



Proteomic analysis of the *Theileria annulata* schizont

M. Witschi^{a,1,2}, D. Xia^{b,1}, S. Sanderson^b, M. Baumgartner^{a,3}, J.M. Wastling^b, D.A.E. Dobbelaere^{a,*}

^aDivision of Molecular Pathobiology, DCR-VPH, Vetsuisse Faculty, University of Bern, CH-3012 Bern, Switzerland

^bDepartment of Infection Biology, Institute of Infection and Global Health & School of Veterinary Science, University of Liverpool, Liverpool L69 7ZJ, UK

ARTICLE INFO

Article history:

Received 4 July 2012

Received in revised form 11 October 2012

Accepted 12 October 2012

Available online 23 November 2012

This publication is dedicated to Dr. Roger Hall (4.10.2012).

Keywords:

Apicomplexa

Detergent

Triton X-114

Mass spectrometry

Proteome

Theileria

Transformation

Transmembrane domain

ABSTRACT

The apicomplexan parasite, *Theileria annulata*, is the causative agent of tropical theileriosis, a devastating lymphoproliferative disease of cattle. The schizont stage transforms bovine leukocytes and provides an intriguing model to study host/pathogen interactions. The genome of *T. annulata* has been sequenced and transcriptomic data are rapidly accumulating. In contrast, little is known about the proteome of the schizont, the pathogenic, transforming life cycle stage of the parasite. Using one-dimensional (1-D) gel LC-MS/MS, a proteomic analysis of purified *T. annulata* schizonts was carried out. In whole parasite lysates, 645 proteins were identified. Proteins with transmembrane domains (TMDs) were under-represented and no proteins with more than four TMDs could be detected. To tackle this problem, Triton X-114 treatment was applied, which facilitates the extraction of membrane proteins, followed by 1-D gel LC-MS/MS. This resulted in the identification of an additional 153 proteins. Half of those had one or more TMD and 30 proteins with more than four TMDs were identified. This demonstrates that Triton X-114 treatment can provide a valuable additional tool for the identification of new membrane proteins in proteomic studies. With two exceptions, all proteins involved in glycolysis and the citric acid cycle were identified. For at least 29% of identified proteins, the corresponding transcripts were not present in the existing expressed sequence tag databases. The proteomics data were integrated into the publicly accessible database resource at EuPathDB (www.eupathdb.org) so that mass spectrometry-based protein expression evidence for *T. annulata* can be queried alongside transcriptional and other genomics data available for these parasites.

© 2012 Australian Society for Parasitology Inc. Published by Elsevier Ltd. Open access under [CC BY license](http://creativecommons.org/licenses/by/3.0/).

1. Introduction

Theileria parasites belong to the Apicomplexa, are transmitted by ticks and cause diseases with significant economic impact in a range of domestic livestock, including large and small ruminants (Bishop et al., 2009). Unique among eukaryotic organisms, several members of the genus *Theileria* possess the ability to transform the cells they infect. Transformation is achieved by interfering with pivotal host cell signalling pathways that regulate proliferation and cell survival (reviewed in Dobbelaere and Baumgartner, 2009; Chaussepied and Langsley, 2011). This is facilitated by the fact that, immediately following entry into the leukocyte, the host cell membrane surrounding the invading sporozoite is dissolved, leaving the developing schizont free in the cytoplasm where it can interact with the host cell cytoskeleton and regulatory compo-

nents of signalling pathways. Analysis of leukocyte gene expression networks provided evidence that the parasite establishes tight control over pathways associated with cellular activation by modulating reception of extrinsic stimuli and by significantly altering the expression outcome of genes targeted by infection-activated transcription factors (Durrani et al., 2012). Furthermore, at each host cell mitosis and cytokinesis, the schizont co-opts the host's astral and central spindle microtubules, ensuring its equal distribution between the two daughter cells (Hulliger et al., 1964; von Schubert et al., 2010). Although significant progress has been made in identifying the host cell pathways that are directly or indirectly targeted by the parasite, little progress has been made in pinning down the parasite proteins that are involved in host cell transformation. This is largely due to the fact that knowledge on the spectrum of *Theileria* schizont proteins that could function as participants in the transformation process is very scarce. As an important step in that direction, the genomes of two transforming *Theileria* parasites, *Theileria parva* and *Theileria annulata*, have been sequenced and annotated, providing a wealth of new information on these organisms (Gardner et al., 2005; Pain et al., 2005; Shiels et al., 2006). Analysis of the *T. parva* schizont transcriptome using Massive Parallel Signature Sequencing (MPSS) provided evidence for the transcription of 2,533 of the 4,036

* Corresponding author. Tel.: +41 31 631 2625; fax: +41 31 631 2658.

E-mail address: dirk.dobbelaere@vetsuisse.unibe.ch (D.A.E. Dobbelaere).

¹ These authors contributed equally to this work.

² Present address: Novartis Campus, WSJ-153.2.12.2, CH-4056 Basel, Switzerland.

³ Present address: University of Zürich, Department of Oncology, Neuro-Oncology Group, University Children's Hospital, August Forel Strasse 1/7, CH-8008 Zürich, Switzerland.

predicted protein coding genes of *T. parva* (Bishop et al., 2005; Shah et al., 2006), 405 of which encoded proteins with a predicted signal peptide, suggesting they could be secreted. Of 3,794 *T. annulata* genes analysed, 628 were predicted to encode proteins with a signal peptide (B. Shiels and W. Weir, personal communication). A more recent study also reported on the evolution and diversity of *T. annulata* secretome genes (Weir et al., 2010) and Shiels et al. (2006) generated a list of *Theileria* proteins containing a predicted signal peptide and which could be involved in manipulating the host cell phenotype. As is the case for many parasites, the majority of the predicted *Theileria* proteins (75%) are classified as hypothetical proteins that show no identity to known proteins in public databases (Gardner et al., 2005; Pain et al., 2005). In the case of *T. annulata*, an expressed sequence tag (EST) library prepared from mRNA isolated from purified schizonts has been generated and 1,407 *T. annulata* genes (representing 37% of predicted genes) could be identified that showed cDNA hits when screened in silico against this library (Pain et al., 2005). As a first step towards characterising the *T. annulata* schizont proteome, a one-dimensional (1-D) gel-based analysis of schizont proteins followed by MS was carried out, using purified parasites as starting material. This first analysis underpins the notion that proteomic analysis is an indispensable tool to study the biology of *Theileria*/host cell interactions.

2. Materials and methods

2.1. Parasite strain, culture conditions and parasite purification

Theileria annulata – infected macrophages (TaC12, strain Ankara; Shiels et al., 1992) were cultured in Leibovitz 15 medium (Gibco, Switzerland) supplemented with 10% FCS (Bioconcept, Switzerland), 10 mM HEPES, pH 7.2 (Merck, Switzerland), 2 mM L-glutamine (Gibco), 70 μ M β -mercaptoethanol (Merck, Switzerland), 100 μ g/ml of penicillin and 100 μ g/ml of streptomycin (Lonza, Switzerland). Schizonts were purified as previously described (Baumgartner et al., 1999). Briefly, TaC12 cells were incubated for 2 h with nocodazole to depolymerize microtubules. Cells were then treated with trypsin-activated aerolysin on ice. After removing excess aerolysin, cells were exposed to a temperature of 37 °C to stimulate toxin-mediated permeabilization of the host cell plasma membrane. Permeabilization was monitored using Trypan blue exclusion. Schizonts were separated from host cell debris using Percoll gradient centrifugation.

