

RESEARCH ARTICLE

Genomic insights into host and parasite interactions during intracellular infection by *Toxoplasma gondii*

Netha Ulahannan^{1,2,✉aa}, Ronald Cutler^{1,✉}, Reanna Doña-Termine^{1,✉}, Claudia A. Simões-Pires^{1,ab}, N. Ari Wijetunga^{1,2,ac}, Matthew McKnight Croken^{1,2,ad}, Andrew D. Johnston^{1,ae}, Yu Kong^{1,af}, Shahina B. Maqbool¹, Masako Suzuki¹, John M. Greally^{1,*}

1 Department of Genetics, Albert Einstein College of Medicine, Bronx, NY, United States of America,

2 Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, NY, United States of America

✉ These authors contributed equally to this work.

aa Current address: Bioreference Laboratories, Elmwood Park, NJ, United States of America

ab Current address: Faculty of Sciences, University of Geneva, Geneva, Switzerland

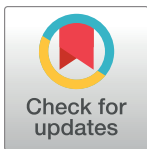
ac Current address: Memorial Sloan Kettering Cancer Center, New York, NY, United States of America

ad Current address: Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, NY, United States of America

ae Current address: Department of Dermatology, Yale School of Medicine, New Haven, CT, United States of America

af Current address: Google Inc., Mountain View, CA, United States of America

* john.greally@einsteinmed.edu



OPEN ACCESS

Citation: Ulahannan N, Cutler R, Doña-Termine R, Simões-Pires CA, Wijetunga NA, Croken MMKnight, et al. (2022) Genomic insights into host and parasite interactions during intracellular infection by *Toxoplasma gondii*. PLoS ONE 17(9): e0275226. <https://doi.org/10.1371/journal.pone.0275226>

Editor: Gregory S. Barsh, HudsonAlpha Institute for Biotechnology, UNITED STATES

Received: July 28, 2022

Accepted: August 2, 2022

Published: September 30, 2022

Copyright: © 2022 Ulahannan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All RNA-seq and ATAC-seq files are available from the GEO database (accession number GSE79612).

Funding: This project was supported by grants from the National Institute of Neurological Disorders and Stroke (T32 NS007098) to Netha Ulahannan, the National Institute of General Medical Sciences (T32 GM007288) to Reanna Doña-Termine, N. Ari Wijetunga and Andrew D. Johnston, and the National Institute on Aging

Abstract

To gain insights into the molecular interactions of an intracellular pathogen and its host cell, we studied the gene expression and chromatin states of human fibroblasts infected with the Apicomplexan parasite *Toxoplasma gondii*. We show a striking activation of host cell genes that regulate a number of cellular processes, some of which are protective of the host cell, others likely to be advantageous to the pathogen. The simultaneous capture of host and parasite genomic information allowed us to gain insights into the regulation of the *T. gondii* genome. We show how chromatin accessibility and transcriptional profiling together permit novel annotation of the parasite's genome, including more accurate mapping of known genes and the identification of new genes and *cis*-regulatory elements. Motif analysis reveals not only the known *T. gondii* AP2 transcription factor-binding site but also a previously-undiscovered candidate TATA box-containing motif at one-quarter of promoters. By inferring the transcription factor and upstream cell signaling responses involved in the host cell, we can use genomic information to gain insights into *T. gondii*'s perturbation of host cell physiology. Our resulting model builds on previously-described human host cell signalling responses to *T. gondii* infection, linked to induction of specific transcription factors, some of which appear to be solely protective of the host cell, others of which appear to be co-opted by the pathogen to enhance its own survival.

(AG057422) to Dr. John M. Grealis. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The infection of a cell by an intracellular pathogen is accompanied by substantial changes in cellular homeostasis, reflected by activation and suppression of cell signaling pathways such as those regulating metabolism, cell division and cell death [1]. Some of these responses are protective for the host, while others represent processes co-opted by the pathogen to enhance its survival and spread. The evolutionary pressures on the host and pathogen are significant, to the point that interactions with pathogens have been proposed to represent the primary selective pressure in human evolution [2]. What we see today in host-pathogen interactions reflects a long history of selection acting on vital cellular processes.

We focus here on human cell infection by *Toxoplasma gondii*. This apicomplexan intracellular parasite is extremely successful, with an estimated one-quarter of humans chronically infected [3], causing symptomatic disease in immunocompromised individuals and following transmission to the developing fetus *in utero*. We now recognize many of the events occurring during *T. gondii* infection of host cells [4], giving insights into why this microorganism is such a successful infectious agent.

The application of genomic assays to cells infected by a eukaryotic pathogen yields rich but complex information. These assays can test transcription and its regulatory control from two organisms simultaneously, with some of the transcriptional and regulatory changes in the genome of the host cell likely to be a protective response to the infection, but other changes potentially advantageous for the pathogen, reflecting the molecular hijacking of host cell physiology. We also obtain a snapshot of the transcription and its regulation in the pathogen's genome. The host loci at which regulatory changes are occurring are highly likely to include those mediating evolutionary selection, including host regulatory loci located distantly from genes in the non-coding majority of the genome, loci that would otherwise be difficult to identify and characterize using sequence information alone.

