

Title:

Transcriptomic profiling of high-density Giardia foci encysting in the murine proximal intestine

¹JK Pham, ¹C Nosala, ²EY Scott, ¹KF Nguyen, ¹KD Hagen, ¹HN Starceovich, and ^{*}¹SC Dawson

¹Department of Microbiology and Molecular Genetics

²Department of Animal Science

One Shields Avenue

UC Davis

Davis, CA 95616

Corresponding author

scdawson@ucdavis.edu

1 **Abstract.** *Giardia* is a highly prevalent, understudied protistan parasite causing significant diarrheal
2 disease worldwide. Its life cycle consists of two stages: infectious cysts ingested from contaminated
3 food or water sources, and motile trophozoites that colonize and attach to the gut epithelium, later
4 encysting to form new cysts that are excreted into the environment. Current understanding of parasite
5 physiology in the host is largely inferred from transcriptomic studies using *Giardia* grown axenically or
6 in co-culture with mammalian cell lines. The dearth of information about the diversity of host-parasite
7 interactions occurring within distinct regions of the gastrointestinal tract has been exacerbated by a
8 lack of methods to directly and non-invasively interrogate disease progression and parasite physiology
9 in live animal hosts. By visualizing *Giardia* infections in the mouse gastrointestinal tract using
10 bioluminescent imaging (BLI) of tagged parasites, we recently showed that parasites colonize the gut in
11 high-density foci that may cause localized pathology to the epithelium. Encystation is also initiated in
12 these foci throughout the entire course of infection, yet how the physiology of parasites within high-
13 density foci in the host gut differs from that of cells in laboratory culture is unclear. Here we use BLI to
14 precisely select parasite samples from high-density foci in the proximal intestine to interrogate *in vivo*
15 *Giardia* gene expression in the host. Relative to axenic culture, we noted significantly higher expression
16 (> 10-fold) of oxidative stress, membrane transporter, and metabolic and structural genes associated
17 with encystation in the high-density foci. These differences in gene expression within parasite foci in
18 the host may reflect physiological changes associated with high-density growth in localized regions of
19 the gut. We also identified and verified six novel cyst-specific proteins, including new components of
20 the cyst wall that were highly expressed in these foci. Our *in vivo* transcriptome data support an
21 emerging view that parasites encyst early in localized regions in the gut, possibly as a consequence of
22 nutrient limitation, and also impact local metabolism and physiology.

1 Introduction

2 *Giardia lamblia* is a zoonotic protozoan parasite causing acute and chronic diarrheal disease in over one
3 billion people worldwide, primarily in areas lacking adequate sanitation and water treatment [1; 2].
4 Giardiasis is a serious disease of children, who may experience substantial morbidity including diarrhea,
5 malnutrition, wasting, and developmental delay [3; 4; 5]. In the United States, giardiasis is the most
6 frequently diagnosed cause of water-borne diarrheal disease, and commonly affects travelers and
7 immunosuppressed individuals [6]. The estimated failure rates of up to 20% for standard treatments [7]
8 and reports of drug resistance [8; 9; 10] underscore the need for new therapeutic treatments for this
9 widespread and neglected diarrheal disease.

10 Giardiasis may be either acute and/or chronic, and infection is generally accompanied by
11 abdominal cramps, gas, nausea, and weight loss. Giardiasis may also result in a severe form of
12 malabsorptive diarrhea presenting as a fatty, watery stool [11]. Trophozoites are not invasive and
13 giardiasis does not produce a florid inflammatory response; however, giardiasis is associated with villus
14 shortening, enterocyte apoptosis, hypermobility, and intestinal barrier dysfunction [12]. The
15 mechanisms by which *Giardia* colonization of the gastrointestinal tract induces diarrheal disease
16 remain unclear.

17 *Giardia* differentiates into two morphological forms during its life cycle: a motile, multi-
18 flagellated trophozoite that colonizes the host small intestine; and an infective cyst that is shed into the
19 environment [13]. Following ingestion by a mammalian host, cysts transform into motile trophozoites
20 as they pass into the gastrointestinal tract. Trophozoites navigate the lumen of the small intestine and
21 attach to the microvilli of the small intestine via the ventral disc, yet do not invade the epithelium [6].
22 Attachment to the gut epithelia allows the parasite to resist peristaltic flow and proliferate in this low

oxygen, nutrient rich environment. Trophozoites are eventually triggered to differentiate into cysts, which exit the host and are disseminated with feces into the environment.

The lack of accessibility to the gastrointestinal tract [14; 15; 16] has limited our understanding of *in situ* *Giardia* physiology in the host. Parasite physiology and metabolism have been inferred from axenic laboratory culture in a non-defined medium [17] or from co-culture with intestinal epithelial cell lines [14; 16; 18]. However, co-incubation studies of trophozoites with intestinal cell lines may not accurately reflect *in vivo* parasite physiology. *Giardia* infects many mammalian hosts, and animal models of giardiasis include adult [19; 20] or suckling mice [21] or adult gerbils [22] infected with either *Giardia lamblia* or *G. muris* isolates [23]. While the adult mouse model of giardiasis is commonly used to evaluate anti-giardial drugs [24], previous *in vivo* parasite studies lack precision due to tissue sampling throughout the intestinal tracts of infected animals without corresponding knowledge of *Giardia* colonization.

Our recent development of methods for non-invasive imaging of bioluminescent *Giardia* parasites permits unprecedented access to real-time host-parasite interactions, allowing us to revise decades-old assumptions about *Giardia* infection and encystation dynamics in living hosts. BLI enables sensitive quantification and live reporting of transcriptional activity of promoter-luciferase fusions [25; 26; 27] but also protein-luciferase fusions [28]. BLI is used commonly to monitor parasitic infection dynamics during malaria, leishmaniasis, trypanosomiasis, and toxoplasmosis [29; 30; 31], as well as bacterial colonization of the intestine [32]. The use of imaging methods to evaluate *in vivo* giardiasis has allowed the precise longitudinal and spatial monitoring of the dynamics of infection, and provides an improved method to evaluate anti-giardial drugs in a relevant animal model of giardiasis.

Using BLI, we recently demonstrated that metabolically active *Giardia* primarily colonizes the

proximal small intestine, rather than the mid-jejunum as has been previously reported [15; 19; 33; 34]. By imaging mice infected with *Giardia* expressing firefly luciferase under the control of encystation-specific [32; 35; 36; 37; 38] promoters, we determined that encystation is initiated early in the course of infection and is correlated to dense sites, or foci. These foci persist in the gastrointestinal tract throughout the course of infection.

Here we interrogate the *in vivo* physiology of high-density *Giardia* foci in the proximal small intestine and compare this to physiology in axenic culture. We observe that also upregulated in parasite metabolism in the foci is shifted toward encystation, and that genes involved in response to oxidative stress are also upregulated. Furthermore, we observe that trophozoites in high-density foci are encysting and exhibit a expression profiles to *in vitro* *Giardia* cells in profile similar to the profile seen during mid-to-late encystation (7-22 hours in encystation media) [39]. Finally, we identify and confirm six new encystation genes that are highly expressed in the foci.

Materials and Methods

Giardia trophozoite and encystation culture conditions

The *G. lamblia* (ATCC 50803) bioluminescent reporter *P_{GDH}-Fluc* [40] strain, all C-terminal GFP-fusion strains, and the wild type WBC6 strain were cultured in modified TYDK medium (also known as TYI-S-33 or Keister's medium [17]) supplemented with bovine bile and 5% adult and 5% fetal bovine serum [56] in sterile 16 ml screw-capped disposable tubes (BD Falcon), and incubated upright at 37°C without shaking. Encystation was induced *in vitro* by decanting TYDK medium from 24-hour cultures (roughly 30% confluent) and replacing with encystation medium modified by the addition of 0.5 grams/liter bovine bile, pH 7.8 [39]. After 24 hours, cysts settled and were harvested for subsequent imaging

analyses.

In vivo and ex vivo bioluminescence imaging of mice with spatial sampling of Giardia colonization

Animal studies were performed with IACUC approval at the University of California, Davis (Scott C. Dawson, PI). Protocols for *in vivo and ex vivo* bioluminescence imaging of *Giardia*-infected mice were as previously described [40]. Specifically, four eight week old, female C57/B6/J mice (Jackson Laboratories) were maintained on *ad libitum* water and alfalfa-free irradiated rodent pellets (Teklad 2918). Water was supplemented with 1 mg/ml ampicillin and neomycin (Teknova) for five days prior to infection [41] to promote parasite colonization, and mice were maintained on antibiotics for the entire experiment. For infections, each animal was gavaged with 1×10^7 *G. lamblia* *P_{GDH}-Fluc* strain trophozoites in 100 μ L phosphate-buffered saline [40].

For non-invasive monitoring of parasite colonization using *in vivo* BLI, mice were sedated, D-luciferin (30 mg/kg) was then injected intraperitoneally, and bioluminescence was imaged using an IVIS Spectrum (PerkinElmer) with an open emission filter [40]. Photons were quantified using an ultra-sensitive CCD camera (IVIS Spectrum) and the resulting heat maps of bioluminescent photon emission intensity were overlaid on still images of anaesthetized animals. To allow the D-luciferin to distribute throughout the body, images were collected with two-minute exposures constantly over 8-10 minutes until the bioluminescent signal stabilized. The exposure time for final image collection ranged from two to five minutes due to differences in the signal strength. Bioluminescence was quantified using LivingImage software to draw a region of interest (ROI) around each mouse, from front paws to anus. BLI data was quantified as total flux (photons/second) for exposure time-independent quantification of signal intensity.

To image the *ex vivo* spatial location and density of *Giardia* in murine gastrointestinal tract,

sedated mice were euthanized on day 3 or day 7 post-gavage by cervical dislocation. The entire gastrointestinal tract was quickly dissected from esophagus to anus and positioned within a plastic Petri dish for *ex vivo* imaging using thirty second exposures [40]. ROI analysis was used to quantify bioluminescence (LivingImage). Total gastrointestinal tract signal was analyzed within a circle encompassing the entirety of the petri dish. The stomach, proximal SI (first half), distal SI (second half), cecum, and large intestine were traced using the free-hand tool and total flux was recorded as stated above (Supplemental Table 1). Spatial imaging of bioluminescence in the entire gastrointestinal tract specified the particular intestinal segments enriched for high-density *Giardia* foci of colonization. Small intestinal segments with such foci were excised for downstream RNA extraction and stored in RNeasy (Life Technologies) at -80°C, or alternatively, were fixed in 10% phosphate-buffered formalin for histology.

Histological sections of these intestinal samples were stained and visualized using light microscopy, confirming the presence of *Giardia* trophozoites in selected intestinal samples [40]. Protocols for preparation of histological slides were the same as those described in Chen et al. [42] and were performed by the UC Davis Anatomic Pathology and IHC Laboratory (Davis, CA). Briefly, a subset of the excised bioluminescent intestinal samples was fixed in 10% phosphate-buffered formalin for 24 hours. Specimens were dehydrated in a series of graded ethanol (70%, 96% and 99%) and embedded in paraffin. Five-micron sections were cut perpendicular to the mucosa surface and the paraffin was cleared from the slides with coconut oil (over 15 min, 60°C). Sections were rehydrated in 99%, 96% and 70% ethanol followed by a 10 minute wash in water and stained with hematoxylin and eosin (HE). All HE slides were visualized via brightfield microscopy using a Leica DMI 6000 wide-field inverted fluorescence microscope with PlanApo 10X and 40X objectives.

Total RNA extraction and RNA-seq of gastrointestinal samples in infected animals

Total RNA was extracted by from *Giardia*-infected intestinal segments by resuspending the tissue in RNA STAT 60 solution (Tel-Test, Inc.) and repeatedly pipetting through a glass pipette until tissue homogenization was complete. All samples were kept on ice during tissue homogenization. Total RNA was purified by phenol-chloroform extraction and isopropanol precipitation. Poly(A) selection was performed with one microgram of total RNA. Illumina sequencing libraries were constructed and ribosomal RNA was depleted (UC Davis DNA Technologies Core Facility, Davis, CA) for downstream Illumina sequencing. 120-180 million 50 bp, single end reads were subsequently generated via Illumina sequencing (HiSeq 2500 system).

Trimming and processing of transcriptome sequences

Raw sequences were trimmed using the sliding window trimmer Sickle [43]. Two pipelines were used for mapping and transcript quantification (Figure 1D). First, the reads were mapped to the *Giardia* assemblage A genome (WBC6, release 5.0) using TopHat (version 2.0.14) [44]. The TopHat output BAM files were processed and assembled into transcripts by Cufflinks (version 2.2.1) [45] and normalized to fragments per kilobase of transcript per million fragments mapped (FPKM, see Supplemental Table 2). Differentially expressed genes were identified using the cuffdiff function of the Cufflinks package (version 2.2.1) [46]. Differentially expressed genes were defined using a threshold cutoff false-discovery rate (FDR) of 5%. The R package cummeRbund was used to generate summary figures for the resulting RNA-seq data, such as heat maps, PCA plots, and gene clusters of differentially expressed genes (via partitioning tools) [46]. The second pipeline mapped the reads back to the *Giardia*

transcriptome ([GiardiaDB-5.0 GintestinalisAssemblageA.gff](#)) with Kallisto [47], using the recommended 30 bootstraps per sample. The companion package, Sleuth [47], was used for differential gene expression analysis, using the Wald test and significance q-value cutoff of 0.05. Transcripts showing concordant significant differential gene expression between Cuffdiff and Sleuth statistical packages were selected as confident DE genes.

For visualization, FPKM values of the 475 confident DE genes were hierarchically clustered. Bi-clustering was performed by Pearson correlation for transcript expression and Spearman correlation of expression profiles associated with each sample. The heatmap (Supplemental Table 3) was plotted using the heatmap.2 function of “gplots” the R software package (<https://CRAN.R-project.org/package=gplots>).

Functional transcriptome annotation, metabolic pathway mapping, and partitioning analyses

Cellular functions were inferred for significant differentially expressed *Giardia* genes using analytical tools associated with the GiardiaDB (<http://giardiadb.org/giardiadb/>) [48], with subsequent manual inspection and curation of homology, functional domains, and cellular process designations according to prior experimental and bioinformatic evidence. GiardiaDB analytical tools were used to compare *in vivo* transcriptomic reads to prior transcriptomic and proteomic experimental studies [34; 49; 50]. In addition, the cummeRbund R package was used for partitioning analysis to predict nine differentially expressed gene clusters using the Jenson-Shannon distance binning methods with the cummeRbund software package [46], and compared to both the *in vivo* foci gene clusters and prior Giardia encystation and Caco-2-associated transcriptome studies [34; 49; 50].