2.2. Generation of whole parasite lysates (WPL), protein digestion and MS

Purified parasites were solubilised in 40 μ l of Lämmli buffer with four cycles of boiling at 95 °C and vortexing for 5 min. Samples were centrifuged at 12,000g for 5 min. The supernatant was run on a 16 cm 12% v/v SDS–polyacrylamide gel at 16 mA for the stacking gel (for ~1 h) and at 24 mA for the separating gel (for ~6 h). Gels were fixed in 40% v/v ethanol, 10% acetic acid overnight at room temperature (RT), rinsed in distilled deionised water and stained with colloidal Coomassie (20% methanol, 0.1% W/v Coomassie Brilliant Blue G250, 1% v/v H₃PO₄ (85%), 10% w/v (NH₄)₂SO₄ and destained in distilled deionised water. Lanes were cut, using a scalpel, into continuous slices of approximately 1 mm thickness and stored in 1% acetic acid at 20 °C. Gel plugs were fully destained in 50 mM NH₄HCO₃, 50% v/v acetonitrile at 37 °C, and incubated at 37 °C with 10 mM DTT, 100 mM NH₄HCO₃ for 30 min and subsequently with 100 mM iodoacetamide, 100 mM NH₄HCO₃ for 1 h in the dark. Gel pieces were dehydrated with acetonitrile and proteolytic in-gel digestion performed at 37 °C by adding 10 μ l of

10 ng/ml of sequencing grade trypsin (in 25 mM NH₄HCO₃). After 45 min, 25 mM NH₄HCO₃ was added to cover the gel pieces and the incubation was carried out overnight at 37 °C. Samples were stored at –20 °C until used for LC-MS/MS.

As the aim of this study was to discover and validate *T. annulata* schizont proteins, rather than comparing proteins or studying differential expression, biological replicates were not produced. On this particular MS platform, running replicates to increase the number of identifications typically results in improvements of no more than 10% and often considerably less. Instead, we focussed on improving the isolation and identification of membrane proteins (see Sections 2.3 and 2.4).

2.3. Triton X-114 extraction of schizont proteins

Purified *T. annulata* schizonts were lysed and homogenised for 1 h at 4 °C in 1 ml of lysis buffer (10 mM Tris–HCl, pH 7.4, 150 mM NaCl, 2% Triton X-114, 1 \times Protease Inhibitor Cocktail). Lysed parasites were centrifuged at 8,800g for 10 min at 0 °C to remove cell debris. The pellet (P) was solubilized in Lämmli buffer. The supernatant was incubated at 37 °C for 10 min to induce phase separation and centrifuged at 3,000g for 3 min. The upper phase (aqueous phase, AP1) was removed and stored on ice. The lower phase (detergent-rich phase, DP) was mixed with 1 ml of buffer A (10 mM Tris–HCl, pH 7.4, 150 mM NaCl, 0.06% Triton X-114, 1 \times Protease Inhibitor Cocktail) incubated at 10 °C for 10 min and subjected to a new phase separation at 37 °C (Supplementary Fig. 1A). The aqueous phase (AP2) was removed and the extraction was repeated once more with 1 ml of buffer A. The three aqueous phases (AP1–3) and the final detergent-rich phase were precipitated by adding 3 vol. of cold acetone at –20 °C for 30 min and then centrifuged at 12,000g for 20 min. Pellets were dried on the bench and solubilized in 20 μ l of Lämmli buffer at 70 °C for 5 min. Samples were run approximately 2 cm into a 12% standard size SDS–polyacrylamide gel (termed short gel). The experiment was repeated with small modifications. Proteins from the detergent-rich phase were solubilized in Lämmli buffer at 37 °C for 15 min with short cycles of vortexing every 5 min. The gel was run at approximately 4 cm to obtain a better protein separation (termed long gel). In Supplementary Table S1 which lists all identified proteins, AP1 refers to the aqueous phase after the first phase separation and DP1 and DP2 refer to the final detergent phase obtained in the first and second experiment, respectively.

2.4. Identification and bioinformatic analysis of proteins

2.4.1. LC-MS/MS analysis

Frozen samples were thawed at RT and centrifuged at 20,000g for 25 min. Thirteen microlitres of supernatant was transferred to a 96 well plate and mixed with 3 μ l of 2.6 M formic acid to reduce evaporation.

The peptide mixtures were then analysed on a LC-MS/MS platform which contained an LTQ ion-trap mass spectrometer (Thermo-Electron, Hemel Hempstead, UK) coupled on-line to a Dionex Ultimate 3000 (Dionex Company, Amsterdam, The Netherlands) HPLC system as previously described (Xia et al., 2008). The data associated with this manuscript may be downloaded from ProteomeCommons.org (<https://proteomecommons.org/data-search.jsp>, Tranche hash:

3idGhZ+gXOGy4LXC9SvjqQShoEJCAY9HcfBWsYFKJmCma05n+6t97jj130XsJybg0Xwn4GSj3GUdS+2wWQkJucvojF4AAAAAAA7Ew==.

2.4.2. Raw data transformation

The resulting MS/MS spectra were submitted to Mascot (Matrix Science, (Perkins et al., 1999)) and searched against *T. annulata* (*Theileria_protos*) and *Bos taurus* (Bovine_IPI) annotated proteins.

Fixed Mascot search parameters were: carbamidomethyl modification of cysteine (C); variable oxidation of methionine (M); peptide tolerance ± 1.5 Da; MS/MS tolerance ± 0.8 Da; +1, +2, +3 peptide charge-state; single missed trypsin cleavage, decoy: yes; report top: auto hits. Proteins with a score below 50 were removed from the list in order to avoid false positive hits. The identified proteins were compared with macroschizont, merozoite and piroplasm EST information on GeneDB (Pain et al., 2005). Lists of all macroschizont, merozoite and piroplasm-expressed genes were obtained with the keyword search for a specific stage.

2.4.3. Transmembrane domain (TMD), signal peptide/anchor and GPI-anchor signal prediction

TMD, signal peptide/anchor and glycosylphosphatidylinositol (GPI)-anchor predictions are annotated for all proteins on GeneDB. TMDs were predicted by TMHMM2.0 (Krogh et al., 2001), signal peptide/anchor by SignalP 2.0 HMM (Nielsen et al., 1997a,b) and GPI-anchor signal by DGPI v2.04. Lists of all proteins with a prediction were downloaded with the Complex/Boolean Query function on GeneDB.