We have previously characterized transcriptional regulatory information in the *T. gondii* genome, showing that it is free of DNA methylation [5] and that post-translational modifications of *T. gondii* histones define *cis*-regulatory loci [6]. In this study, we perform transcriptional and chromatin profiling genome-wide in human fibroblasts prior to and after infection with *T. gondii*, revealing host responses, pathogen genomic regulation during the tachyzoite stage of *T. gondii* life cycle, and a new annotation of genes and *cis*-regulatory loci in the *T. gondii* genome, yielding insights into candidate transcription factor binding sites and the organization of *T. gondii*'s apicoplast genome. We also demonstrate how we can infer from observations of transcriptional perturbation insights into host cell signaling pathway alterations occurring in response to intracellular infection by this common pathogen.

Results

Host transcriptome responses to *Toxoplasma gondii* infection

We performed RNA sequencing (RNA-seq) of human foreskin fibroblasts (HFFs), before and 24 hours after *in vitro* infection with the *T. gondii* RH strain in the tachyzoite stage of the life cycle (Fig 1A). RNA-seq reads were aligned to a reference genome that combined both the human and *T. gondii* genomes simultaneously, an approach that shows more accurate representations in studies of infected host cells [7]. We showed the human host transcriptome to be characterized by hundreds of differentially expressed genes, with more genes increasing than decreasing their expression levels in infected cells (Fig 1B, S1 Table).

The single most markedly upregulated gene was *ADAMTS15*, a member of the A Disintegrin and Metalloprotease with Thrombospondin Motifs (ADAMTS) gene family. These