C-terminal GFP fusion strain construction of select differentially expressed genes

To confirm the cellular localization of novel differentially expressed genes identified in the *in vivo* transcriptome, 15 differentially expressed *Giardia* genes (GL50803_7797, GL50803_4705, GL50803_5258, GL50803_7710, GL50803_15499, GL50803_10822, GL50803_15427, GL50803_24412, GL50803_14748, GL50803_6679, GL50803_88960, GL50803_5515, GL50803_14567, GL50803_14926, GL50803_9354) were GFP-tagged via our laboratory's Gateway cloning pipeline [51]. We also tagged fourteen genes that were more highly expressed in *in vitro* culture. The C-terminal GFP fusion constructs included approximately 200–250 nucleotides upstream of the gene, the gene itself in frame with GFP, and a puromycin resistance cassette [51]. The *Giardia* strain WBC6 was electroporated with 20 µg of GFP-fusion plasmids, and transformed strains were maintained under antibiotic selection (50 µg/ml puromycin) for at least two weeks [51].

Immunostaining and imaging of encysting GFP-tagged strains

To visualize encystation vesicles and the mature cyst, GFP-tagged trophozoites were encysted using established *in vitro* encystation methods [52]. Sterile deionized water was then added to all experimental culture tubes to lyse incompletely encysted cells. Water resistant cysts were stored at 4°C for subsequent imaging. Cyst wall protein 1 (CWP1) was visualized by immunostaining fixed encysting trophozoites attached to coverslips using a 1:200 dilution of anti-CWP1 primary antibody (Waterborne, Inc., New Orleans, LA) and a goat anti-mouse antibody conjugated to Alexa Fluor 594 (Invitrogen) as previously described [52]. Coverslips were mounted onto slides using Prolong Gold Antifade Solution with DAPI (Invitrogen).

Three dimensional stacks of immunostained samples were acquired using the Metamorph image acquisition software (MDS Technologies) with a Leica DMI 6000 wide-field inverted fluorescence microscope with a PlanApo 100X, NA 1.40 oil immersion objective. Serial sections of GFP-tagged strains were acquired at 0.2 μm intervals and deconvolved using Huygens Professional deconvolution software (SVI). Two dimensional maximum intensity projections were created from the 3D stacks for presentation.

Results

Bioluminescent imaging allows precise sampling of regions of the proximal small intestine with high-density Giardia foci for transcriptomics

A cohort of four of mice were inoculated with the constitutive metabolic bioreporter strain *P_{GDH}-FLuc* [40]. *Giardia*-specific bioluminescence was quantified non-invasively in animals from one to seven days post-inoculation (Figure 1A and Methods). Two animals were sacrificed at day 3 p.i., and two more at day 7 p.i. *Ex vivo* imaging of the gastrointestinal tracts revealed dense foci of bioluminescent signal primarily in the proximal and distal small intestine (Figure 1A); bioluminescent flux directly correlates with the number of active *Giardia* parasites [40]. Proximal intestinal segments from each gastrointestinal tract were sampled, and the bioluminescent flux of the *Giardia* foci (5.12×10^5 to 1.5×10^7 photons/second/cm (Figure 1B)) was determined. Representative proximal small intestinal samples (ROI) from each gastrointestinal tract were selected for transcriptomic analyses. The histology of these segments indicated no verifiable host tissue damage (Methods), but did indicate obvious colonization of these regions by *Giardia* trophozoites (Figure 1A and 1C).

Sufficient purified total RNA was obtained from four *in vivo* intestinal samples and was used to

create RNA-seq libraries for sequencing (Methods and Figure 1B). Four additional RNA-seq libraries were created and sequenced from four independent *in vitro* *Giardia* *P_{GDH}-FLuc* cultures grown axenically in standard medium (Methods). Raw sequence reads generated for the RNA-seq samples ranged from 10.9 to 21.3 million (Figure 1B). The majority of RNA-seq reads from the *in vitro* culture (TYDK) datasets (97.7% to 98.7%) were mapped back to the WBC6 genome. Due to the high proportion of host total RNA in the *ex vivo* samples, fewer RNA-seq reads (0.8% to 1.9%) were mapped to the WBC6 genome; however, these numbers were comparable between individual *in vivo* samples and were sufficient for subsequent analyses of differential gene expression (Figure 1B and Methods).

Identification of 475 differentially expressed (DE) genes between in vitro culture and in vivo foci transcriptomes

Transcriptomic comparisons and analyses between *in vitro* and *in vivo* samples were performed using the TopHat, Cufflinks, CummeRbund, Kallisto, and Sleuth packages as outlined in the bioinformatic strategy in Figure 1D and [46]. A comparison of the normalized read (FPKM) profiles of the *in vitro* and *in vivo* transcriptomes indicate that the four *in vitro* profiles were very similar to one another, as were the four *in vivo* foci profiles (see heat map, Figure 2A and Supplemental Table 2). Significant differences in gene expression profiles between the *in vitro* and *in vivo* samples (Figure 2A) were quantified and summarized using two different computational methods (Cuffdiff and Sleuth). Cuffdiff and Sleuth identified 1073 and 1336 differentially expressed genes, respectively, between the *in vitro* and *in vivo* datasets. Four hundred seventy-five differentially expressed genes were concordant between the two methods (Figure 2B). One hundred eighty-seven of the concordant differentially expressed genes had increased expression in the *in vivo* transcriptomes relative to the *in vitro* transcriptomes (Figure 2C).

Two hundred eighty-eight of the concordant differentially expressed genes had decreased expression in the *in vivo* foci transcriptomes relative to *in vitro* transcriptomes. All subsequent analyses focused on the concordant genes with expression levels that varied more than two-fold between the *in vivo* and *in vitro* transcriptomes. These included 187 genes that were more highly expressed *in vivo* foci and 166 that were more highly expressed *in vitro* culture. (Figure 2C and Supplemental Tables 4 and 5). Sixty-one genes that were expressed *in vitro* were not expressed in the *in vivo* foci transcriptomes (Supplemental Table 6).

Statistically significant increases in gene expression ranging from 2.5- to 153-fold were observed for *Giardia* in *in vivo* foci relative to trophozoites cultured *in vitro*. With respect to highly expressed genes, transcripts associated with translation and arginine metabolism were the most abundant in the *in vitro* transcriptomes (Figure 2D and Supplemental Tables 4 and 6). In contrast, transcripts associated with encystation, oxidative stress, or unknown functions are among the most abundant in the *in vivo* foci transcriptomes (Figure 2E). Despite significant differences between *in vivo* and *in vitro* expression profiles, five genes (ornithine carbamoyl transferase (OCT, 10311), ubiquitin (7110), 14-3-3 protein (6430), ribosomal protein L36-1 (16114) and ribosomal protein S24 (10367) are among the top transcribed genes in both transcriptome datasets (Figure 2D and 2E).

Encystation-specific and oxidative stress genes are significantly upregulated in the in vivo foci

Of the 475 differentially expressed genes, 25 genes were expressed more than 10-fold higher in the *in vivo* foci relative to cultured cells; however, only 8 genes were expressed at levels more than 10-fold higher in cultured cells than *in vivo* (Figure 3A and Supplemental Table 4 and 5). The expression of 85 genes was elevated more than 5-fold in foci relative to cultured trophozoites, and there were 29 genes

whose expression was more than 5-fold higher in the *in vitro* transcriptomes relative to the *in vivo* transcriptomes (Supplemental Tables 4 and 5).

In contrast to the genomes of model organisms, many genes within the WBC6 genome lack functional gene annotations associated with *Giardia*-specific processes (e.g., encystation or pyruvate metabolism). Thus, we manually curated the functional annotations of genes whose expression was more than 2-fold higher *in vivo* (187 genes) or *in vitro* (166 genes) and summarized the functional categories associated with these statistically significant differentially expressed genes (Figure 3). Ten of the 25 genes with more than 10-fold higher expression in the *in vivo* foci are associated with encystation-specific processes. These genes encode components of the cyst wall (CWP2, 153X higher *in vivo*), enzymes in the UDP-N-acetylgalactosamine biosynthetic pathway (GL50803_14259, 146X higher *in vivo*), and high cysteine membrane proteins (HCMPs, 10X-20X higher *in vivo*). Eight genes with more than 10-fold higher expression in the *in vivo* foci have unknown or uncharacterized functions (Figure 3A and Supplemental Table 4).

Encystation-specific gene expression in *in vivo* foci becomes even more apparent when we consider the functional annotations of genes with more than 2-fold higher expression in *in vivo* foci relative to *in vitro* culture are summarized. Thirty-seven differentially expressed genes are associated with encystation, 17 with oxidative stress responses, and 15 with membrane transporter functions (Figure 3B and Supplemental Table 4). This is in contrast to genes highly expressed *in vitro* culture, which include cytoskeletal genes and other genes associated with housekeeping functions such as translation, RNA processing, and transcription (Figure 3C and Supplemental Table 5). Different types of genes with membrane trafficking functions are increased in the two transcriptome dataset. About equal numbers of unique variant specific surface proteins (or VSPs, [53]) are differentially expressed

amongst the *in vitro* (7 VSPs) and *in vivo* foci (8 VSPs) transcriptomes.

Lastly, genes with unknown or uncharacterized function are the most abundant category in both the *in vivo* foci transcriptome (75/187) and the *in vitro* transcriptome (82/166) datasets.

In vivo foci display increased expression of genes associated with the UDP-N-acetylgalactosamine, pentose phosphate, and glycolytic and pyruvate pathways

To evaluate the degree to which the differential gene expression we observed is associated with specific metabolic pathways in *Giardia*, we mapped metabolic genes to four catabolic and biosynthetic pathways (Figure 4) using manual curation, as well as GO term and metabolic pathway enrichment tools available from GiardiaDB [48]. Specifically, we mapped the transcriptional upregulation of genes associated with *in vivo* foci to glycolysis and pyruvate metabolism, the pentose phosphate pathway, the UDP-N-acetylgalactosamine pathway, and the arginine dihydrolase pathway (Figure 4). First, we noted that in *in vivo* foci, the three key enzymes of the arginine dihydrolase pathway (OCT, ADI, CK) showed 3.7-6.4X less gene expression in *in vivo* foci than in trophozoites grown in *in vitro* culture.

The carbohydrate component of the cyst is comprised of a D-GalNAc(b1,3)-D-GalNAc homopolymer synthesized from fructose 6-phosphate via the N-acetylgalactosamine (GalNAc) biosynthetic pathway (Figure 4). As has been seen previously during the transcriptional activation of GalNAc pathway enzymes during *in vitro* encystation [54; 55], we found that almost all enzymes in this pathway were expressed between 3.2 to 26X higher in the *in vivo* foci as compared to in the *in vitro* transcriptomes. In fact, the second enzyme in this pathway, glucosamine 6-phosphate *N*-acetyltransferase (GL50803_14651), is one of the top 25 highly expressed genes in *in vivo* foci (Figure 3B and Supp table).

Key enzymes involved in the generation of ATP via glycolysis (GCK, 3.3X, and GAPDH, 13X) and pyruvate metabolism (AcoAS, 3.8X) were also more highly transcribed in the foci. PGM expression was several fold (2.1X) lower in the foci transcriptome relative to *in vitro* transcriptome. Pyruvate metabolism in *Giardia* is sensitive to oxygen concentration; at lower oxygen concentrations, pyruvate is fermented to alanine and ethanol, whereas at higher oxygen concentrations pyruvate is converted to acetate with the concomitant production of ATP [56]. In foci, we observed increased expression of AcoAS, a key enzyme in the higher oxygen pathway resulting in acetate production. We also noted that AAT, involved in alanine production, was less expressed in foci than *in vitro* culture (-2.4X). Lastly, two key enzymes of the pentose phosphate pathway (G6PD, 5.8X and PGD, 8.4X) are more highly expressed in the *in vivo* foci transcriptomes (Figure 4).

Partitioning analysis identifies highly expressed in vivo foci genes associated with encystation

Partitioning analysis sorted nine gene clusters of similar expression levels, and each cluster ranged from between 39 and 218 differentially expressed genes. On average, about 50 percent of the genes identified using the differential expression analyses (Methods) were shared with genes in clusters identified using the partitioning analysis (Figure 5B). The gene clusters that were the most highly transcribed in the foci (cluster A, 15X, cluster B, 5X and cluster C, 3X, respectively) were enriched for encystation, oxidative stress response, and membrane transporter genes (Figure 5C). Cluster A included the most encystation-associated genes (Figure 5B and Supplemental Tables 7 and 8). These three clusters also shared genes in common with several recent studies of *in vitro* encystation genes (Morf et al.), putative cyst-specific genes (Palm et al.), and putative host-induced parasite genes (Ringqvist et al.)

(Figure 5B).

Confirmation of six new cyst-associated proteins expressed at higher levels in the in vivo foci

To evaluate whether some of the genes of unknown function identified using either differential expression or partitioning analyses were associated with encystation, we tagged fifteen genes using C-terminal GFP tags and created GFP-fusion strains [51]. Based on fluorescence microscopy of trophozoites of these GFP fusion strains, thirteen of the fifteen strains had interphase localization in trophozoites (Supplemental Figure 1). Six had localization in cysts (Figure 5D), with three localizing to the cyst wall: GTA-2 (15427), a START domain protein (15499), and one hypothetical protein (4705). Three GFP fusions localized to the interior region of the mature cyst including alpha-5-giardin (7797), a conserved hypothetical protein (10882), and one hypothetical protein (5515). The 4705::GFP and 7797::GFP strains lacked any localization in the trophozoite stage. Fifteen other genes more highly expressed *in vitro* culture were also tagged, and fourteen had localizations in trophozoites (Supplemental Figure 2).