All identified proteins were assigned after the MIPS FunCat (Functional Catalogue) (Ruepp et al., 2004). If possible, proteins were categorised according to the biological process annotation on GeneDB. If no biological process was annotated on GeneDB, but a molecular function was available, the corresponding Gene Ontology (GO) number of the function was sent to MIPS FunCat. If neither a biological process nor a molecular function was available on GeneDB, proteins were classified with the help of annotated motifs/domains (Pfam, InterPro), protein names or homologues. *Theileria annulata* protein sequences were sent to NCBI BLAST to identify homologous proteins. GO annotations from homologues were only trusted if the *E*-value was lower than $1e^{-25}$.

2.4.4. Classification of subcellular localisation

Very few proteins have a predicted subcellular localisation annotated on GeneDB. Therefore all proteins were classified with the help of bioinformatic tools according to a putative localisation. If possible, proteins were classified according to their FunCat annotation since some processes/categories are localisation-specific. Unclassified proteins were then categorised with help from the protein name, homologues and motifs/domains. For the rest of the proteins, prediction programs such as WoLF PSORT (Horton and Nakai, 1999), TargetP (Emanuelsson et al., 2000), Plasmit (Bender et al., 2003), PlasmoAP (Foth et al., 2003) and Pats (Zuegge et al., 2001) were used. To avoid incorrect annotations, all proteins without a clear result were added to the category: not classified/multiple localisation.

3. Results and discussion

3.1. Analysis of whole parasite lysate

The analysis of WPL by 1-D gel LC-MS/MS resulted in 645 non-redundant proteins, with a score of 50 or higher, representing 17% of the 3,792 proteins predicted in the proteome of all life cycle stages of *T. annulata*. A list of all proteins identified in this study is provided in Supplementary Table S1 and can also be accessed via EuPathDB (www.eupathdb.org). These results are in line with observations made for *Cryptosporidium parvum* for which 16% of ~3,900 predicted proteins were identified in the sporozoite life cycle stage (Sanderson et al., 2008). In a similar study on *Toxoplasma*, 939 of ~7,800 (~12%) predicted proteins were identified with 1-DE gel LC-MS/MS (Xia et al., 2008).

For most proteins identified in proteomic studies on *Toxoplasma gondii*, *Plasmodium falciparum* or *Neospora caninum*, corre-

sponding ESTs were found with high frequency (88.2%, 84.2% and 72.6%, respectively). For 372 of the proteins identified (58%) in the present analysis, no corresponding ESTs were found in the macroschizont EST library (GeneDB). For 100 proteins (16%), the corresponding EST was present in libraries prepared from merozoites and/or piroplasms, but not from macroschizonts. For 173 proteins (27%), the corresponding ESTs were lacking altogether. Such discrepancies have also been observed in other apicomplexan parasites where examples exist of readily detected proteins whose corresponding genes display little or no detectable transcription (Wastling et al., 2009). With high-throughput sequencing data now becoming more readily available the situation is much improved; for example, recent mRNA deep sequencing for *T. gondii* and *N. caninum* (Reid et al., 2012) indicates that transcripts can be detected for nearly all proteins identified by MS in these organisms. It is important to note, however, that this does not mean that the correlation between transcript abundance and protein abundance is linear; on the contrary this complex relationship often leads to notably apparent discrepancies reflecting a range of factors such as the rate of protein and/or mRNA turnover (Wastling et al., 2012).

Proteins reported to be expressed by *T. annulata* merozoites or piroplasms could also be detected. The cultures we use consist of continuously proliferating bovine macrophages, which harbour the transforming schizont and have been maintained for many years in the laboratory. Merogony is a stochastic process known to occur at low levels in cultures of *T. annulata*-transformed cells. In established cultures the process is inefficient but merogony can be induced further by increased temperature or different types of stress, eventually resulting in the production of mature merozoites (Shiels et al., 1992; Schmuckli-Maurer et al., 2008).

The presence of piroplasms, which develop solely in red blood cells upon merozoite invasion, can be excluded under the culture conditions that were used. There was also no obvious morphological evidence for the presence of merozoites. The most likely explanation for the presence of proteins encoded by genes reported to be expressed by *T. annulata* merozoites or piroplasms is therefore that some parasites stochastically undergo early merogony.

Interestingly, proteins previously reported to be expressed by *Theileria* sporozoites were also identified. As the schizont-infected cells have been maintained for several years in culture, the presence of sporozoites can be excluded. TA17375 (GeneDB) is a *T. annulata* ortholog of the *T. parva* 'polymorphic antigen precursor' also called p150 (UniProt Accession Number Q27028) (Skilton et al., 1998), a sporozoite protein that contains a proline-rich region and was shown to cross-react with a *T. parva* polymorphic immunodominant molecule (PIM). In this context, p104, a microneme-rhoptry protein expressed in sporozoites, was also identified in Triton X-114 extracts (see Section 3.2). In this case, we posit that these proteins continue to be expressed by the parasite after invasion and differentiation to the schizont stage.

In our analysis of WPL, only 10.9% of the identified proteins contained a predicted TMD, which contrasts with the 22.2% predicted for all proteins. In macroschizonts, 20.9% of the genes for which transcripts could be detected were predicted to contain TMD, indicating that there is no bias towards low expression of such genes for this life cycle stage. A similar observation was made for proteins containing a signal peptide/anchor. Proteins containing a signal peptide are interesting candidates for host pathogen interactions as they may be secreted into the host cell, and parasite proteins anchored in the plasma membrane have the potential to interact with host cell components. Whereas 14.7% of all predicted proteins and 13.1% of the proteins with corresponding macroschizont ESTs are predicted to contain a signal peptide/anchor, this only applies to 7.9% of the proteins found in our analysis. Only

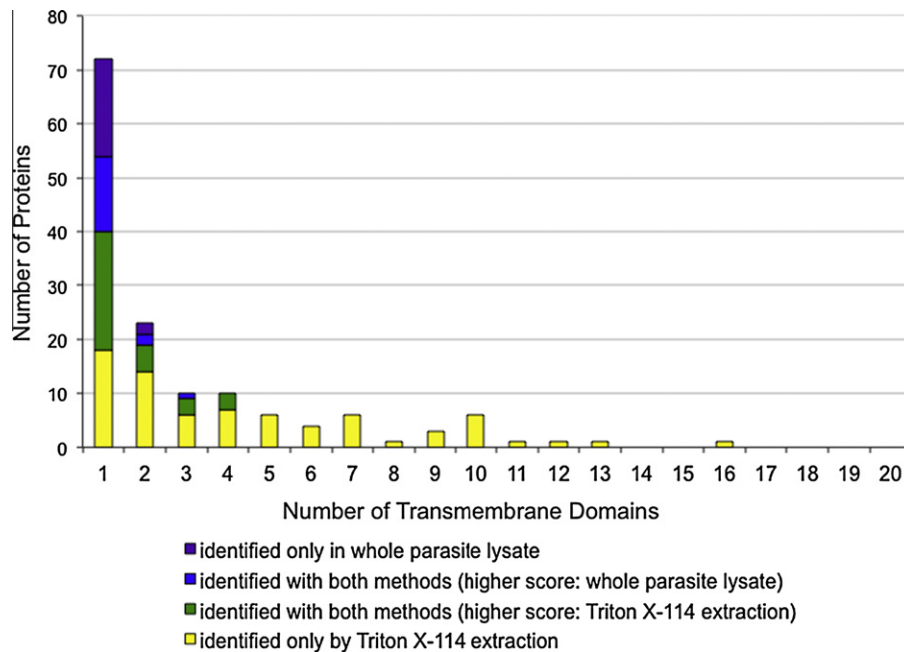


Fig. 1. Triton X-114 extraction enhances the identification of proteins with transmembrane (TM) domains. Seventy-five additional *Theileria annulata* TM proteins were identified and proteins with more than five TM domains were only observed in the Triton X-114 detergent-rich phase.

nine, four and three proteins were detected containing two, three or four TMDs, respectively. The fact that proteins with multiple TMDs appeared to be strongly under-represented in the list of identified proteins detected in WPLs pointed towards a technical explanation. Indeed, in previous proteomics studies on *Toxoplasma* and *Cryptosporidium*, the ratio of identified and predicted TM proteins was similar to that observed in this study and most of the proteins with multi-TM helices were first identified when a multi-dimensional protein identification technology (MudPIT) analysis was performed.

