**Dual host-parasite transcriptomes of apicomplexan *Eimeria falciformis* and its natural mouse host**

Totta Kasemo, Simone Spork, Christoph Dieterich, Richard Lucius, Emanuel Heitlinger

1. Institute of Biology, Humboldt-Universitat zu Berlin

Philippstr. 13, Haus 14, 10115 Berlin, Germany

2. University Hospital Heidelberg - German Center for Cardiovascular Research (DZHK), Analysezentrum III, Im Neuenheimer Feld 669, 69120 Heidelberg, Germany

3. Leibniz Institute of Zoo and Wildlife Research, Alfred-Kowalke-Str. 17, 10315, Berlin, Germany

**ABSTRACT**

Apicomplexan parasites such as *Plasmodium* spp., Toxoplasma gondii and *Eimeria* spp. cause disease in humans, livestock and wild animals. The genus *Eimeria* comprises >1800 monoxenous intracellular parasites believed to be host and niche specific. Most prominent among *Eimeria* are several species which cause losses in poultry industries. *Eimeria* falciformis naturally infects the cecum of mice and thus gives easy access to all lifecycle stages. Completing both asexual and sexual replication in one of the best studied available animal models, this parasite can be used as a model to investigate coccidian infections. However, much is still unknown about the parasite’s basic biology and no in vitro culture has been established for the full lifecycle. We have performed a dual RNA-seq transcriptome study of the full lifecycle in the mouse and of in vitro cultured sporozoites and oocysts. Drastic differences are seen in both parasite and host mRNA abundance at three time-points post infection. Comparisons between immunocompetent and immunocompromised mice show differences in oocyst output as well as transcriptional differences in the mouse. Broad functional gene categories indicating immune reaction and tissue repair are enriched for differently abundant host genes. More specifically TGF-beta, EGF, TNF and IL-1 and IL-6 are examples of genes reacting differently depending on mouse immune status. Much in contrast, parasite transcriptomes are neither different between immune competent and immune deficient mice, nor between naïve and challenge infected mice. Instead, parasite transcriptomes have distinct profiles early and late in infection, characterized largely by biosynthesis and motility, respectively. In sporozoites and oocysts distinct transcriptional profiles can be identified. The use of hosts with different immune competence highlights the role of adaptive and innate immunity in the host and offered a starting point for in-depth analysis of these responses. Availability of transcriptomes of parasites at different stages of their lifecycle in a natural host improves knowledge of the parasite biology and helps positioning *E. falciformis* as a model for Coccidia in a scientifically valuable host animal.

Keywords

Parasite, apicomplexa, RNA-seq, transcriptome, lifecycle, interaction

INTRODUCTION

*Eimeria* falciformis is an intracellular parasite in the phylum Apicomplexa. Among more than 4000 described species of Apicomplexa (Duszynski 2011) prominent pathogens of human are found such as Toxoplasma gondii, causative agent of toxoplasmosis, *Plasmodium* spp., causing malaria, and Cryptosporidium spp, which cause cryptosporidosis. Coccidiosis is a disease in livestock and wildlife caused by coccidian parasites which are dominated by > 1800 species of *Eimeria* (Duszynski 2011). The genus is best known for several species which are problematic for poultry industry (Chapman et al. 2013). A useful model for studying *Eimeria* spp. is *E. falciformis*, which naturally infects wild and laboratory mice, Mus musculus (Heitlinger et al. 2014). This gives access to a broad range of genetically manipulated hosts.

*Eimeria* spp. infect the gut and are highly niche specific. *E. falciformis* has its niche in the caecum and upper part of colon, mainly in the cells of the crypts (Haberkorn 1970; Schmid et al. 2012; Stange 2012). These monoxenous parasites go through asexual (schizogony) and sexual reproduction which results in hosts releasing high numbers of resistant oocysts. When a mouse ingests *E. falciformis* oocysts, one sporulated oocyst releases eight infective sporozoites inside the host, which can infect epithelial crypt cells of the caecum and colon. Within the epithelium, so called merozoite stages form in several rounds of schizogony. Parasite numbers increase drastically during schizogony, which is not completely synchronized, and the exact number of schizogony cycles is either not clear (Haberkorn 1970; Mesfin and Bellamy 1978) or it varies naturally. Haberkorn (1970) reported variation from one to four rounds of schizogony before gamete formation, whereas Mesfin and Bellamy (1978) detected four distinct schizont stages. It therefore appears that the number and exact timing of schizont formation can vary in *E. falciformis* infections. Oocysts form as merozoites and differentiate into gametes, which fuse and form a zygote. Oocysts are first detected in faeces on day six to seven post infection, post infection, and output peaks on day eight to nine post infection After day 13-15 post infection oocysts are no longer detectable (Haberkorn 1970 and this study). Immature, unsporulated oocysts are shed into the environment where they mature into sporulated, infective oocysts (Mesfin and Bellamy 1978).

It is well known that *Eimeria* spp. generally induce protection against reinfection in hosts (M. Elaine Rose 1974; Mesfin and Bellamy 1979; Blagburn and Todd 1984; M. E. Rose, Hesketh, and Wakelin 1992) and that T-cell seem to play a major role (Gadde et al. 2009; Sühwold et al. 2010). In host responses to *E. falciformis* infection of laboratory mice, IFNγ is upregulated (Schmid et al. 2014). It was however shown that even though IFNγ and its receptor influences parasite development, the pathway is not required for clearance of a first infection, or for development of protective immunity in challenge infections. In the IFNγ deficient model which displays larger weight losses and intestinal pathology but also lower oocyst output, the phenotype was recovered by blocking IL-17A and IL-22 signaling. Only blocking IL-22 signaling recovered the oocyst shedding but had no effect on weight loss (Stange et al. 2012). These studies demonstrate the complex relationships between parasite and host and exemplify that increased pathology in the host is not necessarily beneficial for the parasite. Adaptive immunity clearly plays a role in limiting *Eimeria* spp. reinfection, but the effect on the parasite remains poorly understood.

The closely related coccidian parasite *T. gondii* also induces immunity in its definitive host, which only sheds oocysts once in their life (VanWormer et al. 2013; Hartmann et al. 2013). Several genes which are identified as virulence factors in *T. gondii* are also present in *Eimeria* spp. genomes, such as surface antigen genes, SAG-genes and rhoptry proteins, ROPs (Heitlinger et al. 2014; Adam J. Reid et al. 2014). Although the number and functions of these genes are not identical, existing functional knowledge on *T. gondii* orthologues is useful in studying *Eimeria* spp.. Additionally, the close relationship between the parasites and partly shared niche of these species, but the extreme difference in specificity between the generalist *T. gondii* and the highly specific *Eimeria* spp. makes comparisons interesting from an evolutionary perspective.

Existing studies address either host (Schmid et al. 2014) (+the other plos1 paper) or parasite (Periz et al. 2007; Han et al. 2009; Schwarz et al. 2010; Aarthi et al. 2011; Amiruddin et al. 2012; Novaes et al. 2012; Adam James Reid and Berriman 2012; Schmid 2014; Walker et al. 2015; Matsubayashi et al. 2016) transcriptomes in *Eimeria* spp. and other apicomplexan parasite infections. In addition, Reid and Berriman (Reid and Berriman 2012) have analyzed two microarray datasets from independent studies of *Plasmodium* spp. in mouse and in mosquito, respectively (Lovegrove et al. 2006; Xu et al. 2005). In this metaanalysis, the authors correlate host and parasite transcriptional changes to identify interacting host-parasite gene pairs. To our knowledge, no transcriptome studies of apicomplexan parasites simultaneously and symmetrically asses both parasite and host over the full lifecycle.

The genome sequence oof *E. falciformis* (Heitlinger et al. 2014) lays the foundation for transcriptome studies based on high throughput sequencing of RNAs (RNA-seq) and mapping to the parasite genome. The technology allows simultaneous assessment of RNA abundance in host and pathogen. Such approaches have recently been termed dual RNA-seq (Foth et al. 2014; Rosani et al. 2015; Westermann et al. 2016; Fernandes et al. 2016; Li et al. 2016). Applying such dual RNA-seq to an infection of *E. falciformis* in the mouse, we produce host and parasite transcriptomes from the same samples and tissue. We analyze host and parasite mRNA profiles at several time-points post infection and contrast transcriptomes from infections in wild type mice with immunized (challenged), naive and in hosts with strong deficiency in adaptive immune responses. We use these dual transcriptomes to screen for genes potentially involved in host parasite interactions.

RESULTS & DISCUSSION

***E. falciformis* infection dynamics in immune competent and immune deficient mice**

To investigate *E. falciformis* development throughout the lifecycle in its natural mouse host, dual transcriptomes were produced at days 3, 5, and 7 post infection. In contrast to wild-type mice, we also investigated parasite development and mouse transcriptomes in a mouse strain which is severely limited in adaptive immune responses (*Rag1-/-*) at day 5 post infection. To further elucidate host immune responses and parasite sensitivity to host immunity, we also challenge infected all mouse groups and sampled at the same time-points as in naïve mice.

Basic phenotyping of wild-type mice showed drastically decreased oocyst output (Figure 1 a) in immunocompetent challenged hosts compared to naïve animals (Mann–Whitney U test on total output. NMRI; p = 0.004. C57BL6; p = 0.008).

Similarly, a reduction in parasite 18S transcripts determined by quantitative reverse transcription PCR (RT qPCR) was detected (Figure 1b). Therefore, as expected in *Eimeria* spp. infections (Ovington, Alleva, and Kerr 1995), we demonstrate reduced intracellular parasite numbers (18S RT qPCR data) and drastically reduced reproductive success (oocyst shedding) in immune competent, challenge infected mice.

In contrast, no significant difference in reproductive success was observed between first and challenge infection in immune deficient mice (Mann–Whitney U test; p = 0.08). We thus observe a trend towards reduction of total oocyst output also in challenge infections of *Rag1*-/-. This reduction points towards a resistance mechanism independent acquired immunity (T and B cells). Autonomous mechanisms in the epithelium (e.g. Hunn et al. 2011) are imaginable to persist in tissue after a first infection (cite ref) but also mechanisms like physiological changes in epithelial cells might lead to reduction in suitability of a niche for parasites.

Similarly, oocyst numbers in faeces peaked on days 8-9 and all mice, also the immune deficient group, had cleared the infection by day 14. We thereby note that *E. falciformis* infection is self-limiting also in mice without mature T and B cells and it shows the same oocyst shedding kinetics as in immune competent mice.

Asexual replication and increase in parasite numbers of *E. falciformis* intestinal stages is reflected by the percentage of parasite material (reads) sequenced. This data is shown per time-point post infection (Figure 1c) and corresponds well with RT qPCR profile of 18S transcripts (Figure 1b). We thus use dual RNA-seq to analyze the lifecycle of *E. falciformis* and host responses to infection. Transcriptomes of parasites and infected hosts of different immune capacity at early and late stages of infection, and challenge infections are described.

*Parasite and host dual transcriptomes can be assessed in parallel*

We purified mRNA from infected caecum epithelium and demonstrate that even early in infection (3 days post infection) there is sufficient parasite material to detect the parasite mRNA among host mRNAs, and to quantify them by RNA-seq (Figure 1a). Total numbers of sequenced reads as well as reads mapped to either the *E. falciformis* genome or the mouse genome are indicated in Table 1 for all replicates. The number of total (host + parasite) read mappings for individual replicates ranged from 25,362,739 (sample Rag\_1stInf\_0dpi \_rep1) to 139,749,046 with 25% of the reads mapping to *E. falciformis* (NMRI\_1stInf\_5dpi\_rep3).

Multidimensional scaling of host (additional file x) and parasite (additional file y) mRNA abundance patterns confirms the absence of batch effects including batches processed with different sequencing technologies. Concordant patterns and similarities between samples show a large range of read numbers which is a result of different parasite-read proportions, especially for *E. falciformis* at days 3 and 5 post infection. This confirms efficient normalization of read counts (for details on normalization and control of resulting distributions see Methods and additional file X).

Remarkably, on day 7 post infection, the day before oocyst shedding peaks, infected naïve mouse epithelium contained 77% and 92% parasite mRNA, i.e. more mRNA from the parasite than from the host (Figure 1c and Table 1). This can partly be due to transcriptionally active parasites and/or inactive/apoptotic host cells, but is likely due to high parasite numbers.

