansoni | A            |

| Transports      |                                                      | S. mansoni | A            |
|-----------------|-----------------------------------------------------|------------|--------------|
| 256080958       | Multidrug resistance protein                         | S. mansoni | A            |
| 85701472        | Trans-Golgi network vesicle protein 23A              | M. musculus| A            |
| 226478102       | Secretory carrier-associated membrane protein 2      | S. japonicum| A           |
| 124782903       | Phosphatidylinositol transfer protein alpha           | T. asiatica| A            |
| 358336646       | F-type H+-transporting ATPase subunit c               | C. sinensis| A            |
| 226468748       | Voltage-dependent anion-selective channel protein 2   | S. japonicum| A           |
| 392495096       | Sorting nexin SNX11                                  | S. japonicum| A           |

| Translation     |                                                      | E. granulosus | A            |
|-----------------|-----------------------------------------------------|--------------|--------------|
| 148717323       | Elongation factor 1 alpha                            | E. granulosus| A            |
| 148717331       | Elongation factor 1 alpha                            | E. vogeli   | A            |
| 148717335       | Elongation factor 1 alpha                            | E. shiquicus| A            |
| 159138037       | RNA polymerase II elongation factor                   | C. sinensis | A            |
| 358334689       | Elongation factor 2                                  | C. sinensis | A            |

| Transcription   |                                                      | T. gondii   | C. sinensis |
|-----------------|-----------------------------------------------------|------------|-------------|
| 221509352       | Zinc finger (C3HC4 type) protein                     | T. gondii  | C. sinensis |
| 358332148       | Eukaryotic translation initiation factor, TFIIA     | C. sinensis| A            |

| Energy conversion|                                                  | S. mansoni | A            |
|------------------|----------------------------------------------------|------------|--------------|
| 256077755        | ATP synthase beta subunit                          | S. mansoni | A            |
| 226478810        | Putative cytochrome c-type heme lyase (CCHL)     | S. japonicum| A           |

| RNA Processing   |                                                  | C. sinensis| A            |
|------------------|----------------------------------------------------|------------|--------------|
| 358334450        | ATP-dependent RNA helicase FAL1, partial            | C. sinensis| A            |

| Cell cycle       |                                                  | S. mansoni | A            |
|------------------|----------------------------------------------------|------------|--------------|
| 353230502        | Mitotic phosphoprotein 44                          | S. mansoni | A            |

| Others           |                                                  | E. granulosus | A            |
|------------------|----------------------------------------------------|--------------|--------------|
| 5051948          | Antigen B8/1                                       | E. granulosus| A            |
| 7339849          | Immunogenic protein Ts11                           | T. solium   | A            |

*The full names of species can be seen in Table S8.
{Abbreviations: A, antigenic protein; D, drug gene; ST, signal transduction; S, structural; M, molecular chaperone; T, transporters; PI, protease inhibitor.
doi:10.1371/journal.pntd.0003392.t004
Fig. 1. Schematic diagram showing the carbohydrate metabolic pathways involved in the ESPs of EgPSCs transcriptome (Reference from Eisenreich W et al. [19] with some modifications). Glycolysis (GL, purple arrows) and gluconeogenesis (GN, grass green arrows); pentose-phosphate pathway (PPP, broken purple arrows); tricarboxylic acid cycle (TCA, blue circle) other catabolic reactions that occur in the mitochondrion and in the cytosol (black arrows). Anabolic reactions leading to amino acids, nucleotides, and lipids are indicated by broken thick black arrows. Metabolites are marked in black. Enzymes identified in our study are marked in red, while other enzymes are marked in blue. Abbreviations: HK, hexokinase; PFK, phosphofructokinase; FBP, fructose bisphosphatase; PK, pyruvate kinase; PDH, pyruvate dehydrogenase complex; PCK, PEP-carboxylase; PPI, phosphohexose isomerase; TPI, triose phosphate isomerase; PGK, phosphoglycerate kinase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; ALDOA, fructose-biphosphate aldolase; PGK, phosphoglycerate kinase; PGAM, phosphoglycerate mutase; ENO, enolase; G6Pase, glucose-6-phosphatase; GAPDH, glucose-6-phosphatase dehydrogenase; G6PDH, 6-phosphogluconate dehydrogenase. Gln, Glutamine; Asp, aspartic acid; Arg, arginine; Pro, proline; His, histidine; Ala, alanine; Tyr, tyrosine; Cys, cysteine; Ade, adenine; Hyp, hypoxanthine.