3.2. Analysis of Triton X-114 lysates

Considering the paucity of TM proteins, it was decided to bias the protein extraction procedure in favour of membrane proteins and extracted purified schizonts using the detergent Triton X-114 (Bordier, 1981) as described by Cordero et al. (2009), with some minor modifications (Supplementary Fig. S1A). As expected, in 1-D gel analysis the protein pattern of the aqueous and detergent phases differed significantly (not shown). Proteins contained in the detergent-rich phase were subjected to 1-D gel SDS-PAGE and the corresponding lane cut into 30 continuous slices, digested with trypsin and analysed by LC-MS/MS. Using this extraction procedure, 762 proteins with a Mascot score >50 were identified, 459 of which were non-redundant. Among these, 153 had not been detected in WPL and 75 of these were predicted to have at least one TMD. In total 125 TM proteins were found in the detergent enriched fraction, almost twice the number found in the analysis of WPL (Fig. 1) and the relative number of TM proteins was increased from 10.9% (70 of 645 proteins) to 27.2% (125 of 459). Two-thirds of the proteins found in both samples had a higher score in the Triton X-114 extracts. Using Triton X-114 extraction, TA08425 (GeneDB), a *T. annulata* ortholog of the *T. parva* microneme-rhoptry antigen (also called p104; UniProt Accession Number Q962G6) was also identified. This finding is particularly interesting, as p104 was recently found to participate in parasite/host cell microtubule interactions (unpublished data from our laboratory).

Fractionation by Triton X-114 also helped to identify proteins with multiple TMDs (5–16), including different transporters. Inter-

estingly, with a share of 59.5%, proportionately more unclassified proteins were newly identified upon detergent extraction than was the case for WPL (34.9%). Taken together, these findings argue in favour of analysing Triton X-114 extracts in addition to WPL to expand the repertoire of identified proteins, in particular membrane-bound proteins.

3.3. Classification of identified proteins

Using the different lysis and extraction methods, a total of 812 proteins were identified, representing 21.4% of the predicted *T. annulata* proteome (see Supplementary Table S1). Different heat shock proteins, actin, glyceraldehyde-3-phosphate dehydrogenase, enolase and elongation factor belonged to the most readily detected proteins. This is in agreement with observations made in a preliminary analysis of the *T. parva* schizont proteome using a combination of high-resolution 2-D gel electrophoresis and MS (Bishop et al., 2009).

A set of algorithms was used to classify proteins based on predicted subcellular localisation (Fig. 2). Whereas proteins allocated to the cytoplasm, mitochondrion, nucleus and ribosome were prominent, for 29.8% of the proteins a subcellular localisation could not be allocated.

All identified proteins were assigned to a MIPS FunCat Category (Fig. 3). With 38.3%, the unclassified proteins presented the largest category. A large category contains proteins involved in protein translation (14.2%). Other large categories involved proteins that regulate protein folding, modification, processing (together 19.7%) and cellular transport (together 10.2%). Proteins with functions assigned to cellular transport were under-represented in WPL analysis, but increased from 44 to 83 when data obtained from WPL and Triton X-114 extraction were pooled.

3.4. Glycolysis and citric acid cycle

Knowledge on the central carbon metabolism of *Theileria* is very limited compared with other apicomplexan parasites such as *Plasmodium* (Olszewski and Llinas, 2011) and *Toxoplasma* (Polonais and Soldati-Favre, 2010). A comprehensive comparison of apicomplexan

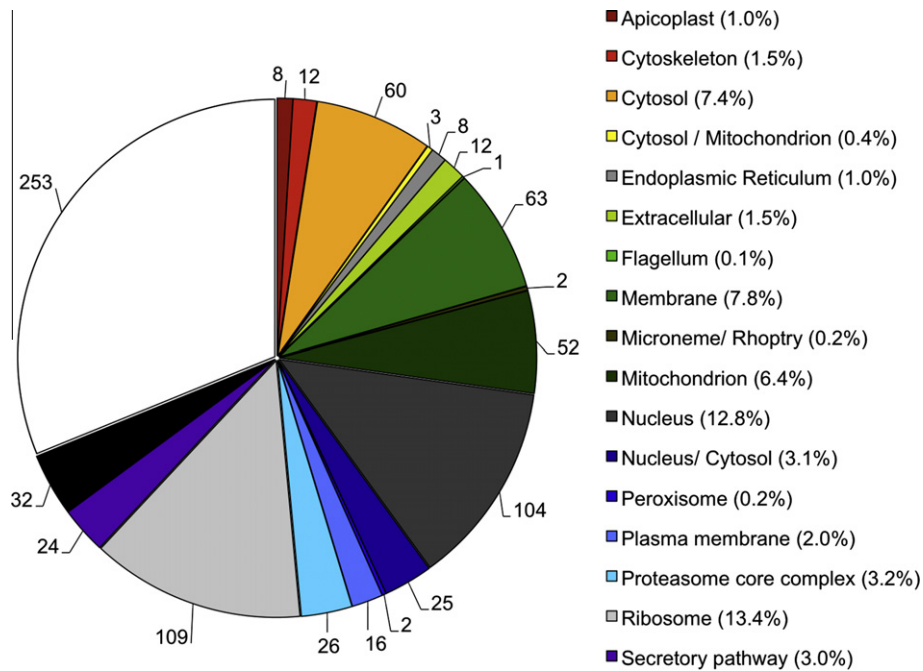


Fig. 2. Subcellular localisation of all identified *Theileria annulata* proteins (whole parasite lysate (WPL) + Triton X-114 extraction). Actual numbers for each category are presented on the pie graph.

metabolic pathways is also available through <http://www.llamp.net>. The glycolysis and citric acid cycle rely on conserved proteins that are found in almost all eukaryotic cells (Voet et al., 2006). The genes for all 10 enzymes of glycolysis are present in the genomes of apicomplexans such as *Toxoplasma*, *Plasmodium*, *Theileria* and *Babesia*. In our analysis, all corresponding enzymes were detected (Fig. 4 and Supplementary Table S2). A transmembrane hexose

transporter was also identified that could contribute to glucose import from the host cell, as shown for *Plasmodium* (Slavic et al., 2010). For these proteins, the average/median score, number of peptide matches and sequence coverage is significantly higher than for all other identified proteins, indirectly indicating that these proteins are abundant. Our observations are in line with the work of Kiama et al. (1999) who measured enzymatic activities