**The mouse transcriptome undergoes broad changes upon *E. falciformis* infection**

We here show that upon infection with *E. falciformis*, which induces weight loss and intestinal pathology in mice (SI xx), the host transcriptome undergoes drastic changes with more than 3000 genes changing their mRNA profile significantly (edgeR; glm likelihood-ratio tests). Statistical testing for differential expression between infected and uninfected mice revealed changes in mRNA abundance becoming more pronounced (both in magnitude and number of genes affected) at later time-points post infection (Table 2 and Figure 2A). 325 mouse mRNAs were considered differently abundant (FDR < 0.01) between controls and 3 days post infection, 1,804 mRNAs between controls and 5 days post infection and 2,711 mRNAs between controls and 7 days post infection. This leads to a combined set of 3,453 unique genes responding to infection. Differentially abundant mRNAs early in infection (3 and 5 days post infection) were not a strict subset of genes differentially abundant later in infection (7 days post infection; Figure 2a), which would be the case if the same genes are regulated throughout infection.

Instead, the transcriptional profile of the mouse changes more fundamentally with different genes changing in abundance late compared to early in infection. Our results for late infection are in agreement with previously published microarray data from mice infected with *E. falciformis* species (Schmid 2014), for which our 7 days post infection data was correlated. The two data-sets show a strong correlation (Spearman's rho = 0.74; SI XX). Considering both biological differences in the experiments such as exact time-points for sampling, and technical differences between the two methods, this correlation confirms the adequacy of using dual RNA-seq for assessing the host transcriptome.

To further analyze the distinct responses early and late in infection, we performed hierarchical clustering on the mouse genes differentially abundant between different time-points post infection (Figure 2b). Three main sample clusters formed (dendrogram of columns at top of Figure 2b). Immune deficient *Rag1-/-* mice, including infected samples, cluster with uninfected samples. The similarity between infected and non-infected *Rag1-/-* samples, confirms the immune deficiency phenotype (failure to react to infection) in these mice and suggest a strong influence of adaptive immune responses on overall transcriptional response. Surprisingly, these patterns indicate that innate immune responses and other B- and T- cell independent processes play a relatively small but detectable role (cluster 4 Figure2) in shaping the mouse transcriptome upon *E. falciformis* infection. However, it is questionable whether this pattern corroborates a host response dominated by adaptive immunity. Adaptive immune responses are a concerted process involving, e.g., immune cells infiltrating tissue (ref general; ref *Eimeria*) likely large enough to be detectable in the transcriptome, whereas an innate, cell autonomous process could rely on few effector molecules with relatively small impact on the overall transcriptome (general ref; toxo e.g. rop ref).

***B and T cell dependent early responses to parasitic infection***

The self-limiting nature of *E. falciformis* infection and host resistance to reinfection (Ovington, Alleva, and Kerr 1995) makes it interesting to analyze transcritpomes of immune competent hosts in depth. For such hosts, clustering of genes and samples based on differentially abundant mRNAs (Figure 2b) clearly distinguishes non-infected controls from infected animals. On days 3 and 5 post infection, two clusters of genes have overall high mRNA abundance (1 and 2). They are also different from controls in having low mRNA abundance in two gene sets which are dominated by high mRNA abundance in non-infected controls (3 and 4). Among these genes which are lower in abundance upon infection, overrepresented GO terms are, e.g., “regulation of” the following: transforming growth factor-β, TGFβ, epidermal growth factor, EGF, tumor necrosis factor, TNF”, “negative regulation of interleukin-1 secretion”, “negative regulation of interleukin-6 secretion”, “Notch receptor processing”, “calcineurin-NFAT signaling cascade”, “Inositol-phosphate mediated signaling”, “autophagy”, “response to mechanical stimulus”, “ammonium transmembrane transport”, “positive regulation of axon guidance”, "lipid metabolic process", "protein complex localization" and "blood coagulation". These terms indicate an induction of IL-1, which is involved, e.g., in T and B cell maturation and IL-6, involved in, e.g., broad acute phase immune responses and T and B cell maturation (The UniProt Consortium 2015: entries for Interleukin-1 alpha, and Interleukin-6). TGFβ is important for wound healing in intestinal epithelium (Beck et al. 2003), and EGF regulates proliferation of epithelial cells and inhibits apoptosis (Suzuki et al. 2010). TNF is dose-dependent and apart from inducing several immune responses is also reported to regulate proliferation of epithelial cells (Kaiser and Polk 1997). IL-6 has also been shown to support tissue repair and inhibit apoptosis after epithelial wounding (Kuhn et al. 2014). IL-6 is also linked to Th17 responses (Park et al. 2005) which play an important role in responses to *E. falciformis* (Stange 2012). Inhibition of Notch signaling in mouse epithelium has been shown to alter the composition of cell-types in the epithelium towards Paneth and Goblet-like cells (VanDussen et al. 2012). Although speculative, several of the GO-terms clusters 1, 2, 3, and 4 can be linked to explain fundamental underlying processes: Inositol signaling can lead to release of calcium and calcineurin-dependent translocation of NFAT to the nucleus and activation of its target genes in T cells, but also many other cell types (reviewed by Macian 2005). In addition, changes in the host epithelium do take place when cells are invaded by, e.g., *E. falciformis*, but also generally by pathogens. It is important to note that the specificity in the GO-terms alone is not enough to determine whether healing is induced or repressed (compare “regulation of” and “negative regulation of”; ref). Further investigation of the role of the processes and molecules highlighted here will contribute to better understanding for responses to intestinal intracellular parasitic infection, intestinal wound healing and other epithelial remodeling and how it is regulated. We do however get insights into the molecules involved in the early responses, which most likely are important for the induction of adaptive immunity later in infection.

GO terms enriched among the mRNAs which become more abundant only early in infection are, e.g., “stem cell population maintenance”, “mRNA processing”, and “cell cycle G2/M transition”, also indicating changes in the epithelium. In addition, terms such as “regulation of response to food” and “negative regulation of appetite” are enriched. A somewhat unexpected term, “photoreceptor cell development”, (and similar terms pointing towards visual perception) is supported, among others, by the gene mechanistic target of rapamycin, mTOR. The mTOR signaling pathway is known as an amino acid sensor (e.g. Wellen and Thompson 2010; Laplante and Sabatini 2013; Gallinetti, Harputlugil, and Mitchell 2013) and it is therefore perceivable that it is involved in the regulation of food responses in *E. falciformis* infections. Weight losses and malnutrition are generally common during parasitic infections (e.g. Preston-Mafham and Sykes 1970; Stephenson, Latham, and Ottesen 2000; Aloisio et al. 2006), also in *Eimeria* spp. (e.g. Sharman et al. 2010; Stange et al. 2012) and weight loss was also seen in this study (SI file xx).

Interestingly, in T- and B- cell deficient hosts, the same four groups of genes described above (1-4, figure 2b) show no differences between infected and non-infected animals. In two of these gene sets, all immune deficient mice - also infected ones - are most similar to non-infected immune competent control animals (1 and 4). These genes are the ones enriched for, e.g., regulation of IL-1 and IL-6, TGFβ, TNF, EGF, Notch receptor, and calcineurin-NFAT signaling (cluster 4) and, e.g., stem cell population maintenance, terms for cell cycle regulation and several RNA-processing related terms (cluster 1). In one gene set (cluster 3), all immune deficient animals’ mRNA profile is most similar to infected immune competent animals. Apparently, T and B cell deficient mice have similarities with infected wild-types for mRNAs linked to GO-terms for neuron differentiation and regulation and several metabolic processes. The pattern in cluster 3 suggests that these mRNAs have a high abundance in healthy wild-type animals in a T or B cell dependent fashion, since levels are low in T and B cell (maturation) deficient mice independent of infection. Such differences in *Rag1*-/- mice [are expected and have been reported (ref)/ are unexpected and have to our knowledge not been reported before.]

***Adaptive immune responses characterize late infection***

Pronounced transcriptional changes occur late in infection. Processes indicated by genes with increased abundance at day 7 post infection reflect the expected onset of an adaptive immune response (cluster 5 in Figure 2b shows general dynamics of these genes). GO terms such as "antigen binding", "immunoglobulin receptor binding", "immune system process", "adaptive immune response", “cytokine production”, and also "innate immune response" and “regulation of apoptotic process” dominate enrichment analyses for this group of genes. Among the same genes, “natural killer cell regulation”, “neutrophil degranulation”, “JAK-STAT signaling”, and IL-1 and interleukin-2, IL-2 productions are enriched biological processes. As late as day 5 post infection, the genes responsible for these enrichments still low on mRNA abundance. This confirms a strong induction of immune responses, particularly adaptive immune responses between days 5 and 7 post infection. The result is well in line with previously described protective immunity against re-infection with *Eimeria* spp. (M. Elaine Rose 1974; Mesfin and Bellamy 1979; Blagburn and Todd 1984; M. E. Rose, Hesketh, and Wakelin 1992; Gadde et al. 2009; Sühwold et al. 2010; Stange et al. 2012; Schmid et al. 2014) .

***Challenge infected animals are less responsive to parasite life cycle stage***

Transcriptomes from three challenge infected samples from early and late secondary infection show a distinct profile of elevated mRNA abundance (3, 5 and 7 days post infection, cluster 6, Figure 2b). The underlying mRNAs are highly enriched for GO terms for RNA processing and splicing as well as terms for histone and chromatin modification. This might suggest that protective immune responses in challenge infected animals are regulated both at the transcriptional and post-transcriptional level. The high abundance of these mRNAs at different time-points post infection in wild-type hosts, and the completely cleared infection in some samples (Table 1; and unexpected clustering of e.g. NMRI\_2ndInf\_7dpi\_rep2) also suggests individual variation in the timing, and possible also strength, of these responses.

Taken together, the mouse epithelial transcriptome changes upon infection by its natural parasite *E. falciformis*. The largest change is detected by comparing un-infected control animals to day 7 post infection samples, which is the day before oocyst shedding peaks. Early in infection, when the parasite reproduces asexually and multiplies within intestinal epithelial cells, in total ~750 unique mouse genes regulated compared to late infection. Late in infection ~1600 genes are uniquely regulated, i.e., these genes were not detected early in infection. Early in infection we identify regulation of a number of cytokines, stem cell population maintenance as well as changes in food response or appetite as altered host processes. Late infection is dominated by adaptive immune responses as well as ongoing innate immune responses. Responses in wild-type challenge infected hosts suggest strong gene expression regulation both at the transcriptional level and in RNA processing. In contrast, all these patterns are missing in T and B cell maturation deficient mice, and in general our analysis detected no or minimal transcriptional response to infection in immune deficient animals.

**Transcriptional differences in the parasite life cycle are independent of mouse immune status**

Not only do we for the first time study transcriptomes for the full parasite lifecycle of *E. falciformis* but we can also assess the parasites’ transcriptome in relation to host immune status. In *Eimeria* spp. to our knowledge nothing is known about parasite counter reactions to immune responses or about differences in these related to host immune status.

In *E. falciformis*, we generally see no differences in the parasite transcriptome between infection in immune competent mice, T and B cell deficient *Rag1-/-* mice, or between naive and challenge infected mice. This is surprising considering the measured differences in oocyst output in the same comparisons (Figure 1a), and the fact that the same comparisons display distinct patterns in mouse transcriptomes.

Major changes in the parasite transcriptome take place between 5 and 7 days post infection (Figure 3a). Between 3 and 5 days post infection differences are smaller:103 mRNAs were differently abundant between early time-points after infection (edgeR likelihood ratio tests on glms; FDR < 0.01), whereas between 3 and 7 days post infection 1399 mRNAs were differentially abundant, and between 5 and 7 days post infection 2084 mRNAs were differentially abundant (Figure 3a). These results motivated us to define 3 and 5 days post infection as "early infection" and 7 days post infection as "late infection" for the parasite transcriptomes. Early and late infection samples were tested for differential abundance compared to sporozoites and sporulated oocysts, resulting in 1697 and 3919 mRNAs, respectively. In order to assess biological relevance of these patterns, we performed hierarchical clustering of the differentially abundant mRNAs described above and applied enrichment analyses for GO terms and “gene family conservation profiles” based on annotations from (Heitlinger et al. 2014).