doi:10.1371/journal.pntd.0003392.g001

Fig. 2. The transcription profiling of putative ESPs in EgPSCs transcriptome. The 20 most abundant ESPs encoded in the transcriptome are shown. Abbreviations: RP-S27Ae, putative small subunit ribosomal protein S27Ae; GAPDH, glyceraldehydes-3-phosphate dehydrogenase; ndk, nucleoside diphosphate kinase B-like; ATP2B, Ca^{2+} transporting ATPase plasma membrane; HSP90α, heat shock protein alpha; 26S p42A, 26S proteasome regulatory complex subunit p42A.

doi:10.1371/journal.pntd.0003392.g002
More importantly, the ESPs obtained by in vitro culture have shown a 92.07% protection rate against a high dose of egg infection in sheep (1,000 eggs per sheep) [100].

Using in silico secretome analysis, we identified 44 antigenic proteins present at high abundance in our set of ESPs (Table 4). Of these, elongation factor 1 alpha, antigen B0/1, myophillin, thioredoxin peroxidase, phosphoglycerate mutase, heat shock protein 90a and actins, were the most abundant. In addition, HSP70, enolase, 14-3-3, phosphoglucomutase, malate dehydrogenase, glutathione S-transferase were also present at high abundance in the set of ESPs (Table S8). These abundant proteins hold enormous potential as diagnostic markers or intervention targets. Indeed, malate dehydrogenase (MDH) has been tested for the immunodiagnosis of E. granulosus, while thioredoxin peroxidase (TPx) has been used for the immunodiagnosis of human CHD [101]. Likewise, the 14-3-3 molecule has been demonstrated to be a candidate vaccine against E. granulosus in mice [12], while recombinant GST protein has been used in the diagnosis of echinococcosis [102].

Proteins that are present at lower levels of abundance might also be relevant as diagnostic markers or target molecules for vaccine development. In this study, these include antigen 5 (Ag5), calreticulin, calcineurin B, thioredoxin, phosphoglucomutase, fructose-bisphosphate aldolase and gp96 (Table S8). Many of these have already shown promise for serodiagnostic purposes. For example, Ag5 is a dominant immunogenic and diagnostic antigen of the E. granulosus metacestode in both adults and PSCs [22]. Similarly, calcineurin B has been previously identified as a candidate for a vaccine or drug target [103]. Surprisingly, the E. granulosus-specific protein domain antigen B (EgAgB) family, which are well known as diagnostic targets, were undetectable in this study. This result was consistent with previous observations that little or no AgB is secreted by in vitro cultured PSCs [17,104]. Previous studies have demonstrated that the germinal layer, but not the PSC, contributes to the primary secretion of AgB [17]. Thus, serological examination based on the AgB antibody would not be useful in early-stage PSC infection as only minute amounts of AgB antibody are produced at that time.

There are currently just two methods for the treatment of hydatid disease: surgery and the use of benzimidazoles, both of which give unsatisfactory results. Hence, novel treatment compounds are urgently needed. In this study, we have identified several secretory drug targets for echinococcosis (Table 4, Table S3), including GPCRs, threonine and tyrosine protein kinase and nuclear hormones, which have been the targets of successful new compounds [69]. Cerebral, hepatic, renal, and pleural hydatids are pathogenic to human, which therefore have become the research focus of CHD. Parkinson et al. [2012] first reported genes with features that reflect physiological adaptations of different parasite stages, including PSCs, and revealed abundant long non-protein coding transcripts, upregulated fermentative pathways, candidate apomucins and a set of phyhelminth-specific gene products, which greatly increased the quality and the quantity of the molecular information regarding E. granulosus [67]. The most newly published genome of the parasite also uncovered several key events of the parasites, including the species-specific genes AgB family, bile salt pathways and Cavβ1 gene variation associated with praziquantel sensitivity [31]. Those studies have provided a molecular understanding of the growth and development of E. granulosus. In this study, we focused on the transcriptome of PSCs, which is the only infective component of the larval stages. We present novel and urgently needed information regarding the components of ESPs released by PSCs and their potential roles in the metabolic adaptation of parasites to their hosts. We suggest that intracellular ESPs are essential to the metabolism of carbohydrates within their hosts and that various molecular chaperones with a high level of expression may play a role in resisting harsh host environments. We also reveal a set of antigenic ESPs that show promise as candidates for vaccine development or in the development of serodiagnostic markers. Such findings will encourage more novel strategies for the treatment and control of CHD.