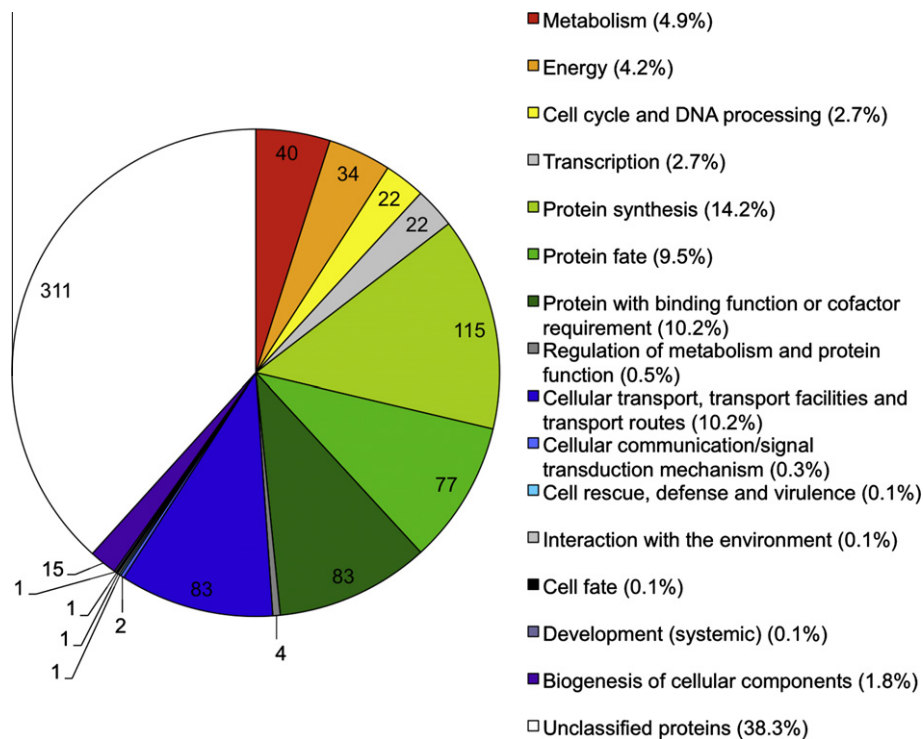


Fig. 3. MIPS Functional Catalogue (FunCat) assignment of all identified *Theileria annulata* proteins (whole parasite lysate (WPL) + Triton X-114 extraction). Actual numbers for each category are presented on the pie graph.