Distinct clusters of genes define early infection (3 and 5 days post infection, cluster 6 in Figure 3c). At those time-points unsynchronized asexual reproduction takes place (Haberkorn 1970; Mesfin and Bellamy 1979) Two separate clusters define late infection (7 days post infection, clusters 2 and 7) in which we assume gametocytes to be present due to the following peak of oocyst shedding (Haberkorn 1970 and Figure 1a). Extracellular samples of sporozoites and oocysts are each defined by distinct gene clusters. Major patters in the parasite transcriptome seem to be determined by lifecycle stages. In contrast, the parasite seems to be "transcriptionally blind" to the host immune status.

*E. falciformis unique genes in sporozoites*

In nature, sporozoites are released from ingested oocysts in the gastrointestinal tract. This process was induced for our samples *in vitro* (see Methods). We find that *E. falciformis* sporozoites are defined by a group of genes (cluster 4, Figure 3b and Additional file 5) which are largely specific to *E. falciformis*, meaning that fewer than expectedorthologues of genes which are abundant in sporozoites are found as gene family members outside of this species (Fisher’s exact test, FET, multiple testing adjusted p-value FDR = 0.002). Interestingly, five out of 12 SAG gene transcripts predicted for *E. falciformis* (Heitlinger et al. 2014) cluster with the mRNAs highly abundant in sporozoites, making them suitable to the name “sporozoite surface antigens” (one original Ef ref). SAG proteins are thought to be involved in host cell attachment and invasion, and possibly in induction of immune responses in other apicomplexan species (Mineo and Kasper 1994; Grimwood and Smith 1996; Cowman and Crabb 2006; Carruthers and Boothroyd 2007; Chow et al. 2011).

In total ten SAGs were detected as differentially abundant in our data, the other five are found in gene clusters 1, 2, 5, 6, and 7, indicating an expression throughout all stages. An expression of particular SAGs in stages other than sporozoites has been reported before in *E. falciformis* (the other Ef SAG ref). The expression of a subset of these mRNAs in sporozoites as an easily accessible lifecycle stage opens for analysis of the function of these genes in *E. falciformis*.

Other proteins which have received attention as important virulence factors in *T. gondii* (Taylor et al. 2006; Saeij et al. 2007; Fleckenstein et al. 2012) and also some attention in *E. tenella* (Talevich and Kannan 2013) are rhoptry kinases (RopKs). Two out of eight predicted RopKs are highly abundant in sporozoites (EfaB\_PLUS\_8664.g829 and EfaB\_PLUS\_15899.g1411 in cluster 4). Orthologues in *T. gondii*, are ROP21 and ROP35. In *E. tenella* sporozoites several RopKs are expressed and have also been shown to be differentially expressed compared to intracellular merozoite stages (Oakes et al. 2013).

GO enrichment data suggests ATP production and biosynthesis processes as dominant features in sporozoites. Metabolic pathway analysis (ToxoDB) of the same genes reveals several pathways which also point towards ATP production, including fatty acid degradation, oxidative phosphorylation and valine, leucine and isoleucine degradation (Additional file 6). The degradation of branched chain amino acids pathway is used in *T. gondii* for energy supply (Oppenheim et al. 2014). In addition, the invasive stage is characterized by "maintenance of protein location in cell" and similar GO terms. Possibly, this is due to control of microneme or rhoptry protein localization as sporozoites prepare for invasion. Sporozoites therefore display a transcriptome indicative of large requirements for ATP and production of known virulence factors such as SAG and RopKs and are characterized by expression of species specific genes.

*Growth processes dominate the transcriptome during asexual reproduction*

After sporozoites have invaded epithelial cells in caecum intracellular asexual reproduction can be initiated. Growth processes and elevated abundance of several RopKs are characteristic for early infection, day 3 and 5 post infection. At these time-points we expect several rounds of schizogony to take place in a somewhat unsynchronized fashion (Haberkorn 1970; Mesfin and Bellamy 1978). In agreement with that expectation, differentially abundant mRNAs can in general not be distinguished between early 3 and 5 days post infection samples (cluster 6 in Figure 3b). The same pattern is seen in mouse data, which suggests mutual influences between host and parasite. Among early infection mRNAs, several GO terms for biosynthetic activity are enriched, e.g., "ribosome biogenesis", "cellular biosynthetic process", "gene expression", "cellular amino acid catabolic process" and “RNA processing”, including terms for “tRNA and ncRNA processing”. Biosynthetic processes are enriched both in early infection and in invasive sporozoites (SI files x and y), but the terms are driven by different genes/mRNAs. Analyzing conservation status in these lifecycle stage defining gene clusters, sporozoite specific genes displayed a species specific profile (low number of orthologues outside *E. falciformis*, see above), whereas early infection mRNAs are enriched (FET,multiple testing adjusted p-value FDR= 0.01) for orthologues shared by 10 apicomplexan species (selected as in Heitlinger et al. 2014).

This could reflect that gene families conserved among these apicomplexan species are involved in asexual reproduction, whereas invasion may be a more species specific event. In early infection high abundance mRNAs we found four out of ten RopKs predicted in *E. falciformis* (Heitlinger et al. 2014). This is the largest number of RopKs in any group of differentially abundant genes in our analysis (EfaB\_PLUS\_24117.g1969, EfaB\_MINUS\_17096.g1521, EfaB\_MINUS\_42996.g2710 and EfaB\_PLUS\_7742.g778) and constitutes a statistically significant enrichment (FET; p=xxx). Three of these have orthologues in *T. gondii*: ROP41, ROP35 and ROP21. In *T. gondii* type II, single deletions of ROP41 or ROP21 results in an intermediate decrease in cyst burden in mouse brain, compared to the parental Δku80 strain (ref ?Howard?). The ROP35 deletion strain produced less, approximately a third, of the cysts compared to the parental strain. (Fox et al. 2016) However, the bioinformatically predicted ROP21 has been shown not to localize to the rhoptry organelles in *T. gondii* (Jones, Wang, and Sibley 2016). We highlight the fact that we base annotation on hidden markov modesl (HMMs) for specific types of kinase domains (Talevich and Kannan 2013) and despite the name these proteins might not be located in rhoptries. Therefore, it is likely that especially the ROP21 orthologue also in *E. falciformis* has a localization different from rhoptries but is better characterized for a kinase function.

Enrichment of replication and growth-related processes highlights the parasite's expansion on 3 and 5 days post infection. This is supported both by previous knowledge about the lifecycle (Haberkorn 1970) and by the increase in parasite derived sequences we measured (Figure 1b and 1c). The mRNAs supporting growth related processes all have low abundance late in infection, on 7 days post infection. This likely reflects the switch from early asexual expansion towards differentiation and sexual reproduction on the time-point one day before oocyst output peaks. Apart from a clear transcriptional profile for growth, the presence of four out of ten RopKs in the early intracellular phase suggests an important role for these proteins. Several RopKs have been identified and partially characterized in other *Eimeria* spp. (Rick, Dubremetz, and Entzeroth 1998; Fetterer et al. 2013) and in *T. gondii* (reviewed in Bradley and Sibley 2007). Our data gives a first overview of expression patterns for *E. falciformis* ROPs and offer a good starting point for functional analysis of these virulence factors in mouse *Eimeria* spp.

*Gametocytes are likely to dominate transcriptome late in infection*

Two *E. falciformis* gene clusters show a distinct profile characterized by high mRNA abundance on 7 days post infection (clusters 2 and 7; Figure 3c). Both clusters display low mRNA abundance in other lifecycle stages, especially in oocysts and sporozoites. Late stage defining mRNAs are underrepresented for orthologues shared among core apicomplexan parasites (Apicomplexa excluding Cryptosporidium; cluster 2; FET corrected for multiple testing; FDR = 0.002) or for genes conserved beyond the Apicomplexa (cluster 7, Additional file 5; FET corrected for multiple testing; FDR <0.0001). Enriched GO terms such as "movement of cell or subcellular component" and "microtubule-based movement" along with terms suggesting ATP production ("ATP generation from ADP") indicate the presence of motile and energy demanding gametocytes in these samples. Other genes support ATP production in this lifecycle stage. For both gene groups that define late infection, different (peptide/nitrogen compound/cellular protein/macromolecule) "biosynthetic process" terms along with "chitin metabolic process" suggest that the parasite is producing building blocks for oocysts and their walls. Similar processes have been observed in *E. tenella* oocysts (Walker et al. 2015). This fits the timing of oocyst output which peaks at 8-9 days post infection, one day after sampling for RNA-seq. In addition, the Apicomplexa specific genes expressed in oocysts are enriched for a number of distinct GO terms such as "blood coagulation". These reflect the presence of Thrombospondin type I domains in the protein products of these mRNAs. Thrombospondin type I domains have been reported in *E. tenella* microneme localizing proteins, MIC, e.g. MIC4 (Tomley et al. 2001). MIC4 mRNA was reported in *E. tenella* sporozoites where it localizes to the apical end, and in late schizonts and late oocyst stages, when sporozoites are forming. Possibly, the mouse parasite's MICs play a role in gamete stages or oocyst formation, or *E. falciformis* prepares for invasion already during oocyst formation.

*Oocysts are characterized by stress responses and differentiation*

Oocysts are the infective stage in the parasite lifecycle. They are shed with feces as unsporulated, “immature”, capsules and in the environment they undergo sporulation - meiotic and mitotic divisions (Duszynski and Duszynski 2011) - and become infective. Overall, the oocyst profile with five out of seven gene clusters characterized by below average abundance of mRNAs reflects the mainly inactive state of oocysts, which survive long times outside the host. However, two groups of mRNA are highly abundant, demonstrating that there is transcriptional activity in oocysts (clusters 1 and 5; Figure 3c). One of the gene sets (cluster 5) is enriched for apicomplexan-shared orthologues (FET corrected for multiple testing; FDR = 0.0005) and GO terms for "DNA repair", "protein modification process" and "cell differentiation". The same cluster is also the only cluster which is enriched for transmembrane domains (FET, FDR < 0.001).

In our study, sporulated oocysts with ~90% purity were used. mRNAs which support GO enrichment for DNA replication and quality control might be contributed by the ~10% unsporulated oocysts in culture and would then reflect, e.g., DNA replication taking place during sporulation. Two candidate RopKs were in a high abundance oocyst cluster (cluster1 and 5, EfaB\_MINUS\_720.g57 and EfaB\_PLUS\_33184.g2393). An orthologue of one of these mRNAs in *T. gondii* is ROP31. Mice infected with *T. gondii* knock-outs for ROP31 had about half the amount of brain cysts compared to the control strain (Fox et al. 2016) but nothing is known about the influence of the orthologeous gene on *Eimeria* spp. infections. Its presence in oocysts might indicate preparation for invasion, or suggest an unknown role in the oocyst stage. The other RopK *T. gondii* orthologue is ROP5…

Our data indicates that oocysts are transcriptionally rather inactive. The abundant mRNAs are largely conserved among apicomplexa and are not *Eimeria* spp. specific. Oocysts from many (but not all) different apicomplexan species can be speculated to face similar challenges in the environment and can therefore be expected to share genes which are expressed in this lifecycle stage. If such conservation can be confirmed, these genes constitute attractive targets for eliminating environmental oocysts. Sporulation blocking medication could then been applied on pasture or in farm housing (ref).

*Evolutionary conservation of different lifecycle stage transcriptomes accross* Coccidia

Working with a model parasite, good understanding of similarities and differences between closely related species is required. Importantly, incorrect assumptions about gene function, localization or stage specificity should be avoided. One way to achieve such understanding is to analyze evolutionary trajectories of gene families and to relate these to other kinds of data (Sidik et al. 2016). In addition to the enrichment analyses for evolutionary conservation categories for gene families in gene expression clusters (figure 3b and table 3) we performed Spearman's correlation analysis between our RNA-seq transcriptomes and RNA-seq data from related parasites.

Two datasets for the economically important chicken parasite *E. tenella* (Adam J. Reid et al. 2014; Walker et al. 2015) and one dataset of the model apicomplexan parasite *T. gondii* (Hehl et al. 2015) were included in the comparison. The latter was used because it is to date the only available dataset for the complete *in vivo* lifecycle of *T. gondii*, and therefore compares well with our data. For all samples from these studies and our data, abundances of orthologous genes were correlated and Spearman’s coefficient was compared (Figure 4).

