International Journal for Parasitology 51 (2021) 809–824

Contents lists available at ScienceDirect

International Journal for Parasitology

journal homepage: www.elsevier.com/locate/ijpara

Multimodal regulation of encystation in *Giardia duodenalis* revealed by deep proteomics

Balu Balan [a,b,*], Samantha J. Emery-Corbin [a,c], Jarrod J. Sandow [a,c], Brendan Robert E. Ansell [a,c], Swapnil Tichkule [a,c], Andrew I. Webb [a,c], Staffan G. Svärd [d], Aaron R. Jex [a,b,c]

[a] Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia
[b] Faculty of Veterinary and Agricultural Science, University of Melbourne, Melbourne, Victoria, Australia
[c] Department of Medical Biology, University of Melbourne, Parkville, Victoria, Australia
[d] Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden

ARTICLE INFO

Article history:
Received 2 September 2020
Received in revised form 24 December 2020
Accepted 7 January 2021
Available online 28 July 2021

Keywords:
Giardia
Encystation
Proteome
Transcriptome-proteome correlation
Metabolism
Lipid transporters
RNA binding proteins

ABSTRACT

Cyst formation in the parasitic protist *Giardia duodenalis* is critical to its transmission. Existing proteomic data quantifies only 17% of coding genes transcribed during encystation and does not cover the complete process from trophozoite to mature cyst. Using high-resolution mass spectrometry, we have quantified proteomic changes across encystation and compared this with published transcriptomic data. We reproducibly identified 3863 (64.5% of *Giardia* proteins) and quantified 3382 proteins (56.5% of *Giardia* proteins) over standard trophozoite growth (TY), during low-bile encystation priming (LB), 16 h into encystation (EC), and at cyst maturation (C). This work provides the first known expanded observation of encystation at the proteomic level and triples the coverage of previous encystation proteomes. One-third (1169 proteins) of the quantified proteome is differentially expressed in the mature cyst relative to the trophozoite, including proteasomal machinery, metabolic pathways, and secretory proteins. Changes in lipid metabolism indicated a shift in lipid species dependency during encystation. Consistent with this, we identified the first, putative lipid transporters in this species, representing the steroidogenic acute regulatory protein–related lipid transfer (StARkin), oxysterol binding protein related protein (ORP/Osh) and glycosphingolipid transfer protein (GLTP) families, and follow their differential expression over cyst formation. Lastly, we undertook correlation analyses of the transcriptome and proteome of trophozoites and cysts, and found evidence of post-transcriptional regulation of key protein classes (RNA binding proteins) and stage-specific genes (encystation markers) implicating translation-repression in encystation. We provide the most extensive proteomic analysis of encystation in *Giardia* to date and the first known exploration across its complete duration. This work identifies encystation as highly coordinated, involving major changes in proteostasis, metabolism and membrane dynamics, and indicates a potential role for post-transcriptional regulation, mediated through RNA-binding proteins. Together our work provides a valuable resource for *Giardia* research and the development of transmission-blocking anti-giardials.
© 2021 The Author(s). Published by Elsevier Ltd on behalf of Australian Society for Parasitology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Giardia duodenalis is a gastrointestinal, parasitic protist that causes approximately 280 million symptomatic cases of diarrhoea (giardiasis) per year worldwide (Lane and Lloyd, 2002). The parasite is one of the earliest diverging species of eukaryotes, and has one of the simplest cell differentiation processes (Svärd et al., 2003). During its life-cycle, *Giardia* transitions from a flagellated and binucleate trophozoite to an environmentally-resistant, tetra-nucleate cyst in a process known as encystation that is required for transmission (Ankarklev et al., 2010). The early phase of encystation is characterized by the formation of encystation-specific vesicles (ESVs) that transport cyst wall proteins (CWPs) to the cell surface (Reiner et al., 1989). The progression of encystation is marked by rolling of the trophozoite margin, fragmentation of the adhesive disc (Palm et al., 2005) and internalization of the flagella, followed by formation of the cyst wall (Lauwaet et al., 2007; Einarsson and Svärd, 2015). Encystation is triggered by changes in environmental factors including bile, cholesterol and pH (Einarsson and Svärd, 2015), but the specific molecular

* Corresponding author at: Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia.
E-mail address: balan.b@wehi.edu.au (B. Balan).

https://doi.org/10.1016/j.ijpara.2021.01.008
0020-7519/© 2021 The Author(s). Published by Elsevier Ltd on behalf of Australian Society for Parasitology.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

mechanisms that induce or regulate encystation are not known. Encystation can be replicated and studied in vitro (Hehl et al., 2000). Upon in vitro stimulation of encystation, with or without low-bile priming (Faso et al., 2013; Einarsson et al., 2016), commitment to cyst formation is irreversible after ~3–6 h in encystation media (Sulemana et al., 2014; Einarsson and Svärd, 2015). After 48 h in encystation media, *Giardia* cysts require a further ~48–72 h to reach maturity (Einarsson and Svärd, 2015).

Several in vitro and in vivo transcriptomic studies (Birkeland et al., 2010; Einarsson et al., 2016; Pham et al., 2017) have provided major insights into encystation in *Giardia*. These highlighted changes in glycolytic, proteolytic and lipid metabolism, surface protein switching and chromatin remodeling during early (1.5–7 h) and late (24–32 h) encystation. Proteomics has provided significant insights into *Giardia* biology (Faso et al., 2013; Emery et al., 2014; Emery et al., 2016; Ma'ayeh et al., 2017; Dubourg et al., 2018), including over the first 12–14 h of encystation (Faso et al., 2013). However, although there is transcriptomic support for 5757 protein coding genes (97.2% of all coding genes) at least within one time point during encystation (Einarsson et al., 2016), only 1063 (18.7%) of these are identified, and 960 (16.9%) quantified (Faso et al., 2013), as proteins. Undoubtedly observed differences between transcriptomic and proteomic detection are due both to the sensitivity of mass-spectrometry equipment on which these first datasets were run (Angel et al., 2012), and the use of timepoints exclusively within early or late encystation. However, the limited repertoire of transcription factors (Teodorovic et al., 2007) and loose chromatin packing (Tůmová et al., 2015) in *Giardia* coincide with "leaky transcription", with ~99% of known open reading frames transcribed and ~20% of the transcriptome constituted by sterile antisense transcripts (Elmendorf et al., 2001; Teodorovic et al., 2007; Birkeland et al., 2010). These observations credit the hypothesis of *Giardia* exercising significant post-transcriptional regulation, as already demonstrated in the variant-specific surface proteins (VSPs) within their antigenic variation (Prucca et al., 2008). This hypothesis is further supported by observations that several RNA-binding proteins (RBPs), which regulate cell differentiation by repressing translation of targeted messenger RNAs (Hentze et al., 2018), are significantly up-regulated in encystation (Gargantini et al., 2012), and play a critical role in transmissive stages of a variety of parasitic protists (Le Roch et al., 2004; Fernández-Moya and Estévez, 2010; Bunnik et al., 2016; Lueong et al., 2016; Vembar et al., 2016; Nandan et al., 2017; Muller et al., 2019).

Here we have performed the first known high-resolution (Aebersold and Mann, 2016), quantitative proteomic analysis of encystation, and provided expanded coverage of this process for the first known time through to cyst maturation. This has facilitated the first known wider correlations with transcriptomic studies (Birkeland et al., 2010; Einarsson et al., 2016; Pham et al., 2017) and indicates that encystation, despite being one of the simplest forms of eukaryotic cell differentiation, is mechanistically complex and tightly regulated, including at the post-transcriptional level.

2. Materials and methods

2.1. Giardia cell culture, induction of in vitro encystation and protein extraction

Giardia duodenalis WB isolate (WB-1B) (Upcroft et al., 1995) trophozoites were maintained in TYI-S-33 medium containing 10% adult bovine serum (Thermo Fisher Scientific, New Zealand origin) and 0.5 g/L of bovine bile (Sigma) as described previously (Emery-Corbin et al., 2018). Encystation was performed using the two-step encystation protocol (Lujan and Svärd, 2011), with low bile medium containing porcine bile (0.1 mg/mL; Sigma) and encystation medium containing porcine bile (0.5 mg/ml; Sigma) at pH 7.8. Cultures were primed in low bile media to early log-phase and sub-confluency, then medium was decanted and replaced with encystation media. Encystation was performed over 48 h, and enriched cysts obtained from these cultures via hypotonic lysis. Biological triplicates of *Giardia* cultures under TY and LB culture conditions, EC, were chilled to detach trophozoites, spun at 680 *g* and all cell material collected. Cysts were enriched after 48 h encystation by removing sedimented material and subjecting it to hypotonic lysis to remove unencysted cells for 24 h at 4 °C in distilled water. Cell lysis, protein extraction and reduction were performed in SDS sample buffer (2.5% SDS (Sigma), 100 mM trisaminomethane (TRIS) (Sigma) and 5 mM DTT (Bio Rad)) at 95 °C for 10 min. Alkylation was performed using 15 mM iodoacetamide (Bio Rad) and proteins precipitated via the methanol-chloroform approach (Wessel and Flügge, 1984). Protein pellets were solubilised in 8 M Urea in 50 mM TRIS (pH 8.8) and the protein concentration quantitated by bicinchoninic acid assay (Pierce). Protein pre-digestion was performed using Lys-C (Wako) for 3 h for 30 °C, followed by Trypsin (Promega) digestion at 37 °C overnight at enzyme to protein ratios of 1:100. Peptides were acidified with trifluoroacetic acid to 1%, and desalted using solid phase extraction with in-house stage tips packed with styrene divinyl benzene (3 M Empore) (Hagen et al., 1990). Peptide extracts were dried by vacuum centrifuge, resuspended in 2% acetonitrile (ACN), 0.1% trifluoroacetic acid, and 0.5% acetic acid for MS.

2.2. LC-Ms/MS

The peptides were separated on an Ultimate 3000 RSLC nanoHPLC system (Dionex Ultimate 3000) coupled with an LTQ Orbitrap Elite (Thermo Scientific, USA) in conjunction with a nano-electrospray ionisation (nano-ESI) source. The peptides were loaded in 0.1% v/v formic acid (solvent A) into an LC system equipped with Acclaim Pepmap nano-trap column (Dionex-C18, 100 Å, 75 μm × 2 cm) and an Acclaim Pepmap RSLC analytical column (Dionex-C18, 100 Å, 75 μm × 50 cm). Peptides were eluted with a non-linear 125 min gradient of 3%-80% solvent B (100% v/v CH_3CN in 0.l% v/v formic acid) at a flow rate of 0.3 uL/ min. Mass spectra from LTQ Orbitrap Elite were acquired in positive mode in a data-dependent manner with full MS scan from m/z 300–1650 in the fourier transform mode, using a lock mass of 401.922718. The top 20 most intense peptides were fragmented by collision-induced dissociation (CID) with normalised collision energy of 30, activation q of 0.25, and fragmented ions were dynamically excluded for 10 s. All data was acquired using Xcalibur software. All the parameters of LC-MS/MS run can be found in Supplementary Data_S1.

2.3. Analysis of LC-MS/MS data

Peptide fragments were identified by database searching with MaxQuant software (v1.5.8.3) (Cox and Mann, 2008) using protein sequence data for the *Giardia intestinalis* WB genome (ATCC 50803) in UniProt (Release February 2018). All isoforms were included, and additional search parameters included using strict trypsin specificity allowing up to two missed cleavages and a minimum required peptide length of seven amino acids. Carbamidomethylation of cysteine was set as a fixed modification, and protein N-terminal acetylation and methionine oxidation were set as variable modifications. For MaxQuant, precursor ion mass error tolerance was set to 4.5 ppm and fragment ions were allowed a mass deviation of 20 ppm. Peptide spectrum matches (PSMs) and protein identifications were filtered using a target-decoy approach at a false discovery rate (FDR) of 1%. Further analysis was performed

using a custom pipeline developed in R, which uses the LFQ intensity and iBAQ values in the MaxQuant output file proteinGroups.txt. Proteins found in at least 50% of the replicates in one group were considered identified and similarly in 50% of the conditions as quantified. Missing values were imputed using a random normal distribution of values with the mean set at mean of the real distribution of values minus 1.8 S.D., and a S.D. of 0.3 times the S.D. of the distribution of the measured intensities. The probability of differential protein expression between groups was calculated using the Limma R package (Ritchie et al., 2015). Probability values were corrected for multiple testing using the Benjamini–Hochberg method. Proteins with −log10 adjusted *P* value >1.3, and log fold-change <−1 or log fold-change >1 were considered differentially expressed. Analysis of the data as well as visualization were performed using Perseus (Tyanova et al., 2016), and graphs were plot using software's RStudio and TIBCO spotfire.

2.4. Gene set enrichment analysis (GSEA)

Gene set enrichment analysis (GSEA) of the differentially expressed proteins was performed using the DAVID bioinformatic resource (Da Wei Huang et al., 2007). Gene set results with $P \leq 0.1$ were considered significant. Fisher's exact test was used to test for enrichment of *Giardia*-specific gene sets between pairs of media conditions using R software. Differences in expression of proteins constituting gene sets that were significant in at least one pairwise comparison by Fisher's exact test, were tested using ANOVA. Significantly, by ANOVA proteins were clustered using k-means and were visually inspected to identify sub-group trends. Figures were generated using ggplot2 (Wickham, 2011).

2.5. Protein–protein interaction networks

A protein–protein interaction network for the most differentially expressed *G. duodenalis* WB isolate proteins was generated using STRING software (v11) (Szklarczyk et al., 2018). Interaction networks with a medium confidence score of 0.4 were visualized, with network edges representing confident networks with continuous lines showing direct interactions, and dashed lines depicting indirect interactions. A Markov Cluster Algorithm (MCL) inflation default score of 3 was used for the clustering of the network.

2.6. Hidden Markov model (HMM) search for lipid transporters and their in-silico structure prediction

Hidden Markov model (HMM) profiles of respective PFAM domains from the protein sequences of already known Lipid transport Proteins (LTPs) (Wong et al., 2019) were retrieved from the PFAM database (https://pfam.xfam.org/). Respective HMM profiles were searched against the *G. duodenalis* proteome from Giardia DB using -hmmsearch command of HMMER (version 3.2.1) (http://hmmer.org/) and proteins identified with an e-value score of $>1^{*}10^{-3}$ were considered significant. In silico three-dimensional (3D) structures of HMMER predicted *Giardia* LTPs were obtained from predictein (http://www.predictein.org) (Ansell et al., 2018) which employs I-TASSER-based de novo structure prediction coupled with a random forest classifier to generate in silico 3D structures for *Giardia* proteins. Protein structures were visualized using University of California San Francisco (UCSF), USA, Chimera software (Pettersen et al., 2004).

2.7. In silico auto-docking

In silico auto-docking was performed between the two predicted lipid transporter protein structures, GL50803_4197, GL50803_14833 in *Giardia* and the lipid molecules Phosphatidylinositol, Phosphatidylserine, respectively. The lipid molecules and active biding sites in *Giardia* proteins GL50803_4197 and GL50803_14833 were located via alignment to homologous crystal structures Human Phosphatidylinositol Transfer Protein Alpha (PDB ID: 1UW5) and *Saccharomyces cerevisiae* oxysterol binding protein 6 (PDB ID: 4B2Z), respectively, and interactions between ligands and *Giardia* proteins compared with respective crystal structures. 3D structures of *Giardia* lipid transporter proteins were curated from the *Giardia* ITASSER study (Ansell et al., 2018), and interactions between their predicted active sites and corresponding lipid ligands were assessed through in silico docking. Auto Dock Tools (ADT), (v1.5.6) was used to prepare the coordinate files for each protein and its predicted lipid ligand, and to add polar hydrogen molecules, partial charges and assign atom types. To build a torsion tree in ADT for lipid molecules, a root atom was also chosen for flexibility and rotatable bonds were defined. A grid box was constructed around the predicted active site in ADT. After these pre-preparations, docking was performed by using Auto Dock-vina version 2 (Trott and Olson, 2010).