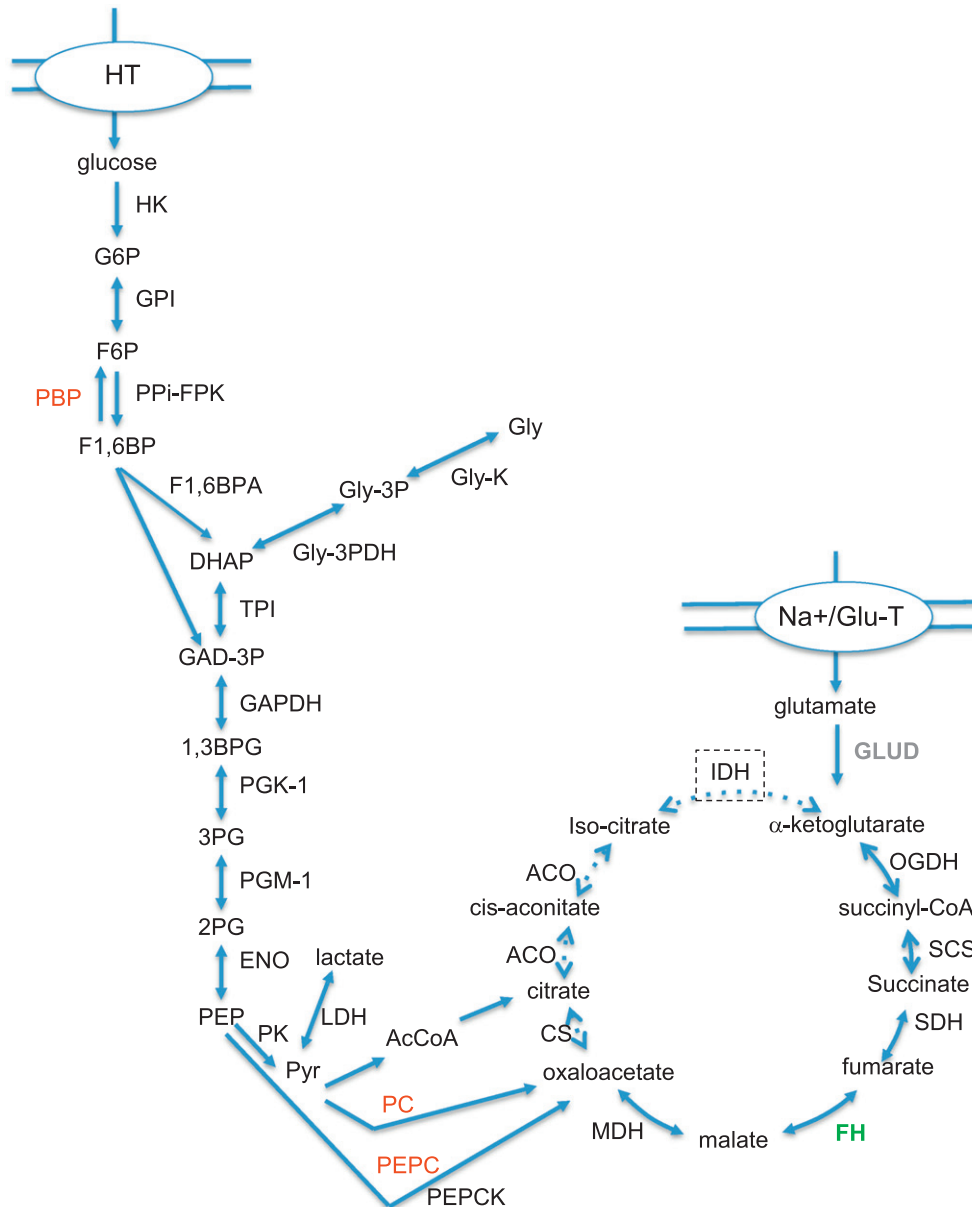


Fig. 4. Almost all *Theileria annulata* proteins involved in glycolysis and citric acid cycle were identified in the study. The glycolysis pathway is shown on the left and the citric acid cycle is shown on the right. Reaction products: G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; F1, 6BP, fructose 1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; Gly-3P, glycerol 3 phosphate; Gly, glycerol; GAD-3P, glyceraldehyde 3-phosphate; 1, 3BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; Pyr, pyruvate; AcCoA, acetyl coenzyme A. Enzymes with identified *Theileria annulata* proteins included in brackets: HK, hexokinase (TA19800); GPI, glucose 6-phosphate-isomerase (TA04045); PPI-FPK, pyrophosphate dependent phosphofructokinase (TA13950); PBP, fructose bisphosphatase; F1, 6BPA, fructose bisphosphate aldolase (TA20060); TPI, triosephosphate isomerase (TA08590); Gly-3PDH, glycerol-3-phosphate dehydrogenase; Gly-K, glycerol kinase; GAPDH, glyceraldehyde phosphate dehydrogenase (TA08145, TA15530); PGK, phosphoglycerate kinase (TA06655); PGM, phosphoglycerate mutase (TA10465); ENO, enolase (TA10425); PK, pyruvate kinase (TA11540, TA10915); LDH, lactate dehydrogenase (TA09590); PC, pyruvate carboxylase; ACO, aconitase (TA17020); IDH, isocitrate dehydrogenase (TA10850); OGDH, oxoglutarate dehydrogenase (TA05275, TA08530); SCS, succinyl coenzyme A synthetase (TA02815, TA10625); SDH, succinate dehydrogenase (TA19430, TA03455); FH, fumarate hydratase; MDH, malate dehydrogenase (TA18100); CS, citrate synthase (TA14450); PEPCK, phosphoenolpyruvate carboxylase; PEPCK, phosphoenolpyruvate carboxykinase (TA20590); GLUD, glutamate dehydrogenase (TA11105) (Voet et al., 2006). Transmembrane transporters: HT, Hexose transporter (TA02480); Na⁺/Glu-T, sodium glutamate symporter (TA10315). Green indicates proteins that were not identified, but predicted, red shows no prediction and grey indicates proteins that were identified and are also likely to be involved. Dashed arrows in the citric acid cycle indicate that only a branch of the cycle is active in the schizont stage.

in order to assess the significance of glycolysis in *T. parva* schizonts. In this work, it was also found that the enzymatic activities of glycerol kinase and glycerol 3-phosphate dehydrogenase were approximately 16 times lower than that of the other enzymes. Together, these results point towards a functional glycolytic pathway in *Theileria* schizonts, with low levels of glycerol catabolism. The fructose bisphosphatase gene, encoding an enzyme required for gluconeogenesis, cannot be found in the *Theileria* genomes. On the other

hand, phosphoenolpyruvate carboxykinase (PEPCK), an enzyme that converts oxaloacetate to PEP, thus supplying citric acid cycle-derived carbon for gluconeogenesis, is expressed. PEP carboxylase, which essentially runs the reverse reaction, however, is absent.

Aerobic glycolysis is linked to the citric acid cycle via acetyl-CoA. Interestingly, not all of the genes encoding subunits of the pyruvate decarboxylase, which catalyses the reaction of pyruvate to acetyl-CoA and CO₂, were found in the *Theileria* genome

(Gardner et al., 2005; Pain et al., 2005). An alternative route to the citric acid cycle is the carboxylation of pyruvate to oxoglutarate, which is catalysed by pyruvate carboxylase, but this enzyme is also not present in *Theileria*. Despite this, a form of citric acid cycle was proposed to be present in *Theileria* (Gardner et al., 2005). All enzymes/subunits predicted from the genome were identified by MS, except for fumarate hydratase which was, however, represented in the EST library. In *T. parva* schizonts, the activities of most of the enzymes of the citric acid cycle were found to be very low, and only a branch of the citric acid cycle involving malate dehydrogenase, fumarase and succinate dehydrogenase appears to be active, potentially operating in the reverse direction of the citric acid cycle to synthesise succinate (Kiama et al., 1999). In our analysis, the number of peptide matches of the citric acid cycle enzymes was much lower compared with glycolytic enzymes. It was suggested that glutamate could be a supplementing intermediate for the citric acid cycle (Gardner et al., 2005). Glutamate can be converted into α -ketoglutarate by glutamate dehydrogenase. The latter was identified with a very high score. Glutamate might thus be a crucial intermediate and this hypothesis is underpinned by the identification of a sodium-glutamate symporter.

Incorporation studies on *Babesia rodhanini* also demonstrated the absence of a complete citric acid cycle (Rickard, 1970). In the case of *Plasmodium*, it has been proposed that, despite the presence of all enzymes, at least the asexual stages do not rely on the citric acid cycle for energy generation (van Dooren et al., 2006). However, recent studies involving transcriptomic, proteomic and metabolomic analyses suggest this model may be too simple (reviewed in Polonais and Soldati-Favre, 2010); the citric acid cycle may be working bidirectionally and be more prominent in different life cycle stages. This also applies to *Toxoplasma*, in which genes for all of the enzymes of the citric acid cycle are present and enzymes appear to be targeted to the mitochondrion (Feige et al., 2008; Xia et al., 2008).

Taken together, it would appear that the anaerobic pathway is the main route of glucose metabolism in *Theileria* macroschizonts, that gluconeogenesis does not occur, glycerol catabolism occurs at low levels and that only a branch of the citric acid cycle is active (Kiama et al., 1999) in a process that does not participate in energy generation in the schizont. Additional studies will be required to determine whether other stages in the *Theileria* life cycle yield energy from oxidative phosphorylation.

3.5. Proteins involved in host/parasite interactions

The schizont is not contained in a parasitophorous vacuole and resides free in the cytoplasm where it interacts with host cell microtubules (Seitzer et al., 2010; von Schubert et al., 2010). A number of proteins have also been proposed to be secreted into the host cell cytoplasm from where they could interfere with host cell signalling pathways or translocate into the nucleus where they could modulate host cell gene expression, potentially contributing to transformation.

TaSP (the *T. annulata* ortholog of *T. parva* PIM), is abundantly expressed on the schizont surface and has been proposed to interact with host cell microtubules (Seitzer et al., 2010); this also applies to TaSE (GeneDB, TA20205). Surprisingly, neither of these proteins could be detected by MS.

Further, members of the family of subtelomere-encoded variable secreted proteins (SVSPs) (Gardner et al., 2005; Pain et al., 2005) and TashATs, which contain AT-hook DNA-binding domains and have been reported to localise to the nucleus of *T. annulata*-transformed cells (Swan et al., 2001), were not detected. With 85 members in *T. parva* and 48 members in *T. annulata*, SVSP genes constitute the largest family observed in these organisms. The general structure of SVSPs consists of a short conserved

N-terminal region, in most cases containing a putative signal peptide for secretion, followed by a QP-rich region, which is predicted to be highly unstructured and a conserved C-terminus that has no significant identity to known proteins. In many SVSPs a nuclear localisation signal was also found. In *T. parva*-transformed cells, a large contingent of the SVSP gene family was found to be expressed at the RNA level (Bishop et al., 2005; Schmuckli-Maurer et al., 2009) and for at least one of the members, protein was also demonstrated using immuno-labelling techniques (Schmuckli-Maurer et al., 2009). TashATs form another intriguing family of parasite-secreted proteins, reported to translocate to the host cell nucleus. Whether these proteins contribute to transformation-specific changes in host cell gene expression is presently unknown. There could be several reasons why none of the SVSPs or TashATs was detected, including low abundance because proteins are systematically released from the schizont or unusual composition preventing detection. It is worth noting, however, that WPL samples for analysis were boiled up to five times in Lämmli buffer before they were subjected to SDS-PAGE. In this context, we recently observed that boiling schizont proteins often results in the formation of precipitates that remain in the stacking gel during SDS-PAGE. It cannot be excluded that this led to reduced detectability of certain proteins.