With the exception of sporozoites (see below), transcriptomes tend to be more strongly correlated between matching lifecycle stages of different species than to other stages from the same parasite. Orthologues in *E. tenella* and *E. falciformis* gamete stages (gametocytes and late infection, respectively) are expressed in highly correlated manner across the two species. Similarly, transcriptomes of *E. tenella* merozoites from both independent studies are most similar to early *E. falciformis* samples, indicating similarity also during asexual reproduction.

Furthermore *E. falciformis* oocyst transcriptomes share a high similarity to unsporulated *E. tenella* oocysts. *E. tenella* sporulated oocysts are most similar to *E. tenella* sporozoites. This might reflect species specificity also in *E. tenella* sporozoites or possibly, differences in the in vitro preparation of these “environmental stages”. Our oocysts clustering with *E. tenella* unsporulated oocysts gives further support to the interpretation that our high mRNA abundance oocyst clusters (Figure 3b) are a result of transcripts contributed by the 10% unsporulated oocysts in our sequenced material.

We have identified groups of genes which are shared among apicomplexa (Figure 3b and table 3) and linked them to asexual reproduction in *E. falciformis*. Our data indicate that genes which are characteristic for intracellular asexual reproductive stages are conserved between *E. tenella* and *E. falciformis*. For considerations on drug or vaccine development further investigation into gene expression and evolutionary conservation between the different poultry infecting *Eimeria* spp. might be of interest. In order to develop vaccines which target several species, the immune system must be activated to target conserved genes (ref). Also in considering potential drug design this type of analysis might be helpful, as it can indicate a) in which stage a targeted gene is prone to change in expression and b) whether expression kinetics can be generalized beyond the species in question.

Overall this analysis highlights evolutionary conservation in gene expression of key developmental factors across Coccidia. Conservation of the underlying promoters and, i.e., transcription factor binding sites needs more investigation. However, it is reassuring that - apart from sporozoite genes - heterologous expression in other parasites with genetic manipulation tools available (first and foremost *T. gondii*; ref) might for many genes result in expression patterns resembling natural expression.

For sporozoites, this analysis confirms the species specificity of mRNA abundance patterns. Transcritpomes of *E. falciformis* sporozoites are more similar to *E. falciformis* early infection samples than to orthologous sporozoite transcriptomes from *E. tenella*. It is possible that the biological challenges faced by extracellular sporozoites in poultry and rodents are more different than intracellular conditions faced by merozoites during asexual reproduction in the two hosts (ref). This might result in fast evolution of virulence factors expressed in these invasive stages to such an extent that homologies in the underlying gene families cannot be detected anymore (Templeton 2007). It is also possible that such virulence factors undergo rapid gene family expansion, as seen in SAGs in *E. falciformis* (Heitlinger et al. 2014), *T. gondii* (Khan et al. 2006) and Neospora caninum, *Eimeria* spp. (Adam J. Reid et al. 2014) or VIR and how-are-these-called-again genes in *Plasmodium* (ref).

**Variability in parasite transcriptomes and response to infection**

In our data, three sampled lifecycle stages; sporozoites, oocysts and 7 days post infection (~gametes), display distinct profiles and replicates cluster together. However, some samples show aberrant expression patterns different between replicates and related samples. For parasite transcriptomes early in infection we explain the lack of separation between day 3 and 5 post infection samples with asynchronous schizogony. This fits well with previous knowledge about schizogony, as dscussed above. In contrast, synchronization upon gametocyte formation around 7 days post infection is expected since the oocyst shedding pattern is highly reproducible (ref). It is worthwhile to consider that such sample differences and similarities might reflect true biological variation. It is perceivable that the overall course and intensity of infection varies with small differences in, e.g., microbiota composition (Singer and Nash 2000; McCabe, Britton, and Parameswaran 2015), host stress levels due to cage mates, draught, differences in light exposure or other factors which do vary also in a controlled animal facility have been suggested (can we cite the editorial highlight in Science?). Variation in the course of infection triggered by such factors could explain the transcriptional profiles we see, with distinct overall patterns but replicate separation in the asexual phase. Furthermore some of the challenge infections are cleared or on their way to be cleared completely. It is perceivable that variation in adaptive immunity leads to complete or almost complete clearance. We suggest that considering such possibilities is important for interpreting results and understanding basic biology of the parasite and, in extension better understand infections in less homogeneous hosts than laboratory mice.

**Correlation between host and parasite mRNAs proposes interacting gene-pairs**

In the context of our dual RNA-seq experiment, variation between individual samples, including replicates, can allow us to identify mRNAs which correlate between host and parasite by their abundance patterns. This analysis follows the rationale that beyond the major (known) experimental groups, individual samples might contain tissue and parasites at a congruent state of infection and thereby provide information on inter-species gene-gene correlation. We thus applied the “ISIGEM method” (A. J. Reid and Berriman 2013) using empirical p-values for cross-species correlation coefficients with the aim to screen for interacting genes in host and parasite. This was performed in order to test whether and how interacting genes behave in relation to genes expressed according to the early-late infection patterns described above.

ISIGEM scores are probabilities of detecting correlations as strong as or stronger than the observed by chance (0 indicates high, 1 indicates low). In our data, ISIGEM scores display a rather flat distribution of number of interactions per gene, similar to that reported by Reid and Berriman (2013) with the difference of an overrepresentation of low values below 0.25 in our study (Figure 5a). Differentially abundant mRNAs between non-infected and infected samples, and between different time-points (same mRNAs as in Figures 2b and 3b) showed a very similar distribution to genes overall (Figure 5a).

We identified 3621 mouse genes and 1323 parasite “candidate interacting genes” defined by an ISIGEM score of zero (highest chance of true interaction). Distributions of the number of candidate interacting partners of these zero-score genes (Figure 5b) highlight a complex interaction network (Figure 5c).

TheCandidate interacting genes from the ISIGEM analysis are strongly enriched for genes which are abundant at 7 days post infection both in mouse (cluster 5, Table 4) and parasite (cluster 2, Table 4). This means that interacting genes identified by the method predominantly correlate for distinct host and parasite responses in the lifecycle, which are seen at 7 days post infection only (Figures 2b and 3b). All other clusters of mRNAs which are differently abundant in the mouse between infection and controls are significantly underrepresented for candidate interacting genes (Table 4).

Comparing the full distributions (Figure 5a) of ISIGEM scores across host and parasite expression clusters, this pattern is confirmed (Figure 5d). For mouse mRNAs abundant only at day 7 post infection, stronger than expected ISIGEM scores (< 0.25) are found with mRNAs abundant only at 7 days post infection in the parasite (expression cluster 2). This means that compared to the null expectation, not only more interacting genes (ISIGEM scores = 0) but also genes with more significant interaction scores (compare Figure 5a and 5d) are recovered in parasite-host gene-pairs which become abundant late in infection.

The method also highlights a lack of correlation between host and parasite mRNA abundance in early stages of infection, seen by an underrepresentation in enrichment tests for zero ISIGEM scores, and less significant ISIGEM scores. This confirms that variation in host immune response (naïve and challenged wild-type and *Rag1-/-* individuals) did not impact parasite gene expression (distributions for *E. falciformis,* esp. clusters 4 and 6 in Figure 5d), as the ISIGEM method did not detect parasite-host correlations in these samples.

We confirm that empirical p-values for correlations can be used to test links between host and parasite mRNA abundance by linking expression clusters observed in either species individually. We found “candidate interacting genes” at stages where adaptive immune reactions on the host side and commitment to sexual reproduction in the parasite prevail. A reduced number of such interactions, compared to null expectations, was observed using parasites in hosts of different immune competence. This supports our proposal of a parasite which is “transcriptomically blind” to the host immune system.

For the *E. falciformis* / mouse interaction we propose that in early infections the parasite follows a genetically predetermined path, fixed in an evolutionary process rather than a plastic response. We emphasize that gene expression is not necessarily a product of plastic host-parasite interactions, especially not in the parasite, but may instead follow genetically determined programs. It remains an open question whether differences in such fixed programs exist in encounters of genetically different host-parasite combinations.

METHODS

*Mice and infection procedure*

Three strains of mice were used in our experiments: NMRI (Charles River Laboratories, Sulzfeld, Germany), C57BL/6 (), and *Rag1-/-* on C57BL/6 background (gift from Susanne Hartmann, FU?). Animal procedures were performed according to the German Animal Protection Laws as directed and approved by the overseeing authority Landesamt fuer Gesundheit und Soziales (Berlin, Germany). Animals where infected as described by Schmid et al., (Schmid et al. 2012), but tapwater was used instead of PBS for administration of oocysts. Briefly, NMRI mice were infected two times, which will be referred to as “first” and “second” or “challenge” infection. For the first infection, 150 sporulated oocysts were administered in 100 µL water by oral gavage. During the first infection of 60 mice, all animals were weighed every day. On day zero, before infection, as well as on 3 days post infection, 5 days post infection and 7 days post infection, caeca from 3-4 sacrificed mice per time-point were collected. Epithelial cells were isolated as described in Schmid et al. (2012) to a 90% purity of epithelial cells. For challenge infection, mice recovered for four weeks before second infection. Recovery was monitored by weighing and visual inspection of fur. For the second infection, 1500 sporulated oocysts were applied by oral gavage in 100µL water. Three mice were used as non-reinfection control, referred to as day 0, second infection. *Rag1-/-* mice and the background C57BL/6 strain control mice were also subjected to first and challenge infections with 10 sporulated oocysts in 100 µL water in both cases. Samples were taken on day 0 (pre-infection control) and day 5 post infection in both first and second infections and treated as described above for NMRI mice.

*Oocyst purification for infection and sequencing*

Sporulated oocysts were purified by flotation from feces stored in potassium dichromate and administered orally in 100 uL tapwater. One *E. falciformis* isolate, *E. falciformis* Bayer Haberkorn 1970, was used for all infections and parasite samples. The strain is maintained through passage in NMRI mice in our facilities as described previously (Schmid et al. 2012).

*Sporozoite isolation*

Sporozoites were isolated from sporocysts by in vitro excystation. For this, sporocysts were incubated at 37°C in DMEM containing 0.04% tauroglycocholate (MP Biomedicals) and 0.25% trypsin (Applichem) for 30 min. Sporozoites were purified in cellulose columns as described by Schmatz et al. (Schmatz 1997).

*RNA extraction*

Total RNA was isolated from infected epithelial cells, sporozoites, and sporulated oocysts using Trizol according to the manufacturer’s protocol (Invitrogen). Purified RNA was used to produce an mRNA library using the Illumina’s TruSeq RNA Sample Preparation guide. Sporozoites were stored in 1 mL Trizol until RNA-isolation. Total RNA was isolated using the PureLink RNA Mini Kit (Invitrogen) and reverse transcribed into cDNA. polyA selection… Simone add?

*Sequencing and quality assessment*

cDNA samples were sequenced by either GAIIX (13 samples) or Illumina Hiseq 2000 (14 samples) as specified in Table 1 (both unstranded). A fastq\_quality\_filter (FASTQ-toolkit, version 0.0.14, available at https://github.com/agordon/fastx\_toolkit.git) was applied to Illumina Hiseq 2000 samples using a phred score of 10. We intentionally did not use a stringent trimming before mapping to genome assemblies as the mapping process itself has been shown to be a superior quality control (MacManes 2014).

*Alignment and reference genomes*

The *Mus musculus* mm10 assembly (Genome Reference Consortium Mouse Build 38, GCA\_000001635.2) was used as reference genome for mapping and corresponding annotations for downstream analyses. The *E. falciformis* genome (Heitlinger et al. 2014) was downloaded from ToxoDB (Gajria et al. 2007). For mapping, mouse and parasite genome files were merged into a combined reference genome, and files including mRNA sequences from both species were aligned against this reference using TopHat2, version 2.0.14, (Trapnell, Pachter, and Salzberg 2009) with –G specified, and a Bowtie2, version 1.1.2, (Langmead and Salzberg 2012). This was done to avoid spurious mapping in ultra-conserved genomic regions. Single-end and pair-end sequence samples were aligned separately with library type 'fr-unstranded' specified for pair-end samples. Bam files (Frazee et al. 2015) were used as input for the function “featureCounts” from of the R package “Rsubread” (Liao, Smyth, and Shi 2014).