2.8. Transcriptome-proteome correlation analysis

Transcriptomic (Einarsson et al., 2016) and proteomic data generated in this study for TY and C stages were compared using transcript and protein absolute correlation (comparing log2(fragments per kilobase of transcript per million mapped reads) (FPKM) and log2(intensity-based absolute quantitation) (iBAQ) relative abundance measures), ranked correlation (based on ranking as inferred from log2(FPKM) and log2(iBAQ) correlations), and semi-quantitative agreement (based on the quintiles of absolute correlation and ranked correlation inferred from log2(FPKM) and log2(iBAQ), and their respective correlation). To assess genes potentially regulated through translational repression, we identified genes that are (i) in the top decile of transcriptional abundance (considering only transcripts with an FPKM > 1) and undetected as a protein in any replicate (per (Muller et al., 2019)) or (ii) transcript:protein pairs identified as statistically significant outliers ($P < 0.05$) based on linear correlation and regression analysis of their relative transcriptomic (log2FPKM) versus proteomic (log2iBAQ) abundance, in which protein abundance was lower than expected based on the corresponding transcript abundance (per Wang et al. (2019)), and (iii) between stages we identified genes among the top decile (5% most up-regulated, based on fold-change) of differential transcription and compared those with protein expression between TY and C to find genes that were negatively correlated in the direction of their transcriptomic versus proteomic change, to identify translationally repressed genes. We also looked for amino acid composition, codon usage and the presence of multi-isoform genes within these transcript:protein comparisons, to assess using the gene models and genomic annotation data for the *G. duodenalis* WB-C6 genome as provided through Giardia DB Release 34 (2017-08-25), to confirm that the variations are not a reflection of specific codon usage or these are not multi-isoform genes. The multi-isoform genes were filtered out from the analysis.

2.9. Data accessibility

All supplementary files have been made available via the Mendeley Data repository (DOI: https://data.mendeley.com/datasets/mhj4v5tz9n/draft?a=f58a2c77-8811-48f1-94d6-629034842355). The supplementary information for this manuscript contains Supplementary Figs. S1–S11, Supplementary Tables S1–S15, Supplementary Data_S1, Supplementary Data_S2 and their descriptions. The search parameters, raw files, and search results from MaxQuant can be accessed via the ProteomeXchange Consortium via

the PRIDE partner repository with the dataset identifier **PXD016863**.

3. Results

We performed high-resolution proteomic analyses across in vitro encystation (Hehl et al., 2000) in *G. duodenalis* WB isolate (WB-1B, Upcroft) (Chen et al., 1994), and compared trophozoites (TY), low-bile primed (LB) trophozoites, 16 h post-induction of encystation (EC) and mature cyst (C) stages (72 h post LB). We reproducibly identified (at least in two replicates per condition) 3863 non-redundant proteins (Supplementary Fig. S1A-B, Supplementary Table S1), representing 64% of coding genes for the WB isolate (Aurrecoechea et al., 2008). This includes detection of 667 proteins for which no prior proteomic data were available (Jedelský et al., 2011; Lauwaet et al., 2011; Faso et al., 2013; Ma'ayeh et al., 2017) and greater than a three-fold increase in proteins quantified during encystation (Supplementary Table S2) (Faso et al., 2013). Principal component analysis (PCA) demonstrated clear separation between life-stages and conditions (Fig. 1A). Approximately 85% of detected proteins were shared between replicates within each condition (Supplementary Fig. S1C). Approximately 82.2% of quantified proteins (2782 proteins) were detected in all replicates and conditions (Fig. 1B).

3.1. Stage and condition-specific proteins associated with encystation

We detected 15, nine, eight and 24 proteins unique in TY, LB, EC and C stages, respectively (Fig. 1B, Supplementary Table S3). TY-specific proteins were mostly annotated as hypothetical (nine of 15) but included a putative RNA methyltransferase (GL_7416), putative NADPH oxidoreductase (GL_17151), NimA related kinases (NEK) (GL_3546), high cysteine membrane protein (HCMP: GL_112135) and phosphatase (GL_36315). Similarly, eight of the nine LB-specific proteins lacked functional annotation. Stage-specific proteins in EC or C were largely previously implicated in encystation (Supplementary Table S3). In EC, these included a palmitoyl transferase (GL_2116) important for sphingolipid biosynthesis, a GlcNAc-PI synthesis protein (GL_113610) required in cyst wall formation (Hernandez et al., 2008; Mendez et al., 2013; Duarte et al., 2019), and a transcription factor (TF) known to regulate encystation, Myb1-like protein (GL_8722) (Sun et al., 2002; Huang et al., 2008). Proteins specific to mature cysts (C) included hypothetical proteins (n = 18), an ATP-dependent DNA helicase (GL_16747), Hop1 (GL_4084), Zinc finger domain protein (GL_103659), a putative Sec20 (GL_5161), a Protein 21.1 (GL_16407) and a variant-specific surface protein (GL_11690). The presence of Hop1 (GL_4084), a part of the synaptonemal complex guiding meiotic recombination (Anuradha and Muniyappa, 2004), and ATP-dependent DNA helicase (GL_16747), is consistent with meiotic events and nuclear division which occur at late encystation or cyst maturation in *Giardia* (Faso et al., 2013; Einarsson and Svärd, 2015; Einarsson et al., 2016).

We identified a further 108 proteins shared between EC and C which were undetected in TY or LB (Fig. 1B; Supplementary Table S3). These were enriched for GPI anchor biosynthetic process (GO:0006506 and ko00563), Biosynthetic process (GO:0009058) and Metabolic pathways (ko01100) (Supplementary Table S3), and include the meiotic markers, Mnd1 (GL_6626) and Rad51 (Gl_1304), as well as serine palmitoyl transferases (GL_14374, GL_23015), ceramide glucosyltransferase (GL_11642) and GPI

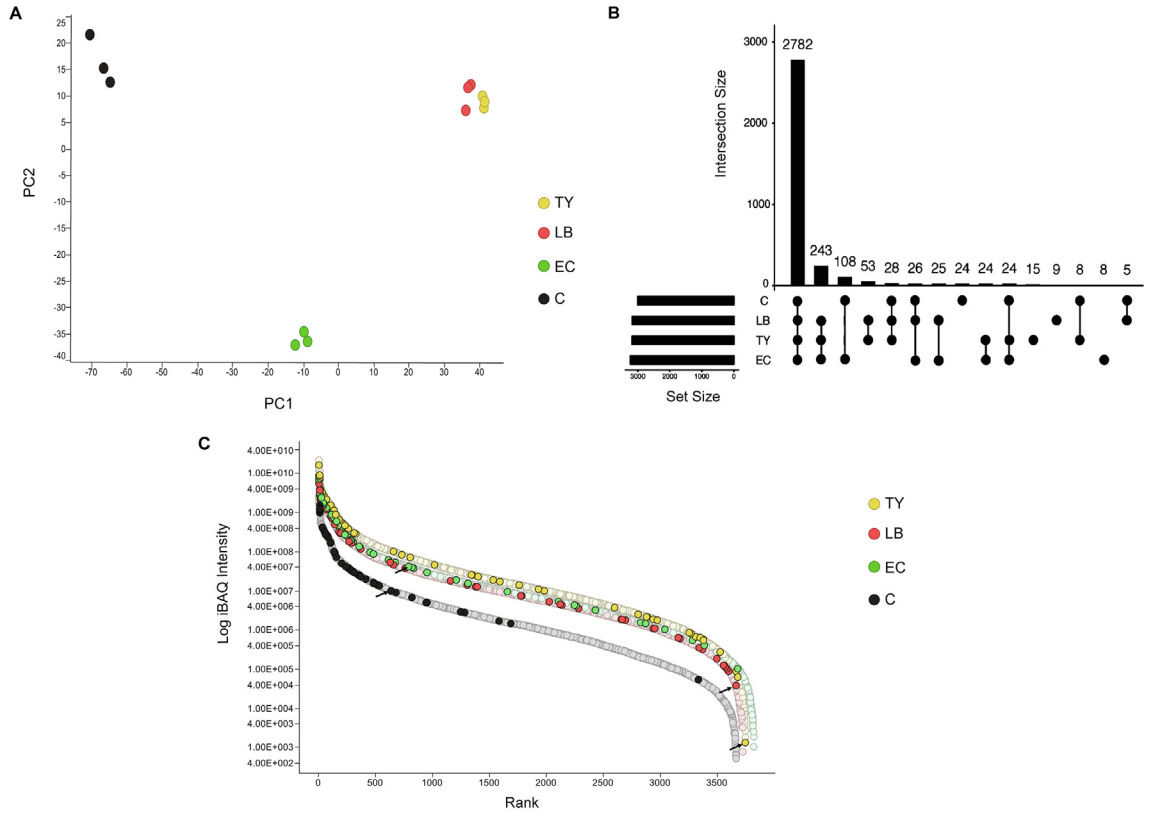

Fig. 1. Deep proteome analysis of *Giardia duodenalis* reproducibly identified 3863 proteins. (A) Principal-component analysis of the first two components of the proteome during encystation with each colour representing a stage in encystation and its replicates. (B) UpSET plot depicting proteins shared across different stages of encystation. (C) Proteins ranked by log iBAQ (Intensity-Based Absolute Quantitation intensity values, with each colour showing each stage in encystation, with the coloured circles in each stage representing encystation markers (depicted in full in Supplementary Fig. S2A–D) with major encystation marker CWP2 (black arrows) shown.

anchor biosynthesis enzymes (GL_17502, GL_16598, GL_3180, GL_12174).

3.2. Changes in ranked abundance of encystation markers

We assessed the progress of in vitro encystation using the mean abundance (iBAQ intensity) of 52 proposed encystation markers (ECMs) (Faso et al., 2013; Einarsson et al., 2016) (Supplementary Table S4) across all experimental conditions (Fig. 1C, Supplementary Fig. S2A–D, Supplementary Table S5), and performed pairwise Fisher permutation tests which revealed higher ranked abundance of this group in EC and C compared with TY and LB ($P = 6.3*10^{-26}$ and $5.2*10^{-29}$, respectively) (Supplementary Table S6). Most ECMs (57.6%) were significantly more highly ranked (within top 500 rank) upon commencement of encystation (Fig. 1C, Supplementary Fig. S2A–D, Supplementary Table S5), with 27 differentially up-regulated in EC and 25 in C compared with LB (Supplementary Fig. S2E). Later into encystation, 17 ECMs were up-regulated and 26 ECMs down-regulated in C compared with EC (Supplementary Fig. S2E). Five ECMs, including two hypothetical proteins (GL_9355, and GL_8377), an aldose reductase (GL_7260), Myb1-like protein marker (GL_8722) and a protein 21.1 (GL_102813), may be more accurately described as EC stage markers given their peak upregulation in EC compared with LB and in C (Supplementary Fig. S2E).

3.3. Differentially expressed proteome suggests complex regulation of encystation

Based on overall differential expressional analysis, few differences were observed between growing (TY) and low-bile primed (LB) trophozoites. Two proteins were up-regulated, GL_15000 (TFIIH P44) and GL_15276 (Glucosamine-6-phosphate isomerase, putative), and three proteins down-regulated in LB relative to TY, GL_41472 (a VSP), GL_16779 (Cathepsin B) and GL_11470 (a VSP with candidate initiator element) (Fig. 2A, Supplementary Table S7). However, in EC 178 proteins were up- and 65 down-regulated compared with LB (Fig. 2B, Supplementary Table S7). Up-regulated proteins in EC were enriched for Gene ontology (GO) terms (Supplementary Table S8) UDP-N-acetylglucosamine biosynthetic process, proteolysis involved in cellular protein catabolic process and lipid metabolic process, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways Amino sugar and nucleotide sugar metabolism and Metabolic pathways. Proteins down-regulated in EC were enriched for cell adhesion, rRNA processing and lipid transport, but no KEGG pathways (Supplementary Table S8). In C relative to EC, 114 proteins were up- and 656 proteins down-regulated (Fig. 2C, Supplementary Table S7). Up-regulated proteins were enriched for the GO term, translation and KEGG pathway term Ribosome. Among down-regulated proteins, C-enriched GO terms were cell redox homeostasis, protein folding, nucleoside metabolic process, ubiquitin-dependent protein catabolic process and enriched KEGG pathway terms were Carbon metabolism, Biosynthesis of antibiotics, Biosynthesis of amino acids, Biosynthesis of secondary metabolites, Metabolic pathways, Proteasome, Glycolysis / Gluconeogenesis, Pentose phosphate pathway and Arginine biosynthesis (Supplementary Table S8). Collectively, down-regulated GO and KEGG terms indicate decreases in protein translation, which is supported by observed decreases in the overall proteome; approximately half (52%) of the down-regulated proteins are at least four-fold lower in C versus EC stages, compared with just 29% up-regulated proteins exceeding this level of difference in C versus EC.

We further compared changes in protein expression between EC or C to TY. This identified 183 proteins up- and 75 proteins down-regulated at EC relative to TY (Fig. 2D, Supplementary Table S7). Up-regulated proteins were enriched for GO terms UDP-N-acetylglucosamine biosynthetic process and lipid metabolic process, and KEGG pathways terms Amino sugar and nucleotide sugar metabolism and Metabolic pathways. Down-regulated proteins were enriched for GO terms, cell adhesion and lipid transport (Supplementary Table S8). The proteomic changes were more pronounced in C compared with TY, with 336 proteins up- and 833 (24.4% of all expressed proteins) down-regulated (Fig. 2E, Supplementary Table S7). Up-regulated proteins in C compared with TY were enriched for GO terms UDP-N-acetylglucosamine biosynthetic process, translation, ER-associated ubiquitin-dependent protein catabolic process, positive regulation of RNA polymerase II transcriptional preinitiation complex assembly and lipid metabolic process and enriched for KEGG pathways Ribosome and Amino sugar and nucleotide sugar metabolism (Supplementary Table S8). Down-regulated proteins were enriched for GO terms including cell redox homeostasis, protein folding and ubiquitin-dependent protein catabolic process, and KEGG pathways including Carbon metabolism, Biosynthesis of antibiotics, Biosynthesis of amino acids, Biosynthesis of secondary metabolites and Metabolic pathways (Supplementary Table S8). Protein-Protein interaction analysis among up-regulated and down-regulated proteins within C relative to TY showed networks previously unmapped in encystation in existing RNA-seq or proteomic studies (Faso et al., 2013; Einarsson et al., 2016; Supplementary Fig. S3A–B).