3.6. Concluding remarks

To our knowledge, we present the first characterisation of the *T. annulata* schizont proteome. This was greatly facilitated by the availability of a method to purify the parasite from the host cell cytoplasm and the application of Triton X-114 extraction for membrane protein isolation. This led to the identification of 812 proteins, most of which had only been predicted to exist based on genome and transcriptome analyses. The identification of proteins that were predicted to be expressed predominantly by the sporozoite, such as p150 and p104, is intriguing. It could point towards functions related to invasion, early establishment of the sporozoite upon entry, as well as long-term persistence of the schizont in the transformed cell. The failure to detect well-characterised proteins such as TaSP, or potentially secreted proteins such as SVSPs or TashATs, is surprising and potentially emphasises the technical limitations of the approach used in this study. With deep sequencing becoming widely accessible and increasingly affordable, and taking into account the continuous developments in quantitative proteomics, both the *Theileria* proteome and transcript datasets are bound to expand and provide valuable tools for a system-wide approach to elucidate this unique host/parasite interaction.

Acknowledgements

We are grateful to Stephan Grimm for technical assistance and Jacqueline Schmuckli-Maurer and Conrad von Schubert for useful advice. This work was supported by grants to D.D. from the Swiss National Science Foundation (Nr 3100A0-116653), The Wellcome Animal Health Initiative (United Kingdom; GR075820MA), the Integrated Consortium on Ticks and Tick-borne Diseases (European Community, ICTTD Nr 510561) and from the Biotechnology and Biological Sciences Research Council (United Kingdom) to J.W.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.ijpara.2012.10.017>.