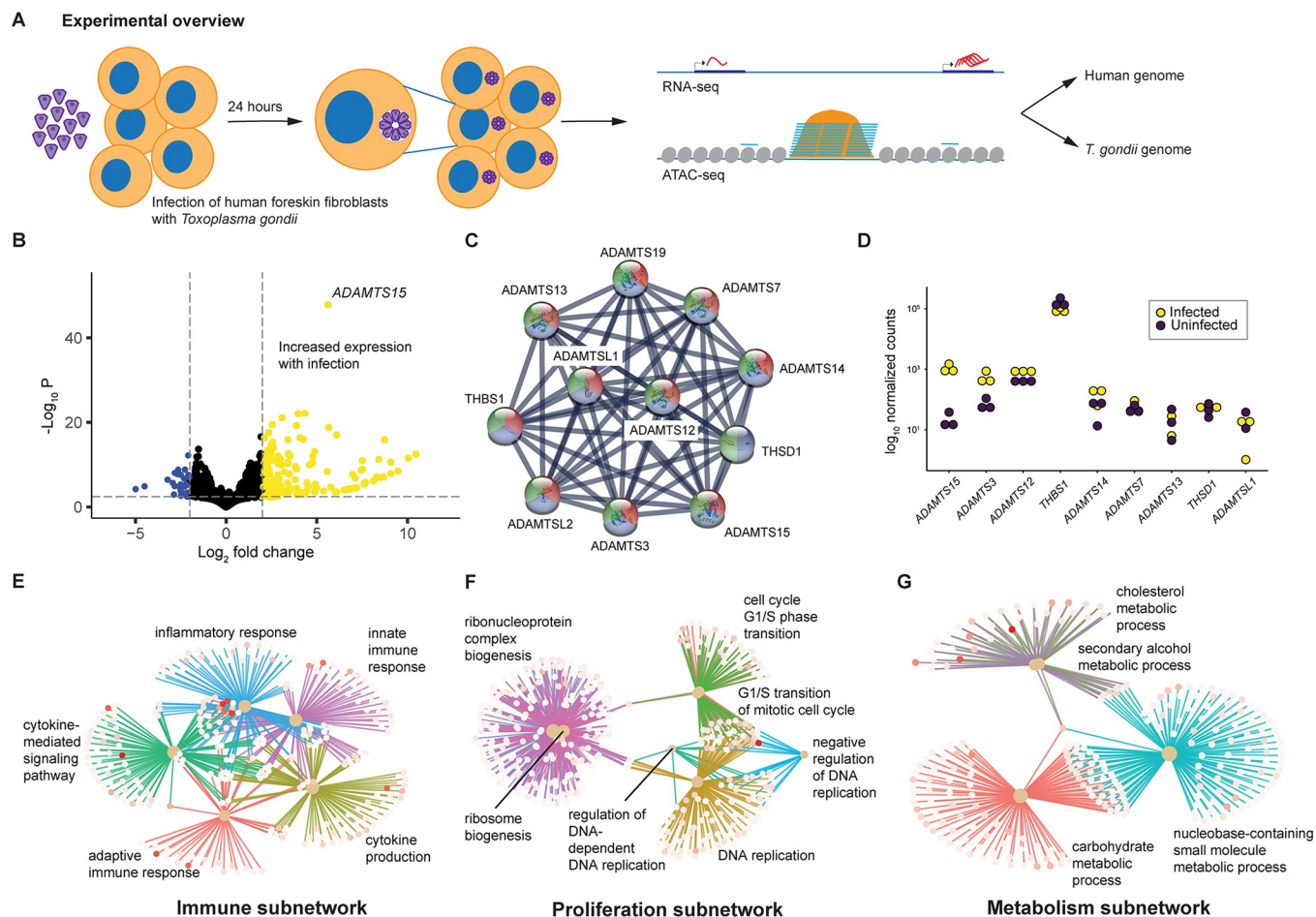


Fig 1. Project overview and host gene transcriptional response to *T. gondii* infection. We show in (A) the overview of the experiments performed, depicting the *in vitro* infection of human fibroblasts with *T. gondii*, with transcriptional (RNA-seq) and chromatin accessibility (ATAC-seq) studies performed before and 24 hours after infection, and alignment of the reads obtained to both the human and *T. gondii* genomes. The RNA-seq results in (B) show more human genes upregulated (yellow) than downregulated (blue) with infection, with ADAMTS15 strongly upregulated. Using network analysis, we extracted genes interacting with ADAMTS15 (C) and showed that other metalloproteases were upregulated with infection (D). (E-G) show the major groups of gene ontologies in the remaining upregulated genes.

<https://doi.org/10.1371/journal.pone.0275226.g001>

metalloproteases function as regulators of extracellular matrix (ECM) composition, potentially enhancing cell migration by removing collagen barriers [8, 9]. By using ADAMTS15 to nucleate a protein-protein interaction analysis (Fig 1C), the RNA-seq data showed four of the nine genes encoding interacting proteins to have increased levels of expression in infected cells (ADAMTS15, ADAMTS3, ADAMTS12, ADAMTS14, Fig 1D). Matrix metalloproteases have been previously found to be overexpressed in response to *T. gondii* infection of astrocytes [10] and dendritic cells [11], potentially permitting the cells to become more mobile, and helping to disseminate the infection [11].

To gain insights into the properties of the broader group of upregulated genes, we performed a gene set enrichment analysis (GSEA), reducing the gene ontology (GO) terms by semantic similarity (S1 Fig, S2 Table). This analysis revealed three main subnetworks, one revealing a host cell immune response, a second the initiation of cell division, and a third reflecting activation of metabolic pathways (Fig 1E–1G). The immune response subnetwork included upregulation of the *NFKB1*, *ORAI1*, and *IL6* genes, indicative of activation of both the innate and adaptive immune responses [12–17]. These were accompanied by activation of

genes related to cytokine production, including *TGFB*, *CXCL6*, *TNFSF15* and *XCL1*, which prompt the recruitment of immune cells to the location of an ongoing infection [12, 18–20]. The cell division subnetwork involved activation of genes relating to ribosome biogenesis and G1/S phase transition, including *PRKDC*, *CASP2*, and *ATR* [21–23]. The third subnetwork is composed of genes involved in metabolic processes, with upregulation of genes such as *CYP27B1*, *DGAT2*, *FASN*, and *PCSK9* [24–27]. These genes are involved in cholesterol and carbohydrate metabolism, nutrients that the auxotrophic *T. gondii* has to acquire from the host cell [28].

Host *cis*-regulatory changes in response to *Toxoplasma gondii* infection

To gain insights into the mechanism for the host transcriptional changes during *T. gondii* infection, we performed the assay for transposase-accessible chromatin (ATAC-seq) [29] in the fibroblasts before and after infection. One infected sample had a pattern of chromatin accessibility more comparable with the uninfected than the other infected samples (S2 Fig), in addition to an average of 28% fewer *T. gondii* reads, and therefore was removed from the comparison with the interpretation that this sample did not achieve a degree of infection comparable to the other replicates. We used DESeq2 [30] to identify the differentially accessible regions (DARs). Loci gaining chromatin accessibility with infection outnumbered those losing accessibility (503 compared with 27, Fig 2A). Of these DARs, only 51 (9.6%) were located within 5 kb of annotated transcription start sites, indicated that chromatin accessibility changes mostly occur at distal regulatory elements (S3 Fig).

To gain insights into the transcription factors (TFs) potentially mediating these changes in chromatin accessibility, we performed motif enrichment analysis using MEME-ChIP [31] on the collection of 100 bp sequences flanking the center of the DARs. The motifs discovered were then compared with known TF binding motifs, revealing enrichment at loci opening chromatin during infection for the AP-1 (JUN-FOS), RELA (an NFκB family member) and RREB1 TFs (Fig 2B). When we performed an analysis of the same DNA sequences using the vertebrate core motif database JASPAR2020, the same motifs were significantly enriched, accompanied by more NFκB family members and a broader group of candidate TFs (S4 Fig and S3 Table).

We asked the question whether these three TFs were distinctively associated with the major groups of differentially expressed genes shown in Fig 1E–1G. To maximize our confidence of an association between a DAR and a gene, we focused on those within 5 kb of an annotated gene transcription start site (TSS), identifying DARs containing only one of these motifs (Fig 2C), and asked whether binding site motifs for each TF was enriched for genes mediating one or more of the cellular properties altered during infection. We show these results in Fig 2D. While all three TF binding motifs were associated with genes mediating immune responses, the RELA (NFκB) motif was enriched at those genes alone, while AP-1 (FOSB::JUNB) motifs were also enriched at genes involved with cell division and metabolic processes, and the motifs for RREB1, which mediates RAS-MAPK signaling [32], are associated with the genes mediating metabolic and extracellular matrix functions. The data could be interpreted to reveal solely host benefits to the NFκB response, but changes that partially benefit the parasite mediated by AP-1 and RREB1.