3.4. Proteomic changes in major functional classes through encystation

In addition to pair-wise comparisons of protein expression between conditions, we performed one-way ANOVA across the four experimental conditions, focusing on differentially expressed gene sets related to encystation (Fig. 3A–J), TFs, epigenetic modifiers (histone modifying enzymes) (Einarsson et al., 2016), kinases, complex cell-surface protein families (HCMPs, HCPs and VSPs), and secreted proteins (Faso et al., 2013; Dubourg et al., 2018). Fischer permutation-based gene set enrichment testing also showed these functional classes were significantly enriched in at least one condition across encystation (Supplementary Table S6). Overall, despite the few differentially expressed proteins in pairwise comparisons, clustering of proteins with divergent expression across all conditions by ANOVA suggested down-regulation of cell surface and secreted proteins upon LB priming (Fig. 3C-D, Supplementary Table S9, Supplementary Figs. S4C-D, S5C-D). This trend was particularly noticeable for VSPs (Supplementary Fig. S4A). Most functional classes had the lowest overall expression in cysts, with the exception of active NEK kinases and lipid metabolic enzymes, which remained stable or increased in overall expression through to cyst maturation. Interestingly, glycolytic and lipid metabolic classes appear to be bimodal in both the TY and C stages based on trend analysis (Supplementary Fig. S5I–J), indicating functional switching within these classes through encystation. Clustering analysis (Supplementary Fig. S4) showed bimodal expression within many of these significant functional classes. EC appears to mark an inflection point, with most members of each gene set either most highly expressed in TY and LB stages and dropping off by C, or markedly increasing in EC and remaining high in C (Fig. 3A–J; Supplementary Table S9, Supplementary Figs. S4A–J, S5A–J).

Einarsson et al. (2016) reported extensive transcriptional regulation during encystation, particularly through differential expression of TFs and chromatin modifying enzymes (CMEs). Fourteen of 16 *Giardia* TFs are transcribed during encystation, however their functions are little understood (Sun et al., 2002; Pan et al., 2009; Morf et al., 2010; Einarsson et al., 2016). Early encystation (<10 h post-induction) coincides with highest transcription of CCR4-NOT, CCAAT-binding, GARP-like, PAX and ARID2 TFs (Einarsson

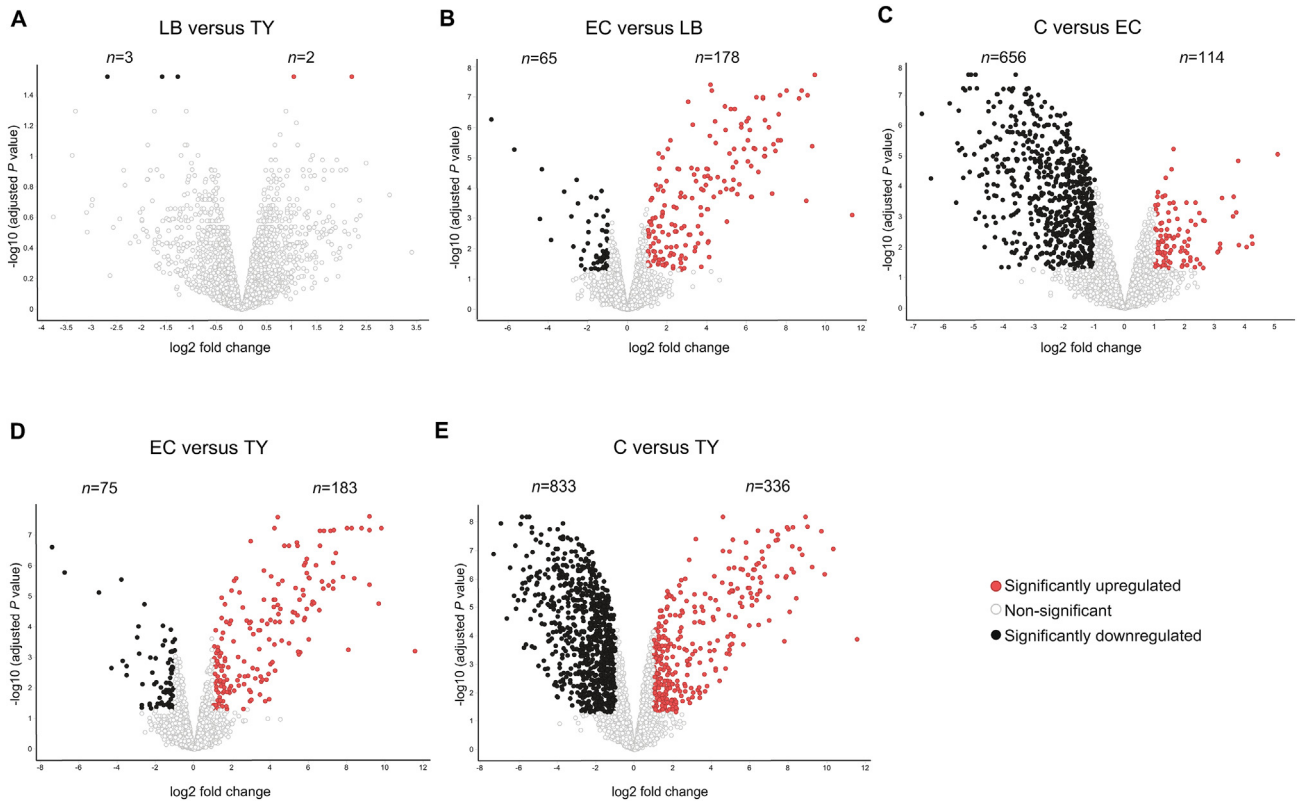

Fig. 2. Differentially expressed proteins across *Giardia duodenalis* encystation. Volcano plots depicting $-\log_{10}$ Benjamini-Hochberg (BH) adjusted *P* values versus log2fold change (FC) across each contrast (A–E). Proteins with $-\log_{10}$ BH-adjusted *P* value >1.3, log2FC > 1 fold and log2FC < −1 were considered statistically differentially expressed. TY, trophozoite; LB, low bile primed trophozoites; EC, 16 h into encystation; C, mature cyst.

et al., 2016), while mid- to late encystation (>22 h post-induction) coincides with an apparent switch towards increased transcription of Myb1-like protein, E2F, TFIIS, NOT4 and WRKY TFs, and reduced transcription of other TFs (Einarsson et al., 2016). Although no TFs were detected in previous encystation proteomes (Faso et al., 2013), we detected Myb1-like (GL_8722), Repressor NOT4Hp (GL_8727), TFIIS (GL_5158), and CCR4-NOT transcription complex subunit 7 (GL_10606) (Fig. 3E, Supplementary Table S9). The Myb-1 like TF (GL_8722) was up-regulated in EC relative to C, consistent with existing transcript data (Einarsson et al., 2016). In contrast, TFIIS (GL_5158) expression was down-regulated in C relative to EC, and the transcriptional repressor NOT4Hp (GL_8727) was down-regulated in EC and C relative to TY and LB, both of which negatively correlate with transcription (Einarsson et al., 2016).

CMEs (Lagunas-Rangel and Bermúdez-Cruz, 2019) are implicated in encystation (Sonda et al., 2010; Lagunas-Rangel and Bermúdez-Cruz, 2019) and drug resistance (Emery et al., 2018). Einarsson et al. (2016) reported peak transcriptional abundance in mid-encystation to cyst maturation of nine CMEs with roles in histone acetylation and methylation. Of the 27 CMEs identified in the *Giardia* genome, 18 were quantified here (Fig. 3F, Supplementary Table S9) and most consistently decreased in expression from TY to C (Fig. 3F, Supplementary Figs. S4F, S5F). Eleven CMEs were differentially expressed during encystation (Fig. 3F, Supplementary Table S9), with three, including a histone acetyltransferase (GL_14753), arginine deiminase (GL_112103) and a Sirtuin deacetylase (GL_10708), down-regulated in C relative to TY, LB and EC. Consistent with previous RNAseq data (Einarsson et al., 2016), Sirtuin deacetylase (GL_16569), was up-regulated in EC and C compared with TY and LB. In addition, several structural maintenance chromosomal (SMC) ATPases, SMC1 (GL_16188), SMC2 (GL_23185) and SMC3 (GL_137745), were also up-regulated in C relative to TY. No CMEs were detected at the proteomic level by Faso et al. (2013), but the authors noted increased expression of histone H2A at 8 and 12 h post-induction of encystation. Interestingly, in our study, histones H2A (GL_14256, GL_27521), H2B (GL_121045, GL_121046) and H4 (GL_135001, GL_135002, GL_132003) are all down-regulated in C relative to all other stages.

Giardia has a minimal core kinome (n = 80) and an expanded NEK kinase family (n = 217) (Manning et al., 2011). The NEK kinome is divided into 53 predicted catalytically active kinases and 138 that lack conserved residues of the catalytic triad and are considered inactive or 'dead' enzymes. All 301 kinases annotated in *G. duodenalis* WB are transcribed in the trophozoite or during encystation (Einarsson et al., 2016) and 223 were quantified as proteins here. Ninety-four kinases, including 27 active and three 'dead' core kinases, and 20 active and 44 'dead' NEKs, were differentially expressed (Fig. 3G–H, Supplementary Table S9). ANOVA results suggested that both active and 'dead' kinases undergo major switching in expression during encystation, with one protein cluster in each group that is progressively down-regulated and another that is progressively up-regulated throughout encystation across LB, EC and C stages (Supplementary Fig. S4G–H). Although this did not result in a statistically significant change in overall expression of active kinases by a Fisher test (i.e., their relative changes in abundance across encystation cancel out), there was a clear, statistically significant down-regulation of 'dead' NEKs in C relative to TY (Supplementary Fig. S5G–H).

Giardia encodes complex families of cysteine-rich, variant surface proteins, including VSPs, HCMPs and high cysteine proteins (HCPs), which are important in oxidative stress defence, antigen-switching and encystation (Davids et al., 2006; Prucca et al., 2008; Emery et al., 2018). Of the 161 VSPs transcribed during

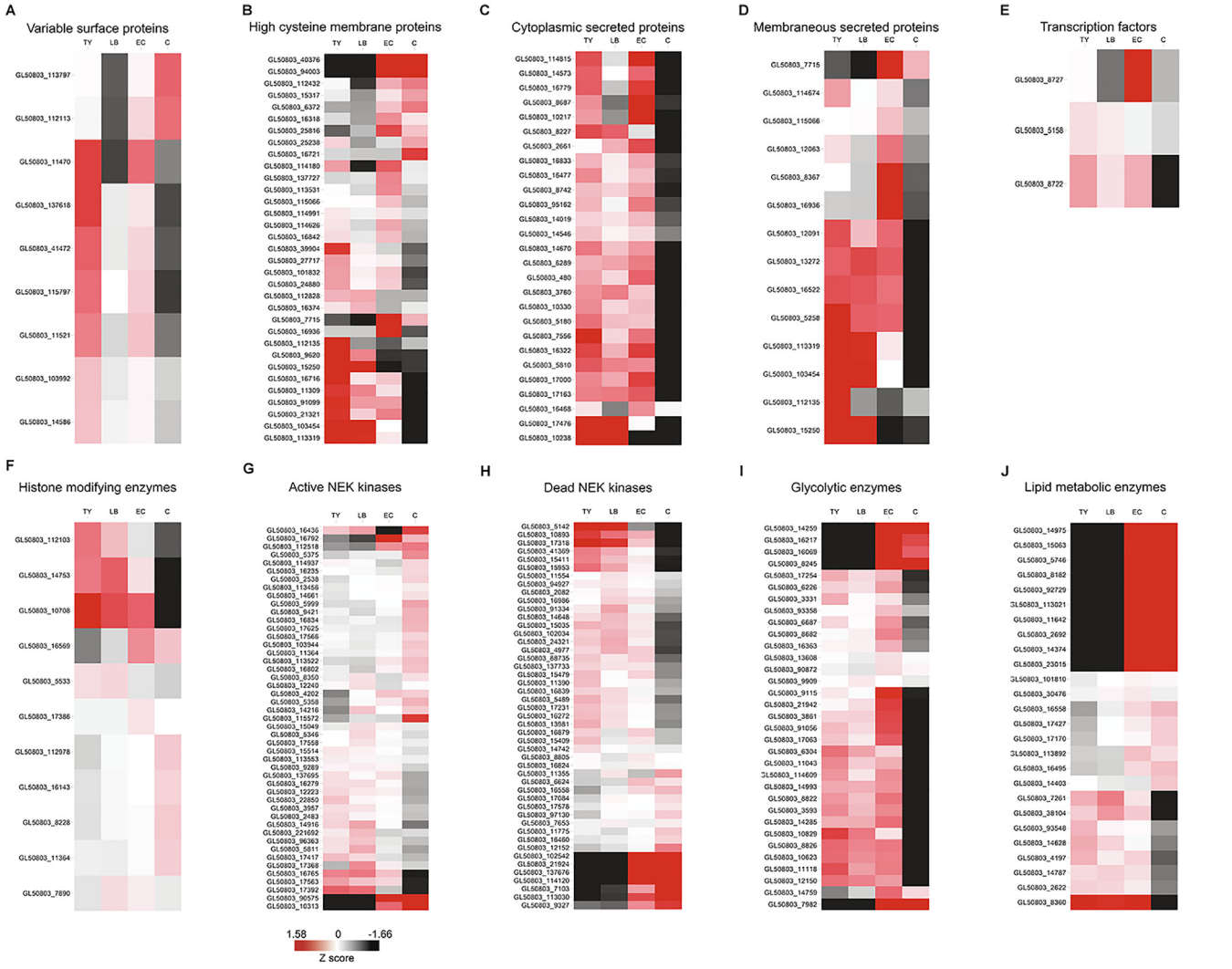

Fig. 3. Differential expression patterns of seven major functional groups across the *Giardia duodenalis* encystation process. Hierarchical cluster analysis of protein expression Z scores computed for significantly differentially expressed proteins identified by a multi-sample ANOVA test Benjamini-Hochberg (BH)-corrected $P < 0.05$) for the eight major functional groups. Each heatmap represents by the order (A) variable surface proteins, (B) high cysteine membrane proteins, (C) cytoplasmic and (D) membraneous secreted proteins, (E) transcription factors, (F) histone modifying enzymes, (G) active and (H) dead NimA related kinase (NEK) kinases, (I) glycolytic enzymes and (J) lipid metabolic enzymes. The colour version is available on-line.

encystation (Einarsson et al., 2016), we quantified 26 and observed nine were differentially expressed. VSP expression was typically highest in the trophozoite stage, with seven up-regulated in TY relative to the other the stages (Fig. 3A, Supplementary Table S9). Two VSPs, GL_41472 and GL_11470, were down-regulated in LB compared with TY, with GL_11470 then up-regulated in EC relative to LB. GL_112113 was up-regulated in both EC and C relative to TY or LB. VSP expression was typically lowest in the mature cyst, however GL_113797 and GL_112113 were up-regulated in C only relative to all other stages (Fig. 3A, Supplementary Table S9). Of 86 HCMPs and HCPs identified at the transcript level (Einarsson et al., 2016), 38 were quantified here, including 28 identified previously in encystation (Faso et al., 2013). Thirty-three of these were differentially expressed (Fig. 3B, Supplementary Table S9). In contrast to VSPs, HCMP/HCP expression was higher overall in EC/C stages compared with TY or LB ($P < 2.63^{*}10^{-8}$) (Supplementary Table S6). Three (GL_137727, GL_113531, GL_16936) were up-regulated in EC relative to LB and another eight were up-regulated in EC and C relative to LB (Fig. 3B, Supplementary Table S9). One HCMP, GL_16721, was up-regulated and 18 HCMP and HCPs down-regulated in C relative to EC (Fig. 3B, Supplementary Table S9).

The early encystation proteome (Faso et al., 2013) implicated changes in secretory machinery during encystation, when encystation secretory vesicles are essential in cyst wall formation (Stefanic et al., 2006; Einarsson and Svärd, 2015). *Giardia* secretory products are not fully defined. However, of the 69 secreted proteins identified by Dubourg et al. (2018), we quantified 55 here, of which 41 (27 cytoplasmic secreted proteins and 14 plasma membrane-bound proteins and hypothetical secreted proteins) were significantly differentially expressed during encystation (Supplementary Table S9). Of these, only GL_16779 (Cathepsin B precursor), which is down-regulated, significantly changed in LB relative to TY (Fig. 3-C-D, Supplementary Table S9). Ten secreted proteins were up-regulated and six down-regulated in EC relative to LB (Fig. 3C-D, Supplementary Table S9). Thirty-six secreted proteins were down-regulated in C relative to EC, with none significantly up-regulated (Fig. 3C-D, Supplementary Table S9).