Discussion

Defining the *in vivo* dynamics of encystation and cyst dissemination is key to understanding *in vivo* host-parasite interactions and *Giardia's* pathogenesis [34; 57; 58]. The *Giardia* cyst is characterized by four nuclei and a thick outer cell wall consisting of cyst wall proteins 1, 2, and 3 and a unique β -1,3-linked N-acetylgalactosamine homopolymer. The structure of the cyst wall allows cysts to persist in the environment until they are ingested by a compatible host [14; 15; 59; 60; 61]. Encystation has been proposed to be triggered by differential environmental chemistry associated with distinct regions of the

gastrointestinal tract (i.e. the distal small intestine and colon), and has thus long been assumed to occur when trophozoites are swept towards the large intestine in mammalian hosts [14; 15; 59]. *In vitro* culture conditions for encystation have been developed that are thought to mimic the high pH and limiting cholesterol that trophozoites encounter in the gut [62]. While landmark *in vitro* studies established that the initiation of encystation is transcriptionally controlled [8; 9; 10], our understanding of the spatiotemporal dynamics of parasite metabolism and differentiation in the host is limited due to the inherent challenges of evaluating *Giardia* physiology *in situ* in the gastrointestinal tract.

Recent work in our laboratory challenges the paradigm that encystation is triggered late in infection via spatially segregated cues in the gastrointestinal tract [40]. By directly imaging *Giardia* infections in the mouse using *in vivo* and *ex vivo* bioluminescent imaging (BLI) of parasites with integrated firefly luciferase bioreporters of metabolism (*P_{GDH}-Fluc* strain) or encystation (*P_{CWP1}-Fluc* and *P_{CWP2}-Fluc* strains), we showed that parasites colonize both the proximal and distal small intestine in high-density foci early in infection. Encystation is also initiated early during infection in these foci, and persists throughout the entire course of infection. Barash et al. also confirmed that trophozoites extracted from proximal small intestines of infected mice displayed encystation-specific vesicles (ESVs), which develop during the differentiation of the trophozoite to the cyst [40]. This work predicted that that trophozoite colonization and encystation occur simultaneously in the host, and showed that encysting trophozoites were enriched in the high-density foci.

Here we used the same BLI *ex vivo* imaging strategy to sample high-density *Giardia* foci in the proximal small intestine of four animals infected with a constitutive luciferase strain *P_{GDH}-Fluc* (Figure 1 and [40]). Despite our deriving the samples from the gastrointestinal tracts of four different animals either three or seven days post *Giardia* infection, the transcriptomic profiles of the *in vivo* foci were

strikingly similar (Figure 2A). Thus, we predict that trophozoites synchronize their physiological responses once they reach a threshold density in regions of the proximal small intestine and are committed to encystation (Figure 2).

In vivo Giardia transcriptional profiles are dominated by encystation

Using two different analytic methods (Figure 2B and 2C), we found that the transcriptome profiles of *Giardia* from *in vivo* foci and log phase axenic cultures were significantly different. Relative to *in vitro* culture, many of the most highly expressed genes in the *in vivo* foci have functions associated with both encystation and oxidative stress response. In contrast, the top *in vitro* expressed genes are associated with translation, cell division, and arginine metabolism as has been previously reported (Figure 2D and 2E and [63; 64]).

We have previously shown that the CWP1 and CWP2 promoter fusions to luciferase are highly expressed in the *in vivo* foci using BLI, and confirmed that trophozoites in high-density foci express numerous encystation specific vesicles (ESVs) [40]. The biogenesis of ESVs that transport cyst wall proteins to the plasma membrane of the trophozoite is a hallmark of encystation. ESVs are synthesized in the endoplasmic reticulum and transported to the periphery of the cell in late encystation. Cyst wall proteins are sorted and transported by novel large vesicles, encystation secretory vesicles (ESVs) and ultimately by peripheral vesicles to the growing cyst wall. Many other components of the cyst have also been suggested by expression or proteomic analyses or confirmed by localization, including protein disulfide isomerases [65], high cysteine membrane proteins (HCMPs) [66], and EGFP family proteins [66].

In support of our assertion that *Giardia* from *in vivo* foci are already committed to encystation

[40], we demonstrate that the transcriptome of *in vivo* foci is defined by the high expression of a number of genes associated with cyst wall biosynthesis. Several aspects of trophozoite differentiation to cysts are transcriptionally regulated [60]. The leucine-rich repeat proteins, cyst wall proteins (CWP1-3) and the enzymes of the GalNAc biosynthetic pathway are all transcriptionally upregulated by the transcription factor Myb2 within 90 minutes of switching trophozoite cultures to encystation medium [67]. GalNAc enzyme transcription has been shown to peak at about 22 hours in encystation medium [39]. In the *in vivo* foci, we noted significantly higher expression of genes associated with encystation, including the transcriptional regulator Myb2, all enzymes of the GalNAc biosynthetic pathway (Figure 4), two known cyst wall proteins (CWP2 and CWP3), 17 HCMPs, and two EGF family proteins [68] (Figures 2, 3, and Supplemental Table 4). CWP1 [69] was identified as differentially expressed by only one of the two methods of analysis we used, and was thus not included in the 475 differentially expressed genes (Methods). Through partitioning analysis and analysis of differential gene expression in the *in vivo* foci, we also identified six new cyst-specific proteins (Figure 5D and Supplemental Tables 7 and 8). Each localizes to the cyst wall or the cyst interior after 24 hours of incubation in encystation medium.

Variant-specific surface proteins (VSPs) are cysteine-rich membrane proteins of variable size that are responsible for antigenic variation and likely contribute to *Giardia's* evasion from the host immune response [70]. The *Giardia* genome encodes over seventy variant surface protein (VSP) genes, yet only one VSP is expressed per cell [71]. VSP expression occurs early during *in vitro* encystation [72], and there are dramatic shifts in expression of VSP genes late in *in vitro* encystation [39]. Expressed VSPs are then transported by ESVs during encystation [6]. Fourteen VSPs were differentially expressed in our analysis, and two VSPs were among the most highly expressed proteins in *in vivo* foci relative to

in vitro culture (Figure 2).

A common symptom of giardiasis is fat malabsorption, with markedly increased lipid concentrations in luminal contents and feces [6; 73; 74]. In addition to suffering from steatorrhea, patients excreting *Giardia* cysts have altered lipid profiles in circulating blood as well as in the gut lumen [75]. *Giardia* directly perturbs lipid metabolism within the host by scavenging the gut lumen for lipids and by excreting novel end products of lipid metabolism [76]; it also contributes to shifts in the local gut microbiota [12] by both excreting novel lipids and influencing the bioavailability of bile acids [73; 77; 78; 79].

Lipids also play key roles in the regulation of encystation [76]. Glucosylceramide transferase-1 (gGlcT1), an enzyme involved in sphingolipid biosynthesis that has also been shown to play a key role in ESV biogenesis and cyst viability [80], is highly expressed in *in vivo* foci. Also highly expressed in foci is acid sphingomyelinase-like phosphodiesterase 3b precursor gSmase 3B. GSMase 3B has been shown to be transcriptionally upregulated during encystation and is proposed to scavenge ceramide from dietary components in the small intestine [80]. We also noted the increased *in vivo* expression of seven genes associated with membrane trafficking that could be involved in ESV biogenesis or transport (Figure 3 and Supplemental Table 4).

Five cysteine proteases (three cathepsin B homologs: 10217, 17516, 16779 and two cathepsin L homologs: 11209, 137680) have roughly between five and 11.5-fold higher expression in the foci (Supplemental Table 2 and 4). Cathepsin B (16779, CP3) and cathepsin L (137680) are known to be upregulated in encysting trophozoites [81]. In *Giardia*, cathepsin cysteine proteases are linked to not only to encystation [82] but also to evasion of innate and adaptive immune responses. In co-culture with rat IEC epithelial cells, one (3099) of the nine cathepsin B proteases genes has significantly

increased expression *Giardia* trophozoites [83]. Cathepsin B proteases secreted by trophozoites are sufficient to prevent inflammatory responses from the host by its degradation of the chemokine CXCL8, inhibiting chemotaxis of pro-inflammatory PMN cells to the site of parasite colonization [84].

In total, we have confirmed the expression of 37 known encystation genes and identified six new cyst-associated proteins in the high-density foci. Because we have identified many highly expressed encystation-associated transcripts in key metabolic, biosynthetic, protease, and membrane trafficking pathways, we suggest that the majority of cells in the high-density foci in the proximal intestine are in mid- to late stages encystation (Figure 6 and Supplemental Table 9). The apparent developmental synchrony of cells in the foci is supported by the strong encystation transcriptional profiles in each sample regardless of the day of infection (Figure 6) or in each animal sampled (Figure 2A) and may be the result of density related cues such as nutrient or lipid deprivation that trigger encystation.

Energy metabolism in the *in vivo* foci

As an anaerobic fermentative microbe, *Giardia* uses glycolysis, pyruvate fermentation and the arginine dihydrolase pathways for energy production ([2] and see Figure 4). Though trophozoites cannot tolerate high oxygen concentrations, they actively consume oxygen [85; 86] and may take advantage of slightly elevated oxygen levels for energetic benefit [6; 63; 87]. Thus, glucose is fermented to alanine, ethanol, acetate and CO₂, but the end product balance of parasite fermentation varies with local oxygen tension and glucose levels, with the highest NAD(P)⁺ or ATP yields occurring at the higher end of tolerated oxygen concentrations [63; 88; 89; 90]. Under strictly anaerobic conditions, pyruvate is converted into alanine and NADP is regenerated. At very low oxygen concentrations (less than 0.25

1 μM), alanine production is inhibited and pyruvate is converted primarily to ethanol with the
2 regeneration of NAD [10]. As oxygen levels increase above 0.25 μM , the ATP producing arm of pyruvate
3 metabolism is activated and acetate and ATP become the primary end products. At these higher oxygen
4 concentrations, it is hypothesized that antioxidant enzymes consume NAD(P)H to neutralize dO_2 and
5 ROS/RNS, leading to regeneration of NAD(P) for glycolysis, allowing pyruvate to be used for ATP
6 production, instead of NAD(P)H regeneration [10].

7 We noted higher expression of genes associated with energy metabolism via glycolysis (GCK,
8 3.3X and GAPDH, 13X) and pyruvate fermentation via pathways associated with higher oxygen tensions
9 (Figure 4). These expression patterns are indicative of redox stress and encystation-specific metabolism
10 associated with the *in vivo* foci. *In vivo*, acetate could be the primary end product of *Giardia*
11 fermentation, rather than ethanol or alanine, due to the relatively increased expression of ACoAS (3.8X)
12 and the decreased expression of AAT (2.4X) we observed (Figure 4). We predict the increased
13 expression of GDH is associated with parasite ammonia assimilation and the cycling of NAD(P) to
14 maintain intracellular redox balance [91].

15 Encysting trophozoites slow their catabolism of glucose for energy and switch to using glucose
16 to synthesize N-acetylgalactosamine for cyst wall synthesis ([55] and Figure 4). GalNAc is synthesized
17 only during encystation via a transcriptionally induced pathway of five enzymes that produces UDP-
18 GalNAc from fructose-6-phosphate that is diverted from ATP production in glycolysis to cyst wall
19 polysaccharide biogenesis. This diversion could result in a net loss of ATP synthesis during encystation.

20 While little is known about the regulation of glycolysis and energy production during
21 encystation, it is thought that additional ATP is generated from the arginine dihydrolase pathway
22 (ADiHP) (Figure 4). This pathway yields ATP via the conversion of arginine to ammonia and citrulline,

with the substrate-level phosphorylation of citrulline yielding ornithine and carbamoylphosphate, and resulting in NH_3 , CO_2 as end products [90]. Because *Giardia* and epithelial cells use arginine for growth, the parasite and host are proposed to influence the severe malnutrition and increased susceptibility to villus shortening observed during *Giardia* infection [92; 93]. Ornithine carbamoyltransferase (OCT) is one of the most highly expressed enzymes in culture and in the *in vivo* foci, yet we observed less expression of key enzymes in the ADiHP relative to *in vitro* culture (Figure 4). This is counter to the prevailing notion that encysting trophozoites increase arginine catabolism through the ADiHP to offset energy lost from slowing glycolysis [64].

Lastly, trophozoites also lack purine and pyrimidine synthesis, relying solely on salvage pathways to obtain these metabolites from the host and commensal microbiota [94; 95; 96]. We also discovered that several key enzymes of the pentose phosphate pathway (G6PD 5.8X and PGD 8.4X) have increased expression in the *in vivo* foci transcriptome (Figure 4). A key product of the pentose phosphate pathway (PPP) is NADPH, which plays an integral role in reductive biosynthesis as well as in the defense against oxidative stress and the maintenance of cellular redox. In *Giardia* and in *Trichomonas*, the G6PD and 6PGL enzymes are fused and are directly involved in the oxidative branch of the PPP in generating reduced NADPH (Figure 4 and [97]). This suggests that in the foci, there is increased cycling of reducing equivalents with CO_2 production and the increased production of D-ribose-5P as a key substrate for nucleotide and amino acid biosynthesis and pyrimidine/purine salvage pathways.

Oxidative stress responses are increased in the *in vivo* foci

Dynamic changes in the gut oxygen concentrations affect the formation of nitric oxide and reactive oxygen species (ROS). *Giardia* attaches to the villi of the small intestine, where oxygen levels range up

to 60 μ M [6]. Anti-oxidant defense is thought to be mediated by the thioredoxin-thioredoxin reductase system, as *Giardia* lacks superoxide dismutase, catalase, and the glutathione-glutathione reductase. This system includes an FAD containing NADPH-dependent disulfide reductase, a thioredoxin, a thioredoxin reductase (TrxR) and a thioredoxin-peroxidase [98]. As *Giardia* cysts do consume oxygen [86], the thioredoxin-thioredoxin reductase system may play a role in the maintenance of oxidation/reduction homeostasis in both trophozoites and cysts.

Giardia gene expression is modulated by oxygen and other stressors, and a microaerophilic environment may improve *Giardia's* energy metabolism. Oxygen depletion thus may reduce the efficiency of glycolysis *in vitro*, and byproducts from antioxidant enzymes are proposed in *Giardia* to promote glycolysis and energy production [6]. In oxidative environments, *Giardia* expresses NADH oxidase and flavodiiron protein that metabolize oxygen to form water [99; 100]. Nitrosative stresses can have differing metabolic effects, however; while high micromolar concentrations of nitric oxide are toxic [101], lower levels inhibit cell proliferation, encystation, and excystation [102]. Resolution of nitric oxide stress by *Giardia* is believed to be mediated by activity of its flavohemoglobin protein [103], which catalyzes the oxidation of nitric oxide to nitrate using oxygen as a co-substrate [104; 105; 106; 107].