The tachyzoite transcriptome of *Toxoplasma gondii* during human host cell infection

Our RNA-seq dataset in infected cells included transcripts from the *T. gondii* genome. We could therefore use the ME49 reference genome for *T. gondii* to align RNA-seq reads, detecting

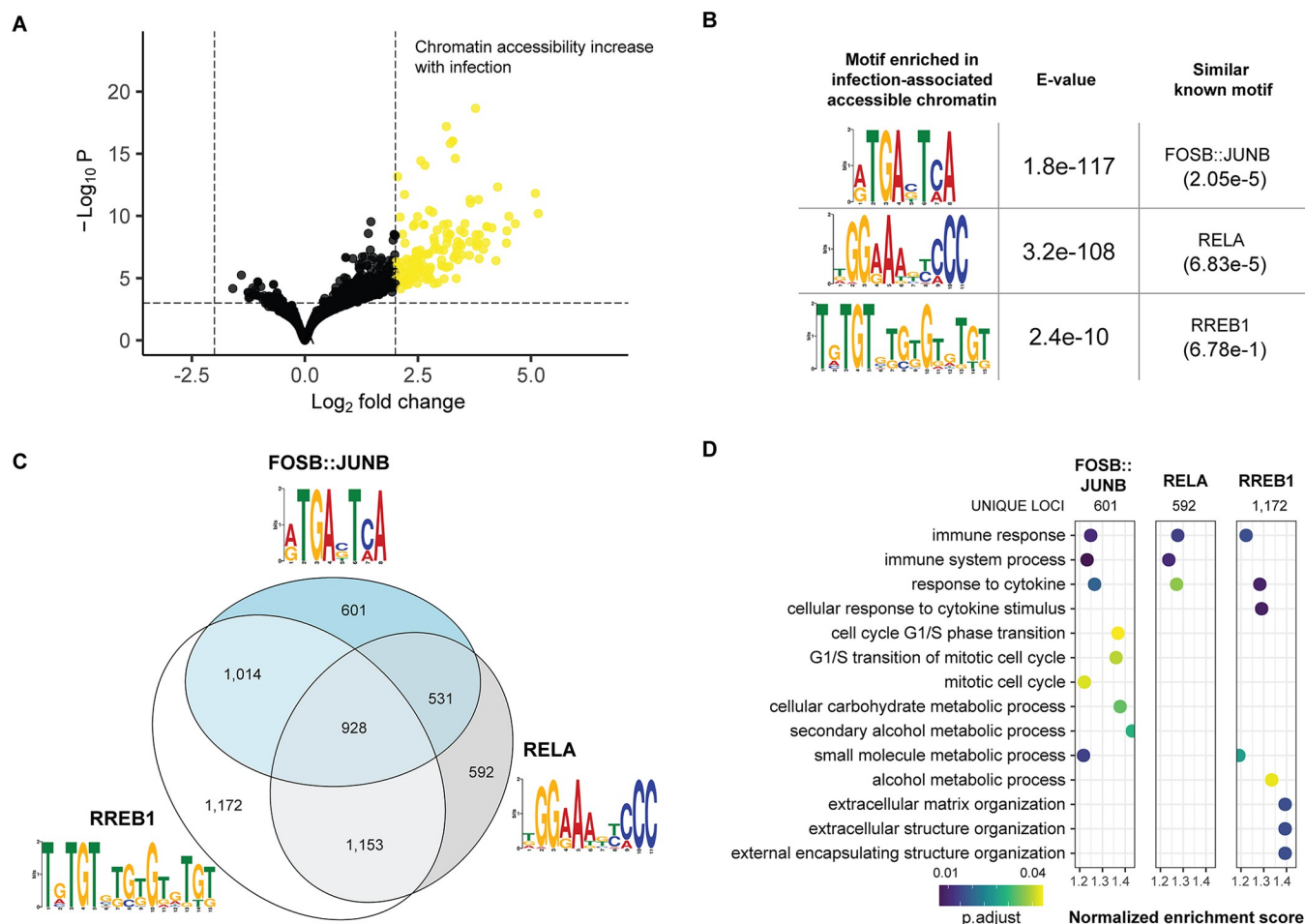


Fig 2. Inferred transcription factors mediating host cell chromatin accessibility changes. In (A) we show the response to *T. gondii* infection in the human genome is marked by increased accessibility of chromatin at numerous loci. The results in (B) describe the motifs enriched in these loci that open their chromatin, corresponding to the binding sites of three known TFs. The results in (C) show the numbers of expressed genes containing one or more of the TF binding site motifs within 5 kb of their transcription start site. For those genes with only one motif, testing their ontological properties in (D) revealed the common property of induction of immune responses, but also other properties that are likely to be more favorable to the pathogen.