3.5. Shift from glycolysis to lipid metabolism during encystation

There are significant metabolic changes during encystation (Jarroll et al., 2011), including in glycolysis and lipid metabolism (Einarsson et al., 2016). We quantified all 37 glycolytic enzymes

previously detected as transcripts (Einarsson et al., 2016), of which 33 were differentially expressed. Although there were no significant differences in glycolytic enzyme expression between LB and TY, eight were up-regulated in EC relative to LB, and most glycolytic enzymes associated with ATP production (Bar-Even et al., 2012) were down-regulated in C relative to EC (Fig. 3I, Supplementary Table S9, Supplementary Figs. S4I, S5I). In contrast, enzymes from the glucosamine biosynthesis pathway required to synthesise Uridine-5′-diphospho-N-acetylgalactosamine (UDP-GalNAc) for the cyst wall (Kwiatkowska-Semrau et al., 2015) (GL_14759, GL_8245, GL_14259, GL_16069, GL_16217, GL_7982) remain up-regulated through EC and C (Fig. 3I, Supplementary Table S9).

Exposure to exogenous lipids such as fatty acids, bile, cholesterol, and sphingolipids promote encystation (Hernandez et al., 2008; Sonda et al., 2008; Štefanić et al., 2010; Yichoy et al., 2011; Mendez et al., 2015; Wong et al., 2019). However, *Giardia* has limited lipid biosynthetic capacity (Gibson et al., 1999; Das et al., 2001) and is largely dependent on phospholipid synthesis via cytidine diphosphate diacylglycerol (CDP-DAG) or the fatty acid remodelling pathway (Das et al., 2001; Yichoy et al., 2009), although it may have some additional, uncharacterised, biosynthetic capacity (Ellis et al., 1996; Yichoy et al., 2009). Einarsson et al. (2016) noted extensive transcriptional changes in lipid metabolism in late encystation, and lipid metabolic pathway and GO terms are enriched in EC and C stages here (see above). There has been no systematic description of proteomics changes in these pathways during encystation or at cyst maturation until now. We explored expression of 44 lipid and fatty acid metabolic genes in *Giardia* (Yichoy et al., 2011), 38 of which were quantified in our proteome (Supplementary Table S9). These genes belong to six functional pathways within lipid metabolism; phospholipid, fatty acid, sterol, neutral lipid, sphingolipids and signalling lipids. Seven of the 11 phospholipid metabolic enzymes in the *Giardia* genome were quantified in the proteome. Among these, phosphatidylserine synthase (GL_17427) and phospholipid-transporting ATPase IA (GL_8182) were up-regulated in EC and C relative to TY and LB, and phospholipid-transporting ATPase IIB (GL_38104) and PI transfer protein alpha isoform (GL_4197) were down-regulated in C relative to TY, LB or EC (Fig. 3J, Supplementary Table S9). Eleven out of 12 fatty-acid metabolising enzymes from *Giardia* (Yichoy et al., 2011) were quantified across encystation, with four significantly up-regulated in both EC and C relative to TY and LB (Fig. 3I, Supplementary Table S9). Among sphingolipid metabolic enzymes, serine palmitoyl transferases (GL_14374, GL_23015) and ceramide glucosyltransferase (GL_11642) were identified only in EC and C (Fig. 1C, Supplementary Table S3), with ceramide glucosyltransferase (GL_11642) and serine palmitoyl transferase 2 (GL_23015, GL_14374) significantly up-regulated in EC and C relative to TY or LB. The only known cholesterol metabolic enzymes in *Giardia* are lecithin-cholesterol acyl transferases (GL_5746 and GL_16286). Of these, GL_5746 was up-regulated in both EC and C relative to TY or LB (Fig. 3I, Supplementary Table S9). Finally, 12

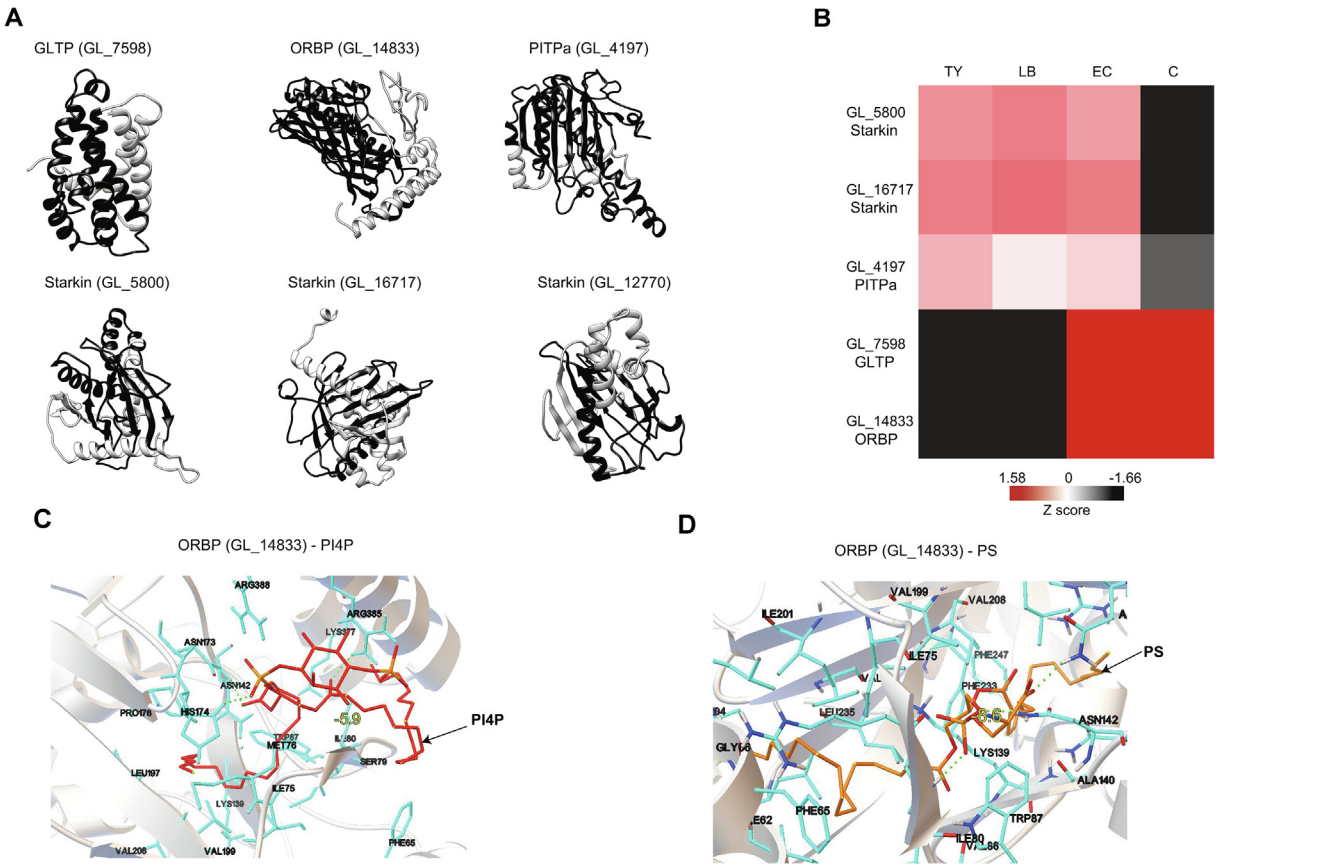

Fig. 4. In silico predicted three-dimensional (3D) structures of *Giardia duodenalis* lipid transporters, their differential protein expression, and respective interactions. (A) iTasser-based high confidence (Root-mean-square deviation (RMSD) score >0.7) in silico predicted 3D structures of *Giardia* lipid transporters identified by HMMER analysis belonging to Oxysterol binding proteins (GL_14833), steroidogenic acute regulatory protein–related lipid transfer (StARkin) (GL_5800, GL_16717 and GL_12770) and glycolipid transport protein (GL_7598) families. (B) Hierarchical cluster analysis of protein expression Z scores computed for significantly differentially expressed lipid transporters identified by a multi-sample ANOVA test (BH-corrected $P < 0.05$) across encystation. Computationally docked structure of oxysterol binding protein (GL_14833) (grey) bound with (C) phosphatidyl inositol 4 phosphate (PI4P) and (D) phosphatidylserine (PS). Binding amino acid residues, hydrogen bonds formed, and respective free energy are shown for each molecule with the corresponding protein. TY, trophozoite; LB, low bile primed trophozoites; EC, 16 h into encystation; C, mature cyst. The colour version is available on-line.

of the 13 signalling lipid genes in the *Giardia* genome were quantifiably identified in our proteome. Phosphatidylinositol 4-kinase (GL_16558) was up-regulated in both EC and C relative to TY or LB. Three phosphatidyl inositol-4-phosphate 5-kinases and a type II inositol-1,4,5-trisphosphate 5-phosphatase (GL_14787) were down-regulated in C compared with EC, LB or TY and neutral lipid Phospholipase B in C relative to EC (Fig. 3I, Supplementary Table S9). Clustering of lipid metabolism enzymes, significant by ANOVA, indicated two clusters, one which gradually increased from TY through to C, and the other which gradually decreased in expression from TY to C (Supplementary Figs. S4J, S5J).

3.6. Putative PS – PI4P shuttling in Giardia during encystation

Considering the limited capacity of *Giardia* for de novo lipid synthesis (Yichoy et al., 2011) and the dramatic shift in lipid metabolic enzymes during encystation, we hypothesised that *Giardia* must have lipid transporter proteins (LTPs). We identified six putative LTPs in *Giardia* through in silico sequence-based domain mining and 3D protein structural modelling (Supplementary Data_S2). These include four members of the Starkin superfamily, including three StART proteins (GL_5800, GL_16717 and GL_12770) and one (GL_4197) phosphatidyl inositol transfer protein (PITP), one (GL_7598) glycolipid transfer protein (GLTP) and one (GL_14833) of the oxysterol binding protein related (ORP) protein families (Fig. 4A) (Supplementary Data_S2). We quantified all LTPs, except GL_12770, during encystation (Fig. 4B). Overall, all putative Starkin superfamily LTPs (GL_4197, GL_5800 and GL_16717) are down-regulated in C relative to all other stages. In contrast, the GLTP (GL_7598) and ORP (GL_14833), were up-regulated during EC and remained elevated in C relative to TY or LB (Fig. 4B). GL_14833 is structurally homologous (exact match prediction score = 0.858) to osh6p (Supplementary Data_S2) (Maeda et al., 2013), a bispecific transporter exchanging phosphatidylinositol-4-phosphate (PI4P) and phosphatidylserine (PS) between the endoplasmic reticulum (ER) and plasma membrane (PM) (Maeda et al., 2013; von Filseck et al., 2015). We performed computational docking which confirms GL_14833 interactions with 18:0/20:4 PI4P and PS in silico (Fig. 4C, D) (Supplementary Data_S2). This together with the presence of novel SACI homologs (GL_8589 and GL_16356) and PI4 kinase (GL_16558) which we identified in silico (Supplementary Fig. S6; Supplementary Fig. S7; Supplementary Data_S2), and differential upregulation of all these components specifically in EC and C relative to TY or LB (Fig. 5A) points to the possibility of an active PI4P-PS shuttling in *Giardia* (Fig. 5B).

3.7. Giardia duodenalis transcriptome-proteome correlation

Although prior studies (Wang et al., 2007; Einarsson et al., 2016) implicate transcriptional and epigenetic regulation in encystation, *Giardia* has loose chromatin condensation (Tůmová et al., 2015), few transcription factors (Teodorovic et al., 2007) and stochastic transcription (Franzén et al., 2013). This prompted us to investigate additional potential mechanisms of gene expression regulation. First, to account for biological variation between strains WB1B (proteomics) and WBC6 (transcriptomics), we analysed the relationship between transcript abundance in the respective trophozoite stages (Ansell et al., 2015; Einarsson et al., 2016), and found relatively strong correlation (R = 0.88, $P < 2.2e^{-16}$) (Supplementary Fig. S8). Following that, we focused on the TY and C stages, for which both RNA-seq (Einarsson et al., 2016) and high-resolution proteomic data are available. We considered Pearson correlations between RNA transcription (FPKM) and protein expression (iBAQ) for all genes, and within transcriptional quintiles. Correlations between RNA and protein for TY were R = 0.6 ($P < 2.2e^{-16}$) but notably weaker for C (R = 0.37, $P < 2.2e^{-16}$) (Fig. 6-A-B, Supplementary Fig. S9A-B). When divided into transcriptional quintiles, both TY and C showed similarly weak correlation between transcript and protein abundance among the lower (first to fourth) quintiles, whereas the correlation in the top (fifth) quintile diverged notably between TY (R = 0.61, $P < 2.2e^{-16}$) and C (R = 0.16, $P = 1.9e^{-05}$) (Supplementary Fig. S10A-B). In addition to lower correlation between biotypes in the cyst, we noted a significant difference in the number of transcribed genes not detected as proteins in the top transcriptional quintile in cysts (n = 272/1156; 23.5%) relative to trophozoites (n = 141/1162; 12.1%, chisq $P < 10E-11$) (Supplementary Fig. S11A-B, Supplementary Table S10).

In addition to protein abundance, we considered other potential technical causes for under or undetected proteins among the top decile of most abundant transcripts. Amino acid composition may impact accurate identification of peptides by LC-MS in some instances. In *Giardia*, CWP1 and CWP2, which are LYS and ARG rich, are reportedly difficult to detect in proteomic datasets (Stefanic et al., 2006). Cyst wall proteins, including CWP1 and CWP2, were abundant in our dataset, indicating against this issue. Nonetheless, we compared amino acid composition of under/undetected proteins among the top-decile transcripts relative to readily detected proteins in the same transcript decile. We found no evidence that under-expressed or undetected proteins were notably enriched for LYS or ARG, or differed significantly in amino acid composition relative to well-detected proteins from the same transcript decile. Similarly, we compared codon usage among these groups and found no significant differences. In addition, we looked at the behaviour of 50 *Giardia* genes, reported previously as being naturally codon optimized (Lafay and Sharp, 1999). Of these, 45 were detected in our dataset, but they showed no association with highly expressed versus under-expressed or undetected proteins among top-decile transcripts, and were seen to be distributed across the scale of transcript and protein relative abundance (Supplementary Table S11) in both TY and C. Finally, in considering under-represented proteins, we considered the possibility that they represent multi-copy genes for which the transcripts have been applied to one paralogue and the proteins another. *Giardia* encodes few multi-copy genes (Morrison et al., 2007; Franzén et al., 2013; Einarsson et al., 2016). Nonetheless, we manually curated all genes to ensure that under-expression was not an artefact of the RNA-seq reads being attributed to one isoform and the peptides to another.