Transcriptomic analyses of the high-density foci suggest that they are susceptible to and are responding to localized oxidative and nitrosative stresses [99; 100]. In particular, we observed the increased expression of genes associated with oxidative or nitrosative stress and detoxification, including NADPH oxidoreductase (17150, 15004), nitroreductase (FMN) family protein (8377), and flavohemoglobin (15009) at levels that were 2.5-8.5 fold higher than *in vitro* culture. In pathogenic bacteria such as *E. coli* and *Salmonella*, flavohemoglobin activity is significantly increased *in vivo*

intestinal infections, coincident with host nitric oxide production [108; 109]. In response to nitrosative stress (in the form of exposure to >5 mM nitrite), *in vitro* cultures displayed enhanced flavohemoglobin expression, and trophozoites exhibited more efficient metabolism of nitric oxide [110]. Indeed many of the top 25 expressed genes in the *in vivo* foci (e.g. NADPH oxidoreductase) are associated with oxidative or nitrosative stress responses [111], which may explain how *Giardia* can colonize the proximal part of the small intestine, a fairly aerobic portion of the intestinal tract [110].

Thioredoxin associated proteins also aid in the detoxification of oxidative stress, and five thioredoxin domain proteins (9827, 6289, 2388, 8064, 9335) were also highly expressed in the *in vivo* foci (Figure 2 and 3). One thioredoxin reductase (TrxR, 9827) has disulfide reductase activity and converts oxidized thioredoxin to its reduced form. Reduced thioredoxin is, in turn, used to prevent protein misfolding that can occur under oxidative stress conditions [8]. *Giardia's* TrxR also has NADPH oxidase activity, and during conversion of thioredoxin to its reduced form, NADP is generated from NADPH [16].

Due to these transcriptomic patterns, we predict that *Giardia* trophozoites in high-density *in vivo* foci are more likely to be subject to host-induced oxidative stressors than *Giardia* growing at lower density. Parasites in the high-density foci may take advantage of a more oxidized environment to switch to more energetically favorable pyruvate fermentation.

Density dependence of encystation and the consequences of encystation on host pathology

The physiology of parasites within the high-density foci in the host gut clearly differs from that of cells in laboratory culture or in co-culture with cell lines. The observed encystation- and oxidative stress-specific responses in gene expression within parasite foci may reflect physiological changes associated

with high-density growth in localized regions of the gut. We recently showed that increased density in culture results in the upregulation of encystation-specific promoters in *Giardia* [40]. Density-associated differentiation also occurs other protists, such as the amoeba *Dictyostelium discoideum*, wherein the accumulation of a factor secreted by starved cells leads to initiation of multicellular aggregation and assembly of a fruiting body [17]. Differentiation of *Trypanosoma brucei* is also dependent on cell density. When the slender, proliferative bloodstream form reaches a specific threshold density, accumulation of secreted ‘stumpy induction factor’ induces differentiation to the stumpy, transmissive form that is taken up by the tsetse fly [18]. At high-density in the mouse, *Giardia* responds in a way that promotes differentiation and transmission.

The similarity of transcriptional profiles between all sampled foci indicates that trophozoites have already committed to encystation, suggesting they have received the inducing encystation stimulus within the last 12 hours [58]. We hypothesize that once parasites reach threshold high densities in discrete regions of the gut, they could trigger localized host immune responses and redox shifts, resulting in the observed increase in oxidative stress and encystation responses that define transcription in the foci. Current hypotheses about how giardiasis induces host symptoms have focused primarily on specific damage induced by parasite attachment to the host gut epithelium or on the production of anti-microbial metabolites by the mammalian host.

Lastly, we have recently reported that *Giardia* infection results in a dysbiosis throughout the gastrointestinal tract in mice, characterized by a shift in the diversity of commensal microbiota towards more aerobic Proteobacteria species [112]. As we find that encystation-specific metabolism occurs early and consistently during *Giardia* infection in the host, it is likely that nutrient deprivation and waste production during encystation would affect the nutrition of the host and associated commensal

microbiome. Further investigation of *Giardia*-host-microbiome interactions should thus emphasize the study of such interactions *in situ* in order to vet laboratory culture and co-culture studies.

ACKNOWLEDGEMENTS

Bioluminescent imaging was performed at the Center for Molecular and Genomic Imaging (CMGI), University of California, Davis with gracious help and training from Jennifer Fung and Charles Smith. The authors especially acknowledge Sarah Guest for the *Giardia* life cycle drawings. This study was supported by NIH R01AI077571 and NIH R21 AI119791 awards to SCD.

Figure and Table Legends

Figure 1. Bioluminescent imaging allows precise sampling of *Giardia* foci in the proximal small intestine for transcriptomic analyses.

A cohort of four mice were infected with one million trophozoites by gavage of the *P_{GDH}-FLuc* *Giardia* strain [40]. Whole body images were non-invasively collected for parasite bioluminescence from the constitutive *P_{GDH}-FLuc* bioreporter (A), and mice were subsequently sacrificed at 3 or 7 days post infection (p.i.). Image overlays display the *P_{GDH}-FLuc* bioluminescence intensity, with the highest signal intensity shown in red and the lowest in blue. The colored radiance scale indicates the photon flux (p/s/cm²/sr) for each intestinal segment. *Ex vivo* imaging defined four foci in the proximal small intestine (yellow numbered boxed regions) subsequently excised for transcriptomic sequencing and analysis (A). In the schematic, regions of the gastrointestinal tract are marked as: S = stomach, pSI = proximal small intestine, dSI = distal small intestine, CEC = cecum, and LI = large intestine, with the positions of the *Giardia* foci (ROI1-ROI4) within the excised fragments in red. In B, the enrichment of *Giardia* trophozoites (white arrow) in a representative bioluminescent small intestinal sample was verified using light microscopy of histological sections. Panel C summarizes the total transcriptomic reads for each *in vitro* culture replicate (TYDK) and for each transcriptomic sample (ROI1-ROI4) from the proximal small intestine (pSI). The total reads include those from the host (mouse) before the *Giardia* reads were computationally mapped the *Giardia* genome. The total flux (p/s/cm²/sr) was quantified for that particular sample. The overall transcriptomic analysis is summarized in panel D.

Figure 2. Summary of differentially expressed *in vitro* and *in vivo* foci transcriptomes using Cuffdiff and Sleuth.

Panel A depicts a heat map of normalized FPKM comparing RNA-seq datasets from the *Giardia* grown *in vitro* culture (TYDK) to *Giardia* transcriptomes associated with the *in vivo* foci excised from the proximal small intestine (pSI) (see Figure 1). Two computational methods of identifying differentially expressed transcripts in the *in vitro* (TYDK) and *in vivo* (foci) datasets have some overlap (B). Overlap in both upregulated and downregulated genes in the *in vivo* transcriptomes was calculated and shown in (C). The asterisk indicates upregulated genes and downregulated genes that were identified using both methods. The top 25 expressed genes (FPKM) present in the *in vitro* (D) and *in vivo* foci transcriptomes (E) are also ranked. Underlined gene names are those that occur in the top 25 expressed genes in both the *in vitro* and the *in vivo* transcriptome datasets.

Figure 3. Encystation and oxidative stress genes are the top differentially expressed genes associated with the *in vivo* foci transcriptomes.

In panel A, genes that are greater than 10-fold differentially expressed in the *in vivo* foci transcriptomes relative to *in vitro* culture are ranked and colored according to cellular process (e.g., encystation). Genes differentially expressed greater than two-fold are ranked according to the number that are associated with a particular cellular process in the *in vivo* foci transcriptome (B) and the *in vitro* (TYDK) transcriptome (C).

Figure 4. Key enzymes in the glycolytic, pentose phosphate, pyruvate, and UDP-N-acetyl-

galactosamine pathways are upregulated in the *in vivo* *Giardia* foci.

Diagrammatic representation of differentially expressed enzymes associated with *in vivo* *Giardia* energy and biosynthetic pathways. Red shading denotes increased expression of *in vivo* relative to *in vitro* transcripts, and blue shading denotes decreased expression. White shading indicates no significant differential expression between *in vivo* foci and *in vitro* (TYDK) transcriptomes. Oxygen concentrations associated with different branches of the pyruvate metabolic pathway are noted [56]. The enzyme abbreviations, names, GiardiaDB (50803) ORFIDs, and Enzyme Commission numbers are:

Glycolysis: ACYP, acylphosphatase (7871), EC 3.6.1.7; ENO, enolase (11118), EC 4.2.1.11; FBA, fructose-bisphosphate aldolase (11043), EC 4.1.2.13; GAPDH, glyceraldehyde-3-phosphate dehydrogenase (17043), EC 1.2.1.12; GPI, glucose-6-phosphate isomerase (9115), EC 5.3.1.9; GCK, glucokinase (8826), EC 2.7.1.2; PFK, phosphofructokinase (14993), EC 2.7.1.90; PGAM, phosphoglycerate mutase (8822), EC 5.4.2.1; PGK, phosphoglycerate kinase (90872), EC 2.7.2.3; PGM, phosphoglucomutase (17254), EC 5.4.2.2; PK, pyruvate kinase (17143,3206), EC 2.7.1.40; POD, pyruvate:orthophosphate dikinase (9909), EC 2.7.9.1; TPI, triose phosphate isomerase (93938), EC 5.3.1.1.

Pyruvate metabolism: ATA, aspartate transaminase (91056), EC 2.6.1.1; FeADH, Fe-alcohol dehydrogenase (3861, 3593) EC 1.1.1.1; ACoAS, acetyl-CoA synthetase (13608); AAT, alanine aminotransferase (12150, 16353), EC 2.6.1.2; biADH (bifunctional alcohol/aldehyde dehydrogenase E (93358), EC 1.1.1.1; GDH, NADP-specific glutamate dehydrogenase (21942) EC 1.4.1.3, 1.4.1.4; Fd, ferredoxin (9662, 27266,10329); HYD, hydrogenase (6304), EC 1.12.7.2; MAL, malic enzyme (14285), EC 1.1.1.38; MDH, malate dehydrogenase (3331), EC 1.1.1.37; PEPCK, phosphoenolpyruvate

carboxykinase (10623), EC 4.1.1.32; PFOR, pyruvate:ferredoxin oxidoreductase (17063, 114609);

Pentose Phosphate Pathway (PPP): G6PD-6PGL, glucose-6-phosphate-1-dehydrogenase (8682), EC 1.1.1.49; PGD, phosphogluconate dehydrogenase (14759), EC 1.1.1.44;

UDP-N-acetylgalactosamine (GalNac) biosynthetic pathway: GNPDA, glucosamine-6-phosphate deaminase (8245), EC 3.5.99.6; GNPAT, glucosamine 6-phosphate *N*-acetyltransferase (14651), EC 2.3.1.4; PGM3, phosphoacetylglucosamine mutase (16069), EC 5.4.2.10; UAE, UDP-*N*-acetylglucosamine 4-epimerase (7982), EC 5.1.3.2; UAP, UDP-*N*-acetylglucosamine diphosphorylase (16217), EC 2.7.7.23.

Arginine Dihydrolase Pathway: ADI, arginine deaminase (112103), EC 3.5.3.6; ARG-S, arginyl-tRNA synthetase (10521), EC 6.1.1.19; CK, carbamate kinase (16453), EC 2.7.2.2; NOS, nitric oxide synthase (91252), EC 1.14.13.39; OCD, ornithine cyclodeaminase (2452), EC 4.3.1.12; OCT, ornithine carbamoyltransferase (10311), EC 2.1.3.3; ODC, ornithine decarboxylase (94582), EC 4.1.1.17; PRO-S, prolyl-tRNA synthetase (15983), EC 6.1.1.15;

Figure 5. Identification and confirmation of novel proteins expressed highly *in vivo* that localize to cysts.

Differentially expressed genes from the *in vivo* foci and *in vitro* culture datasets were partitioned into nine gene clusters (A-I) based on JS distance (Methods). Many of the 475 differentially expressed genes

classified using both CUFFDIFF and SLEUTH (Methods) are highly represented in the clusters from this study (B). Other host-induced or encystation-specific transcriptome studies (Ringqvist et al [113], Morf et al [34], Palm et al [50]) have fewer genes represented in the clusters than those identified by partitioning analysis in the foci (B). Many genes in clusters A, B, and C are associated with encystation or oxidative stress (C). In D, C-terminal GFP fusions of genes from cluster B and C also are associated with the cyst wall or the interior of the cyst as visualized using colocalization immunofluorescent images with the cyst wall protein 1 (CWP-1) antibody.

Figure 6. Differentially expressed genes in *in vivo* foci are similar to mid-late encystation time points in *in vitro* encystation transcriptomes.

In vitro and *in vivo* differentially expressed genes from this study are compared to *in vitro* encystation transcriptome of Einarsson et al. Panel A compares the 475 differentially expressed genes in this study (red = upregulated in foci; blue = downregulated in foci) to *in vitro* culture transcriptome (TYDK) and the transcriptomes of four time points of an *in vitro* encystation transcriptome study. Percentages indicate the number of similarly differentially expressed genes in the *in vivo* foci relative to that *in vitro* encystation time point. Panel B summarizes processes upregulated in *in vitro* culture, which is primarily trophozoite growth and division (above) as contrasted with cellular processes upregulated in the *in vivo* foci (green) that are more similar to those genes upregulated during in mid-to-late *in vitro* encystation.