*Differential mRNA abundance, data normalization and sample exclusions*

After import of data to R, mouse and parasite data was separated using transcript IDs and analyzed, including normalization, separately. For each species, count data was normalized using the R-package edgeR (version 3.14.0; cite) with the upperquartile normalization method. Briefly, genes with below a an overall of 3000 (mouse) and 100 (E.falciformis) reads summed over all samples (libraries) were removed and normalization factors were calculated for the 75% quantile for each library. This normalization is suitable for densities of mapping read counts following a negative binomial distribution (SI XX). We excluded samples NMRI\_2nd\_3days post infection \_rep1 and NMRI\_2nd\_5days post infection \_rep2 due to low parasite contribution (0.012% and 0.023%) to the overall transcriptome. Technically, this exclusion made it possible to obtain parasite read counts in agreement with a negative binomial distribution (SI x). It is likely that the number of reads in the excluded samples would have been insufficient to fully normalize these datasets to those with the highest parasite contributions. Both excluded samples are from challenge infection and it is likely that the infected mice were immune to re-infection. One additional sample (NMRI\_1stInf\_0days post infection \_rep1) was excluded because the uninfected control showed unexpected mapping of reads to the *E. falciformis* genome (0.033%). As samples and individual replicates were sequenced in batches to different depth and using different instrumentation (Table 1) we performed multidimensional scaling of samples as quality controls using the function provided in the package EdgeR (additional files xyz).

*Testing of differentially abundant mRNAs and hierarchical clustering*

We used EdgeR (v 3.16.2 )to fit generalized linear models (GLMs with a negative binomial link function) for each gene (glmFit) and to perform likelihood ratio tests for models with or without a focal factor (glmLRT) using the “alternat design matrix” approach specifying focal contrasts individually. Tested contrasts comprised for the mouse a) infections at each time-point vs. uninfected controls b) corresponding time-points between different mouse strains and c) corresponding time-points and mouse strains for first and challenge infection. For the parasite contrasts were set between a) all different stages of the lifecycle, as well as b) and c) as above (see also results, table 2).

Mouse mRNAs responding to infection or differently abundant at different time-points of infection (0 vs “any days post infection ” or “any days post infection ” vs “any days post infection ”; see Table 2) and *E. falciformis* genes showing differences between any lifecycle stage (oocysts vs sporozoites, or any of those vs. “any days post infection ” or “any days post infection ” vs. “any days post infection ”) were selected and used for hierarchical clustering. Hierarchical clustering was performed using the complete linkage method based on Euclidean distances between Z-scores (mRNA abundance values scaled for differences from mean of each gene in units of standard deviations).

*Enrichment tests*

Gene Ontology (GO) enrichment analysis was performed using the R-package topGO with the “weight01” algorithm and Fisher's exact tests. We additionally performed a correction for multiple testing on the returned p-values (p.adjust using the BH-method). Similarly, a Fisher's exact test and corrections for multiple testing were also used to test for an overrepresentation of transcripts with a signal sequence for entering the secretory pathway or with transmembrane domains (as inferred using Signal P) which are predicted for the *E. falciformis* genome (Heitlinger et al. 2014).

Evolutionary conservation of gene families was analyzed based from categories from Heitlinger et al. (2014) which are as follows: i) *E. falciformis* specific, ii) specific to the genus *Eimeria*, compiled by an analysis of *E. falciformis*, E. maxima and *E. tenella*, iii) Coccidia: *Eimeria* plus *T. gondii* and Neospora caninum, iv) Coccidia plus Babesia microti, Theileria annulata, *Plasmodium* falciparum and *Plasmodium* vivax v) the same apicomplexan parasites plus Cryprosporidium hominis, vi) universally conserved in the eukaryote super-kingdom inferred from an analysis of Saccharomyces cerevisiae and Arabidopsis thaliana. These categories were tested for overrrepresentation in gene-sets with particular patterns described in the text using Fisher's exact-tests and the resulting p-values were again corrected for multiple testing.

*Correlation analysis of apicomplexan transcriptomes*

Transcriptome datasets from Reid et al (2014), Walker et al. (2014) and Hehl et al (2012) were downloaded from ToxoDB (Gajria et al. 2007). Orthologues between *E. falciformis*, *E. tenella* and *T. gondii* were compiled as defined in Heitlinger et al. (2014) and only 1:1:1 otholog triplets were retained for analysis, as multi-paralog gene-families might contain member showing divergent evolution of gene-expression due to neo/sub functionalization. Spearman’s correlation coefficients for expression over different samples in all studies and over different species represented by their orthologues were determined. Hierarchical clustering with complete linkage was used to cluster resulting correlations coefficients.

*Correlation analysis between host and parasite*

The ISIGEM method (A. J. Reid and Berriman 2013) was used to assess correlations of host and parasite transcripts. This method avoids spurious correlations for host-parasite gene pairs remaining e.g. unchanged over many conditions. It does so by estimating emmpircal p-values (ISIGEM-scores), which correspond to the proportion of cases in which a random shuffeling of samples produces better correlations than those observed in the given samples between host-parasite pairs. Similar to Reid and Berriman we used 5x10⁵ replicates of random shuffles and did not correct for multiple testing in our set of 79,764,454 tested correlations between 13,343 mouse, and 5978 parasite genes.

All analyses were performed in R (R Development Core Team 2008). Complete scripts are available at [https://github.com/derele/Ef\_RNAseq.git tagged as version 1.0](https://github.com/derele/Ef_RNAseq.git%20tagged%20as%20version%201.0).

**COMPETING INTERESTS**

The authors declare that they have no competing interests.

**AUTHOR CONTRIBUTIONS**

Animal and parasite experiments: Simone Spork, experimental design: Richard Lucius, Simone Spork, Emanuel Heitlinger, RNA sequecing: Christoph Dieterich, data anylsis: Emanuel Heitlinger, Totta Kasemo, text: all?

**ACKNOWLEDGEMENTS**

This project was funded by..... Totta Kasemo's PhD project is funded by the German Research Foundation (DFG) program GRK4026: Parasite Infections: From Experimental Models to Natural Systems.

**Tables**

**Table 1** Summary of data sorted per sample. Sequencing method, experimental batch, total reads sequenced, reads mapping to mouse genome, reads mapping to *E. falciformis* genome, percentage of *E. falciformis* reads, and number of *E. falciformis* genes with detected transcripts are shown.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Sample** | **Sequencing method** | **Batch** | **Total reads** | **Reads mapping mouse** | **Reads mapping**  **E. falciformis** | **Percentage**  **E. falciformis** | **# E. falciformis genes** |
| NMRI\_2ndInf\_0dpi\_rep1 | GAII | 2 | 108,937,797 | 70,489,674 | 247 | 0.0004 | 1 |
| Rag\_1stInf\_0dpi\_rep1 | hiseq | 3 | 25,362,793 | 18,853,850 | 443 | 0.0023 | 2 |
| C57BL6\_1stInf\_0dpi\_rep1 | hiseq | 3 | 35,731,249 | 25,119,348 | 457 | 0.0018 | 2 |
| C57BL6\_1stInf\_0dpi\_rep2 | hiseq | 3 | 47,085,959 | 34,377,133 | 608 | 0.0018 | 2 |
| Rag\_1stInf\_0dpi\_rep2 | hiseq | 3 | 46,556,156 | 35,233,327 | 676 | 0.0019 | 2 |
| NMRI\_2ndInf\_0dpi\_rep2 | hiseq | 3 | 58,122,244 | 40,794,245 | 3,406 | 0.0083 | 51 |
| NMRI\_2ndInf\_3dpi\_rep1 | hiseq | 3 | 57,934,016 | 40,544,287 | 4,803 | 0.0118 | 95 |
| NMRI\_2ndInf\_5dpi\_rep2 | hiseq | 3 | 63,965,539 | 48,289,181 | 10,941 | 0.0227 | 407 |
| NMRI\_1stInf\_0dpi\_rep1 | GAII | 1 | 82,364,585 | 55,176,243 | 17,954 | 0.0325 | 701 |
| NMRI\_2ndInf\_3dpi\_rep2 | hiseq | 3 | 65,548,826 | 46,171,909 | 29,548 | 0.0640 | 1,580 |
| NMRI\_2ndInf\_7dpi\_rep2 | hiseq | 3 | 67,487,466 | 51,722,265 | 40,091 | 0.0775 | 1,836 |
| Rag\_1stInf\_5dpi\_rep1 | hiseq | 3 | 38,651,359 | 29,982,453 | 63,024 | 0.2098 | 2,548 |
| Rag\_1stInf\_5dpi\_rep2 | hiseq | 3 | 34,779,832 | 25,297,803 | 99,000 | 0.3898 | 2,828 |
| C57BL6\_1stInf\_5dpi\_rep1 | hiseq | 3 | 40,904,388 | 29,319,604 | 185,969 | 0.6303 | 4,173 |
| Rag\_2ndInf\_5dpi\_rep1 | hiseq | 3 | 50,049,848 | 37,093,621 | 192,856 | 0.5172 | 4,167 |
| C57BL6\_1stInf\_5dpi\_rep2 | hiseq | 3 | 29,511,368 | 18,062,349 | 215,696 | 1.1801 | 3,823 |
| C57BL6\_2ndInf\_5dpi\_rep1 | hiseq | 3 | 35,148,432 | 25,660,184 | 262,909 | 1.0142 | 4,563 |
| NMRI\_1stInf\_3dpi\_rep1 | GAII | 1 | 73,236,430 | 49,993,358 | 394,384 | 0.7827 | 5,220 |
| NMRI\_1stInf\_3dpi\_rep2 | GAII | 2 | 160,709,694 | 117,791,044 | 413,051 | 0.3494 | 4,862 |
| NMRI\_1stInf\_5dpi\_rep2 | GAII | 2 | 119,902,722 | 76,419,774 | 794,570 | 1.0290 | 5,333 |
| NMRI\_2ndInf\_5dpi\_rep1 | GAII | 2 | 230,773,955 | 143,186,486 | 1,846,840 | 1.2734 | 5,533 |
| NMRI\_2ndInf\_7dpi\_rep1 | hiseq | 3 | 70,366,762 | 41,467,146 | 8,634,201 | 17.2335 | 5,875 |
| NMRI\_1stInf\_5dpi\_rep1 | GAII | 2 | 76,702,168 | 47,037,087 | 8,669,701 | 15.5631 | 5,700 |
| NMRI\_sporozoites\_rep2 | GAII | 0 | 19,551,681 | 8,656 | 11,470,604 | 99.9246 | 5,513 |
| NMRI\_1stInf\_5dpi\_rep3 | GAII | 0 | 191,099,180 | 83,735,624 | 27,839,458 | 24.9513 | 5,784 |
| NMRI\_1stInf\_7dpi\_rep1 | GAII | 1 | 66,505,514 | 3,310,666 | 39,400,884 | 92.2488 | 5,932 |
| NMRI\_sporozoites\_rep1 | GAII | 1 | 67,325,397 | 4,334 | 43,774,401 | 99.9901 | 5,825 |
| NMRI\_oocysts\_rep1 | GAII | 1 | 68,859,802 | 3,805 | 49,653,065 | 99.9923 | 5,695 |
| NMRI\_oocysts\_rep2 | GAII | 0 | 151,090,783 | 18,524 | 71,019,860 | 99.9739 | 5,777 |
| NMRI\_1stInf\_7dpi\_rep2 | GAII | 1 | 139,749,046 | 21,699,324 | 73,539,445 | 77.2159 | 5,943 |

\* Sample names are given with information separated by underscore as follows: 1) mouse strain, 2) first or second infection, 3) day post infection (days post infection ), and 4) replicate number.  
\*\* Percentage of *E. falciformis* reads is provided as percentage of total (host plus parasite) reads.