Although the coverage of the transcriptome data in this study was not deep as the genome-wide study [31,67], these findings are novel and hold importance for understanding the mechanisms of parasite metabolic adaptations within their hosts. Overall, this study adds supplementary knowledge regarding the genomics of E. granulosus, and deepens our understanding of host-parasite interactions.

Supporting Information
S1 Figure Genotype identification of E. granulosus. (A) PCR amplification. M, DNA maker; Cytb, 601 bp; Cox1, 885 bp. (B) Sequence alignment of the cytochrome b (cytb) gene. Bases that differ are marked with red boxes.
(TIF)
S2 Figure Length distribution of singletons and isoforms of the EgPSCs transcriptome.
(TIF)
S3 Figure Gene ontology (GO) analysis of the EgPSCs transcriptome. BLASTP against SwissProt and GO mapping of identified proteins (performed with BLAST2GO) [61].
(TIF)
S4 Figure Distribution of the KOG functional categories of the proteins identified from the EgPSCs transcriptome. Percentages and numbers of proteins in each functional category are indicated in the sectors of the circle. KOG functional categories: (A) RNA processing and modification; (B) Chromatin structure and dynamics; (C) Energy production and conversion; (D) Cell cycle control, cell division, chromosome partitioning; (E) Amino acid transport and metabolism; (F) Nucleotide transport and metabolism; (G) Carbohydrate transport and metabolism; (H) Coenzyme transport and metabolism; (I) Lysosome transport and metabolism; (J) Translation, ribosomal structure and biogenesis; (K) Transcription; (L) Replication, recombination and repair; (M) Cell wall/membrane/envelope biogenesis; (N) Cell motility; (O) Posttranslational modification, protein turnover, chaperones; (P) Inorganic ion transport and metabolism; (Q) Secondary metabolites biosynthesis, transport and catabolism; (R) General function prediction only; (S) Function unknown; (T) Signal transduction mechanisms; (U) Intracellular trafficking, secretion, and vesicular transport; (V) Defense mechanisms; (W) Extracellular structures; (X) Nuclear structure; (Z) Cytoskeleton. The number of proteins in the graphic might exceed the total of predicted ESP because some were grouped in more than one functional category.
(TIF)
Domain analysis of ESPs was then carried out using InterProScan.

S3 Table Domains associated with carbohydrate metabolism in the ESP.

S4 Table Domains associated with carbohydrate metabolism in the ESP.