Supplemental Figures

Supplemental Figure 1: Subcellular localization of C-terminal GFP fusion proteins of selected genes with increased expression in the *in vivo* foci

Supplemental Figure 2: Subcellular localization of C-terminal GFP fusion proteins of selected genes with increased expression in *in vitro* axenic culture

1 Supplemental Table Legends

2

3 Supplemental Table 1: Quantification of whole body and *ex vivo* bioluminescent imagine at day 3 and
4 day 7 p.i. for transcriptomic sampling of foci

5 Supplemental Table 2 FPKM of *in vivo* and *in vitro* transcriptomes

6 Supplemental Table 3 Overall heatmap of all FPKM

7 Supplemental Table 4 Heat map table unregulated

8 Supplemental Table 5 Heat map table - down regulated

9 Supplemental Table 6 Heat map table - *in vitro* only

10 Supplemental Table 7 Cuffdiff partitioning analysis into nine clusters

11 Supplemental Table 8 Annotated nine clusters

12 Supplemental Table 9 Comparison of *in vivo* foci and *in vitro* encystation transcriptomes

13

14

15

References

- [1] L. Savioli, H. Smith, and A. Thompson, *Giardia* and *Cryptosporidium* join the 'Neglected Diseases Initiative'. Trends Parasitol 22 (2006) 203-208.
- [2] H. Troeger, H.J. Epple, T. Schneider, U. Wahnschaffe, R. Ullrich, G.D. Burchard, T. Jelinek, M. Zeitz, M. Fromm, and J.D. Schulzke, Effect of chronic *Giardia lamblia* infection on epithelial transport and barrier function in human duodenum. Gut 56 (2007) 328-35.
- [3] H.L. DuPont, *Giardia*: both a harmless commensal and a devastating pathogen. The Journal of clinical investigation 123 (2013) 2352-4.
- [4] L.A. Bartelt, J. Roche, G. Kolling, D. Bolick, F. Noronha, C. Naylor, P. Hoffman, C. Warren, S. Singer, and R. Guerrant, Persistent *G. lamblia* impairs growth in a murine malnutrition model. The Journal of clinical investigation 123 (2013) 2672-84.
- [5] S. Solaymani-Mohammadi, and S.M. Singer, *Giardia duodenalis*: the double-edged sword of immune responses in giardiasis. Exp Parasitol 126 (2010) 292-7.
- [6] R.D. Adam, Biology of *Giardia lamblia*. Clin. Microbiol. Rev. 14 (2001) 447-475.
- [7] P. Upcroft, and J.A. Upcroft, Drug targets and mechanisms of resistance in the anaerobic protozoa. Clinical microbiology reviews 14 (2001) 150-164.
- [8] K.M. Land, and P.J. Johnson, Molecular basis of metronidazole resistance in pathogenic bacteria and protozoa. Drug Resist Updat 2 (1999) 289-294.
- [9] L.M. Barat, and P.B. Bloland, Drug resistance among malaria and other parasites. Infect Dis Clin North Am 11 (1997) 969-87.
- [10] J. Upcroft, N. Samarawickrema, D. Brown, and P. Upcroft, Mechanisms of metronidazole resistance

1 in *Giardia* and *Entamoeba*. Abstracts of the Interscience Conference on Antimicrobial Agents
2 and Chemotherapy 36 (1996) 47.

3 [11] C. Nosala, and S.C. Dawson, The Critical Role of the Cytoskeleton in the Pathogenesis of *Giardia*.
4 Curr Clin Microbiol Rep 2 (2015) 155-162.

5 [12] M.C. Halliez, and A.G. Buret, Extra-intestinal and long term consequences of *Giardia duodenalis*
6 infections. World journal of gastroenterology : WJG 19 (2013) 8974-85.

7 [13] M.F. Heyworth, Immunological aspects of *Giardia* infections. Parasite 21 (2014) 55.

8 [14] H.D. Lujan, M.R. Mowatt, L.G. Byrd, and T.E. Nash, Cholesterol starvation induces differentiation of
9 the intestinal parasite *Giardia lamblia*. Proc Natl Acad Sci U S A 93 (1996) 7628-33.

10 [15] H.D. Lujan, M.R. Mowatt, and T.E. Nash, Mechanisms of *Giardia lamblia* differentiation into cysts.
11 Microbiol Mol Biol Rev 61 (1997) 294-304.

12 [16] H.D. Lujan, M.R. Mowatt, and T.E. Nash, The molecular mechanisms of *Giardia* encystation.
13 Parasitology Today 14 (1998) 446-450.

14 [17] D.B. Keister, Axenic culture of *Giardia lamblia* in TYI-S-33 medium supplemented with bile. Trans. R.
15 Soc. Trop. Med Hyg. 77 (1983) 487-488.

16 [18] P.M. Inge, C.M. Edson, and M.J. Farthing, Attachment of *Giardia lamblia* to rat intestinal epithelial
17 cells. Gut 29 (1988) 795-801.

18 [19] L.G. Byrd, J.T. Conrad, and T.E. Nash, *Giardia lamblia* infections in adult mice. Infection and
19 immunity 62 (1994) 3583-5.

20 [20] S.M. Singer, Control of Giardiasis by Interleukin-17 in Humans and Mice--Are the Questions All
21 Answered? Clin Vaccine Immunol 23 (2016) 2-5.

22 [21] G. Mayrhofer, R.H. Andrews, P.L. Ey, M.J. Albert, T.R. Grimmond, and D.J. Merry, The use of suckling

mice to isolate and grow Giardia from mammalian faecal specimens for genetic analysis.

Parasitology 105 (Pt 2) (1992) 255-63.

[22] F.D. Rivero, A. Saura, C.G. Prucca, P.G. Carranza, A. Torri, and H.D. Lujan, Disruption of antigenic variation is crucial for effective parasite vaccine. Nature medicine 16 (2010) 551-7, 1p following 557.

[23] A. Aggarwal, and T.E. Nash, Comparison of two antigenically distinct Giardia lamblia isolates in gerbils. Am J Trop Med Hyg 36 (1987) 325-32.

[24] Y. Miyamoto, and L. Eckmann, Drug Development Against the Major Diarrhea-Causing Parasites of the Small Intestine, Cryptosporidium and Giardia. Front Microbiol 6 (2015) 1208.

[25] J. Luo, A.H. Lin, E. Masliah, and T. Wyss-Coray, Bioluminescence imaging of Smad signaling in living mice shows correlation with excitotoxic neurodegeneration. Proc Natl Acad Sci U S A 103 (2006) 18326-31.

[26] D.K. Welsh, and S.A. Kay, Bioluminescence imaging in living organisms. Current opinion in biotechnology 16 (2005) 73-8.

[27] R. Weissleder, and V. Ntziachristos, Shedding light onto live molecular targets. Nature medicine 9 (2003) 123-8.

[28] A.C. Stacer, S. Nyati, P. Moudgil, R. Iyengar, K.E. Luker, A. Rehemtulla, and G.D. Luker, NanoLuc reporter for dual luciferase imaging in living animals. Mol Imaging 12 (2013) 1-13.

[29] J.Q. Reimao, C.T. Trinconi, J.K. Yokoyama-Yasunaka, D.C. Miguel, S.P. Kalil, and S.R. Uliana, Parasite burden in Leishmania (Leishmania) amazonensis-infected mice: validation of luciferase as a quantitative tool. Journal of microbiological methods 93 (2013) 95-101.

[30] J.P. Saeij, J.P. Boyle, M.E. Grigg, G. Arrizabalaga, and J.C. Boothroyd, Bioluminescence imaging of

- 1 Toxoplasma gondii infection in living mice reveals dramatic differences between strains.
- 2 Infection and immunity 73 (2005) 695-702.
- 3 [31] S. D'Archivio, A. Cosson, M. Medina, T. Lang, P. Minoprio, and S. Goyard, Non-invasive in vivo study
- 4 of the Trypanosoma vivax infectious process consolidates the brain commitment in late
- 5 infections. PLoS Negl Trop Dis 7 (2013) e1976.
- 6 [32] M. Hutchens, and G.D. Luker, Applications of bioluminescence imaging to the study of infectious
- 7 diseases. Cellular microbiology 9 (2007) 2315-22.
- 8 [33] F.D. Gillin, D.S. Reiner, and S.E. Boucher, Small-intestinal factors promote encystation of Giardia
- 9 lamblia in vitro. Infection and immunity 56 (1988) 705-7.
- 10 [34] L. Morf, C. Spycher, H. Rehrauer, C.A. Fournier, H.G. Morrison, and A.B. Hehl, The transcriptional
- 11 response to encystation stimuli in Giardia lamblia is restricted to a small set of genes. Eukaryot
- 12 Cell 9 (2010) 1566-76.
- 13 [35] N. Andreu, A. Zelmer, and S. Wiles, Noninvasive biophotonic imaging for studies of infectious
- 14 disease. FEMS microbiology reviews 35 (2011) 360-94.
- 15 [36] S. Wiles, K.M. Pickard, K. Peng, T.T. MacDonald, and G. Frankel, In vivo bioluminescence imaging of
- 16 the murine pathogen Citrobacter rodentium. Infection and immunity 74 (2006) 5391-6.
- 17 [37] K.J. Rhee, H. Cheng, A. Harris, C. Morin, J.B. Kaper, and G. Hecht, Determination of spatial and
- 18 temporal colonization of enteropathogenic E. coli and enterohemorrhagic E. coli in mice using
- 19 bioluminescent in vivo imaging. Gut microbes 2 (2011) 34-41.
- 20 [38] S. Gourguechon, and W.Z. Cande, Rapid tagging and integration of genes in Giardia intestinalis.
- 21 Eukaryotic Cell 10 (2011) 142-5.
- 22 [39] E. Einarsson, K. Troell, M.P. Hoepfner, M. Grabherr, U. Ribacke, and S.G. Svard, Coordinated

1 Changes in Gene Expression Throughout Encystation of *Giardia intestinalis*. PLoS Negl Trop Dis
2 10 (2016) e0004571.

3 [40] N. Barash, Nosala, C, Pham, JK, McInally, SG, Gourguechon, S., McCarthy-Sinclair, B, and SC
4 Dawson, *Giardia* colonizes and encysts in high density foci in the murine small intestine.
5 mSphere in press (2017).

6 [41] S. Solaymani-Mohammadi, and S.M. Singer, Host immunity and pathogen strain contribute to
7 intestinal disaccharidase impairment following gut infection. J Immunol 187 (2011) 3769-75.

8 [42] P.M. Chen, H. Gregersen, and J.B. Zhao, Advanced glycation end-product expression is upregulated
9 in the gastrointestinal tract of type 2 diabetic rats. World journal of diabetes 6 (2015) 662-72.

10 [43] F.J. Joshi NA, Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version
11 1.33) [Software]. , 2011.

12 [44] C. Trapnell, L. Pachter, and S.L. Salzberg, TopHat: discovering splice junctions with RNA-Seq.
13 Bioinformatics 25 (2009) 1105-11.

14 [45] C. Trapnell, B.A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M.J. van Baren, S.L. Salzberg, B.J. Wold,
15 and L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated
16 transcripts and isoform switching during cell differentiation. Nature biotechnology 28 (2010)
17 511-5.

18 [46] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn,
19 and L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with
20 TopHat and Cufflinks. Nature protocols 7 (2012) 562-78.

21 [47] N.L. Bray, H. Pimentel, P. Melsted, and L. Pachter, Near-optimal probabilistic RNA-seq
22 quantification. Nature biotechnology 34 (2016) 525-7.

- [48] C. Aurrecochea, J. Brestelli, B.P. Brunk, J.M. Carlton, J. Dommer, S. Fischer, B. Gajria, X. Gao, A. Gingle, G. Grant, O.S. Harb, M. Heiges, F. Innamorato, J. Iodice, J.C. Kissinger, E. Kraemer, W. Li, J.A. Miller, H.G. Morrison, V. Nayak, C. Pennington, D.F. Pinney, D.S. Roos, C. Ross, C.J. Stoeckert, Jr., S. Sullivan, C. Treatman, and H. Wang, GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens Giardia lamblia and Trichomonas vaginalis. Nucleic Acids Res 37 (2009) D526-30.
- [49] E. Ringqvist, L. Avesson, F. Söderbom, and S.G. Svärd, Transcriptional changes in Giardia during host-parasite interactions. International journal for parasitology 41 (2011) 277-85.
- [50] D. Palm, M. Weiland, A.G. McArthur, J. Winiecka-Krusnell, M.J. Cipriano, S.R. Birkeland, S.E. Pacocha, B. Davids, F. Gillin, E. Linder, and S. Svard, Developmental changes in the adhesive disk during Giardia differentiation. Mol Biochem Parasitol 141 (2005) 199-207.
- [51] S.C. Dawson, and S.A. House, Imaging and analysis of the microtubule cytoskeleton in Giardia. Methods Cell Biol 97 (2010) 307-39.
- [52] M.L. Carpenter, Z.J. Assaf, S. Gourguechon, and W.Z. Cande, Nuclear inheritance and genetic exchange without meiosis in the binucleate parasite Giardia intestinalis. Journal of cell science 125 (2012) 2523-32.
- [53] T.E. Nash, Surface antigenic variation in Giardia lamblia. Molecular microbiology 45 (2002) 585-90.
- [54] P.T. Macechko, P.A. Steimle, D.G. Lindmark, S.L. Erlandsen, and E.L. Jarroll, Galactosamine-synthesizing enzymes are induced when Giardia encyst. Mol Biochem Parasitol 56 (1992) 301-9.
- [55] E.L. Jarroll, P.T. Macechko, P.A. Steimle, D. Bulik, C.D. Karr, H. van Keulen, T.A. Paget, G. Gerwig, J. Kamerling, J. Vliegenthart, and S. Erlandsen, Regulation of carbohydrate metabolism during Giardia encystment. J Eukaryot Microbiol 48 (2001) 22-6.

- 1 [56] D.G. Lindmark, Energy metabolism of the anaerobic protozoon *Giardia lamblia*. *Mol Biochem*
2 *Parasitol* 1 (1980) 1-12.
- 3 [57] C. Faso, S. Bischof, and A.B. Hehl, The proteome landscape of *Giardia lamblia* encystation. *PloS one*
4 8 (2013) e83207.
- 5 [58] A. Sulemana, T.A. Paget, and E.L. Jarroll, Commitment to cyst formation in *Giardia*. *Microbiology*
6 160 (2014) 330-9.
- 7 [59] F.D. Gillin, S.E. Boucher, S.S. Rossi, and D.S. Reiner, *Giardia lamblia*: the roles of bile, lactic acid, and
8 pH in the completion of the life cycle in vitro. *Exp Parasitol* 69 (1989) 164-74.
- 9 [60] T. Lauwaet, B.J. Davids, D.S. Reiner, and F.D. Gillin, Encystation of *Giardia lamblia*: a model for other
10 parasites. *Current opinion in microbiology* 10 (2007) 554-9.
- 11 [61] C. Faso, and A.B. Hehl, Membrane trafficking and organelle biogenesis in *Giardia lamblia*: use it or
12 lose it. *Int J Parasitol* 41 (2011) 471-80.
- 13 [62] F.D. Gillin, D.S. Reiner, M.J. Gault, H. Douglas, S. Das, A. Wunderlich, and J.F. Sauch, Encystation and
14 expression of cyst antigens by *Giardia lamblia* in vitro. *Science* 235 (1987) 1040-3.
- 15 [63] M.R. Edwards, P.J. Schofield, W.J. O'Sullivan, and M. Costello, Arginine metabolism during culture
16 of *Giardia intestinalis*. *Mol Biochem Parasitol* 53 (1992) 97-103.
- 17 [64] B. Stadelmann, M.C. Merino, L. Persson, and S.G. Svard, Arginine consumption by the intestinal
18 parasite *Giardia intestinalis* reduces proliferation of intestinal epithelial cells. *PloS one* 7 (2012)
19 e45325.
- 20 [65] B.J. Davids, K. Mehta, L. Fesus, J.M. McCaffery, and F.D. Gillin, Dependence of *Giardia lamblia*
21 encystation on novel transglutaminase activity. *Mol Biochem Parasitol* 136 (2004) 173-80.
- 22 [66] B.J. Davids, D.S. Reiner, S.R. Birkeland, S.P. Preheim, M.J. Cipriano, A.G. McArthur, and F.D. Gillin, A

new family of giardial cysteine-rich non-VSP protein genes and a novel cyst protein. PLoS One 1
(2006) e44.