<https://doi.org/10.1371/journal.pone.0275226.g002>

3 or more reads in ≥ 2 replicates at 6,653 genes (75% of all ME49-annotated genes), of which 6,191 genes (93%) showed expression in all three replicates (S4 Table). Focusing on this last group of consistently expressed genes, we used hierarchical clustering to identify groups with high (391 genes), medium (4,201 genes), or low (2,061 genes) levels of expression (Fig 3A). Using the gene ontology (GO) annotations of *T. gondii* ME49 genes from ToxoDB (toxodb.org), we found that highly expressed genes are enriched for protein translation processes, medium expression genes for proteosome and catabolic protein processes, and low expression genes for RNA processing and genomic activity (Fig 3B, S5 Table). Overall, our enrichment analysis reveals a hierarchy of biological functions of differentially-expressed genes during the tachyzoite stage of *T. gondii* related to their roles during infection.

Cis-regulatory loci in the tachyzoite *Toxoplasma gondii* genome

Using the ATAC-seq dataset from the infected cells, we were able to profile the loci with accessible chromatin genome-wide in the *T. gondii* genome, using the *T. gondii* ME49 reference genome for alignment, representing a follow up to our prior microarray ChIP study of 650 kb

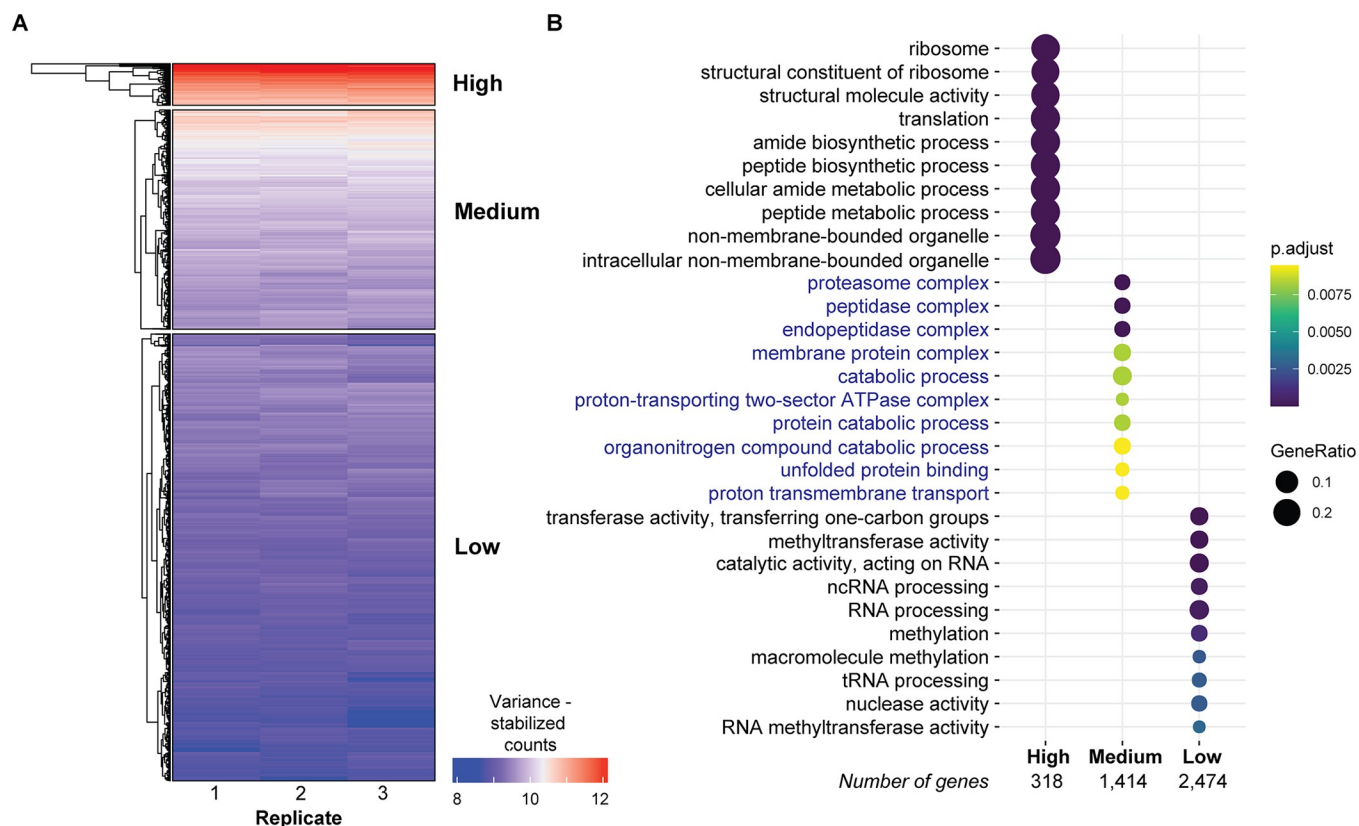


Fig 3. Distinct properties of *T. gondii* genes expressed during tachyzoite stage. The *T. gondii* transcriptome in its tachyzoite stage while infecting human fibroblasts can be classified into three groups of genes based on their expression levels (A). In (B) we show the ontological properties of the genes in each of these three categories.

<https://doi.org/10.1371/journal.pone.0275226.g003>

of the *T. gondii* genome [6], and representing one of only a few genome-wide studies to date of transcriptional regulation in *T. gondii* [33–37].