In eukaryotic cells there is growing evidence for post-transcriptional regulation during stage transitions/development, in particular translational repression, through which messenger RNAs are captured and held in stasis for later translation (Le Roch et al., 2004; Tarun et al., 2008; Hentze et al., 2018; Becker et al., 2018; Garfoot et al., 2019; Wang et al., 2019). In *Plasmodium* and *Toxoplasma*, translational repression primes the cell for transmission between hosts (Vembar et al., 2016; Lindner et al., 2019; Muller et al., 2019). *Giardia* encodes homologs of known translational repressors (Williams and Elmendorf, 2011; Gargantini et al., 2012), including 32 SF2 helicases (DEAD Box, DEAH helicases and Ski2), 11 eukaryotic initiation factors, seven eukaryotic elongation and five Pumilio homology proteins (Puf) (Williams and Elmendorf, 2011). SF2 helicases have been implicated in regulating *Giardia* cyst formation (Gargantini et al., 2012). We quantified 46 translational repressors, including 25 SF2 helicases, three PUF proteins, six elongation and 12 initiation factors. Eight of these were differentially expressed (Fig. 6E, Supplementary Table S12). Gargantini et al. (2012) observed significant up-regulation in most SF2 helicases during encystation (at 16 h post-induction; corresponding to the EC stage here). Interestingly, these observations are not reflected in our protein expression data nor in more recent RNA-seq data (Einarsson et al., 2016). Three SF2 helicases were

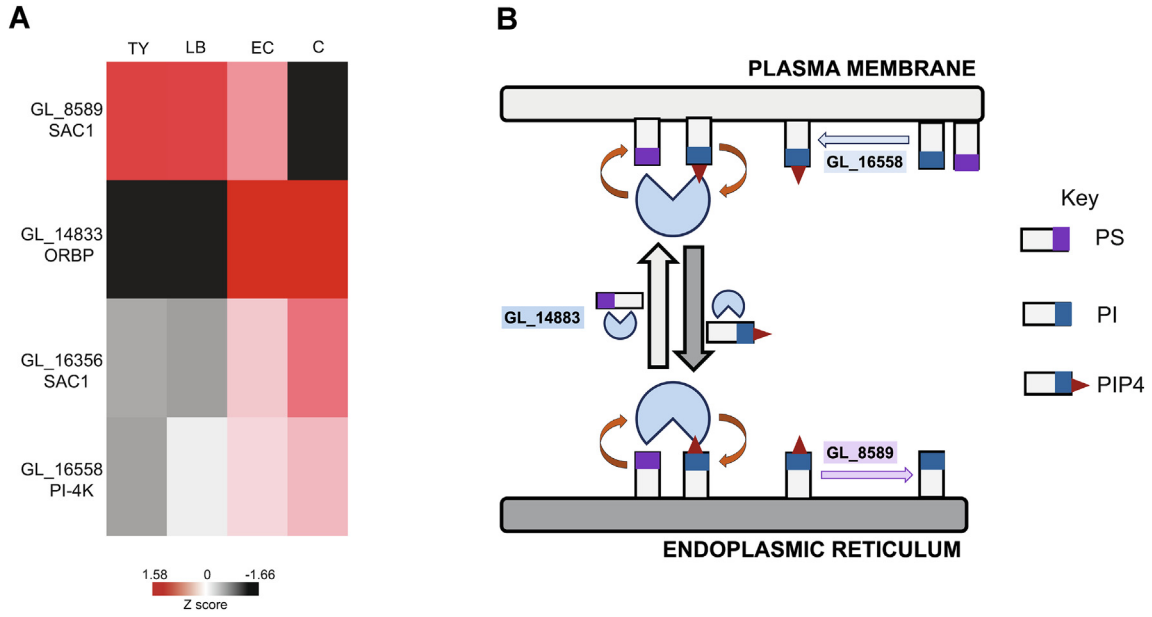

Fig. 5. Proposed Phosphatidylserine/Phosphatidylinositol 4-phosphate (PS/PI4P) exchange cycle at endoplasmic reticulum-plasma membrane contact sites in *Giardia*. Heat map depicting hierarchically clustered protein expression Z scores of in silico identified *Giardia* homologues of oxysterol binding protein (GL_14833), PI4 kinase (GL_16558) and the PI4P phosphatase (GL_8589/GL_16356) (Benjamini-Hochberg (BH) corrected $P < 0.05$) (A). The proposed model for PS/PI4P exchange cycles employing oxysterol binding protein (GL_14833), PI4 kinase (GL_16558) and the PI4P phosphatase (GL_8589/GL_16356) together maintaining the PI4P gradient necessary for efficient counter-exchange (B). TY, trophozoite; LB, low bile primed trophozoites; EC, 16 h into encystation; C, mature cyst. The colour version is available on-line.

differentially expressed during encystation in our study and all were down-regulated compared with TY and LB stages (Fig. 6E). DEAD box helicases, GL_13220 and GL_95898, were down-regulated in EC and C relative to LB, and in C relative to TY and LB, respectively. GL_17387, a DEAH helicase, was down-regulated in C relative to LB (Fig. 6E). Similarly, among eukaryotic initiation factors (EIF), EIF2G (GL_2970) was down-regulated in C relative to LB and EC, EIF5A (GL_14614) was down-regulated in C relative to LB, and EIF1A (GL_8708) in C relative to TY (Fig. 6E). In contrast, a Pumilio homology protein (GL_17590) was up-regulated in EC and C relative to LB, and in C relative to TY, and the elongation factor 2 (GL_12055) in C relative to TY (Fig. 6E). These data suggest complex changes in expression of translational repressors in encystation, mainly PUF proteins, in particular, known to play a key role in regulating transmissive stages in other parasitic protists (Miao et al., 2010; Wang et al., 2018). This led us to further explore which transcripts might be translationally repressed during encystation.

To identify genes that might be under post-transcriptional regulation in *Giardia*, we focused on those that were transcribed but undetected, or under-abundant as proteins. Genes present as proteins but not transcripts, or over-abundant as proteins relative to their transcription can be found in Supplementary Table S13 but are not dealt with further here. Due to the lower sensitivity of MS-based proteomics (Timp and Timp, 2020), for each stage we focused on genes found in the top decile of transcription or readily detectable as transcript and protein but differing in relative abundance. For TY, 55 (9.8%) of the top decile (n = 560) of the most highly transcribed genes were not detected as proteins in any replicate (Supplementary Table S14). In addition, 141 genes in TY were significantly ($P < 0.05$) under-expressed relative to their transcription (Fig. 6A, Supplementary Table S15). For C, 140 (25 %) of the top decile (n = 560) most highly transcribed genes were undetectable as proteins in any replicate (Supplementary Table S14) and 100 genes were under-expressed relative to their transcription (Fig. 6B, Supplementary Table S15). We also looked at dynamic changes in transcription versus expression by ranking genes based on fold-change between TY and C (Fig. 6C). Again, to minimize the impact of LC-MS sensitivity, we took the top 10% of genes by fold-change (~620 genes) based on both RNA-Seq and proteomic data, between TY and C stages. Approximately one-third of these genes changed in diverging directions between TY and C at the transcript versus protein level. Of these, 110 genes were transcriptionally over-represented in terms of fold-change, but under-represented at protein level in TY compared with C. Another 106 were transcriptionally over-represented in terms of fold-change, but translationally under-represented in C compared with TY (Fig. 6D, Supplementary Table S13).

Following that, we looked for functional enrichment in these putatively, translationally repressed genes. For TY, undetected proteins were enriched for the GO term "ATP hydrolysis coupled proton transport" (GO:1902600: $P = 5.3*10^{-2}$) and the KEGG pathway "oxidative phosphorylation" (gla00190: $P = 5.3*10^{-2}$). Under-expressed proteins for TY were primarily hypothetical proteins (n = 85), protein 21.1 (n = 8), VSPs (n = 11) and NEK kinases (n = 7) (Supplementary Table S15) but were nonetheless enriched for the KEGG pathways for "nucleotide excision repair" (gla03420) and "basal TFs" (gla03022). Interestingly, four out of the 52 ECMs, including cyst wall protein 2 (GL_5435), UDP-glucose-4-epimerase (GL_7982), HCMP (GL_40376) and ankyrin repeat protein 1 (GL_24412), as well as epigenetic modifier histone acetyltransferase MYST2 (GL_2851) and transcription factors (GL_13512, GL_4328), are among the under-expressed protein subset in TY (Fig. 6B, Supplementary Table S15). For C, undetected proteins in the top decile of transcripts were enriched for the GO term "spliceosomal complex assembly" (GO:1902600: $P = 8.7*10^{-2}$) and KEGG pathway "spliceosome" (gla00190: $P = 5.2*10^{-3}$). The 100 under-expressed proteins in C were also comprised primarily of hypothetical proteins (n = 54), NEK kinases (n = 3), VSPs (n = 2) and protein 21.1 s (n = 5) (Supplementary Table S15) but were enriched for the GO term "developmental process", and included histone modifiers (GL_10666, GL_14753), ubiquitin system components (GL_13708, GL_4083) and post-transcriptional regulators including an RRM_domain containing protein (93463)

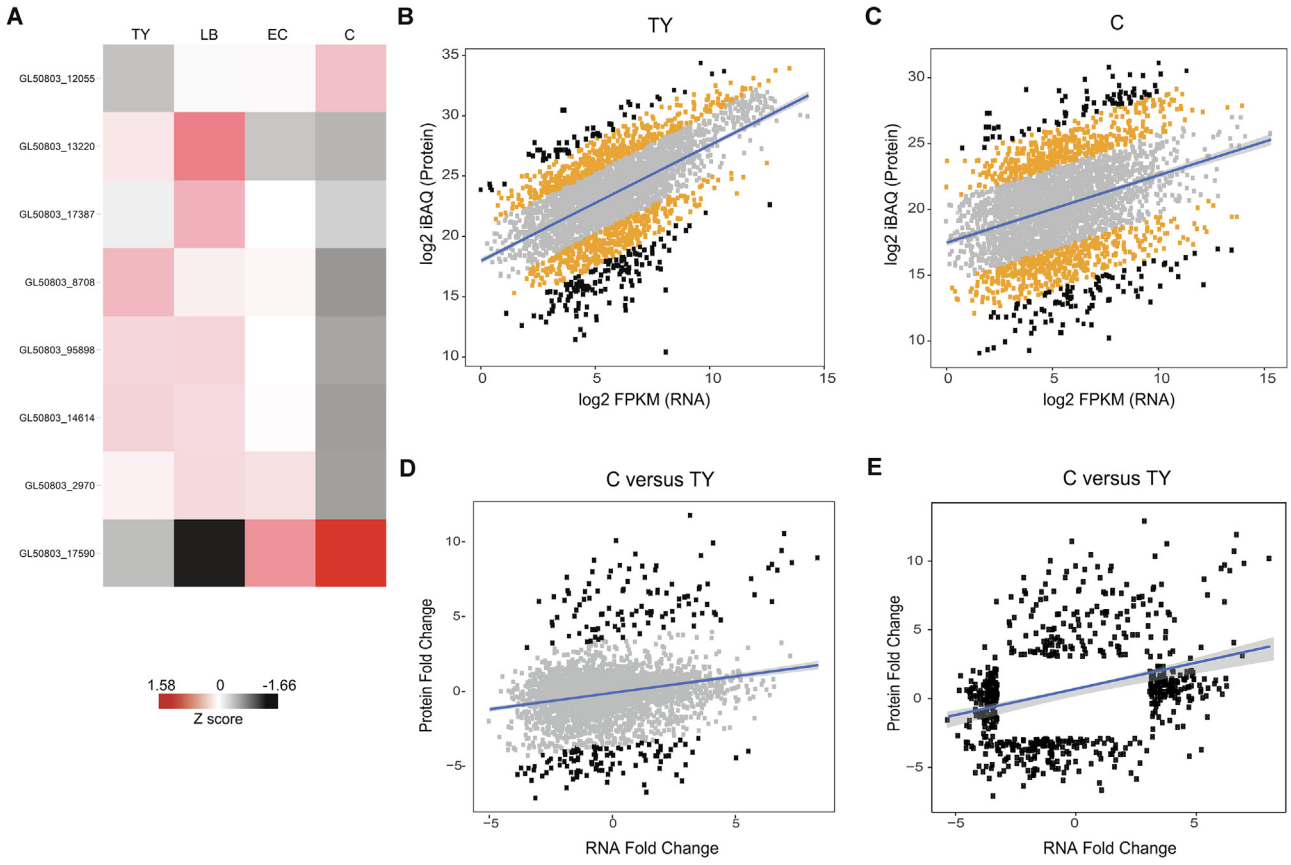

Fig. 6. Transcriptome-proteome correlation analysis within the *Giardia duodenalis* trophozoite and cyst. Scatter plot depicting transcriptome-proteome correlation analysis of log2 (fragments per kilobase of transcript per million mapped reads) (FPKM) from Einnarsson et al. (2016) (X axis) with log2(intensity-based absolute quantitation) (iBAQ) intensity (Y axis) from this study across TY (A) and C (B), with residualized outlier transcript-protein pairs depicted in (black; $P < 0.05$). Fold change correlations between TY and C are displayed in C, with statistically significant residualized outliers displayed in black ($P < 0.05$). The top and bottom 10% of genes ($n = 620$) from C is displayed in D. Heat map depicting hierarchically clustered protein expression Z scores of significantly differentially expressed translational repressors identified by multi-sample ANOVA test (Benjamini-Hochberg (BH) -corrected $P < 0.05$) (E). TY, trophozoite; LB, low bile primed trophozoites; EC, 16 h into encystation; C, mature cyst. The colour version is available on-line.

and a tRNA-methyltransferase (GL_8691) (Fig. 6C, Supplementary Table S15). Among divergently expressed proteins, those translationally under-represented in TY were enriched for GO terms "translation" and "fatty acid biosynthesis", and KEGG pathways "nucleoside excision repair", "ribosome" and "DNA replication" (Supplementary Table S13), and those under-represented in C were enriched for the GO term "cell redox homeostasis".

4. Discussion

During its simple biphasic life cycle *Giardia* massively reconfigures its gene expression without cytokinesis, resulting in a highly infectious, metabolically quiescent and environmentally resilient cyst (Einarsson and Svärd, 2015). Understanding molecular mechanisms leading to encystation, maturation and survival in the environment, as well as transmission to and infection of a new host, can provide important targets for novel anti-giardial drugs, as well as broad insights into regulatory processes driving cell differentiation at the base of the eukaryotic ancestral tree.

We performed the most high-resolution proteomic analysis to date of encystation in *Giardia*, including new data covering early changes to the trophozoite during low-bile stimulation of in vitro encystation, and the first proteomic data from encystation to cyst maturation. Despite its importance in stimulating encystation in vitro, and presumed relevance in vivo, low-bile priming had a limited effect on the trophozoite proteome at the individual gene level. However, we did note significant reductions in overall expression of cell-surface proteins (including VSP and HCMPs) and secreted proteins. Cholesterol, sphingolipids and other fatty acids play an important role in *Giardia*'s secretory and cell-surface expression systems (Yichoy et al., 2011). It may be that low bile contributes to encystation by stimulating a shift in the lipid-dependent secretion system, which is activated in encystation and essential for cyst formation (Luján et al., 1995; Hehl and Marti, 2004). Alternatively, our observations may be a consequence of withholding cholesterol in vitro and may be a sign of cholesterol starvation, rather than a mechanism of encystation. It is possible also that low-bile priming initiates transcriptional changes that are not detectable as proteins until encystation is initiated. Low bile conditions may also stimulate changes at the metabolic or post-translational level, which are not measurable in our data set and are limited by a lack of a defined media for in vitro cultivation of this parasite (Jex et al., 2020).