**Table 2** Number of mRNAs significantly differentially abundant in contrasts listed for *E. falciformis* and mouse as indicated. Empty cells indicate that comparison is not applicable to that species. “1st“ and “2nd“ indicates infection number.

|  |  |  |
| --- | --- | --- |
| **Sample(s) comparison** | ***E. falciformis*, FDR < 0.01** | **Mouse, FDR < 0.01** |
| NMRI 7dpi vs ctrl |  | 2,711 |
| NMRI 5dpi vs ctrl |  | 1,804 |
| NMRI 3 dpi vs 7 dpi | 1,399 | 1,322 |
| BL6 5 dpi vs ctrl |  | 919 |
| NMRI 7 dpi 1st vs NMRI 7 dpi 2nd | 0 | 857 |
| NMRI 5 dpi vs NMRI 7 dpi | 2,084 | 732 |
| *Rag1*-/- vs BL6 |  | 362 |
| NMRI 3 dpi vs ctrl |  | 325 |
| BL6 5 dpi 1st vs BL6 5 dpi 2nd | 0 | 175 |
| *Rag1*-/- 5 dpi vs ctrl |  | 42 |
| NMRI 3 dpi 1st vs NMRI 3 2nd | 1 | 18 |
| NMRI 3 dpi vs NMRI 5 dpi | 103 | 0 |
| NMRI 5 dpi vs oocysts | 3,691 |  |
| Sporozoites vs oocysts | 3,532 |  |
| NMRI 3 dpi vs oocysts | 3,303 |  |
| NMRI 7 dpi vs oocysts | 3,202 |  |
| NMRI 7 dpi vs sporozoites | 2,663 |  |
| NMRI 5 dpi vs sporozoites | 1,726 |  |
| NMRI 3 dpi vs sporozoites | 1,705 |  |
| NMRI vs BL6 | 13 |  |

**Table 3** Enrichments and underrepresentation ofspecies or species group orthologues in *E. falciformis* gene clusters (Figure 3b).Odds ratios higher than one indicate enrichment and smaller than one indicate underrepresentation. Conservation categories were chosen as previously described (Heitlinger et al. 2014).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***E. falciformis* cluster** | **Conservation category** | **Odds ratio** | **p-value** | **FDR** |
| Cluster 2 | Conserved | 0.67 | 9.03E-06 | 1.90E-04 |
| Cluster 4 | Conserved | 0.72 | 2.44E-04 | 1.71E-03 |
| Cluster 7 | Conserved | 1.72 | 1.11E-10 | 4.65E-09 |
| Cluster 2 | ApicomplexaC | 0.45 | 1.84E-04 | 1.71E-03 |
| Cluster 5 | ApicomplexaC | 1.86 | 3.76E-05 | 5.26E-04 |
| Cluster 4 | *E. falciformis* | 3.05 | 2.38E-04 | 1.71E-03 |
| Cluster 1 | *Eimeria* | 0.68 | 1.83E-03 | 9.59E-03 |
| Cluster 6 | Apicomplexa | 1.46 | 1.11E-03 | 6.64E-03 |

**Table 4** Enrichments of candidate interacting genes in mRNA clusters from figures 2b and 3b for mouse and *E. falciformis*. Odds ratios > 1 indicate overrepresentation and < 1 indicates underrepresentation. Number of genes in clusters are listed (Genes per cluster) as well as p values and corrected p-values (FDR) by FET.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Genes per cluster** | **% interacting** | **Odds ratio** | **P value** | **Cluster** | **FDR** |
| 206 | 38.83 | 1.72 | 2.59E-04 | Mouse 1 | 3.62E-04 |
| 179 | 13.97 | 0.43 | 2.92E-05 | Mouse 2 | 1.02E-04 |
| 255 | 16.86 | 0.54 | 1.16E-04 | Mouse 3 | 2.03E-04 |
| 295 | 21.36 | 0.72 | 2.43E-02 | Mouse 4 | 2.43E-02 |
| 220 | 58.64 | 3.91 | 5.77E-23 | Mouse 5 | 4.04E-22 |
| 276 | 17.03 | 0.55 | 8.92E-05 | Mouse 6 | 2.03E-04 |
| 106 | 13.21 | 0.41 | 8.91E-04 | Mouse 7 | 1.04E-03 |
| 566 | 19.43 | 0.84 | 1.10E-01 | *E.falciformis* 1 | 1.64E-01 |
| 748 | 31.68 | 1.77 | 8.90E-11 | *E.falciformis* 2 | 6.23E-10 |
| 210 | 33.33 | 1.80 | 1.31E-04 | *E.falciformis* 3 | 3.07E-04 |
| 731 | 13.68 | 0.52 | 9.61E-10 | *E.falciformis* 4 | 3.36E-09 |
| 779 | 23.11 | 1.07 | 4.88E-01 | *E.falciformis* 5 | 4.88E-01 |
| 1,344 | 20.54 | 0.89 | 1.17E-01 | *E.falciformis* 6 | 1.64E-01 |
| 691 | 20.41 | 0.89 | 2.62E-01 | *E.falciformis* 7 | 3.06E-01 |

**Figures**

Figure 1. Oocyst output and changes in intensity of *E. falciformis* infection in mouse. A) Oocyst counts in first and challenge infection are shown for different mouse strains. Immune competent mice (NMRI and C57BL/6) have reduced oocyst counts in challenge infection, whereas immune incompetent Rag1-/- mice do not. B) RT-qPCR data of *E. falciformis* 18S in NMRI mice displays an increase in parasite material over the course of infection. Less parasite material is detected in challenge infected mice. C) The percentage of parasite mRNA detected by RNA-seq increases during infection. More material is detected in naïve mice compared to challenge infected mice, as is also demonstrated in B. Sporozoites and oocysts are shown with 100% parasite material.

Figure 2. a) Venn diagram visualizing the overlap between genes showing differential abundance (FDR < 0.01; EdgeR glm likelihood-ratio tests) between uninfected controls and different time-points after infection and ii) between different time-points and all those genes reacting to infection. b) Hierarchical clustering of expression profiles for these mRNAs performed on Euklidean distances using complete linkage. Cluster cut-offs were set to identify gene-sets with profile interpretable in relation to the parasite lifecycle and between mice of different immune competence.

Figure 3. Differential abundance of *E. falciformis* mRNAs and clustering of abundance profiles. a) Overlap between genes showing differential abundance (FDR < 0.01; EdgeR glm likelihood-ratio tests) between intracellular stages at 3 days post infection, 5 days post infection and 7 days post infection. b) Hierarchical clustering of expression profiles for these differentially abundant mRNAs performed on Euclidean distances using complete linkage. Cluster cut-offs were set to identify gene-sets with profile interpretable in relation to the parasite lifecycle.

Figure 4: Correlations of our mRNA abundance data with orthologues from published RNA-seq transcriptomes from other coccidia: *E. tenella* (Reid et al. 2014, Walker et al. 2014) and *T. gondii* (Hehl et al. 2015) are shown. Correlations coffeicients (Spearman's rho) were clustered using complete linkage. *T. gondii* and *Eimeria* spp. “late infection” samples cluster together. *E. falciformis* early infection samples cluster with *E. tenella* merozoites. *E. falciformis* sporozoites cluster with *E. falciformis* early infection, whereas oocysts cluster with *E. tenella* unsporulated oocysts.

Figure 5: Pairwise interaction of host and parasite genes deduced from correlated expressions. a) The distributions of ISIGM scores (A. J. Reid and Berriman 2013) used to infer potential interaction between host-parasite genes pairs are shown for general genes in our study, and for those with significant differences in abundance in tests between different experimental condidions (EdgeR glmLRT). Lower values represent more significant interactions. b) Distribution of the number of “candidate interactions” detected when using a threshold of zero for mouse genes and parasite genes and c) a visualization of interaction links in a network. d) ISIGEM interaction scores for combinations of host and parasite mRNA abundance clusters. An overabundance of lower scores can be observed for corresponding host-parasite clusters of mRNAs abundant at 7 days post infection.

**ADDITIONAL FILES**

**Additional file 1**: Raw and normalized counts

Raw counts of read mappings to the *E. falciformis* and mouse genome for individual samples in our study. Normalized counts forseparately for the host and parasite mappings (three compressed csv files).

**Additional file 2:** Results of statistical tests (edgeR) -Focal contrast, fold-changes, likelihood ratio in/excluding this difference in models, p-values , and false discovery rates (adjusted p-values) are given for all tested contrasts (one compressed csv file).

**Additional file 3:** Additional methods and results

Document containing additional figures and summary tables (pdf).

**Additional file 4**: Results of enrichment analyses (topGO)

**Additional file 5**: Results of enrichment analysis for categories of differently conserved gene orthologues

**Additional file 6**: Result of metabolic pathway enrichment of sporozoite defining genes

Tables listing all tested gene sets and resulting significant GO terms.

Additional file X (figure): Differential abundance of mouse mRNAs and clutering of abundance profiles – Log2 fold-changes obtained from a comparison of RNAseq data at 7days post infection vs. unifected controls are plotted against log2 fold changes obtained in microarray study (Schmid et al. 2014) at 6days post infection vs. similar controls.

REFERENCES

Aarthi, S., Gopal Dhinakar Raj, M. Raman, Damer Blake, Chandra Subramaniam, and Fiona Tomley. 2011. ‘Expressed Sequence Tags from Eimeria Brunetti--Preliminary Analysis and Functional Annotation’. *Parasitology Research* 108 (4): 1059–62. doi:10.1007/s00436-010-2182-6.

Aloisio, Fabio, Giovanni Filippini, Pietro Antenucci, Elvio Lepri, Giovanni Pezzotti, Simone M. Cacciò, and Edoardo Pozio. 2006. ‘Severe Weight Loss in Lambs Infected with Giardia Duodenalis Assemblage B’. *Veterinary Parasitology* 142 (1–2): 154–58. doi:10.1016/j.vetpar.2006.06.023.

Amiruddin, Nadzirah, Xin-Wei Lee, Damer P. Blake, Yutaka Suzuki, Yea-Ling Tay, Lik-Sin Lim, Fiona M. Tomley, Junichi Watanabe, Chihiro Sugimoto, and Kiew-Lian Wan. 2012. ‘Characterisation of Full-Length cDNA Sequences Provides Insights into the Eimeria Tenella Transcriptome’. *BMC Genomics* 13 (January): 21. doi:10.1186/1471-2164-13-21.

Beck, Paul L., Ian M. Rosenberg, Ramnik J. Xavier, Theodore Koh, Josée F. Wong, and Daniel K. Podolsky. 2003. ‘Transforming Growth Factor-Beta Mediates Intestinal Healing and Susceptibility to Injury in Vitro and in Vivo through Epithelial Cells’. *The American Journal of Pathology* 162 (2): 597–608.

Blagburn, B. L., and K. S. Todd. 1984. ‘Pathological Changes and Immunity Associated with Experimental Eimeria Vermiformis Infections in Mus Musculus’. *The Journal of Protozoology* 31 (4): 556–61.

Bradley, Peter J., and L. David Sibley. 2007. ‘Rhoptries: An Arsenal of Secreted Virulence Factors’. *Current Opinion in Microbiology* 10 (6): 582–87. doi:10.1016/j.mib.2007.09.013.

Carruthers, Vern, and John C Boothroyd. 2007. ‘Pulling Together: An Integrated Model of Toxoplasma Cell Invasion’. *Current Opinion in Microbiology*, Host-microbe interactions: bacteria, 10 (1): 83–89. doi:10.1016/j.mib.2006.06.017.

Chapman, H. David, John R. Barta, Damer Blake, Arthur Gruber, Mark Jenkins, Nicholas C. Smith, Xun Suo, and Fiona M. Tomley. 2013. *A Selective Review of Advances in Coccidiosis Research*. 1st ed. Vol. 83. Elsevier Ltd. http://dx.doi.org/10.1016/B978-0-12-407705-8.00002-1.

Chow, Yock-Ping, Kiew-Lian Wan, Damer P. Blake, Fiona Tomley, and Sheila Nathan. 2011. ‘Immunogenic Eimeria Tenella Glycosylphosphatidylinositol-Anchored Surface Antigens (SAGs) Induce Inflammatory Responses in Avian Macrophages’. *PLoS ONE* 6 (9). doi:10.1371/journal.pone.0025233.

Cowman, Alan F., and Brendan S. Crabb. 2006. ‘Invasion of Red Blood Cells by Malaria Parasites’. *Cell* 124 (4): 755–66. doi:10.1016/j.cell.2006.02.006.