Within the ATAC-seq sample, a mean of 34.6% of reads across the replicates mapped to the *T. gondii* ME49 genome. The reads exhibited the expected fragment size distribution characteristic of nucleosome free and nucleosomally-organized fragments (S5A Fig), with the exception of reads originating from the apicoplast chromosome, where there is no evidence of nucleosomal organization (S5B Fig).

A total of 4,672 unique peaks were called across all replicates, with the fraction of reads located within peaks (FrIP score [38]) averaging 61.9%, and an average of 94.9% for peak concordance between replicates. Applying ChIP-R [39], we defined 1,631 peaks as reproducible and focused on these for subsequent analyses (Fig 4A). The majority of these peaks (96.1%) were located within 5 kb of an annotated genomic feature, with 91.1% located <1 kb from a transcription start site. When we compared the distribution of ATAC-seq peaks to annotated transcription start sites (TSS), we found a skewing upstream (Fig 4B), reflecting many gene annotations for which the TSS is located substantially downstream from the open chromatin region where the start of the RNA-seq reads is also apparent (examples in Fig 4C).

Some ATAC-seq peaks were not located close to annotated TSS. We explored the representation of RNA-seq reads near this subset of peaks and performed K-nearest neighbors (KNN) clustering to define 5 clusters (Fig 5A). Inspecting these loci, we found examples of what appear to be new, unannotated genes (Fig 5B), unannotated anti-sense transcripts (Fig 5C), the use of alternative TSS within bodies of annotated genes (as previously noted by Markus