References

- Baumgartner, M., Tardieux, I., Ohayon, H., Gounon, P., Langsley, G., 1999. The use of nocodazole in cell cycle analysis and parasite purification from *Theileria parva*-infected B cells. *Microb. Infect.* 1, 1181–1188.
- Bender, A., van Dooren, G.G., Ralph, S.A., McFadden, G.I., Schneider, G., 2003. Properties and prediction of mitochondrial transit peptides from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 132, 59–66.
- Bishop, R., Shah, T., Pelle, R., Hoyle, D., Pearson, T., Haines, L., Brass, A., Hulme, H., Graham, S.P., Taracha, E.L., Kanga, S., Lu, C., Hass, B., Wortman, J., White, O., Gardner, M.J., Nene, V., de Villiers, E.P., 2005. Analysis of the transcriptome of the protozoan *Theileria parva* using MPSS reveals that the majority of genes are transcriptionally active in the schizont stage. *Nucleic Acids Res.* 33, 5503–5511.
- Bishop, R.P., Odongo, D.O., Mann, D.J., Pearson, T.W., Sugimoto, C., Haines, L.R., Glass, E., Jensen, K., Seitzer, U., Ahmed, J.S., Graham, S.R., de Villiers, E.P., 2009. *Theileria*. In: Nene, V., Kole, C. (Eds.), *Genome Mapping and Genomics in Animal-Associated Microbes*. Springer-Verlag Berlin, Heidelberg, pp. 191–231.
- Bordier, C., 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* 256, 1604–1607.
- Chaussepied, M., Langsley, G., 2011. *Theileria*-induced leukocyte transformation: an example of oncogenic addiction? In: Kaufmann, S.H.E., Rouse, B.T., Sacks, D.L. (Eds.), *The Immune Response to Infection*. ASM Press, Washington, DC, pp. 537–546.
- Cordero, E.M., Nakayasu, E.S., Gentil, L.G., Yoshida, N., Almeida, I.C., da Silveira, J.F., 2009. Proteomic analysis of detergent-solubilized membrane proteins from insect-developmental forms of *Trypanosoma cruzi*. *J. Proteome Res.* 8, 3642–3652.
- Dobbelaere, D., Baumgartner, M., 2009. *Theileria*. In: Schaible, U.E., Haas, A. (Eds.), *Intracellular Niches of Microbes: A Pathogens Guide Through the Host Cell*. Wiley-VCH Verlag, Weinheim, pp. 613–632.
- Durrani, Z., Weir, W., Pillai, S., Kinnaird, J., Shiels, B., 2012. Modulation of activation-associated host cell gene expression by the apicomplexan parasite *Theileria annulata*. *Cell. Microbiol.* 14, 1434–1454.
- Emanuelsson, O., Nielsen, H., Brunak, S., von Heijne, G., 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016.
- Fleige, T., Pfaff, N., Gross, U., Bohne, W., 2008. Localisation of gluconeogenesis and tricarboxylic acid (TCA)-cycle enzymes and first functional analysis of the TCA cycle in *Toxoplasma gondii*. *Int. J. Parasitol.* 38, 1121–1132.
- Foth, B.J., Ralph, S.A., Tonkin, C.J., Struck, N.S., Fraunholz, M., Roos, D.S., Cowman, A.F., McFadden, G.I., 2003. Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science* 299, 705–708.
- Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., Ren, Q., Paulsen, I.T., Pain, A., Berriman, M., Wilson, R.J., Sato, S., Ralph, S.A., Mann, D.J., Xiong, Z., Shallom, S.J., Weidman, J., Jiang, L., Lynn, J., Weaver, B., Shoaibi, A., Domingo, A.R., Wasawo, D., Crabtree, J., Wortman, J.R., Haas, B., Angiuoli, S.V., Creasy, T.H., Lu, C., Suh, B., Silva, J.C., Utterback, T.R., Feldblyum, T.V., Perlea, M., Allen, J., Nierman, W.C., Taracha, E.L., Salzberg, S.L., White, O.R., Fitzhugh, H.A., Morzaria, S., Venter, J.C., Fraser, C.M., Nene, V., 2005. Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* 309, 134–137.
- Horton, P., Nakai, K., 1999. PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24, 34–36.
- Hulliger, L., Wilde, K.H., Brown, C.G., Turner, L., 1964. Mode of multiplication of *Theileria* in cultures of bovine lymphocytic cells. *Nature* 203, 728–730.
- Kiama, T.N., Kiara, J.K., Konji, V.N., Musoke, A.J., 1999. Enzymes of glucose and glycerol catabolism in in vitro-propagated *Theileria parva* schizonts. *Vet. J.* 158, 221–227.
- Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580.
- Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997a. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1–6.
- Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997b. A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int. J. Neural Syst.* 8, 581–599.
- Olszewski, K.L., Llinas, M., 2011. Central carbon metabolism of *Plasmodium* parasites. *Mol. Biochem. Parasitol.* 175, 95–103.
- Pain, A., Renaud, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W., Kerhornou, A., Aslett, M., Bishop, R., Bouchier, C., Cochet, M., Coulson, R.M., Cronin, A., de Villiers, E.P., Fraser, A., Fosker, N., Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F., Larke, N., Lord, A., Maser, P., McKellar, S., Mooney, P., Morton, F., Nene, V., O'Neil, S., Price, C., Quail, M.A., Rabinowitsch, E., Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T., Squares, R., Squares, S., Tivey, A., Walker, A.R., Woodward, J., Dobbelaere, D.A., Langsley, G., Rajandream, M.A., McKeever, D., Shiels, B., Tait, A., Barrell, B., Hall, N., 2005. Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*. *Science* 309, 131–133.
- Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.
- Polonais, V., Soldati-Favre, D., 2010. Versatility in the acquisition of energy and carbon sources by the Apicomplexa. *Biol. Cell* 102, 435–445.
- Rickard, M.D., 1970. Carbohydrate metabolism in *Babesia rodhaini*: investigations with 14C-labeled substrates and enzyme assays. *Exp. Parasitol.* 28, 512–520.
- Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Guldener, U., Mannhaupt, G., Munsterkotter, M., Mewes, H.W., 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* 32, 5539–5545.
- Reid, A.J., Vermont, S.J., Cotton, J.A., Harris, D., Hill-Cawthorne, G.A., Könen-Waisman, S., Latham, S.M., Mourier, T., Norton, R., Quail, M.A., Sanders, M., Shanmugam, D., Sohal, A., Wasmuth, J.D., Brunk, B., Grigg, M.E., Howard, J.C., Parkinson, J., Roos, D.S., Trees, A.J., Berriman, M., Pain, A., Wastling, J.M., 2012. Comparative genomics of the apicomplexan parasites *Toxoplasma gondii* and *Neospora caninum*: Coccidia differing in host range and transmission strategy. *PLoS Pathog.* 8(3), e1002567. <http://dx.doi.org/10.1371/journal.ppat.1002567>.
- Sanderson, S.J., Xia, D., Prieto, H., Yates, J., Heiges, M., Kissinger, J.C., Bromley, E., Lal, K., Sinden, R.E., Tomley, F., Wastling, J.M., 2008. Determining the protein repertoire of *Cryptosporidium parvum* sporozoites. *Proteomics* 8, 1398–1414.
- Schmuckli-Maurer, J., Shiels, B., Dobbelaere, D.A., 2008. Stochastic induction of *Theileria annulata* merogony in vitro by chloramphenicol. *Int. J. Parasitol.* 38, 1705–1715.
- Schmuckli-Maurer, J., Casanova, C., Schmied, S., Affentranger, S., Parvanova, I., Kang'a, S., Nene, V., Katzer, F., McKeever, D., Müller, J., Bishop, R., Pain, A., Dobbelaere, D.A., 2009. Expression analysis of the *Theileria parva* subtelomere-encoded variable secreted protein gene family. *PLoS One* 4 (3), e4839. <http://dx.doi.org/10.1371/journal.pone.0004839>.
- Seitzer, U., Gerber, S., Beyer, D., Dobschanski, J., Kullmann, B., Haller, D., Ahmed, J.S., 2010. Schizonts of *Theileria annulata* interact with the microtubuli network of their host cell via the membrane protein TaSP. *Parasitol. Res.* 106, 1085–1102.
- Shah, T., de Villiers, E., Nene, V., Hass, B., Taracha, E., Gardner, M.J., Sansom, C., Pelle, R., Bishop, R., 2006. Using the transcriptome to annotate the genome revisited: application of massively parallel signature sequencing (MPSS). *Gene* 366, 104–108.
- Shiels, B., Kinnaird, J., McKellar, S., Dickson, J., Miled, L.B., Melrose, R., Brown, D., Tait, A., 1992. Disruption of synchrony between parasite growth and host cell division is a determinant of differentiation to the merozoite in *Theileria annulata*. *J. Cell Sci.* 101, 99–107.
- Shiels, B., Langsley, G., Weir, W., Pain, A., McKellar, S., Dobbelaere, D., 2006. Alteration of host cell phenotype by *Theileria annulata* and *Theileria parva*: mining for manipulators in the parasite genomes. *Int. J. Parasitol.* 36, 9–21.
- Skilton, R.A., Bishop, R.P., Wells, C.W., Spooner, P.R., Gobright, E., Nkongue, C., Musoke, A.J., Macklin, M., Iams, K.P., 1998. Cloning and characterization of a 150 kDa microspherule antigen of *Theileria parva* that is immunologically cross-reactive with the polymorphic immunodominant molecule (PIM). *Parasitology* 117, 321–330.
- Slavic, K., Straschil, U., Reiningger, L., Doerig, C., Morin, C., Tewari, R., Krishna, S., 2010. Life cycle studies of the hexose transporter of *Plasmodium* species and genetic validation of their essentiality. *Mol. Microbiol.* 75, 1402–1413.
- Swan, D.G., Stern, R., McKellar, S., Phillips, K., Oura, C.A., Karagenc, T.I., Stadler, L., Shiels, B.R., 2001. Characterisation of a cluster of genes encoding *Theileria annulata* AT hook DNA-binding proteins and evidence for localisation to the host cell nucleus. *J. Cell Sci.* 114, 2747–2754.
- van Dooren, G.G., Stimmer, L.M., McFadden, G.I., 2006. Metabolic maps and functions of the *Plasmodium* mitochondrion. *FEMS Microbiol. Rev.* 30, 596–630.
- Voet, D., Voet, J.G., Pratt, C.W., 2006. *Fundamentals of Biochemistry: Life at the Molecular Level*. Wiley, Hoboken, NJ.
- von Schubert, C., Xue, G., Schmuckli-Maurer, J., Woods, K.L., Nigg, E.A., Dobbelaere, D.A., 2010. The transforming parasite *Theileria* co-opts host cell mitotic and central spindles to persist in continuously dividing cells. *PLoS Biol.* 8 (9), e1000499. <http://dx.doi.org/10.1371/journal.pbio.1000499>.
- Wastling, J.M., Xia, D., Sohal, A., Chaussepied, M., Pain, A., Langsley, G., 2009. Proteomes and transcriptomes of the Apicomplexa—where's the message? *Int. J. Parasitol.* 39, 135–143.
- Wastling, J.M., Armstrong, S.D., Krishna, R., Xia, D., 2012. Parasites, proteomes and systems: has Descartes' clock run out of time? *Parasitology* 139, 1103–1118.
- Weir, W., Karagenc, T., Baird, M., Tait, A., Shiels, B.R., 2010. Evolution and diversity of secretome genes in the apicomplexan parasite *Theileria annulata*. *BMC Genomics* 11, 42.
- Xia, D., Sanderson, S.J., Jones, A.R., Prieto, J.H., Yates, J.R., Bromley, E., Tomley, F.M., Lal, K., Sinden, R.E., Brunk, B.P., Roos, D.S., Wastling, J.M., 2008. The proteome of *Toxoplasma gondii*: integration with the genome provides novel insights into gene expression and annotation. *Genome Biol.* 9, R116.
- Zuegge, J., Ralph, S., Schmucker, M., McFadden, G.I., Schneider, G., 2001. Deciphering apicoplast targeting signals—feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. *Gene* 280, 19–26.