References

1. Carmona D, Sánchez-Serrano LP, Barbero-Martínez I (2008) Echinococcus granulosus infection in Spain. Zoonoses Public Health 55: 156–165.
2. Battrell G (2009) Echinococcosis: costs, losses and social consequences of a neglected zoonosis. Vet Res Commun 33: 47–52.
3. Moro PL, Gilman RH, Verastegui M, Bern C, Silva B, et al. (1999) Human tylosisidosis in the central Andes of Peru: evolution of the disease over 3 years. Clin Infect Dis 29: 807–812.
4. McManus DP, Zhang W, Li J, Bartley PB (2003) Echinococcosis. Lancet 362: 1295–1304.
5. Craig PS, McManus DP, Lightowlers MW, Chabalgoity JA, Garcia HH, et al. (2007) Prevention and control of cystic echinococcosis. Lancet Infect Dis 7: 385–394.
6. Li J, Isio A, Penguio R, Sako Y, Chen X, et al. (2011) Post-treatment follow-up study of abdominal cystic echinococcosis in uibetan communities of northwest Sichuan Province, China. PLoS Negl Trop Dis 5: e1364.
7. Moro PL, Gilman RH, Verastegui M, Bern C, Silva B, et al. (1999) Human tylosisidosis in the central Andes of Peru: evolution of the disease over 3 years. Clin Infect Dis 29: 807–812.
8. McManus DP, Gray DJ, Zhang W, Yang Y (2012) Diagnosis, treatment, and management of echinococcosis. BMJ 344: e3066.
9. Monteiro KM, de Carvalho MO, Zaha A, Ferreira HB (2010) Proteomic analysis of the Echinococcus granulosus metacestode during infection of its intermediate host. Proteomics, 10: 2085–1099.
10. Thompson RCA, L symmetry AJ. (1995) Echinococcus and Hydatid Disease. CAB International, Wallingford, 1–50.
11. Prenel PE, Castro S, Allemendi D, Bruni SS, Palma SD, et al. (2014) Enhanced echinococcal and clinical efficacy of albendazole formulated as solid dispersions in experimental cystic echinococcosis. Vet Parasitol 203:30–96.
12. Li ZJ, Wang YN, Wang Q, Zhao W (2012) Echinococcus granulosus 14-3-3 protein: a potential vaccine candidate against challenge with Echinococcus granulosus in mice. Biomed Environ Sci 25: 352–358.
13. Dzik JM (2006) Molecules released by helmith parasites involved in host colonization. Acta Biochim Pol 53: 33–64.
14. Kannan BH, Chain BM (2006) Modulation of dendritic cell differentiation and cytokine secretion by the hydatid cyst fluid of Echinococcus granulosus. Immunology 118: 271–278.
15. Rijan R, Buttari B, Profumo E, Ortona E, Delunardo F, et al. (2007) Echinococcus granulosus antigen B impairs human dendritic cell differentiation and polarizes immature dendritic cell maturation towards a Th2 cell response. Infect Immun 75: 1667–1678.
16. Nono JK, Pletinckx K, Lutz MB, Brehm K (2012) Excretory/Secretory-products of Echinococcus multilocularis larvae induce apoptosis and tologenic properties in dendritic cells in vitro. PLoS Negl Trop Dis 6: 1516–1529.
17. Virginia VG, Monteiro KM, Drumond F, de Carvalho MO, Vargas DM, et al. (2012) Excretory/secretory products from a vitro-cultured Echinococcus granulosus protoscoleces. Mol Biochem Parasitol 201: 15–22.
18. Roseneitz MC, Camicia F, Kamenetzky L, Muzalin PM, Gutierrez AM (2006) Identification and intra-specific variability analysis of secreted and membrane-bound proteins from Echinococcus granulosus. Parasitol Res 98: S63–S7.
19. El-Ebriri W, Hesseman J, Rudel T, Goebel W (2013) Metabolic host responses to infection by intracellular bacterial pathogens. Front Cell Microbiol 3: 24.
20. Aziz A, Zhang W, Li J, Loukas A, McManus DP, et al. (2011) Proteomic characterization of Echinococcus granulosus hydatid cyst fluid from sheep, cattle and humans. J Proteomics 74: 1560–1572.
21. Chemale G, van Rossum AJ, Jefferies JR, Barrett J, Brophy PM, et al. (2003) Proteomic analysis of the larval stage of the parasite Echinococcus granulosus: the causative agent of cystic hydatid disease. Proteomics 3: 1633–1636.
22. Cui SJ, Xu LL, Zhang T, Xu M, Yao J, et al. (2013) Proteomic characterization of larval and adult developmental stages in Echinococcus granulosus reveals novel insight into host-parasite interactions. J Proteom 84: S5–S7.
23. Moreto Y, Goss PP, Tam M, Segura M, Vilañparamú R, et al. (2011) Proteomic analysis of excretory-secretory products of Heligmosomoides polygyrus assessed with next-generation sequencing transcriptomic information. PLoS Negl Trop Dis 5: e1370.
24. Cacace MS, Mulholland J, Young ND, Kasmy M, Horak P, Aziz A, Hofmann A, Loukas A, Gasser RB. A deep exploration of the transcriptome and ‘excretory/secretory’ proteome of adult Fasciiculatus magna. Mol Protom 2012, 11(1): 1340–1353.
25. Xiao N, Qu J, Nakao M, Nakaya K, Yamashiki A, et al. (2007) Short report: Identification of Echinococcus species from a yak in the Qinghai-Tibet plateau region of China. Am J Trop Med Hyg 69: 445–446.
26. Luo X, Cheng L, Xu P, Li G, Walch A, et al. (2013) Transcriptomic analysis of crucian carp (Carassius auratus), an important aquaculture and hypoxia-tolerant species. PLoS One 8: e62308.
27. Stefanni S, Bettencourt R, Pinheiro M, Moro GD, Bonjirni L, et al. (2014) Transcriptome of the Deep-Sea Black Scabbardfish, Aphanopus carbo (Periformes: Trichiuridae): Tissue-specific expression patterns and candidate genes associated to depth adaptation. Int J Genomics 2014: 267492.
28. Li R, Yu C, Li Y, Lam TW, Yu SM, et al. (2009) SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics 25: 1966–1967.
29. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5: 621–628.
30. Tsai IJ, Zarowiecki M, Holroyd N, Giacarriugio A, Sanchez-Flores A, et al. (2013) The genomes of four tapeworm species reveal adaptations to parasitism. Nature 496: 57–63.
31. Zheng H, Zhang W, Zhang L, Zhang Z, Li J, et al. (2013) The genome of the hydatid tapeworm Echinococcus granulosus. Nat Genet 45: 1160–1175.
32. Young ND, Jex A R, Li B, Liu S, Yang L, et al. (2012) Whole-genome sequence of Schistosoma haematobium. Nat Genet 44: 221–225.
33. Zhou Yan, Zheng Huajun, Chen Xiangyi, Zhang Lei, Wang Kai, et al The Genome of Schistosoma japonicum Genome Sequencing and Functional Analysis Consor-
Vercauteren I, Geldhof P, Peelaers I, Claerebout E, Berx G, et al. (2003) Proteomic profile of the adult excretory-secretory products from protoscoleces of Echinococcus granulosus: its biological and biotechnological implications. BMC Genomics 13: S10.