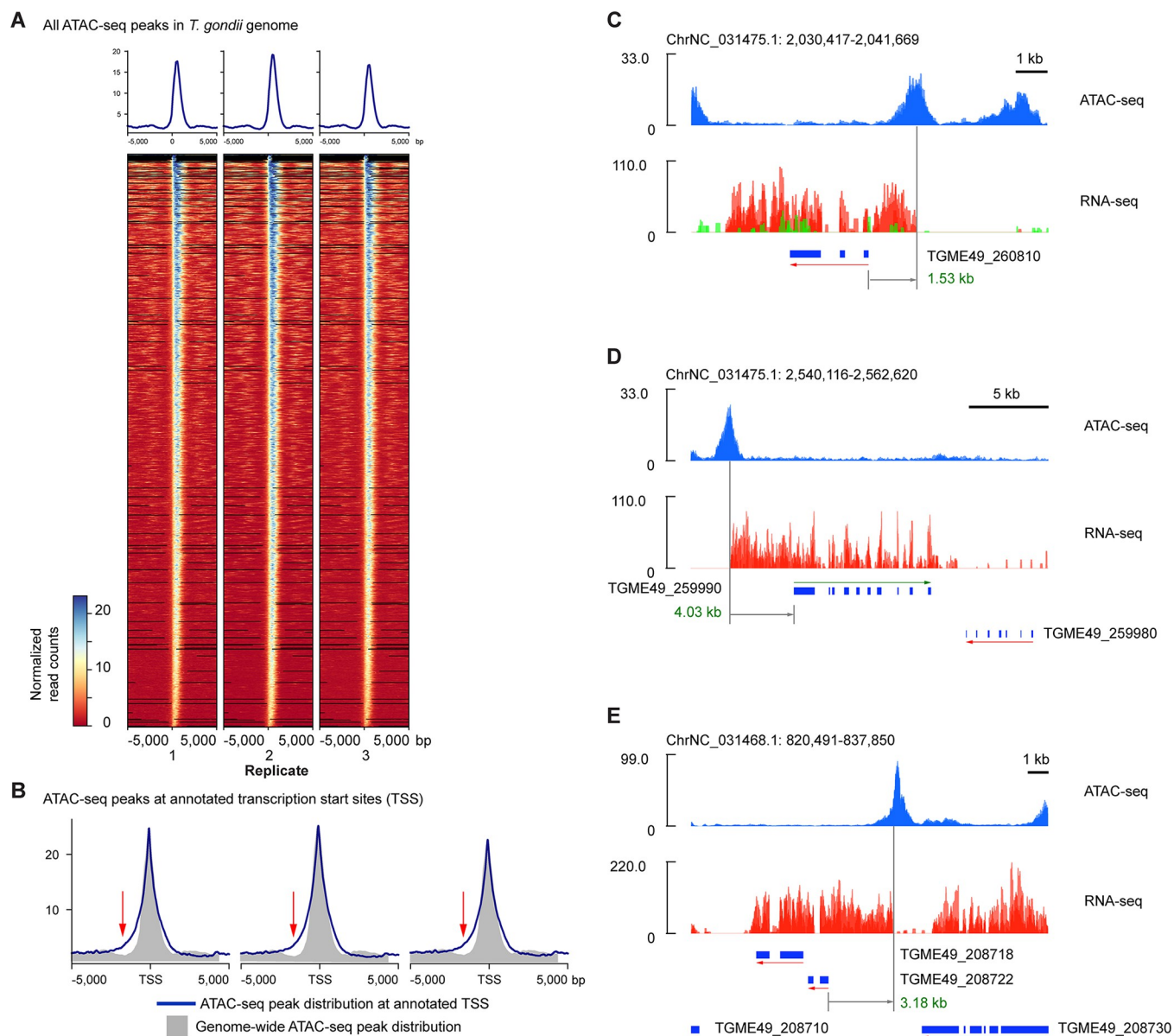


Fig 4. Chromatin accessibility profiles in the *T. gondii* genome indicate incomplete annotations of genes. The chromatin accessibility pattern at all of the ATAC-seq peaks in the *T. gondii* genome at the tachyzoite stage of its life cycle is shown in (A). When the subset within ± 5 kb of annotated *T. gondii* transcription start sites (TSS) is plotted with the associated genes represented to the right of the TSS (B, blue line), the patterns shows a skewing relative to the distributions of all peaks from panel (A) (gray distribution). Examples of genes are shown to represent how the annotated TSS can be several kilobases (kb) from the upstream ATAC-seq peak and the start of the RNA-seq reads, with examples of 1.53 kb (C) and 4.03 kb (D) differences illustrated. In (E) we gain the additional insight that the two annotated genes are only associated with a single upstream ATAC-seq peak, indicating that there is only one transcript at this locus and the genes could be combined into a single annotation.

<https://doi.org/10.1371/journal.pone.0275226.g004>

and colleagues [35], Fig 5D) and distal *cis*-regulatory elements (Fig 5E). The combination of ATAC-seq and RNA-seq appears to permit more detailed annotation of the genome than would each assay on its own.

Candidate *Toxoplasma gondii* transcription factor binding motifs

The chromatin accessibility dataset provided a novel opportunity to define candidate TF binding sites involved in *T. gondii* transcriptional regulation. We used *de novo* motif enrichment