Duszynski, Donald W., and Donald W. Duszynski. 2011. ‘Eimeria, Eimeria’. In *eLS, eLS*. John Wiley & Sons, Ltd, John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0001962.pub2/abstract, http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0001962.pub2/abstract.

Fernandes, Maria Cecilia, Laura A. L. Dillon, Ashton Trey Belew, Hector Corrada Bravo, David M. Mosser, and Najib M. El-Sayed. 2016. ‘Dual Transcriptome Profiling of Leishmania-Infected Human Macrophages Reveals Distinct Reprogramming Signatures’. *mBio* 7 (3): e00027-16. doi:10.1128/mBio.00027-16.

Fetterer, Raymond H., Ryan S. Schwarz, Katarzyna B. Miska, Mark C. Jenkins, Ruth C. Barfield, and Charles Murphy. 2013. ‘Characterization and Localization of an Eimeria-Specific Protein in Eimeria Maxima’. *Parasitology Research* 112 (10): 3401–8. doi:10.1007/s00436-013-3518-9.

Fleckenstein, Martin C., Michael L. Reese, Stephanie Könen-Waisman, John C. Boothroyd, Jonathan C. Howard, and Tobias Steinfeldt. 2012. ‘A Toxoplasma Gondii Pseudokinase Inhibits Host IRG Resistance Proteins’. *PLOS Biology* 10 (7): e1001358. doi:10.1371/journal.pbio.1001358.

Foth, Bernardo J., Isheng J. Tsai, Adam J. Reid, Allison J. Bancroft, Sarah Nichol, Alan Tracey, Nancy Holroyd, et al. 2014. ‘Whipworm Genome and Dual-Species Transcriptome Analyses Provide Molecular Insights into an Intimate Host-Parasite Interaction’. *Nature Genetics* 46 (7): 693–700. doi:10.1038/ng.3010.

Fox, Barbara A., Leah M. Rommereim, Rebekah B. Guevara, Alejandra Falla, Miryam Andrea Hortua Triana, Yanbo Sun, and David J. Bzik. 2016. ‘The Toxoplasma Gondii Rhoptry Kinome Is Essential for Chronic Infection’. *mBio* 7 (3): e00193-16. doi:10.1128/mBio.00193-16.

Frazee, Alyssa C, Geo Pertea, Andrew E Jaffe, Ben Langmead, Steven L Salzberg, and Jeffrey T Leek. 2015. ‘Ballgown Bridges the Gap between Transcriptome Assembly and Expression Analysis’. *Nature Biotechnology* 33 (3): 243–246. doi:10.1038/nbt.3172.

Gadde, U., H. D. Chapman, T. R. Rathinam, and G. F. Erf. 2009. ‘Acquisition of Immunity to the Protozoan Parasite Eimeria Adenoeides in Turkey Poults and the Peripheral Blood Leukocyte Response to a Primary Infection’. *Poultry Science* 88 (11): 2346–52. doi:10.3382/ps.2009-00320.

Gajria, B., A. Bahl, J. Brestelli, J. Dommer, S. Fischer, X. Gao, M. Heiges, et al. 2007. ‘ToxoDB: An Integrated Toxoplasma Gondii Database Resource’. *Nucleic Acids Research* 36 (Database): D553–56. doi:10.1093/nar/gkm981.

Gallinetti, Jordan, Eylul Harputlugil, and James R Mitchell. 2013. ‘Amino Acid Sensing in Dietary-Restriction-Mediated Longevity: Roles of Signal-Transducing Kinases GCN2 and TOR.’ *The Biochemical Journal* 449 (1): 1–10. doi:10.1042/BJ20121098.

Grimwood, Jane, and Judith E. Smith. 1996. ‘Toxoplasma Gondii: The Role of Parasite Surface and Secreted Proteins in Host Cell Invasion’. *International Journal for Parasitology* 26 (2): 169–73. doi:10.1016/0020-7519(95)00103-4.

Haberkorn, A. 1970. ‘Die Entwicklung vonEimeria falciformis (Eimer 1870) in der weißen Maus (Mus musculus)’. *Zeitschrift für Parasitenkunde* 34 (1): 49–67. doi:10.1007/BF00629179.

Han, Hong Y., Jiao J. Lin, Qi P. Zhao, Hui Dong, Lian L. Jiang, Mei Q. Xu, Shun H. Zhu, and Bing Huang. 2009. ‘Identification of Differentially Expressed Genes in Early Stages of Eimeria Tenella by Suppression Subtractive Hybridization and cDNA Microarray’. *Journal of Parasitology* 96 (1): 95–102. doi:10.1645/GE-2221.1.

Hartmann, Katrin, Diane Addie, Sándor Belák, Corine Boucraut-Baralon, Herman Egberink, Tadeusz Frymus, Tim Gruffydd-Jones, et al. 2013. ‘Toxoplasma Gondii Infection in Cats ABCD Guidelines on Prevention and Management’. *Journal of Feline Medicine and Surgery* 15 (7): 631–37. doi:10.1177/1098612X13489228.

Hehl, Adrian B., Walter U. Basso, Christoph Lippuner, Chandra Ramakrishnan, Michal Okoniewski, Robert A. Walker, Michael E. Grigg, Nicholas C. Smith, and Peter Deplazes. 2015. ‘Asexual Expansion of Toxoplasma Gondii Merozoites Is Distinct from Tachyzoites and Entails Expression of Non-Overlapping Gene Families to Attach, Invade, and Replicate within Feline Enterocytes’. *BMC Genomics* 16: 66. doi:10.1186/s12864-015-1225-x.

Heitlinger, Emanuel, Simone Spork, Richard Lucius, and Christoph Dieterich. 2014. ‘The Genome of Eimeria Falciformis - Reduction and Specialization in a Single Host Apicomplexan Parasite.’ *BMC Genomics* 15 (1): 696. doi:10.1186/1471-2164-15-696.

Hunn, Julia P., Carl G. Feng, Alan Sher, and Jonathan C. Howard. 2011. ‘The Immunity-Related GTPases in Mammals: A Fast-Evolving Cell-Autonomous Resistance System against Intracellular Pathogens’. *Mammalian Genome : Official Journal of the International Mammalian Genome Society* 22 (1–2): 43–54. doi:10.1007/s00335-010-9293-3.

Jones, Nathaniel G., Qiuling Wang, and L. David Sibley. 2016. ‘Secreted Protein Kinases Regulate Cyst Burden during Chronic Toxoplasmosis’. *Cellular Microbiology*, January, n/a-n/a. doi:10.1111/cmi.12651.

Kaiser, G. C., and D. B. Polk. 1997. ‘Tumor Necrosis Factor Alpha Regulates Proliferation in a Mouse Intestinal Cell Line’. *Gastroenterology* 112 (4): 1231–40.

Khan, Asis, Ulrike Böhme, Krystyna A. Kelly, Ellen Adlem, Karen Brooks, Mark Simmonds, Karen Mungall, et al. 2006. ‘Common Inheritance of Chromosome Ia Associated with Clonal Expansion of Toxoplasma Gondii’. *Genome Research* 16 (9): 1119–25. doi:10.1101/gr.5318106.

Kuhn, Kristine A., Nicholas A. Manieri, Ta-Chiang Liu, and Thaddeus S. Stappenbeck. 2014. ‘IL-6 Stimulates Intestinal Epithelial Proliferation and Repair after Injury’. *PloS One* 9 (12): e114195. doi:10.1371/journal.pone.0114195.

Langmead, Ben, and Steven L Salzberg. 2012. ‘Fast Gapped-Read Alignment with Bowtie 2’. *Nature Methods* 9 (4): 357–59. doi:10.1038/nmeth.1923.

Laplante, Mathieu, and David M Sabatini. 2013. ‘mTOR Signaling in Growth Control and Disease’. *Cell* 149 (2): 274–293. doi:10.1016/j.cell.2012.03.017.mTOR.

Li, Yuan, Sheena Shah-Simpson, Kwame Okrah, A. Trey Belew, Jungmin Choi, Kacey L. Caradonna, Prasad Padmanabhan, et al. 2016. ‘Transcriptome Remodeling in Trypanosoma Cruzi and Human Cells during Intracellular Infection’. *PLOS Pathog* 12 (4): e1005511. doi:10.1371/journal.ppat.1005511.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. ‘featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features’. *Bioinformatics* 30 (7): 923–30. doi:10.1093/bioinformatics/btt656.

Lovegrove, Fiona E., Lourdes Peña-Castillo, Naveed Mohammad, W. Conrad Liles, Timothy R. Hughes, and Kevin C. Kain. 2006. ‘Simultaneous Host and Parasite Expression Profiling Identifies Tissue-Specific Transcriptional Programs Associated with Susceptibility or Resistance to Experimental Cerebral Malaria’. *BMC Genomics* 7: 295. doi:10.1186/1471-2164-7-295.

Macian, Fernando. 2005. ‘NFAT Proteins: Key Regulators of T-Cell Development and Function’. *Nature Reviews. Immunology* 5 (6): 472–84. doi:10.1038/nri1632.

MacManes, Matthew D. 2014. ‘On the Optimal Trimming of High-Throughput mRNA Sequence Data’. *Frontiers in Genetics* 5. doi:10.3389/fgene.2014.00013.

Matsubayashi, Makoto, Fumiya Kawahara, Takeshi Hatta, Junya Yamagishi, Takeharu Miyoshi, Anisuzzaman, Kazumi Sasai, Takashi Isobe, Kiyoshi Kita, and Naotoshi Tsuji. 2016. ‘Transcriptional Profiles of Virulent and Precocious Strains of Eimeria Tenella at Sporozoite Stage; Novel Biological Insight into Attenuated Asexual Development’. *Infection, Genetics and Evolution* 40 (June): 54–62. doi:10.1016/j.meegid.2016.02.021.

McCabe, Laura, Robert A. Britton, and Narayanan Parameswaran. 2015. ‘Prebiotic and Probiotic Regulation of Bone Health: Role of the Intestine and Its Microbiome’. *Current Osteoporosis Reports* 13 (6): 363–71. doi:10.1007/s11914-015-0292-x.

Mesfin, G. M., and J. E. Bellamy. 1978. ‘The Life Cycle of Eimeria Falciformis Var. Pragensis (Sporozoa: Coccidia) in the Mouse, Mus Musculus’. *The Journal of Parasitology* 64 (4): 696–705.

Mesfin, G. M., and J. E. C. Bellamy. 1979. ‘Effects of Acquired Resistance on Infection with Eimeria Falciformis Var. Pragensis in Mice’. *Infection and Immunity* 23 (1): 108–14.

Mineo, J. R., and L. H. Kasper. 1994. ‘Attachment of Toxoplasma Gondii to Host Cells Involves Major Surface Protein, SAG-1 (P-30)’. *Experimental Parasitology* 79 (1): 11–20. doi:10.1006/expr.1994.1054.

Novaes, Jeniffer, Luiz Thibério L. D. Rangel, Milene Ferro, Ricardo Y. Abe, Alessandra P. S. Manha, Joana C. M. de Mello, Leonardo Varuzza, Alan M. Durham, Alda Maria B. N. Madeira, and Arthur Gruber. 2012. ‘A Comparative Transcriptome Analysis Reveals Expression Profiles Conserved across Three Eimeria Spp. of Domestic Fowl and Associated with Multiple Developmental Stages’. *International Journal for Parasitology* 42 (1): 39–48. doi:10.1016/j.ijpara.2011.10.008.

Oakes, Richard D., Dominic Kurian, Elizabeth Bromley, Chris Ward, Kalpana Lal, Damer P. Blake, Adam J. Reid, et al. 2013. ‘The Rhoptry Proteome of Eimeria Tenella Sporozoites’. *International Journal for Parasitology*, International Meeting on Apicomplexan Parasites in Farm Animals (ApiCOWplexa) 2012, 43 (2): 181–88. doi:10.1016/j.ijpara.2012.10.024.

Oppenheim, Rebecca D., Darren J. Creek, James I. Macrae, Katarzyna K. Modrzynska, Paco Pino, Julien Limenitakis, Valerie Polonais, et al. 2014. ‘BCKDH: The Missing Link in Apicomplexan Mitochondrial Metabolism Is Required for Full Virulence of Toxoplasma Gondii and Plasmodium Berghei’. *PLOS Pathogens* 10 (7): e1004263. doi:10.1371/journal.ppat.1004263.