Hewitson JP, Harcus YM, Curwen RS, Dowle AA, Atmadja AK, et al. (2008) Proteomic analysis of the excretory-secretory products of Schistosoma haematobium cercarial secretions. Mol Cell Proteomics 8: 1236–1251.

Huang J, Ding Y, Wu J, et al. (2011) Novel matrix metalloproteinase from the parasitic nematode Echinococcus granulosus is involved in metacestode growth and parasite-host cell interactions. Int J Parasitol 40: 1563–1574.

Kovaleva ES, Masler EP, Skantar AM, Chitwood DJ (2004) Novel matrix metalloproteinase from the parasitic nematode Echinococcus granulosus for parasite biology and host adaptation. PLoS Negl Trop Dis 6: e1897.

Kraeuter I, Bajpai S, Smith SK, Nisbet AJ, Meikle LI, Inglis NF, Sales J, et al. (2009) Proteomic analysis of excretory/secretory products released by Haemonchus contortus in vitro. Int J Parasitol 39: 1693–1701.

Liu F, Cui SJ, Hu W, Feng Z, Wang QZ, et al. (2009) Excretory/secretory proteome of the adult development stage of human blood fluke, Schistosoma japonicum. Mol Cell Proteomics 8: 1295–1304.

Liu W, Joo HK, Lee MS, Cho SH, Cheon HI, et al. (2009) Identification of a serodiagnostic antigen, legumain, by immunoproteomic analysis of excretory-secretory products of Clonorchis sinensis adult worms. Proteomics 9: 3066–3075.

Momeni R, Gasser RB, Mitreva M, Ranganathan S (2012) An analysis of the transcriptome of Teladorsagia circumcincta: its biological and biotechnological implications. BMC Genomics 13: S10.

Monge R, Gasser RB, Mitreva M, Ranganathan S (2012) The secretome of the filarial parasite, Brugia malayi. Proteomic profile of the adult excretory-secretory products. Mol Biochem Parasitol 160: 8–21.

Morgan WR, Hayes RS, Ashton JD, Sundaramoorthy S, Wilson RA (2006) Identification of novel proteases and immunomodulators in the secretions of schistosome cercaria that facilitate host entry. Mol Cell Proteomics 5: 835–844.

Nishida S, Takeda H, Takahashi T, Tani M, et al. (2009) Identification of excretory-secretory products from the adult development stage of the blood fluke, Schistosoma mansoni. Mol Cell Proteomics 8: 1256–1261.

Ortona E, Margutti P, Delunardo F, Nobili V, Profumo E, et al. (2005) Identification of secreted cysteine proteases from the parasitic nematode Hymenolepis diminuta detected by biotinylated inhibitors. Infect Immun 74: 1919–1923.

Park J, Zhou HJ, Shen Y, Wu Y, Wang YX, et al. (2013) Surveillance on the status of immune cells after Echinococcus granulosus protoscoleces infection in Balb/c mice. PLoS One 8: e69746.

Prasad WW, Wu M, McManus DP (2006) Recent advances in the immunology and serodiagnosis of echinococcosis. Vet Parasitol 141: 1–10.

Price CS, Field J, Porter K, Lobo J, Ditonello J, et al. (2011) Universal stress proteins are important for oxidative and acid stress resistance and growth of Listeria monocytogenes EGD-e in vitro and in vivo. PLoS One 6: e24963.

Reed RS, Atkinson JP, Needham DL, Winstanley PA, Harris JH, et al. (2006) An analysis of the transcriptome of the blood fluke Schistosoma mansoni. Nucleic Acids Res 34: 87. 10.