[67] C.H. Sun, D. Palm, A.G. McArthur, S.G. Svard, and F.D. Gillin, A novel Myb-related protein involved
in transcriptional activation of encystation genes in *Giardia lamblia*. Molecular microbiology 46
(2002) 971-84.

[68] N. Chen, J.A. Upcroft, and P. Upcroft, A new cysteine-rich protein-encoding gene family in *Giardia*
duodenalis. Gene (Amsterdam) 169 (1996) 33-38.

[69] J.A. Ebner, S.D. Heusser, E.M. Schraner, A.B. Hehl, and C. Faso, Cyst-Wall-Protein-1 is fundamental
for Golgi-like organelle neogenesis and cyst-wall biosynthesis in *Giardia lamblia*. Nat Commun 7
(2016) 13859.

[70] C.G. Prucca, and H.D. Lujan, Antigenic variation in *Giardia lamblia*. Cellular microbiology 11 (2009)
1706-15.

[71] W. Li, A.A. Saraiya, and C.C. Wang, Experimental verification of the identity of variant-specific
surface proteins in *Giardia lamblia* trophozoites. MBio 4 (2013) e00321-13.

[72] C.G. Prucca, I. Slavin, R. Quiroga, E.V. Elias, F.D. Rivero, A. Saura, P.G. Carranza, and H.D. Lujan,
Antigenic variation in *Giardia lamblia* is regulated by RNA interference. Nature 456 (2008) 750-4.

[73] S. Das, T. Stevens, C. Castillo, A. Villasenor, H. Arredondo, and K. Reddy, Lipid metabolism in
mucous-dwelling amitochondriate protozoa. Int J Parasitol 32 (2002) 655-75.

[74] P. Katelaris, F. Seow, and M. Ngu, The effect of *Giardia lamblia* trophozoites on lipolysis in vitro.
Parasitology 103 Pt 1 (1991) 35-9.

[75] D. Bansal, H.S. Bhatti, and R. Sehgal, Altered lipid parameters in patients infected with *Entamoeba*
histolytica, *Entamoeba dispar* and *Giardia lamblia*. British journal of biomedical science 62

1 (2005) 63-5.

2 [76] T.L. Mendez, A. De Chatterjee, T. Duarte, J. De Leon, L. Robles-Martinez, and S. Das, Sphingolipids,
3 Lipid Rafts, and Giardial Encystation: The Show Must Go On. *Curr Trop Med Rep* 2 (2015) 136-
4 143.

5 [77] M. Yichoy, T.T. Duarte, A. De Chatterjee, T.L. Mendez, K.Y. Aguilera, D. Roy, S. Roychowdhury, S.B.
6 Aley, and S. Das, Lipid metabolism in Giardia: a post-genomic perspective. *Parasitology* 138
7 (2011) 267-78.

8 [78] S. Das, C.D. Schteingart, A.F. Hofmann, D.S. Reiner, S.B. Aley, and F.D. Gillin, Giardia lamblia:
9 evidence for carrier-mediated uptake and release of conjugated bile acids. *Exp Parasitol* 87
10 (1997) 133-41.

11 [79] F.D. Gillin, M.J. Gault, A.F. Hofmann, D. Gurantz, and J.F. Sauch, Biliary lipids support serum-free
12 growth of Giardia lamblia. *Infection and immunity* 53 (1986) 641-5.

13 [80] S. Sonda, S. Stefanic, and A.B. Hehl, A Sphingolipid Inhibitor Induces a Cytokinesis Arrest and Blocks
14 Stage Differentiation in Giardia lamblia. *Antimicrobial agents and chemotherapy* 52 (2008) 563-
15 9.

16 [81] K.N. DuBois, M. Abodeely, J. Sakanari, C.S. Craik, M. Lee, J.H. McKerrow, and M. Sajid, Identification
17 of the major cysteine protease of Giardia and its role in encystation. *J Biol Chem* 283 (2008)
18 18024-31.

19 [82] M.C. Touz, M.J. Nores, I. Slavin, C. Carmona, J.T. Conrad, M.R. Mowatt, T.E. Nash, C.E. Coronel, and
20 H.D. Lujan, The activity of a developmentally regulated cysteine proteinase is required for cyst
21 wall formation in the primitive eukaryote Giardia lamblia. *J Biol Chem* 277 (2002) 8474-81.

22 [83] S.Y. Ma'ayeh, and P.T. Brook-Carter, Representational difference analysis identifies specific genes in

the interaction of *Giardia duodenalis* with the murine intestinal epithelial cell line, IEC-6. *Int J Parasitol* 42 (2012) 501-9.

[84] J.A. Cotton, A. Bhargava, J.G. Ferraz, R.M. Yates, P.L. Beck, and A.G. Buret, *Giardia duodenalis* cathepsin B proteases degrade intestinal epithelial interleukin-8 and attenuate interleukin-8-induced neutrophil chemotaxis. *Infection and immunity* 82 (2014) 2772-87.

[85] T.A. Paget, P. Manning, and E.L. Jarroll, Oxygen uptake in cysts and trophozoites of *Giardia lamblia*. *J Eukaryot Microbiol* 40 (1993) 246-50.

[86] T.A. Paget, M.L. Kelly, E.L. Jarroll, D.G. Lindmark, and D. Lloyd, The effects of oxygen on fermentation in *Giardia lamblia*. *Mol Biochem Parasitol* 57 (1993) 65-71.

[87] T.A. Paget, M.H. Raynor, D.W. Shipp, and D. Lloyd, *Giardia lamblia* produces alanine anaerobically but not in the presence of oxygen. *Mol Biochem Parasitol* 42 (1990) 63-7.

[88] P.J. Schofield, M. Costello, M.R. Edwards, and W.J. O'Sullivan, The arginine dihydrolase pathway is present in *Giardia intestinalis*. *Int J Parasitol* 20 (1990) 697-9.

[89] P.J. Schofield, M.R. Edwards, and P. Kranz, Glucose metabolism in *Giardia intestinalis*. *Mol Biochem Parasitol* 45 (1991) 39-47.

[90] P.J. Schofield, M.R. Edwards, J. Matthews, and J.R. Wilson, The pathway of arginine catabolism in *Giardia intestinalis*. *Mol Biochem Parasitol* 51 (1992) 29-36.

[91] J. Yee, M.R. Mowatt, P.P. Dennis, and T.E. Nash, Transcriptional analysis of the glutamate dehydrogenase gene in the primitive eukaryote, *Giardia lamblia*. Identification of a primordial gene promoter. *J Biol Chem* 275 (2000) 11432-9.

[92] L.A. Bartelt, and R.B. Sartor, Advances in understanding *Giardia*: determinants and mechanisms of chronic sequelae. *F1000Prime Rep* 7 (2015) 62.

- 1 [93] L.L. Ventura, D.R. Oliveira, J.C. Viana, J.F. Santos, M.V. Caliri, and M.A. Gomes, Impact of protein
2 malnutrition on histological parameters of experimentally infected animals with *Giardia lamblia*.
3 *Exp Parasitol* 133 (2013) 391-5.
- 4 [94] E.L. Jarroll, P.J. Muller, E.A. Meyer, and S.A. Morse, Lipid and carbohydrate metabolism of *Giardia*
5 *lamblia*. *Mol Biochem Parasitol* 2 (1981) 187-96.
- 6 [95] S.M. Aldritt, P. Tien, and C.C. Wang, Pyrimidine salvage in *Giardia lamblia*. *J Exp Med* 161 (1985)
7 437-45.
- 8 [96] C.C. Wang, and S. Aldritt, Purine salvage networks in *Giardia lamblia*. *J Exp Med* 158 (1983) 1703-
9 12.
- 10 [97] N.A. Stover, T.A. Dixon, and A.R. Cavalcanti, Multiple independent fusions of glucose-6-phosphate
11 dehydrogenase with enzymes in the pentose phosphate pathway. *PLoS One* 6 (2011) e22269.
- 12 [98] Z. Faghiri, and G. Widmer, A comparison of the *Giardia lamblia* trophozoite and cyst transcriptome
13 using microarrays. *BMC Microbiol* 11 (2011) 91.
- 14 [99] D.M. Brown, J.A. Upcroft, and P. Upcroft, A H₂O-producing NADH oxidase from the protozoan
15 parasite *Giardia duodenalis*. *European journal of biochemistry / FEBS* 241 (1996) 155-61.
- 16 [100] A. Di Matteo, F.M. Scandurra, F. Testa, E. Forte, P. Sarti, M. Brunori, and A. Giuffrè, The O₂-
17 scavenging flavodiiron protein in the human parasite *Giardia intestinalis*. *J Biol Chem* 283 (2008)
18 4061-8.
- 19 [101] P.D. Fernandes, and J. Assreuy, Role of nitric oxide and superoxide in *Giardia lamblia* killing.
20 *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e*
21 *biologicas / Sociedade Brasileira de Biofisica ... [et al.]* 30 (1997) 93-9.
- 22 [102] L. Eckmann, F. Laurent, T.D. Langford, M.L. Hetsko, J.R. Smith, M.F. Kagnoff, and F.D. Gillin, Nitric

oxide production by human intestinal epithelial cells and competition for arginine as potential determinants of host defense against the lumen-dwelling pathogen *Giardia lamblia*. *J Immunol* 164 (2000) 1478-87.

[103] S. Rafferty, B. Luu, R.E. March, and J. Yee, *Giardia lamblia* encodes a functional flavohemoglobin. *Biochem Biophys Res Commun* 399 (2010) 347-51.

[104] A. Ilari, and A. Boffi, Structural studies on flavohemoglobins. *Methods Enzymol* 436 (2008) 187-202.

[105] P.R. Gardner, Nitric oxide dioxygenase function and mechanism of flavohemoglobin, hemoglobin, myoglobin and their associated reductases. *Journal of inorganic biochemistry* 99 (2005) 247-66.

[106] P.R. Gardner, A.M. Gardner, L.A. Martin, and A.L. Salzman, Nitric oxide dioxygenase: an enzymic function for flavohemoglobin. *Proc Natl Acad Sci U S A* 95 (1998) 10378-83.

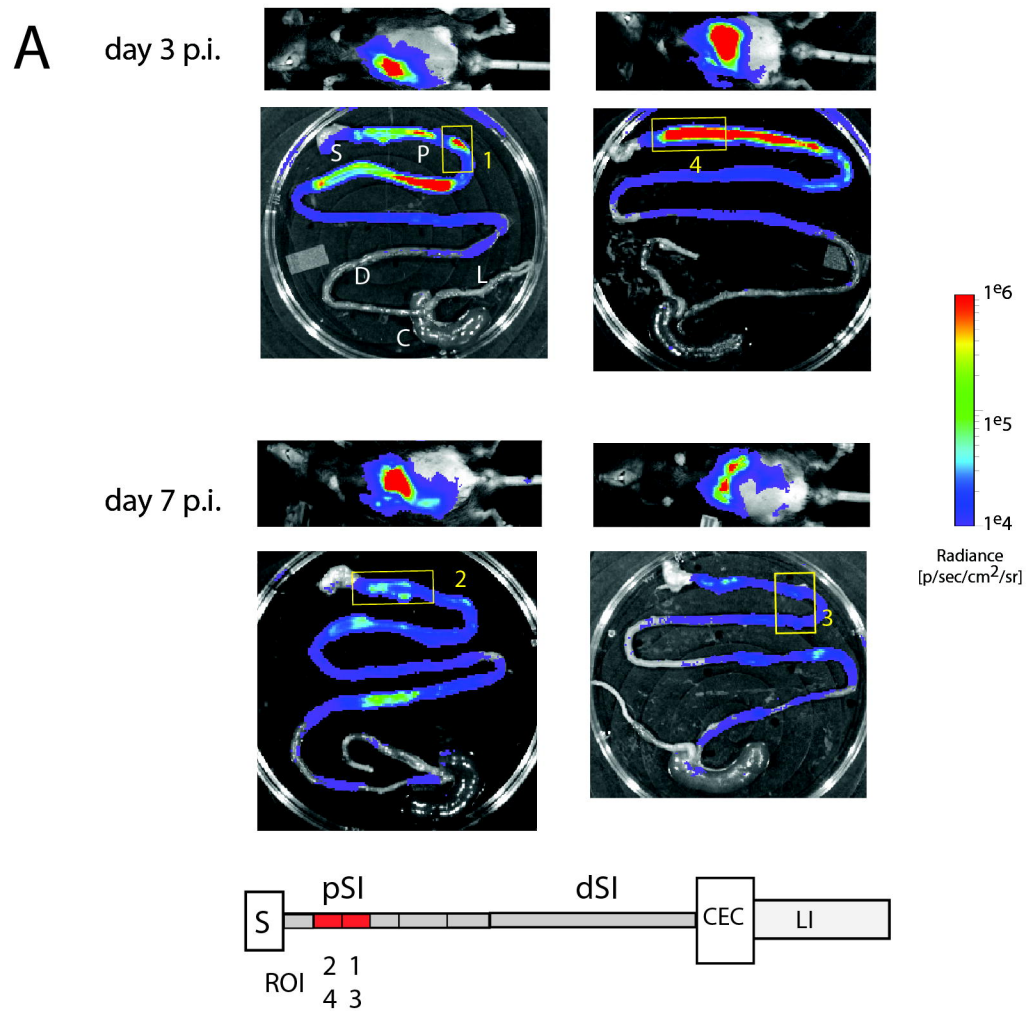
[107] A. Hausladen, A.J. Gow, and J.S. Stamler, Nitrosative stress: metabolic pathway involving the flavohemoglobin. *Proc Natl Acad Sci U S A* 95 (1998) 14100-5.