In contrast, consistent with previous RNA-seq data (Einarsson et al., 2016), 16 h post-induction of encystation (EC) corresponded to major proteomic changes. Faso et al. (2013) noted a similar significant shift in protein expression 4 h post-induction of encystation. However, this early stage of encystation was not assessed here. Proteomic changes at EC included down-regulation of cell adhesion proteins correlating with cytoskeletal rearrangements, leading to the loss of the ventral disc, flagella and epithelial adhesion (Lauwaet et al., 2007; Einarsson and Svärd, 2015). Amidst this cytoskeletal re-arrangement, we also observed down-regulation of

VSPs, which play an important role in antigenic variation (Prucca et al., 2008; Ankarklev et al., 2010) at EC. The trend of EC as a tipping point can be observed in many of the major protein functional classes, with most proteins abundantly expressed either in growing or encysting trophozoites, or in encysting and encysted cells. EC also marks a change in the abundance of *Giardia*'s major transcription factors, with the emergence of MYB1-like protein (GL_8722) as the major transcription factor in EC here and in the transcriptome (Einarsson et al., 2016). MYB1-like protein controls key markers of encystation, including cyst wall proteins, CWP1 (GL_5638), CWP2 (GL_5435) and CWP3 (GL_2421), and glucose-6-phosphate isomerase (GL_9115) (Huang et al., 2008), although its emergence appears specific and short-lived. Its transcription increases over the first 22 h of encystation (Einarsson et al., 2016) and its protein expression is detected in early encystation (Sun et al., 2002). It is then down-regulated transcriptionally (Einarsson et al., 2016) and here at the protein level in the mature cyst. EC also coincides with significant changes in many regulatory processes at the chromatin, transcriptional and post-transcriptional level.

Cyst maturation appears to be marked by major changes to the cell-surface membrane, synthesis of the cyst wall, and an overall down-regulation of protein synthesis. VSP expression was greatly reduced in the cyst. VSPs are post-transcriptionally silenced, such that most are transcribed simultaneously, but only one is expressed on the cell surface at a time (Prucca et al., 2008). During in vitro culture, it is common for different trophozoites to express different VSPs (Franzén et al., 2013) as was observed in Faso et al. (2013). Here, although multiple VSPs were expressed within a stage, one, GL_16501, was by far the most abundant in trophozoites (ranked 33 out of the 3863 iBAQ quantified proteins) but fell dramatically through to cyst maturation. As expression of GL_16501 fell, it was replaced by GL_113797 (VSP-5 or TSA 417), which by cyst maturation was the dominant variant (ranked 210 out of the 3863 iBAQ quantified proteins). TSA 417 is, consistently, the primary VSP up-regulated during encystation and appears to play a key role in cyst wall formation (Meng et al., 1993; Svärd et al., 1998). The drop in VSP expression during encystation coincides with increased expression of HCMPs/HCPs, including up-regulation of the most widely studied HCMP, HCNcP (GL_40376), in EC/C stages compared with TY. GL_40376 is up-regulated during encystation and localises to ESVs and the mature cyst wall (Davids et al., 2006). The function of HCMPs in the cyst is not known, however cysteine-rich proteins feature prominently in cyst walls in *Giardia*, in for example CWP1-3 (Luján et al., 1995; Chatterjee et al., 2010), and in other protists including *Toxoplasma* (Possenti et al., 2010) and *Cryptosporidium* (Templeton et al., 2004). In *Toxoplasma*, cysteine-rich mucins are glycosylated with GalNAc to form the backbone of the cyst wall (Tomita et al., 2017). GalNAc is also an essential component of the *Giardia* cyst wall (Jarroll et al., 1989; Chatterjee et al., 2010).

Finally, *Giardia* undergoes nuclear division and DNA replication without cytokinesis during encystation, forming an 8 N (4*2N) precyst, which replicates its DNA again, without cell division, to form a 16 N (4*4N) mature cyst (Einarsson and Svärd, 2015). Karyotype observations in *Giardia* indicate that the cyst is the only life-cycle stage in which significant change in chromatin packing occurs (Sonda et al., 2010; Lagunas-Rangel and Bermúdez-Cruz, 2019). Consistent with this, we identified changes in expression of key SMC ATPases in the mature cyst compared with the trophozoite. Encystation also coincides with changes in chromatin proteins for histone acetylation and methylation, and which influence gene expression and cell differentiation (Sonda et al., 2010; Lagunas-Rangel and Bermúdez-Cruz, 2019). Furthermore, transcriptomic data (Einarsson et al., 2016) indicates major up-regulation of key histone methylation, acetylation and deacetylation enzymes in mid- to late-encystation. Analyses of histone modification enzymes themselves are less clear with only four HMEs detected in our study. However, they suggest reduction in cysts of histone acetylation, which is associated with closed chromatin (Sonda et al., 2010).

Encystation and cyst maturation are marked throughout by major changes in metabolic and biosynthetic pathways. Bivalent switching within glucose and lipid metabolism appears a particularly essential feature of encystation. The substantial compositional changes of the *Giardia* proteome at EC must be energetically expensive, and coincide with a major up-regulation of glycolysis at EC, which is supported by the encystation transcriptome (Einarsson et al., 2016). This also coincides with up-regulation of the glycolytic shunt required for GalNAc biosynthesis (Einarsson and Svärd, 2015; Einarsson et al., 2016), which is a primary constituent of the cyst wall. *Giardia* cysts presumably enter a semi-quiescent state upon maturation in preparation for an extended period in the external environment. Consistent with this, glycolytic enzymes required for ATP generation are down-regulated to their lowest expression levels in the mature cyst. In contrast, the expression of GalNAc biosynthesis remains elevated at levels of expression comparable to that seen in EC. These data suggest that GalNAc remains important even at cyst maturation, possibly for maintenance of the cyst wall, and indicates that the cyst remains metabolically active to some extent (Paget et al., 1989).

We see a similar bivalent switching in lipid metabolism and composition of the *Giardia* cell surface during encystation. This observation is supported also by prior transcriptomic data (Einarsson et al., 2016). Altered cholesterol availability and transport is a key factor in these changes. *Giardia* acquires cholesterol from the host or culture media (Yichoy et al., 2011), and cholesterol deprivation in low-bile priming appears critical to in vitro induction of encystation (Luján et al., 1996). Our data shows increased expression of *Giardia* lecithin–cholesterol acyltransferase during encystation which would convert free cholesterol into cholesteryl ester, and deplete cholesterol in the cell membrane (Peelman et al., 1998). Encystation also coincides with the down-regulation of two putative Starkin lipid transporters (GL_5800 and GL_16717) which are structural homologues of cholesterol transporters in humans. Together, these suggest cholesterol depletion as a key process in the plasma membrane during encystation. This altered lipid metabolism has implications beyond cholesterol. Lipids and fatty acids are thought to play a key role in *Giardia* encystation through regulating growth and differentiation (Yichoy et al., 2011), redistribution of lipid rafts (Mendez et al., 2015), and formation of encystation secretory vesicles (Mendez et al., 2013). Inhibition of sphingolipid metabolism, specifically glucosylceramide synthesis, blocks endocytosis, production of ESVs and cell cycle progression, resulting in reduced numbers and viability of cysts in vitro (Štefanić et al., 2010). The sphingolipid metabolic enzymes serine palmitoyl transferases 1 (GL_23015), serine palmitoyl transferases 2 (GL_14374) and giardial ceramide glucosyltransferase 1 (GL_11642) are transcribed only in encysting cells (Mendez et al., 2013). Each of these enzymes are up-regulated at the protein level from EC onward in our study. We also find up-regulated expression of other key fatty acid metabolic enzymes, including long chain fatty acid coA ligase (GL_15063), Acetyl-CoA carboxylase (gACC)/pyruvate carboxylase fusion protein, putative (GL_113021) and fatty acid elongase 1 (GL_92729) in EC and at cyst maturation. Multiple studies have found these enzymes relevant in lipid raft formation (De Chatterjee et al., 2015; Mendez et al., 2015), ESV biogenesis (Hernandez et al., 2008; Sonda et al., 2008) and cyst viability (Mendez et al., 2013).

Major lipid and fatty acid species described in *Giardia* are primarily understood in the context of cyst wall formation. However,

lipid rafts are microdomains within the cellular membrane that act as sensors to direct signalling in response to external and internal stimuli (Mendez et al., 2015). Sphingolipids and cholesterols are the major components of these lipid rafts (Mendez et al., 2015). It is possible that changes in lipid metabolism alter the composition of lipid rafts or the plasma membrane, triggering signalling cascades that directly or indirectly influence encystation (Luján et al., 1996). Within this study, we have provided the first in silico evidence for probable lipid transporters (LTPs) in *Giardia* including the Starkin superfamily, and Glycolipid Transfer Protein (GLTP) and Oxysterol Binding Protein Related protein families. In silico docking-based interaction studies, coupled with the shift from expression of Starkin LTPs in TY to GLTP and Oxysterol Binding Protein in the cyst indicates changes in lipid dependency of *Giardia* during encystation. The presence of Oxysterol Binding Protein, a bispecific transporter that binds to phosphatidyl serine (PS) and PI4P, together with the presence of SAC1-like phosphatase and PI4P kinase, points to the possibility of a PS-PI4P exchange cycle creating a lipid gradient between the endoplasmic reticulum and the plasma membrane. Consistent with this, we see increased expression of each of these proteins in EC and at C. We also observed much lower levels of enzymes regulating signalling lipids in the mature cyst, including Phosphatidyl inositol-4-Phosphate 5 kinase (PI4P5K) (GL_2622, GL_7261) and Type II inositol-1,4,5-trisphosphate 5-phosphatase precursor (GL_14787) and up-regulation in EC and C compared to TY or LB of Phosphatidylinositol-glycan biosynthesis, class O protein (gPIG) (GL_14975). While prokaryotes lack phosphatidyl-inositol signalling systems (Clarke et al., 2019), these data suggest that this system is present in deep-branching eukaryotes and appears to be important in encystation.

Our data suggest that encystation in *Giardia* is tightly regulated. There is, however, a paradox that this parasite mounts specific and well-regulated molecular responses (Ansell et al., 2015, 2016; Emery et al., 2018), yet lacks transcriptional controllers, having a minimal nuclear condensome and few transcription factors (Elmendorf et al., 2001; Teodorovic et al., 2007). Post-transcriptional regulation is important in *Giardia*, with RNA silencing required in VSP expression (Prucca et al., 2008), for example. Post-transcriptional regulation can be mediated also through the suppression of mRNAs by RNA binding proteins (RBPs). Similar to prokaryotes (Redder et al., 2015), *Giardia* encodes numerous SF2 helicases within its genome, including DEAD box, DEAH and Ski2 helicases (Gargantini et al., 2012). *Giardia* is also one of the deepest-branching known eukaryotes with pumilio (PUF) domain proteins (Williams and Elmendorf, 2011), a major family of RBPs that regulate cellular differentiation and fating through translational repression in eukaryotes (Wang et al., 2018).

A previous study suggested *Giardia*'s SF2 helicases are transcriptionally enriched in and may regulate encystation (Gargantini et al., 2012). However, these data are at odds with more recent RNA-seq data (Einarsson et al., 2016) and our proteomic data here, in which most SF2 helicases are down-regulated through encystation. In contrast, encystation coincided with up-regulated expression of a PUF protein (GL_17590) with closest structural homology to PUF3 in *S. cerevisiae* (PDB_ID: 3K49). Although the functions of the *Giardia* PUFs are unknown, the Yeast PUF3 regulates mRNA stability and mediates translational repression of mitochondrial and oxidative stress proteins (Houshmandi and Olivas, 2005; Lapointe et al., 2018). Among parasites, PUF proteins and other RBPs are critical to life-cycle progression, particularly in transmissive stages (Vembar et al., 2016). Cyst maturation also coincides with down-regulation of eukaryotic initiation factors (EIF)-2G (GL_2970), 5A (GL_14614) and 1A (GL_8708). In *S. cerevisiae*, eIF5A depletion results in 30% reduction in global protein synthesis (Henderson and Hershey, 2011). We observed that the proteomic expression was consistently down-regulated in the mature cyst. However, whether reduced levels of protein expression in cysts are related to EIF expression cannot be inferred from our data.

As a first step in identifying potentially repressed transcripts, we compared transcription and expression in the trophozoite and cyst stage, using approaches based on previous studies exploring translation repression (Muller et al., 2019; Wang et al., 2019). There was a clear difference in the number of putatively repressed transcripts in the mature cyst (25% of top decile transcripts) compared with the trophozoite (9.8%). This is consistent with the cyst being the transmissive stage which, based on research in other parasitic protists (Vembar et al., 2016; Lindner et al., 2019; Muller et al., 2019), may be pre-loaded with repressed transcripts required for early infection and development in the new host. These differences could not be explained based on WB1B versus WBC6 strain and laboratory differences, LC-MS sensitivity, or the amino acid composition or codon usage of the missing or under-represented proteins. In most cases in cysts, undetected proteins corresponding to top decile transcripts were readily detected in other stages here, or in previously published studies within Giardia DB (Aurrecoechea et al., 2008) confirming their detectability by LC-MS methods. Overall, our data implicates translation repression in encystation, potentially mediated through RNA-binding proteins. We acknowledge that this is an intriguing, yet circumstantial, exploration that requires empirical data beyond the scope of our study including through direct transcript and RBP pulldowns (Castello et al., 2013; Castello et al., 2016; Hentze et al., 2018), co-localisation of putatively repressed transcripts and regulatory RBPs in P-granules (Updike and Strome, 2010; Lee et al., 2020) and functional perturbation of RBPs through CRISPRi (Jex et al., 2020) or other approaches.

In both stages, putatively repressed transcripts were functionally enriched for complex gene families (NEKs, VSPs and protein 21.1 s). As noted, RNA silencing plays a key role in the expression of VSPs in *Giardia* (Prucca et al., 2008). In *Plasmodium falciparum*, complex gene families, particularly *var* genes, are epigenetically controlled at the chromatin level (Deitsch and Dzikowski, 2017). It is possible that in the presence of limited chromatin regulation and without complex transcription factors (which are both believed to have evolved after RNA-level control in eukaryotes (Sharp, 2009; Cojocaru and Unrau, 2017)), *Giardia* regulates its complex gene families at a post-transcriptional level using both RNA silencing and translational repression. Although the function of protein 21.1 proteins is not well understood, they encode a large number of ankyrin repeat domains, which mediate protein–protein interactions, and tend to be found adjacent to NEK kinases in the *Giardia* genome (Manning et al., 2011) and colocalise with NEKs in the cytoskeleton (Gadelha et al., 2017), suggesting a possible role in regulating or interacting with NEKs. Interestingly, four of 52 known encystation markers, cyst wall protein 2 (GL_5435), UDP-glucose-4-epimerase (GL_7982), HCMP (GL_40376) and ankyrin repeat protein 1 (GL_24412), are highly transcribed in *Giardia* trophozoites, but under-expressed as proteins. These are required for encystation and are clearly detectable as proteins at later encystation time points in our study. Their abundance as, it would seem, untranslated transcripts in the trophozoite may assist *Giardia* in rapidly forming cysts when required (e.g., under times of heavy stress or imminent danger). With cyst maturation, putatively repressed transcripts were functionally enriched for spliceosomal proteins. Intriguingly, *Giardia* has cis-intron splicing (Xue et al., 2018) and trans-splicing, in which exons from distinct mRNA transcripts are combined through post-transcriptional splicing (Kamikawa et al., 2011). The two most well characterised of *Giardia*'s trans-spliced genes are *hsp90* (Nageshan et al., 2014; Iyer et al., 2019) and dynein heavy chain, beta (*dhc-β*; Roy et al.

(2012)). Both of these are encystation markers that are up-regulated in the mature cyst here and in a prior study (Nageshan et al., 2014). Whether trans-splicing is associated with encystation in, and what additional transcripts may be processed in, this way is not known. What role translational repression might play in regulating spliceosomal proteins is also not clear.