Roe CA, Mansfield SJ, Garg G, Ranganathan S (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. BMC Bioinformatics 12: S14.

Sugita Y, Yoshida H, Nakamura S, Fujita Y, et al. (2012) Identification of a novel egg antigen: immunological properties of the recombinant protein and identification of a T-cell epitope. Infect Immun 80: 3305–3309.

Suzuki Y, Zhou HJ, Shen Y, Wu Y, Wang YX, et al. (2013) Surveillance on the status of immune cells after Echinococcus granulosus protoscoleces infection in Balb/c mice. PLoS One 8: e69746.
98. Patterson GI, Padgett RW (2000) TGF beta-related pathways. Roles in Caenorhabditis elegans development. Trends Genet 16: 27–33.
99. Beall MJ, Pearce EJ (2002) Transforming growth factor-beta and insulin-like signaling pathways in parasitic helminthes. Int J Parasitol 32: 399–404.
100. Ren P, Lim CS, Johnson R, Albert PS, Pilgrim D, et al. (1996) Control of C. elegans larval development by neuronal expression of a TGF-beta homolog. Science 274: 1389–1391.
101. Sze JY, Victor M, Loer C, Shi Y, Ruvkun G (2000) Food and metallic signaling defects in a Caenorhabditis elegans serotonin-synthesis mutant. Nature 403: 560–564.
102. Nisbet AJ, Redmond DL, Matews JB, Watkins C, Yaga R, et al. (2008) Stage-specific gene expression in Teladorsagia circumcincta (Nematode: Strongylida) infective larvae and early parasitic stages. Int J Parasitol 38: 829–838.
103. Gomez-Escobar N, van den Biggelaar A, Maizels R (1997) A member of the TGF-beta receptor gene family in the parasitic nematode Brugia pahangi. Gene 199: 101–109.
104. Gomez-Escobar N, Gregory WF, Maizels RM (2000) Identification of tgh-2, a filarial nematode homolog of Caenorhabditis elegans daf-7 and human transforming growth factor beta, expressed in microfilarial and adult stages of Brugia malayi. Insect Immunol 68: 6402–6410.
105. Crook M, Thompson FJ, Grant WN, Viney ME (2005) Daf-7 and the development of Strongylus ratti and Parastrongylus triquarius. Mol Biochem Parasitol 139: 213–223.
106. Hemer S, Konrad C, Spiroto M, Koziol U, Schaack D, et al. (2014) Host insulin stimulates Echinococcus multilocularis insulin signaling pathways and larval development. BMC Biol 12: 5.
107. Carmona D, Cardona GA (2013) Canine echinococcosis: global epidemiology and genotypic diversity. Acta Trop 128: 441–460.
108. Heath DD, Robinson C, Shakes T, Huang Y, Gulnur T, et al. (2012) Vaccination of bovines against Echinococcus granulosus (cystic echinococcosis). Vaccine 30: 3076–3081.
109. Alvarez-Rojas CA, Gauci CG, Lightowlers MW (2013) Antigenic differences between the EG95-related proteins from Echinococcus granulosus G1 and G6 genotypes: implications for vaccination. Parasite Immunol 35: 99–102.
110. Zhu XQ, Dou LQ, Sun XQ, Shi XH, Wang XH, et al. (1991) 21: 8–10. Studies on the secretory/excretory proteins of Echinococcus granulosus—the immunogenicity of protoscoleex secretory/excretory proteins. Chin J Vet Sci Tech 21: 8–10.
111. Marjutti P, Ortona E, Delunardo F, Tagliani A, Profumo E, et al. (2008) Thioredoxin peroxidase from Echinococcus granulosus: a candidate to extend the antigenic panel for the immunodiagnosis of human cystic echinococcosis. Diagn Microbiol Infect Dis 60: 279–283.
112. Hariri A, Garcia G, Arbidhi P, Pascoich L, Chalar C, et al. (2010) Biochemical analysis of a recombinant glutathione transferase from the cestode Echinococcus granulosus. Acta Trop 114: 31–36.
113. Hu M, Su Z, Yin Y, Li J, Wei Q (2012) Calcineurin B subunit triggers innate immunity and acts as a novel Engerix-B HBV vaccine adjuvant. Vaccine 30: 4719–4727.
114. Zhang W, Li J, Jones MK, Zhang Z, Zhao L, et al. (2010) The Echinococcus granulosus antigen B gene family comprises at least 10 unique genes in five subclasses which are differentially expressed. PLoS Negl Trop Dis 4: e784.