[108] V. Roos, and P. Klemm, Global gene expression profiling of the asymptomatic bacteriuria *Escherichia coli* strain 83972 in the human urinary tract. *Infection and immunity* 74 (2006) 3565-75.

[109] F. Sebbane, N. Lemaitre, D.E. Sturdevant, R. Rebeil, K. Virtaneva, S.F. Porcella, and B.J. Hinnebusch, Adaptive response of *Yersinia pestis* to extracellular effectors of innate immunity during bubonic plague. *Proc Natl Acad Sci U S A* 103 (2006) 11766-71.

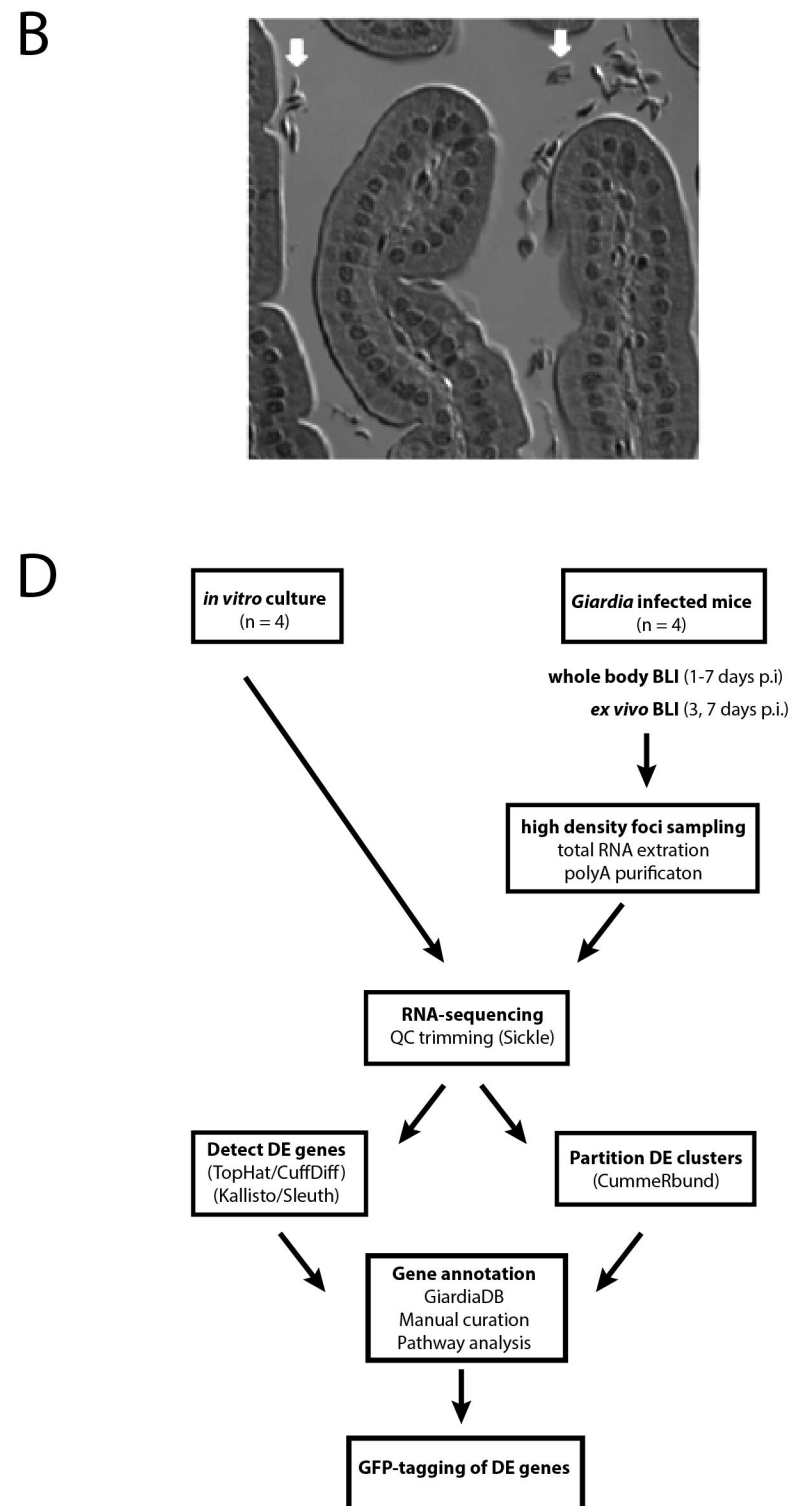
[110] D. Mastronicola, F. Testa, E. Forte, E. Bordi, L.P. Pucillo, P. Sarti, and A. Giuffre, Flavohemoglobin and nitric oxide detoxification in the human protozoan parasite *Giardia intestinalis*. *Biochem Biophys Res Commun* 399 (2010) 654-8.

- 1 [111] S.Y. Ma'ayeh, L. Knorr, and S.G. Svard, Transcriptional profiling of Giardia intestinalis in response to
2 oxidative stress. Int J Parasitol 45 (2015) 925-38.
- 3 [112] N. Barash, Maloney, JG, Singer, SM, and SC Dawson, Giardia alters commensal microbial diversity
4 throughout the murine gut. Infection and immunity (2017) in press.
- 5 [113] K. Roxstrom-Lindquist, E. Ringqvist, D. Palm, and S. Svard, Giardia lamblia-induced changes in
6 gene expression in differentiated Caco-2 human intestinal epithelial cells. Infection and
7 immunity 73 (2005) 8204-8.
- 8



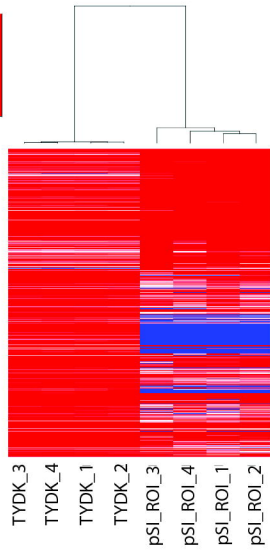
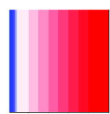
C

sample	day p.i.	total reads	Giardia reads	Total Flux
in_vitro_Giardia_TYDK_1	n.a.	18018527	11687074	n.a.
in_vitro_Giardia_TYDK_2	n.a.	16762838	16543819	n.a.
in_vitro_Giardia_TYDK_3	n.a.	21282183	20798237	n.a.
in_vitro_Giardia_TYDK_4	n.a.	16792678	16477881	n.a.
in_vivo_pSI_ROI_1	3	10880724	125414	2.77E+06
in_vivo_pSI_ROI_4	3	10017244	209854	1.50E+07
in_vivo_pSI_ROI_2	7	11669725	188434	2.31E+06
in_vivo_pSI_ROI_3	7	11226893	86076	5.12E+05



A

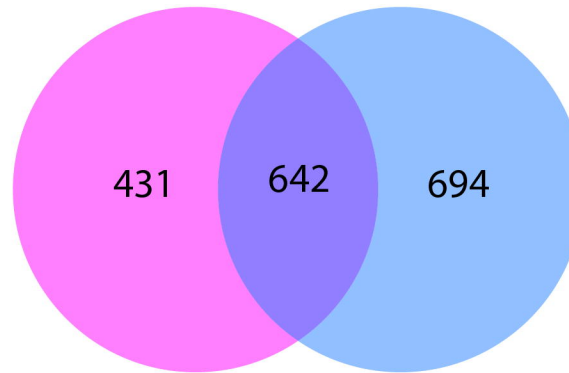
normalized FPKM

*in vitro*
(TYDK)*in vivo*
(foci)

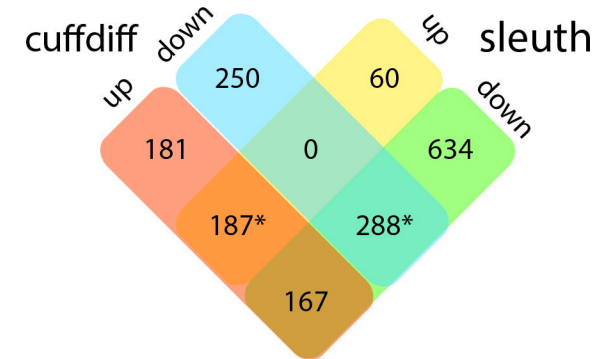
B

cuffdiff DE

sleuth DE



C



D

Top 25 differentially expressed genes (TYDK)

GiardiaDB	FPKM	annotation	cellular process
GL50803_10311	11794	ornithine carbamoyltransferase	arginine metabolism
GL50803_7110	11345.3	ubiquitin	proteasome
GL50803_3570	11039.2	ribosomal protein S28	translation
GL50803_9810	8077.48	ribosomal protein L44	translation
GL50803_14457	7153.68	ribosomal L29e protein	translation
GL50803_14620	7042.32	ribosomal protein S6	translation
GL50803_16114	6117.49	ribosomal protein L36-1	translation
GL50803_10367	5779.27	ribosomal protein S24	translation
GL50803_14321	5626.14	ribosomal protein L30	translation
GL50803_4239	5308.94	hypothetical protein	unknown
GL50803_14091	5105.58	ribosomal protein L14	translation
GL50803_8462	4845.23	ribosomal protein L27	translation
GL50803_17054	4808.81	acidic ribosomal protein P0	translation
GL50803_10919	4705.52	ribosomal protein S10B	translation
GL50803_13561	4000.37	ribosomal protein L9	translation
GL50803_6430	3703.33	14-3-3 protein	intracellular signalling
GL50803_10780	3603.7	ribosomal protein S27	translation
GL50803_33862	3575.4	ribosomal protein S12	translation
GL50803_5947	3530.06	ribosomal protein L35a	translation
GL50803_7870	3488.21	ribosomal protein L23A	translation
GL50803_10428	3212.27	ribosomal protein L10	translation
GL50803_112103	3205.57	arginine deiminase	arginine metabolism
GL50803_16453	3114.43	carbamate kinase	arginine metabolism
GL50803_14938	3110.26	ribosomal protein L12	translation
GL50803_14699	3019.58	ribosomal protein S23	translation

E

Top 25 differentially expressed genes (*in vivo foci*)

GiardiaDB	FPKM	annotation	cellular process
GL50803_25013	39321.1	high cysteine protein	unknown
GL50803_86908	13593.7	hypothetical protein	unknown
GL50803_6430	11496	14-3-3 protein	intracellular signalling
GL50803_5435	11050.2	cyst wall protein 2	encystation
GL50803_5810	10666.9	FMN binding domain containing protein	oxidative stress
GL50803_14670	9544.56	protein disulfide isomerase PDI3	oxidative stress
GL50803_7719	6786.52	hypothetical protein	unknown
GL50803_15427	6016.92	GTA-2 antigen	unknown
GL50803_7110	5756.46	ubiquitin	proteasome
GL50803_21942	5563.28	NADP-specific glutamate dehydrogenase	pyruvate metabolism
GL50803_13864	5221.63	HSP90	oxidative stress
GL50803_13747	5187.45	C4 protein	unknown
GL50803_114787	4868.7	alpha-7.3 giardin	unknown
GL50803_15832	4743.68	aminoacyl-histidine dipeptidase	amino acid metabolism
GL50803_6289	4696.71	thioredoxin domain containing protein	oxidative stress
GL50803_10217	4685.11	cathepsin B precursor (CP1)	protease
GL50803_98054	4382.04	heat shock protein HSP 90-alpha	oxidative stress
GL50803_14259	4314.15	glucose 6-phosphate N-acetyltransferase	encystation
GL50803_17516	3912.31	cathepsin B precursor	protease
GL50803_10882	3836.3	conserved hypothetical protein	unknown
GL50803_17150	3724	NADPH oxidoreductase	oxidative stress
GL50803_10367	3423.34	ribosomal protein S24	translation
GL50803_10311	3169.7	ornithine carbamoyltransferase	arginine metabolism
GL50803_16114	3003.37	ribosomal protein L36-1	translation
GL50803_16779	2998.53	cathepsin B precursor (CP3)	protease

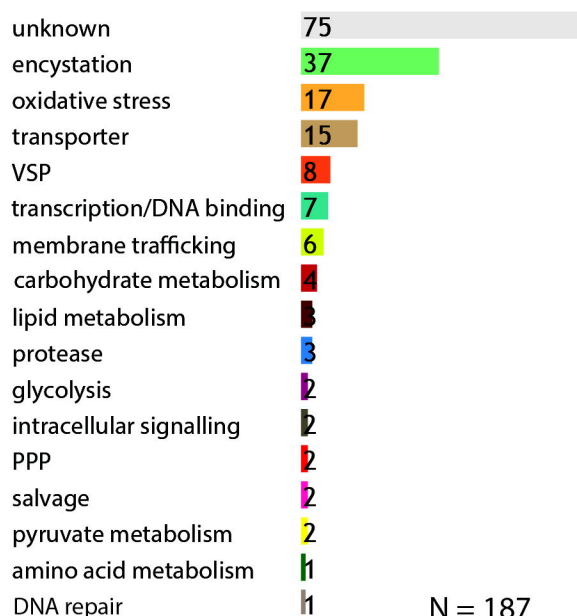
A

Greater than 10 fold differentially expressed in *in vivo* foci relative to *in vitro* culture



B

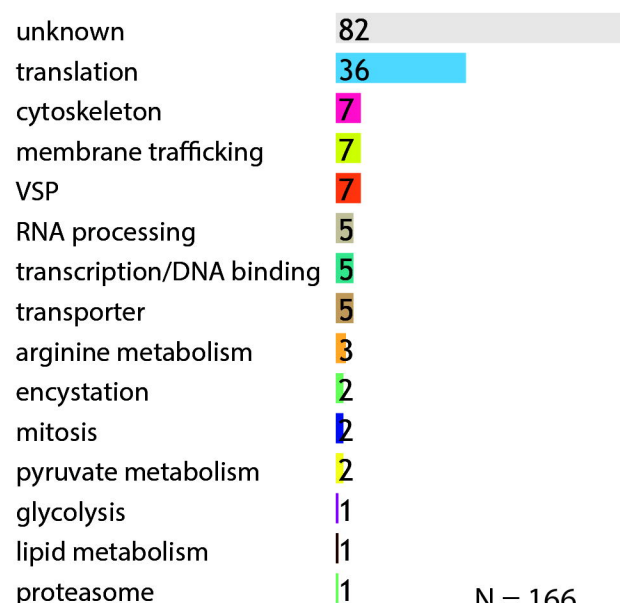
2-fold and greater *in vivo* (foci) cellular processes