Overall, we provide the first known comprehensive proteomic study across the complete encystation process in *Giardia*. These data show that the transition from a metabolically active trophozoite to an environmentally resilient cyst is complex and highly regulated. In addition to a major rearrangement of the cytoskeletal system, encystation appears to involve major changes in the composition of the cell-surface, as well as key changes in metabolic, biosynthetic and regulatory pathways. We provide the first known annotation of lipid transporters in *Giardia* and find evidence for an active PS – PI4P lipid shuttling system in encystation. Finally, our data suggests that post-transcriptional regulation, including translational repression involving, in particular, *Giardia*'s PUF proteins, may have a key role in encystation, which requires further study.

Acknowledgements

B. Balan's PhD scholarship was funded by Melbourne International Research Scholarship, University of Melbourne, Australia. We gratefully acknowledge the Walter and Eliza Hall Research Institute, Australia, for the top up scholarship. AR Jex acknowledges the Australian National Health and Medical Research Foundation Career Development Fellowship program (APP1126395). SJ Emery-Corbin was funded by a Jack Brockhoff Foundation (Australia) Early Career Grant (ID JBF 4184, 2016). We also acknowledge funding from the Victorian State Government Operational Infrastructure Support, Australia and the Australian Government National Health and Medical Research Council Independent Research Institute Infrastructure Support Scheme. Proteomics analysis was performed at the Mass Spectrometry and Proteomics Facility, Bio21 Institute, University of Melbourne. We also acknowledge Dr. Chin Seng Ang (Mass Spectrometry and Proteomics Facility, Bio21 Institute, University of Melbourne) for providing technical support. We also thank Amrita Vijay for all the expert opinions given regarding this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpara.2021.01.008.

References

Aebersold, R., Mann, M., 2016. Mass-spectrometric exploration of proteome structure and function. Nature 537, 347.

Angel, T.E., Aryal, U.K., Hengel, S.M., Baker, E.S., Kelly, R.T., Robinson, E.W., Smith, R. D., 2012. Mass spectrometry-based proteomics: existing capabilities and future directions. Chem. Soc. Rev. 41, 3912–3928.

Ankarklev, J., Jerlström-Hultqvist, J., Ringqvist, E., Troell, K., Svärd, S.G., 2010. Behind the smile: cell biology and disease mechanisms of *Giardia* species. Nat. Rev. Microbiol. 8, 413.

Ansell, B.R.E., McConville, M.J., Baker, L., Korhonen, P.K., Young, N.D., Hall, R.S., Rojas, C.A.A., Svärd, S.G., Gasser, R.B., Jex, A.R., 2015. Time-dependent transcriptional changes in axenic *Giardia duodenalis* trophozoites. PLoS Negl. Trop. Dis. 9, e0004261.

Ansell, B.R., McConville, M.J., Baker, L., Korhonen, P.K., Emery, S.J., Svärd, S.G., Gasser, R.B., Jex, A.R., 2016. Divergent transcriptional responses to physiological and xenobiotic stress in *Giardia duodenalis*. Antimicrob. Agents Chemother. 60, 6034–6045.

Ansell, B.R., Pope, B.J., Georgeson, P., Emery-Corbin, S.J., Jex, A.R., 2018. Annotation of the *Giardia* proteome through structure-based homology and machine learning. GigaScience 8, giy150.

Anuradha, S., Muniyappa, K., 2004. Meiosis-specific yeast Hop1 protein promotes synapsis of double-stranded DNA helices via the formation of guanine quartets. Nucleic Acids Res. 32, 2378–2385.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Carlton, J.M., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., 2008. GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens *Giardia lamblia* and *Trichomonas vaginalis*. Nucleic Acids Res. 37, D526–D530.

Bar-Even, A., Flamholz, A., Noor, E., Milo, R., 2012. Rethinking glycolysis: on the biochemical logic of metabolic pathways. Nat. Chem. Biol. 8, 509.

Becker, K., Bluhm, A., Casas-Vila, N., Dinges, N., Dejung, M., Sayols, S., Kreutz, C., Roignant, J.-Y., Butter, F., Legewie, S., 2018. Quantifying post-transcriptional regulation in the development of *Drosophila melanogaster*. Nat. Comm 9, 1–14.

Birkeland, S.R., Preheim, S.P., Davids, B.J., Cipriano, M.J., Palm, D., Reiner, D.S., Svärd, S.G., Gillin, F.D., McArthur, A.G., 2010. Transcriptome analyses of the *Giardia lamblia* life cycle. Mol. Biochem. Parasitol. 174, 62–65.

Bunnik, E.M., Batugedara, G., Saraf, A., Prudhomme, J., Florens, L., Le Roch, K.G., 2016. The mRNA-bound proteome of the human malaria parasite *Plasmodium falciparum*. Genome Biol. 17, 147.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., Hentze, M.W., 2013. System-wide identification of RNA-binding proteins by interactome capture. Nat. Protoc. 8, 491.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., Hentze, M.W., 2016. Comprehensive identification of RNA-binding proteins by RNA interactome capture. In: Dassi, E. (Ed.), Post-Transcriptional Gene Regulation. Humana Press, New York, NY, pp. 131–139.

Chatterjee, A., Carpentieri, A., Ratner, D.M., Bullitt, E., Costello, C.E., Robbins, P.W., Samuelson, J., 2010. *Giardia* cyst wall protein 1 is a lectin that binds to curled fibrils of the GalNAc homopolymer. PLoS Path. 6.

Chen, N., Upcroft, J.A., Upcroft, P., 1994. Physical map of a 2 Mb chromosome of the intestinal protozoan parasite *Giardia duodenalis*. Chromosome Res. 2, 307–313.

Clarke, B.P., Logeman, B.L., Hale, A.T., Luka, Z., York, J.D., 2019. A synthetic biological approach to reconstitution of inositide signaling pathways in bacteria. Adv. Biol. Regul. 73, 100637.

Cojocaru, R., Unrau, P.J., 2017. Origin of life: transitioning to DNA genomes in an RNA world. eLife 6, e32330.

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. 26, 1367–1372.

Da Wei Huang, B.T.S., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., Stephens, R., Baseler, M.W., Lane, H.C., Lempicki, R.A., 2007. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 8, R183.

Das, S., Castillo, C., Stevens, T., 2001. Phospholipid remodeling/generation in *Giardia*: the role of the Lands cycle. Trends Parasitol. 17, 316–319.

Davids, B.J., Reiner, D.S., Birkeland, S.R., Preheim, S.P., Cipriano, M.J., McArthur, A.G., Gillin, F.D., 2006. A new family of giardial cysteine-rich non-VSP protein genes and a novel cyst protein. PLoS One 1.

De Chatterjee, A., Mendez, T.L., Roychowdhury, S., Das, S., 2015. The assembly of GM1 glycolipid-and cholesterol-enriched raft-like membrane microdomains is important for giardial encystation. Infect. Immun. 83, 2030–2042.

Deitsch, K.W., Dzikowski, R., 2017. Variant gene expression and antigenic variation by malaria parasites. Annu. Rev. Microbiol. 71, 625–641.

Duarte, T.T., Ellis, C.C., Grajeda, B.I., De Chatterjee, A., Almeida, I.C., Das, S., 2019. A targeted mass spectrometric analysis reveals the presence of a reduced but dynamic sphingolipid metabolic pathway in an ancient protozoan, Giardia lamblia. Front. Cell.Infect.Micro. 9.

Dubourg, A., Xia, D., Winpenny, J.P., Al Naimi, S., Bouzid, M., Sexton, D.W., Wastling, J.M., Hunter, P.R., Tyler, K.M., 2018. *Giardia* secretome highlights secreted tenascins as a key component of pathogenesis. GigaScience 7, giy003.

Einarsson, E., Svärd, S.G., 2015. Encystation of *Giardia intestinalis*—a journey from the duodenum to the colon. Curr. Trop. Med. Rep. 2, 101–109.

Einarsson, E., Troell, K., Hoeppner, M.P., Grabherr, M., Ribacke, U., Svärd, S.G., 2016. Coordinated changes in gene expression throughout encystation of *Giardia intestinalis*. PLoS Negl. Trop. Dis. 10, e0004571.

Ellis, J.E., Wyder, M.A., Jarroll, E.L., Kaneshiro, E.S., 1996. Changes in lipid composition during in vitro encystation and fatty acid desaturase activity of *Giardia lamblia*. Mol. Biochem. Parasitol. 81, 13–25.

Elmendorf, H.G., Singer, S.M., Nash, T.E., 2001. The abundance of sterile transcripts in *Giardia lamblia*. Nucleic Acids Res. 29, 4674–4683.

Emery, S.J., van Sluyter, S., Haynes, P.A., 2014. Proteomic analysis in *Giardia duodenalis* yields insights into strain virulence and antigenic variation. Proteomics 14, 2523–2534.

Emery, S.J., Mirzaei, M., Vuong, D., Pascovici, D., Chick, J.M., Lacey, E., Haynes, P.A., 2016. Induction of virulence factors in *Giardia duodenalis* independent of host attachment. Sci. Rep. 6, 20765.

Emery, S.J., Baker, L., Ansell, B.R., Mirzaei, M., Haynes, P.A., McConville, M.J., Svärd, S. G., Jex, A.R., 2018. Differential protein expression and post-translational modifications in metronidazole-resistant *Giardia duodenalis*. GigaScience 7, giy024.

Emery-Corbin, S.J., Vuong, D., Lacey, E., Svärd, S.G., Ansell, B.R., Jex, A.R., 2018. Proteomic diversity in a prevalent human-infective *Giardia duodenalis* sub-species. Int. J. Parasitol. 48, 817–823.

Faso, C., Bischof, S., Hehl, A.B., 2013. The proteome landscape of *Giardia lamblia* encystation. PLoS One 8, e83207.