Ovington, K. S., L. M. Alleva, and E. a. Kerr. 1995. ‘Cytokines and Immunological Control of Eimeria Spp.’ *International Journal for Parasitology* 25 (11): 1331–1351. doi:10.1016/0020-7519(95)00069-E.

Park, Heon, Zhaoxia Li, Xuexian O Yang, Seon Hee Chang, Roza Nurieva, Yi-Hong Wang, Ying Wang, et al. 2005. ‘A Distinct Lineage of CD4 T Cells Regulates Tissue Inflammation by Producing Interleukin 17’. *Nature Immunology* 6 (11): 1133–41. doi:10.1038/ni1261.

Periz, Javier, Andrew C. Gill, Lawrence Hunt, Philip Brown, and Fiona M. Tomley. 2007. ‘The Microneme Proteins EtMIC4 and EtMIC5 of Eimeria Tenella Form a Novel, Ultra-High Molecular Mass Protein Complex That Binds Target Host Cells’. *Journal of Biological Chemistry* 282 (23): 16891–98. doi:10.1074/jbc.M702407200.

Preston-Mafham, R. A., and A. H. Sykes. 1970. ‘Changes in Body Weight and Intestinal Absorption during Infections with Eimeria Acervulina in the Chicken’. *Parasitology* 61 (03): 417. doi:10.1017/S0031182000041263.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna, Austria. http://www.R-project.org.

Reid, A. J., and Matthew Berriman. 2013. ‘Genes Involved in Host–parasite Interactions Can Be Revealed by Their Correlated Expression’. *Nucleic Acids Research* 41 (3): 1508–18. doi:10.1093/nar/gks1340.

Reid, Adam J., Damer P. Blake, Hifzur R. Ansari, Karen Billington, Hilary P. Browne, Josephine Bryant, Matt Dunn, et al. 2014. ‘Genomic Analysis of the Causative Agents of Coccidiosis in Domestic Chickens’. *Genome Research* 24 (10): 1676–85. doi:10.1101/gr.168955.113.

Reid, Adam James, and Matthew Berriman. 2012. ‘Genes Involved in Host–parasite Interactions Can Be Revealed by Their Correlated Expression’. *Nucleic Acids Research*, December, gks1340. doi:10.1093/nar/gks1340.

Rick, B., J. F. Dubremetz, and R. Entzeroth. 1998. ‘A Merozoite-Specific 22-kDa Rhoptry Protein of the Coccidium Eimeria Nieschulzi (Sporozoa, Coccidia) Is Exocytosed in the Parasitophorous Vacuole upon Host Cell Invasion’. *Parasitology Research* 84 (4): 291–96.

Rosani, U., L. Varotto, S. Domeneghetti, G. Arcangeli, A. Pallavicini, and P. Venier. 2015. ‘Dual Analysis of Host and Pathogen Transcriptomes in Ostreid Herpesvirus 1-Positive Crassostrea Gigas’. *Environmental Microbiology* 17 (11): 4200–4212. doi:10.1111/1462-2920.12706.

Rose, M. E., P. Hesketh, and D. Wakelin. 1992. ‘Immune Control of Murine Coccidiosis: CD4+ and CD8+ T Lymphocytes Contribute Differentially in Resistance to Primary and Secondary Infections’. *Parasitology* 105 ( Pt 3) (December): 349–54.

Rose, M. Elaine. 1974. ‘Immune Responses in Infections with Coccidia: Macrophage Activity’. *Infection and Immunity* 10 (4): 862–71.

Saeij, J. P. J., S. Coller, J. P. Boyle, M. E. Jerome, M. W. White, and J. C. Boothroyd. 2007. ‘Toxoplasma Co-Opts Host Gene Expression by Injection of a Polymorphic Kinase Homologue’. *Nature* 445 (7125): 324–27. doi:10.1038/nature05395.

Schmatz, D M. 1997. ‘The Mannitol Cycle in Eimeria.’ *Parasitology* 114 Suppl: S81–S89.

Schmid, Manuela. 2014. ‘Eimeria Falciformis Infection of Mouse Cells Identifies Host Determinants of Parasite Development’. Humboldt-Universität zu Berlin.

Schmid, Manuela, Emanuel Heitlinger, Simone Spork, Hans-Joachim Mollenkopf, Richard Lucius, and Nishith Gupta. 2014. ‘Eimeria Falciformis Infection of the Mouse Caecum Identifies Opposing Roles of IFNγ-Regulated Host Pathways for the Parasite Development’. *Mucosal Immunology* 7 (4): 969–82. doi:10.1038/mi.2013.115.

Schmid, Manuela, Maik J. Lehmann, Richard Lucius, and Nishith Gupta. 2012. ‘Apicomplexan Parasite, Eimeria Falciformis, Co-Opts Host Tryptophan Catabolism for Life Cycle Progression in Mouse’. *Journal of Biological Chemistry* 287 (24): 20197–20207. doi:10.1074/jbc.M112.351999.

Schwarz, Ryan S, Raymond H Fetterer, George H Rosenberg, and Katarzyna B Miska. 2010. ‘Coccidian Merozoite Transcriptome Analysis from Eimeria Maxima in Comparison to Eimeria Tenella and Eimeria Acervulina.’ *The Journal of Parasitology* 96 (1): 49–57. doi:10.1645/GE-2253.1.

Sharman, P. A., N. C. Smith, M. G. Wallach, and M. Katrib. 2010. ‘Chasing the Golden Egg: Vaccination Against Poultry Coccidiosis’. *Parasite Immunology*, February. doi:10.1111/j.1365-3024.2010.01209.x.

Sidik, Saima M., Diego Huet, Suresh M. Ganesan, My-Hang Huynh, Tim Wang, Armiyaw S. Nasamu, Prathapan Thiru, et al. 2016. ‘A Genome-Wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes’. *Cell* 166 (6): 1423–1435.e12. doi:10.1016/j.cell.2016.08.019.

Singer, Steven M., and Theodore E. Nash. 2000. ‘The Role of Normal Flora in Giardia Lamblia Infections in Mice’. *Journal of Infectious Diseases* 181 (4): 1510–12. doi:10.1086/315409.

Stange, Jörg. 2012. ‘Studies on Host-Pathogen Interactions at Mucosal Barrier Surfaces Using the Murine Intestinal Parasite Eimeria Falciformis - Deutsche Digitale Bibliothek’. http://www.deutsche-digitale-bibliothek.de/item/DDAKP5LJSJBAPPALDVQ5Y52YV3AG7NCL.

Stange, Jörg, Matthew R Hepworth, Sebastian Rausch, Lara Zajic, Anja A Kühl, Catherine Uyttenhove, Jean-Christophe Renauld, Susanne Hartmann, and Richard Lucius. 2012. ‘IL-22 Mediates Host Defense against an Intestinal Intracellular Parasite in the Absence of IFN-γ at the Cost of Th17-Driven Immunopathology.’ *Journal of Immunology (Baltimore, Md. : 1950)* 188 (5): 2410–8. doi:10.4049/jimmunol.1102062.

Stephenson, L. S., M. C. Latham, and E. A. Ottesen. 2000. ‘Malnutrition and Parasitic Helminth Infections’. *Parasitology* 121 (S1): S23–38. doi:10.1017/S0031182000006491.

Sühwold, Anke, Carlos Hermosilla, Torsten Seeger, Horst Zahner, and Anja Taubert. 2010. ‘T Cell Reactions of Eimeria Bovis Primary- and Challenge-Infected Calves’. *Parasitology Research* 106 (3): 595–605. doi:10.1007/s00436-009-1705-5.

Suzuki, Atsushi, Sayaka Sekiya, Eriko Gunshima, Setsuko Fujii, and Hideki Taniguchi. 2010. ‘EGF Signaling Activates Proliferation and Blocks Apoptosis of Mouse and Human Intestinal Stem/Progenitor Cells in Long-Term Monolayer Cell Culture’. *Laboratory Investigation; a Journal of Technical Methods and Pathology* 90 (10): 1425–36. doi:10.1038/labinvest.2010.150.

Talevich, Eric, and Natarajan Kannan. 2013. ‘Structural and Evolutionary Adaptation of Rhoptry Kinases and Pseudokinases, a Family of Coccidian Virulence Factors’. *BMC Evolutionary Biology* 13: 117. doi:10.1186/1471-2148-13-117.

Taylor, S., A. Barragan, C. Su, B. Fux, S. J. Fentress, K. Tang, W. L. Beatty, et al. 2006. ‘A Secreted Serine-Threonine Kinase Determines Virulence in the Eukaryotic Pathogen Toxoplasma Gondii’. *Science* 314 (5806): 1776–80. doi:10.1126/science.1133643.

Templeton, Thomas J. 2007. ‘Whole-Genome Natural Histories of Apicomplexan Surface Proteins’. *Trends in Parasitology* 23 (5): 205–12. doi:10.1016/j.pt.2007.03.001.

The UniProt Consortium. 2015. ‘UniProt: A Hub for Protein Information’. *Nucleic Acids Research* 43 (D1): D204–12. doi:10.1093/nar/gku989.

Tomley, Fiona M., Karen J. Billington, Janene M. Bumstead, Julie D. Clark, and Paul Monaghan. 2001. ‘EtMIC4: A Microneme Protein from Eimeria Tenella That Contains Tandem Arrays of Epidermal Growth Factor-like Repeats and Thrombospondin Type-I Repeats’. *International Journal for Parasitology*, Life in Vacuoles-A strategy for Parasite Survival, 31 (12): 1303–10. doi:10.1016/S0020-7519(01)00255-7.

Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. ‘TopHat: Discovering Splice Junctions with RNA-Seq’. *Bioinformatics* 25 (9): 1105–11. doi:10.1093/bioinformatics/btp120.

VanDussen, Kelli L., Alexis J. Carulli, Theresa M. Keeley, Sanjeevkumar R. Patel, Brent J. Puthoff, Scott T. Magness, Ivy T. Tran, et al. 2012. ‘Notch Signaling Modulates Proliferation and Differentiation of Intestinal Crypt Base Columnar Stem Cells’. *Development (Cambridge, England)* 139 (3): 488–97. doi:10.1242/dev.070763.

VanWormer, Elizabeth, Heather Fritz, Karen Shapiro, Jonna A. K. Mazet, and Patricia A. Conrad. 2013. ‘Molecules to Modeling: Toxoplasma Gondii Oocysts at the Human-Animal-Environment Interface’. *Comparative Immunology, Microbiology and Infectious Diseases* 36 (3): 217–31. doi:10.1016/j.cimid.2012.10.006.

Walker, Robert a, Philippa a Sharman, Catherine M Miller, Christoph Lippuner, Michal Okoniewski, Ramon M Eichenberger, Chandra Ramakrishnan, et al. 2015. ‘RNA Seq Analysis of the Eimeria Tenella Gametocyte Transcriptome Reveals Clues about the Molecular Basis for Sexual Reproduction and Oocyst Biogenesis’. *BMC Genomics* 16 (1): 1–20. doi:10.1186/s12864-015-1298-6.

Wellen, Kathryn E., and Craig B. Thompson. 2010. ‘Cellular Metabolic Stress: Considering How Cells Respond to Nutrient Excess’. *Molecular Cell* 40 (2): 323–32. doi:10.1016/j.molcel.2010.10.004.

Westermann, Alexander J., Konrad U. Förstner, Fabian Amman, Lars Barquist, Yanjie Chao, Leon N. Schulte, Lydia Müller, Richard Reinhardt, Peter F. Stadler, and Jörg Vogel. 2016. ‘Dual RNA-Seq Unveils Noncoding RNA Functions in Host–pathogen Interactions’. *Nature* 529 (7587): 496–501. doi:10.1038/nature16547.

Xu, Xiaojin, Yuemei Dong, Eappen G. Abraham, Anna Kocan, Prakash Srinivasan, Anil K. Ghosh, Robert E. Sinden, et al. 2005. ‘Transcriptome Analysis of Anopheles Stephensi-Plasmodium Berghei Interactions’. *Molecular and Biochemical Parasitology* 142 (1): 76–87. doi:10.1016/j.molbiopara.2005.02.013.