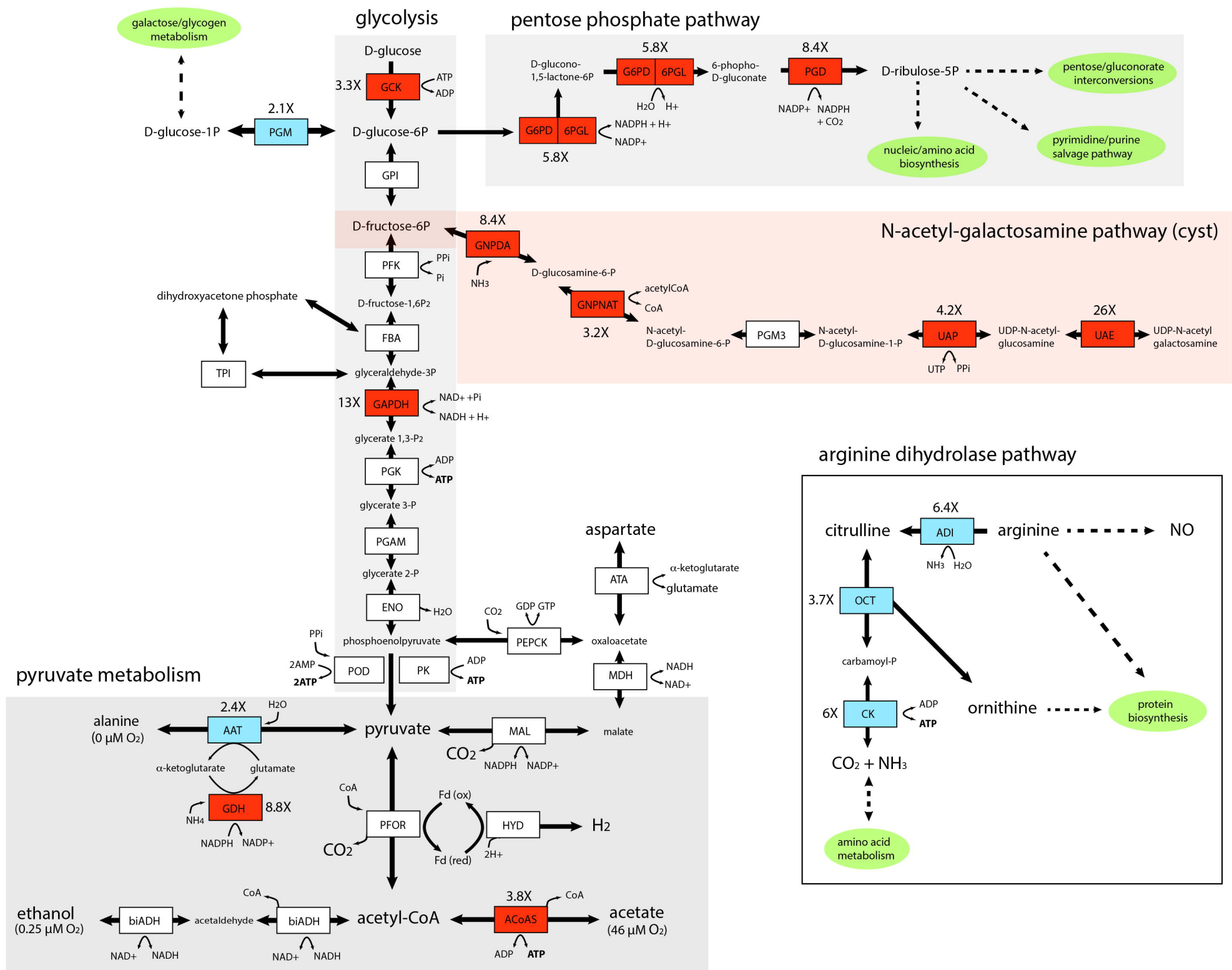
N = 187

C

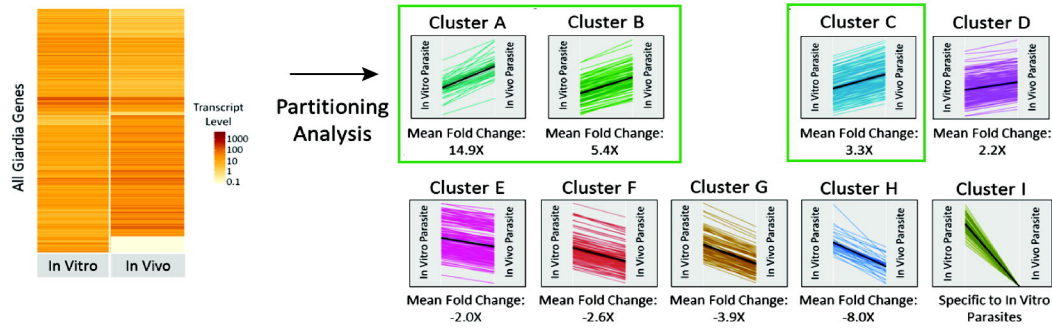
2-fold and greater *in vitro* (TYDK) cellular processes



N = 166



A



B

mean fold change (in vivo: in vitro)	cluster	total	475 DE	% in cluster	not-induced Ringqvist et al	encystation Morf et al	cyst Palm et al	
14.91	A	50	33	66.00%	3	1	7	up in foci
5.37	B	104	76	73.08%	12	4	18	
3.28	C	164	65	39.63%	6	1	25	
2.22	D	218	13	5.96%	4	0	15	down in foci
-1.99	E	176	113	64.20%	0	0	10	
-2.63	F	151	65	43.05%	0	0	16	
-3.94	G	101	31	30.69%	0	0	12	
-7.96	H	39	18	46.15%	0	0	8	
in vitro specific	I	70	61	87.14%	0	0	6	

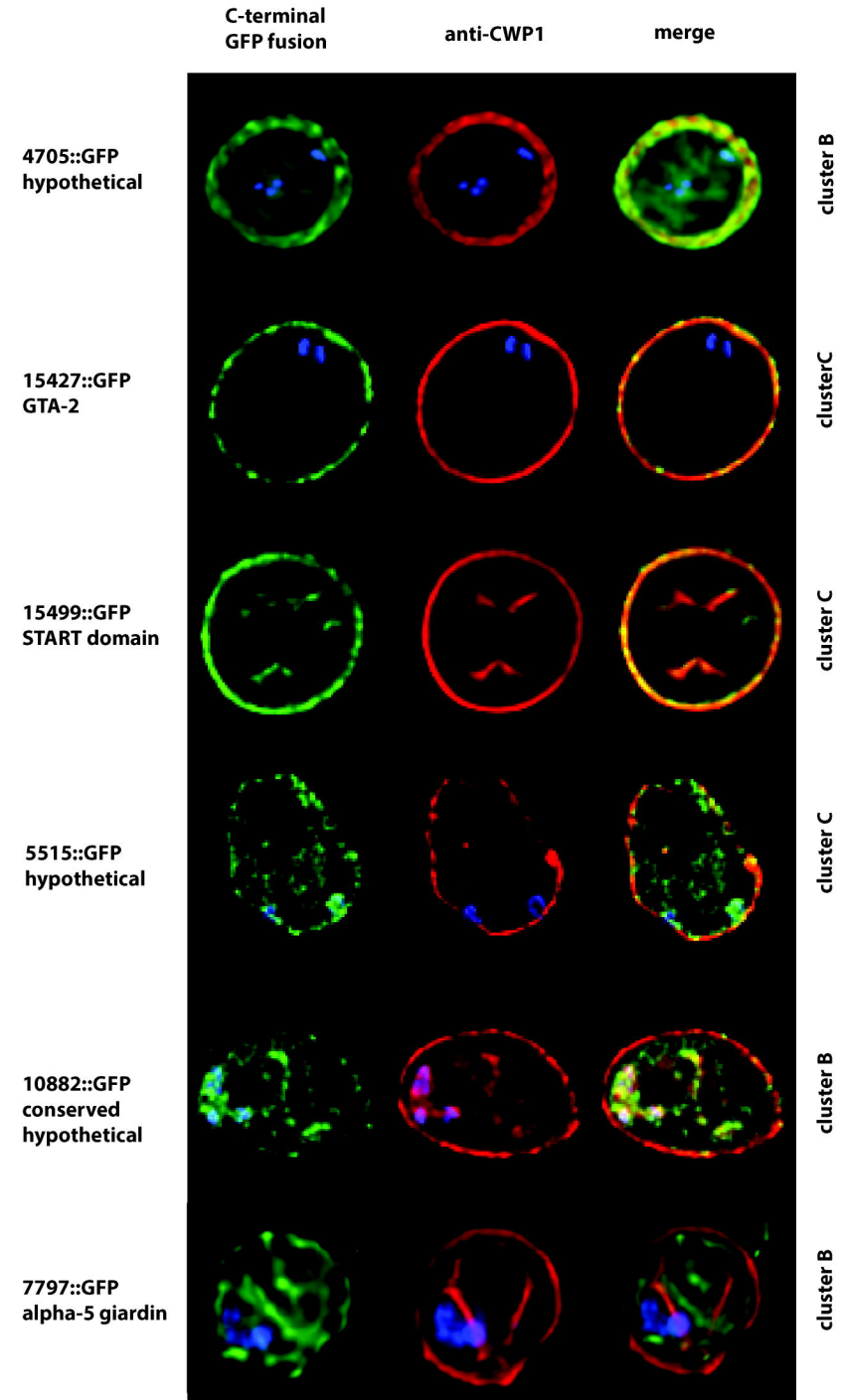
C

cluster	process	%
A	encystation	60.00
A	unknown	22.00
A	VSP	10.00
A	transporter	4.00
A	oxidative stress	2.00
A	transcription/DNA binding	2.00

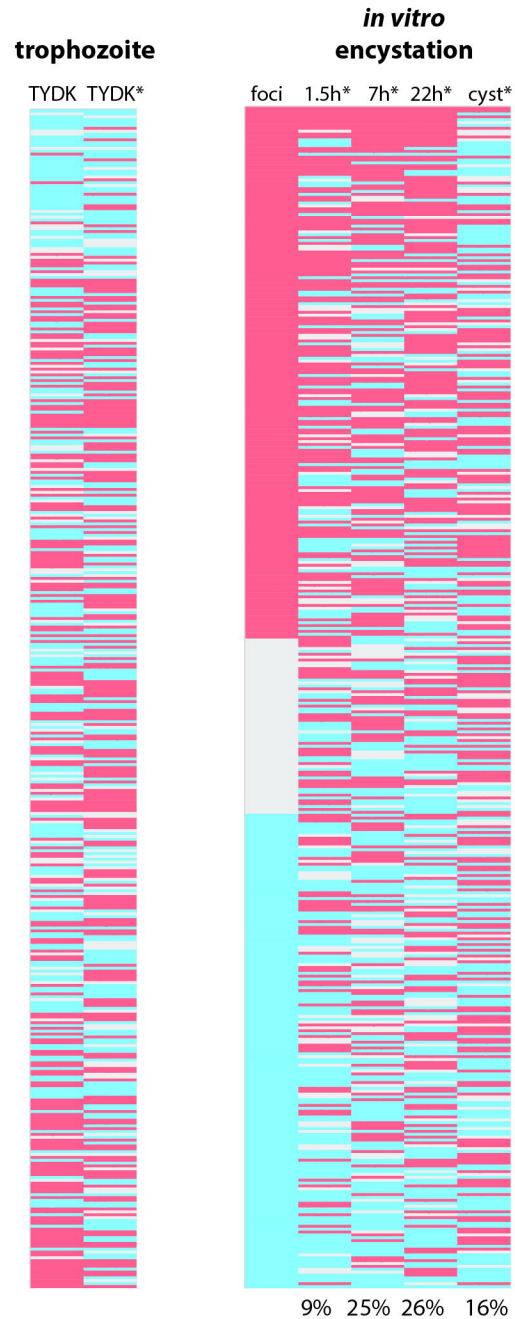
cluster	process	%
B	unknown	35.58
B	not in 475DE	23.08
B	encystation	8.65
B	protease	3.85
B	transcription/DNA binding	3.85
B	transporter	3.85
B	VSP	3.85
B	carbohydrate metabolism	2.88
B	membrane trafficking	2.88
B	oxidative stress	2.88
B	encystation	1.92
B	PPP	1.92
B	amino acid metabolism	0.96
B	intracellular signalling	0.96
B	lipid metabolism	0.96
B	purine/pyrimidine salvage	0.96
B	pyruvate metabolism	0.96

cluster	process	%
C	not in 475DE	59.76
C	unknown	18.90
C	transporter	6.10
C	oxidative stress	5.49
C	encystation	3.66
C	lipid metabolism	1.22
C	membrane trafficking	1.22
C	transcription/DNA binding	1.22
C	carbohydrate metabolism	0.61
C	DNA repair	0.61
C	pyruvate metabolism	0.61
C	VSP	0.61

D

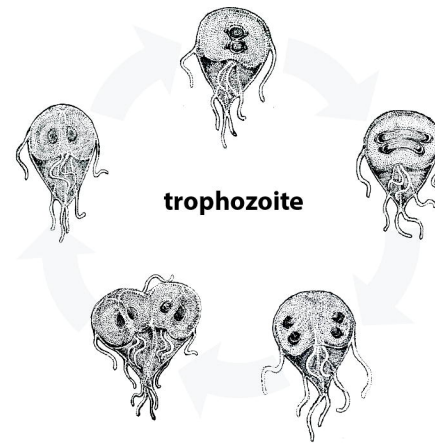


A



*Einarsson et al

B

in vitro culture (TYDK)

translation
transcription
cytoskeleton
membrane trafficking
arginine diHP
pyruvate metabolism
mitosis