Fernández-Moya, S.M., Estévez, A.M., 2010. Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosomatid protozoan parasites. Wiley Interdiscip. Rev. RNA 1, 34–46.
Franzén, O., Jerlström-Hultqvist, J., Einarsson, E., Ankarklev, J., Ferella, M., Andersson, B., Svärd, S.G., 2013. Transcriptome profiling of *Giardia intestinalis* using strand-specific RNA-seq. PLoS Comp. Biol. 9, e1003000.
Gadelha, A., Benchimol, M., de Souza, W., 2017. The Cytoskeleton of *Giardia intestinalis*, Current Topics in Giardiasis. In: Rodriguez-Morales, A.J. (Ed.), InTech. Rijeka, Croatia, pp. 63–86.
Garfoot, A.L., Wilson, G.M., Coon, J.J., Knoll, L.J., 2019. Proteomic and transcriptomic analyses of early and late-chronic *Toxoplasma gondii* infection shows novel and stage specific transcripts. BMC Genomics 20, 859.
Gargantini, P.R., Serradell, M.C., Torri, A., Lujan, H.D., 2012. Putative SF2 helicases of the early-branching eukaryote *Giardia lamblia* are involved in antigenic variation and parasite differentiation into cysts. BMC Microbiol. 12, 284.
Gibson, G.R., Ramirez, D., Maier, J., Castillo, C., Das, S., 1999. *Giardia lamblia*: incorporation of free and conjugated fatty acids into glycerol-based phospholipids. Exp. Parasitol. 92, 1–11.
Hagen, D.F., Markell, C.G., Schmitt, G.A., Blevins, D.D., 1990. Membrane approach to solid-phase extractions. Anal. Chim. Acta 236, 157–164.
Hehl, A.B., Marti, M., Kohler, P., 2000. Stage-specific expression and targeting of cyst wall protein–green fluorescent protein chimeras in *Giardia*. Mol. Biol. Cell 11, 1789–1800.
Hehl, A.B., Marti, M., 2004. Secretory protein trafficking in *Giardia intestinalis*. Mol. Microbiol. 53, 19–28.
Henderson, A., Hershey, J.W., 2011. Eukaryotic translation initiation factor (eIF) 5A stimulates protein synthesis in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. 108, 6415–6419.
Hentze, M.W., Castello, A., Schwarzl, T., Preiss, T., 2018. A brave new world of RNA-binding proteins. Nat. Rev. Mol. Cell Biol. 19, 327.
Hernandez, Y., Shpak, M., Duarte, T.T., Mendez, T.L., Maldonado, R.A., Roychowdhury, S., Rodrigues, M.L., Das, S., 2008. Novel role of sphingolipid synthesis genes in regulating giardial encystation. Infect. Immun. 76, 2939–2949.
Houshmandi, S.S., Olivas, W.M., 2005. Yeast Puf3 mutants reveal the complexity of Puf-RNA binding and identify a loop required for regulation of mRNA decay. RNA 11, 1655–1666.
Huang, Y.-C., Su, L.-H., Lee, G.A., Chiu, P.-W., Cho, C.-C., Wu, J.-Y., Sun, C.-H., 2008. Regulation of cyst wall protein promoters by Myb2 in *Giardia lamblia*. J. Biol. Chem. 283, 31021–31029.
Iyer, V., Chettiar, S.T., Grover, M., Rajyaguru, P., Nageshan, R.K., Tatu, U., 2019. *Giardia lamblia* Hsp90 pre-mRNA s undergo self-splicing to generate mature RNA in an in vitro trans-splicing reaction. FEBS Lett. 593, 433–442.
Jarroll, E.L., Manning, P., Lindmark, D.G., Coggins, J.R., Erlandsen, S.L., 1989. Giardia cyst wall-specific carbohydrate: evidence for the presence of galactosamine. Molecular and biochemical parasitology 32 (2–3), 121–131.
Jarroll, E.L., van Keulen, H., Paget, T.A., Lindmark, D.G., 2011. *Giardia* Metabolism. In: Luján, H.D., Svärd, S. (Eds.), *Giardia*: A Model Organism. Springer Vienna, Vienna, pp. 127–137.
Jedelský, P.L., Doležal, P., Rada, P., Pyrih, J., Šmíd, O., Hrdý, I., Šedinová, M., Marcinčiková, M., Voleman, L., Perry, A.J., 2011. The minimal proteome in the reduced mitochondrion of the parasitic protist *Giardia intestinalis*. PLoS One 6, e17285.
Jex, A.R., Svärd, S., Hagen, K.D., Starcevich, H., Emery, S.J., Balan, B., Nosala, C., Dawson, S.C., 2020. Recent advances in functional research in *Giardia intestinalis*. In: Ortega-Pierres, M.G. (Ed.), *Giardia* and Giardiasis-Part B. Academic Press, London, United Kingdom, pp. 97–137.
Kamikawa, R., Inagaki, Y., Tokoro, M., Roger, A.J., Hashimoto, T., 2011. Split introns in the genome of *Giardia intestinalis* are excised by spliceosome-mediated trans-splicing. Curr. Biol. 21, 311–315.
Kwiatkowska-Semrau, K., Czarnecka, J., Wojciechowski, M., Milewski, S., 2015. Heterogeneity of quaternary structure of glucosamine-6-phosphate deaminase from *Giardia lamblia*. Parasitol. Res. 114, 175–184.
Lane, S., Lloyd, D., 2002. Current trends in research into the waterborne parasite Giardia. Critical reviews in microbiology 28 (2), 123–147.
Lafay, B., Sharp, P.M., 1999. Synonymous codon usage variation among *Giardia lamblia* genes and isolates. Mol. Biol. Evol. 16, 1484–1495.
Lagunas-Rangel, F.A., Bermúdez-Cruz, R.M., 2019. Epigenetics in the early divergent eukaryotic Giardia duodenalis: An update. Biochimie 156, 123–128.
Lapointe, C.P., Stefely, J.A., Jochem, A., Hutchins, P.D., Wilson, G.M., Kwiecien, N.W., Coon, J.J., Wickens, M., Pagliarini, D.J., 2018. Multi-omics reveal specific targets of the RNA-binding protein Puf3p and its orchestration of mitochondrial biogenesis. Cell Syst. 6. 125–135 e126.
Lauwaet, T., Davids, B.J., Reiner, D.S., Gillin, F.D., 2007. Encystation of *Giardia lamblia*: a model for other parasites. Curr. Opin. Microbiol. 10, 554–559.
Lauwaet, T., Smith, A.J., Reiner, D.S., Romijn, E.P., Wong, C.C., Davids, B.J., Shah, S. A., Yates 3rd, J.R., Gillin, F.D., 2011. Mining the *Giardia* genome and proteome for conserved and unique basal body proteins. Int. J. Parasitol. 41, 1079–1092.
Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J., 2004. Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. Genome Res. 14, 2308–2318.
Lee, C.-Y.-S., Putnam, A., Lu, T., He, S., Ouyang, J.P.T., Seydoux, G., 2020. Recruitment of mRNAs to P granules by condensation with intrinsically-disordered proteins. Elife 9, e52896.
Lindner, S.E., Swearingen, K.E., Shears, M.J., Walker, M.P., Vrana, E.N., Hart, K.J., Minns, A.M., Sinnis, P., Moritz, R.L., Kappe, S.H., 2019. Transcriptomics and proteomics reveal two waves of translational repression during the maturation of malaria parasite sporozoites. Nat. Comm 10, 1–13.
Lueong, S., Merce, C., Fischer, B., Hoheisel, J.D., Erben, E.D., 2016. Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. Mol. Microbiol. 100, 457–471.
Luján, H.D., Mowatt, M.R., Conrad, J.T., Bowers, B., Nash, T.E., 1995. Identification of a novel *Giardia lamblia* cyst wall protein with leucine-rich repeats Implications for secretory granule formation and protein assembly into the cyst wall. J. Biol. Chem. 270, 29307–29313.
Luján, H.D., Mowatt, M.R., Byrd, L.G., Nash, T.E., 1996. Cholesterol starvation induces differentiation of the intestinal parasite *Giardia lamblia*. Proc. Natl. Acad. Sci. 93, 7628–7633.
Lujan, H.D., Svärd, S. (Eds.), 2011. *Giardia*: A model organism. Springer Wien New York; Vienna, AT, Austria.
Ma'ayeh, S.Y., Liu, J., Peirasmaki, D., Hörnaeus, K., Lind, S.B., Grabherr, M., Bergquist, J., Svärd, S.G., 2017. Characterization of the Giardia intestinalis secretome during interaction with human intestinal epithelial cells: the impact on host cells. PLoS Negl. Trop. Dis. 11, e0006120.
Maeda, K., Anand, K., Chiapparino, A., Kumar, A., Poletto, M., Kaksonen, M., Gavin, A.-C., 2013. Interactome map uncovers phosphatidylserine transport by oxysterol-binding proteins. Nature 501, 257.
Manning, G., Reiner, D.S., Lauwaet, T., Dacre, M., Smith, A., Zhai, Y., Svard, S., Gillin, F. D., 2011. The minimal kinome of *Giardia lamblia* illuminates early kinase evolution and unique parasite biology. Genome Biol. 12, R66.
Mendez, T.L., De Chatterjee, A., Duarte, T.T., Gazos-Lopes, F., Robles-Martinez, L., Roy, D., Sun, J., Maldonado, R.A., Roychowdhury, S., Almeida, I.C., 2013. Glucosylceramide transferase activity is critical for encystation and viable cyst production by an intestinal protozoan, *Giardia lamblia*. J. Biol. Chem. 288, 16747–16760.
Mendez, T.L., De Chatterjee, A., Duarte, T., De Leon, J., Robles-Martinez, L., Das, S., 2015. Sphingolipids, lipid rafts, and giardial encystation: the show must go on. Curr. Trop. Med. Rep. 2, 136–143.
Meng, T., Hetsko, M., Gillin, F., 1993. Antigenic switching of TSA 417, a trophozoite variable surface protein, following completion of the life cycle of *Giardia lamblia*. Infect. Immun. 61, 5394–5397.
Miao, J., Li, J., Fan, Q., Li, X., Li, X., Cui, L., 2010. The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite *Plasmodium falciparum*. J. Cell Sci. 123, 1039–1049.
Morf, L., Spycher, C., Rehrauer, H., Fournier, C.A., Morrison, H.G., Hehl, A.B., 2010. The transcriptional response to encystation stimuli in *Giardia lamblia* is restricted to a small set of genes. Eukaryotic Cell 9, 1566–1576.
Morrison, H.G., McArthur, A.G., Gillin, F.D., Aley, S.B., Adam, R.D., Olsen, G.J., Best, A. A., Cande, W.Z., Chen, F., Cipriano, M.J., 2007. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. Science 317, 1921–1926.
Muller, I., Jex, A.R., Kappe, S.H., Mikolajczak, S.A., Sattabongkot, J., Patrapuvich, R., Lindner, S., Flannery, E.L., Koepfli, C., Ansell, B., 2019. Transcriptome and histone. epigenome of *Plasmodium vivax* salivary-gland sporozoites point to tight regulatory control and mechanisms for liver-stage differentiation in relapsing malaria. Int. J. Parasitol. 49, 501–513.
Nageshan, R.K., Roy, N., Ranade, S., Tatu, U., 2014. Trans-spliced heat shock protein 90 modulates encystation in *Giardia lamblia*. PLoS Negl. Trop. Dis. 8, e2829.
Nandan, D., Thomas, S.A., Nguyen, A., Moon, K.M., Foster, L.J., Reiner, N.E., 2017. Comprehensive identification of mRNA-binding proteins of *Leishmania donovani* by interactome capture. PLoS One 12, e0170068.
Paget, T.A., Jarroll, E.L., Manning, P., Lindmark, D.G., Lloyd, D., 1989. Respiration in the cysts and trophozoites of *Giardia muris*. Microbiology 135, 145–154.
Palm, D., Weiland, M., McArthur, A.G., Winiecka-Krusnell, J., Cipriano, M.J., Birkeland, S.R., Pacocha, S.E., Davids, B., Gillin, F., Linder, E., 2005. Developmental changes in the adhesive disk during *Giardia* differentiation. Mol. Biochem. Parasitol. 141, 199–207.
Pan, Y.-J., Cho, C.-C., Kao, Y.-Y., Sun, C.-H., 2009. A novel WRKY-like protein involved in transcriptional activation of cyst wall protein genes in *Giardia lamblia*. J. Biol. Chem. 284, 17975–17988.
Peelman, F., Vinaimont, N., Vanloo, B., Labeur, C., Rosseneu, M., Verhee, A., Verschelde, J.L., Vanderckove, J., Tavernier, J., Seguret-Mace, S., 1998. A proposed architecture for lecithin cholesterol acyl transferase (LCAT): identification of the catalytic triad and molecular modeling. Protein Sci. 7, 587–599.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.
Pham, J.K., Nosala, C., Scott, E.Y., Nguyen, K.F., Hagen, K.D., Starcevich, H.N., Dawson, S.C., 2017. Transcriptomic profiling of high-density *Giardia* foci encysting in the murine proximal intestine. Front. Cell. Infect. Microbiol. 7, 227.

Possenti, A., Cherchi, S., Bertuccini, L., Pozio, E., Dubey, J., Spano, F., 2010. Molecular characterisation of a novel family of cysteine-rich proteins of *Toxoplasma gondii* and ultrastructural evidence of oocyst wall localisation. Int. J. Parasitol. 40, 1639–1649.

Prucca, C.G., Slavin, I., Quiroga, R., Elías, E.V., Rivero, F.D., Saura, A., Carranza, P.G., Luján, H.D., 2008. Antigenic variation in *Giardia lamblia* is regulated by RNA interference. Nature 456, 750.

Redder, P., Hausmann, S., Khemici, V., Yasrebi, H., Linder, P., 2015. Bacterial versatility requires DEAD-box RNA helicases. FEMS Microbiol. Rev. 39, 392–412.

Reiner, D., Douglas, H., Gillin, F., 1989. Identification and localization of cyst-specific antigens of *Giardia lamblia*. Infect. Immun. 57, 963–968.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47.

Roy, S.W., Hudson, A.J., Joseph, J., Yee, J., Russell, A.G., 2012. Numerous fragmented spliceosomal introns, AT–AC splicing, and an unusual dynein gene expression pathway in *Giardia lamblia*. Mol. Biol. Evol. 29, 43–49.

Sharp, P.A., 2009. The centrality of RNA. Cell 136, 577–580.

Sonda, S., Štefanić, S., Hehl, A.B., 2008. A sphingolipid inhibitor induces a cytokinesis arrest and blocks stage differentiation in *Giardia lamblia*. Antimicrob. Agents Chemother. 52, 563–569.

Sonda, S., Morf, L., Bottova, I., Baetschmann, H., Rehrauer, H., Caflisch, A., Hakimi, M. A., Hehl, A.B., 2010. Epigenetic mechanisms regulate stage differentiation in the minimized protozoan *Giardia lamblia*. Mol. Microbiol. 76, 48–67.

Stefanic, S., Palm, D., Svärd, S.G., Hehl, A.B., 2006. Organelle proteomics reveals cargo maturation mechanisms associated with Golgi-like encystation vesicles in the early-diverged protozoan *Giardia lamblia*. J. Biol. Chem. 281, 7595–7604.

Štefanić, S., Spycher, C., Morf, L., Fabriàs, G., Casas, J., Schraner, E., Wild, P., Hehl, A.B., Sonda, S., 2010. Glucosylceramide synthesis inhibition affects cell cycle progression, membrane trafficking, and stage differentiation in *Giardia lamblia*. J. Lipid Res. 51, 2527–2545.

Sulemana, A., Paget, T.A., Jarroll, E.L., 2014. Commitment to cyst formation in *Giardia*. Microbiology 160, 330–339.

Sun, C.H., Palm, D., McArthur, A.G., Svärd, S.G., Gillin, F.D., 2002. A novel Myb-related protein involved in transcriptional activation of encystation genes in *Giardia lamblia*. Mol. Microbiol. 46, 971–984.

Svärd, S.G., Meng, T.C., Hetsko, M.L., McCaffery, J.M., Gillin, F.D., 1998. Differentiation-associated surface antigen variation in the ancient eukaryote *Giardia lamblia*. Mol. Microbiol. 30, 979–989.

Svärd, S.G., Hagblom, P., Palm, J.D., 2003. *Giardia lamblia*—a model organism for eukaryotic cell differentiation. FEMS Microbiol. Lett. 218, 3–7.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., 2018. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47, D607–D613.

Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., Daly, T.M., Bergman, L.W., Kappe, S.H., 2008. A combined transcriptome and proteome survey of malaria parasite liver stages. Proc. Natl. Acad. Sci. 105, 305–310.

Templeton, T.J., Lancto, C.A., Vigdorovich, V., Liu, C., London, N.R., Hadsall, K.Z., Abrahamsen, M.S., 2004. The *Cryptosporidium* oocyst wall protein is a member of a multigene family and has a homolog in *Toxoplasma*. Infect. Immun. 72, 980–987.

Teodorovic, S., Walls, C.D., Elmendorf, H.G., 2007. Bidirectional transcription is an inherent feature of *Giardia lamblia* promoters and contributes to an abundance of sterile antisense transcripts throughout the genome. Nucleic Acids Res. 35, 2544–2553.

Timp, W., Timp, G., 2020. Beyond mass spectrometry, the next step in proteomics. Sci. Adv. 6, eaax8978.

Tomita, T., Sugi, T., Yakubu, R., Tu, V., Ma, Y., Weiss, L.M., 2017. Making home sweet and sturdy: *Toxoplasma gondii* ppGalNAc-Ts glycosylate in hierarchical order and confer cyst wall rigidity. MBio 8, e02048–02016.

Trott, O., Olson, A.J., 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. 31, 455–461.

Tůmová, P., Uzlíková, M., Wanner, G., Nohýnková, E., 2015. Structural organization of very small chromosomes: study on a single-celled evolutionary distant eukaryote *Giardia intestinalis*. Chromosoma 124, 81–94.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., Cox, J., 2016. The Perseus computational platform for comprehensive analysis of (prote) omics data. Nat. Methods 13, 731.

Upcroft, J.A., Boreham, P.F., Campbell, R.W., Shepherd, R.W., Upcroft, P., 1995. Biological and genetic analysis of a longitudinal collection of *Giardia* samples derived from humans. Acta Tropica 60, 35–46.

Updike, D., Strome, S., 2010. P granule assembly and function in *Caenorhabditis elegans* germ cells. J. Androl. 31, 53–60.

Vembar, S.S., Droll, D., Scherf, A., 2016. Translational regulation in blood stages of the malaria parasite *Plasmodium* spp.: systems-wide studies pave the way. Wiley Interdiscip. Rev. RNA 7, 772–792.

von Filseck, J.M., Čopič, A., Delfosse, V., Vanni, S., Jackson, C.L., Bourguet, W., Drin, G., 2015. Phosphatidylserine transport by ORP/Osh proteins is driven by phosphatidylinositol 4-phosphate. Science 349, 432–436.

Wang, D., Eraslan, B., Wieland, T., Hallström, B., Hopf, T., Zolg, D.P., Zecha, J., Asplund, A., Li, L.h., Meng, C., 2019. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. Mol. Syst. Biol. 15.

Wang, M., Ogé, L., Perez-Garcia, M.-D., Hamama, L., Sakr, S., 2018. The PUF protein family: overview on PUF RNA targets, biological functions, and post transcriptional regulation. Int. J. Mol. Sci. 19, 410.

Wang, C.-H., Su, L.-H., Sun, C.-H., 2007. A novel ARID/Bright-like protein involved in transcriptional activation of cyst wall protein 1 gene in *Giardia lamblia*. J. Biol. Chem. 282, 8905–8914.

Wessel, D.M., Flügge, U., 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Anal. Biochem. 138, 141–143.

Wickham, H., 2011. ggplot2. Wiley Interdiscip. Rev. Comp. Stat. 3, 180–185.

Williams, C.W., Elmendorf, H.G., 2011. Identification and analysis of the RNA degrading complexes and machinery of *Giardia lamblia* using an in silico approach. BMC Genomics 12, 586.

Wong, L.H., Gatta, A.T., Levine, T.P., 2019. Lipid transfer proteins: the lipid commute via shuttles, bridges and tubes. Nat. Rev. Mol. Cell Biol. 20, 85–101.

Xue, M., Chen, B., Ye, Q., Shao, J., Lyu, Z., Wen, J., 2018. Sense-antisense gene overlap is probably a cause for retaining the few introns in *Giardia* genome and the implications. Biol. Direct 13, 1–7.

Yichoy, M., Nakayasu, E.S., Shpak, M., Aguilar, C., Aley, S.B., Almeida, I.C., Das, S., 2009. Lipidomic analysis reveals that phosphatidylglycerol and phosphatidylethanolamine are newly generated phospholipids in an early-divergent protozoan, *Giardia lamblia*. Mol. Biochem. Parasitol. 165, 67–78.

Yichoy, M., Duarte, T., De Chatterjee, A., Mendez, T., Aguilera, K., Roy, D., Roychowdhury, S., Aley, S., Das, S., 2011. Lipid metabolism in *Giardia*: a post-genomic perspective. Parasitology 138, 267–278.