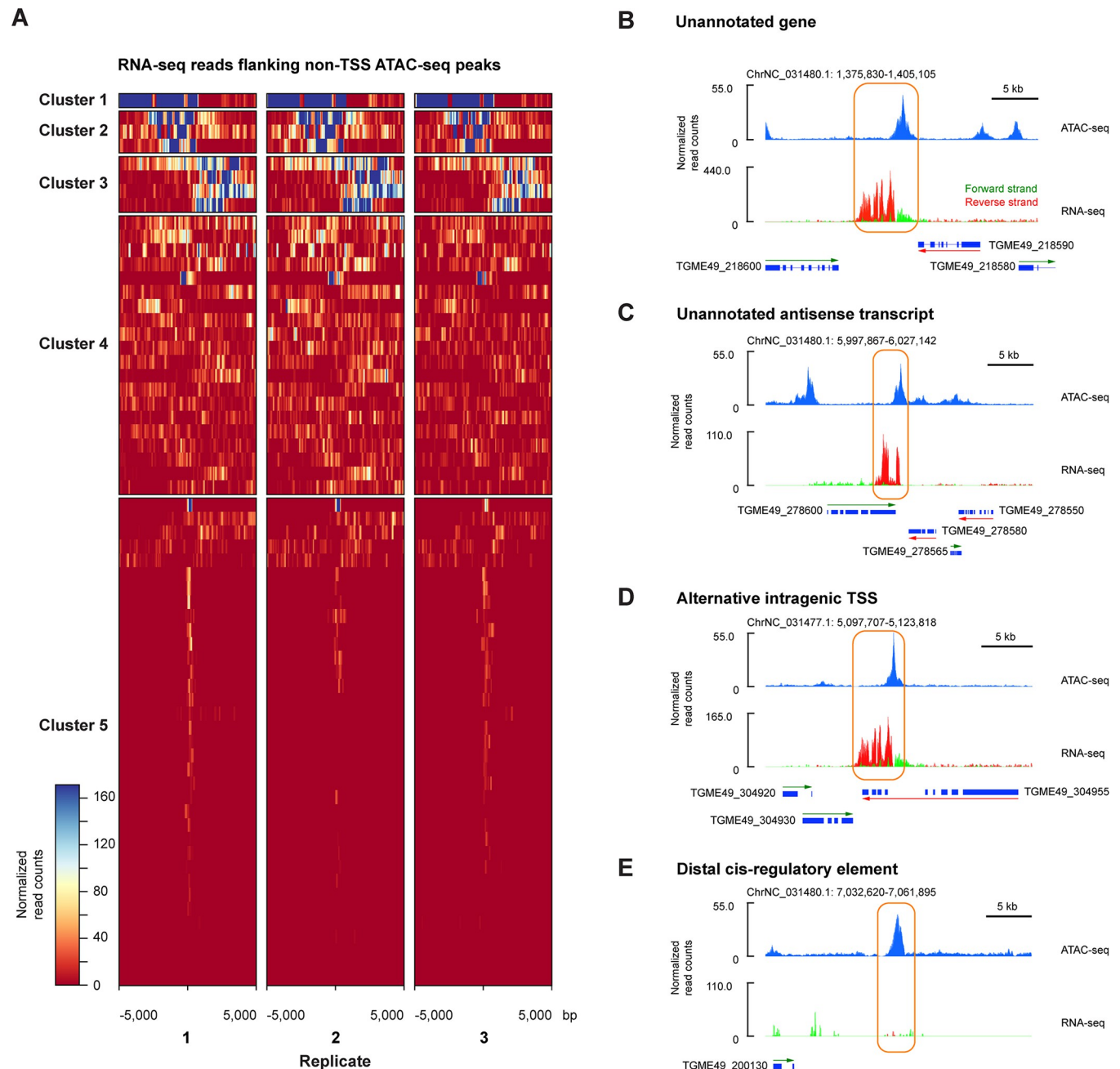


Fig 5. Chromatin accessibility and transcriptional profiling identifies new features in the *T. gondii* genome. The gene expression and chromatin accessibility information combine to allow new *T. gondii* genomic annotations. In (A) we show the RNA expression at ATAC-seq peaks that are not within 5 kb of an annotated TSS. We sorted these peaks by clustering the characteristics of the RNA-seq data nearby, with the top clusters (1–3 especially) showing evidence for adjacent transcripts. Examples of the loci are shown on the right, including what appears to be a new gene (B), an antisense transcript from the 3' end of an annotated gene (C), the use of an alternative intragenic promoter (D) and a locus of open chromatin with no nearby genes or RNA expression, potentially representing a distal *cis*-regulatory element.

<https://doi.org/10.1371/journal.pone.0275226.g005>

analysis with MEME-ChIP [31] and identified three significant ($e\text{-value} < 1e^{-20}$) motifs (Fig 6A).

The most abundant *T. gondii* motif that we identified is the purine-rich AAAAAGGAAGAAA sequence, located at 267 open chromatin regions. The next most abundant motif that we

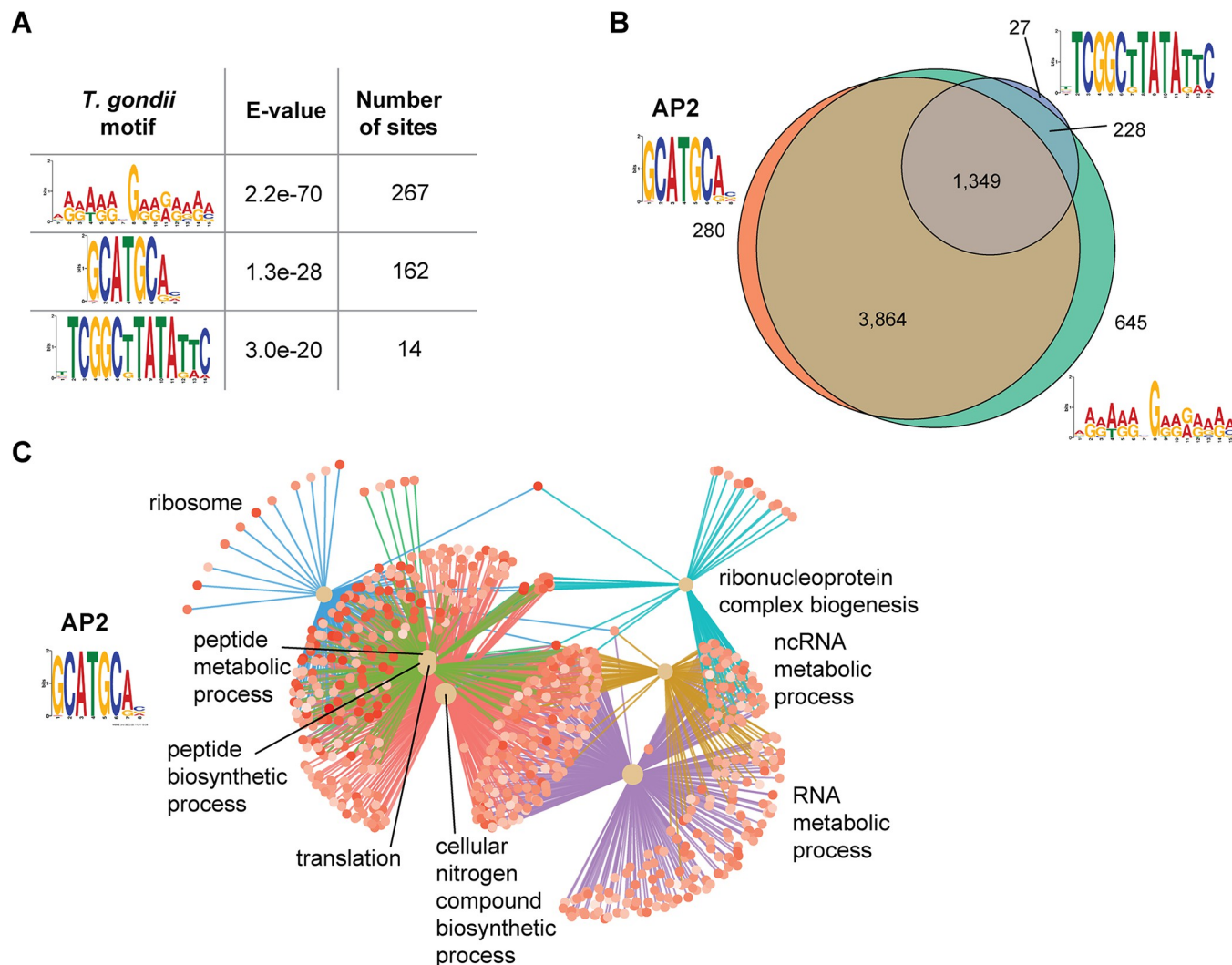


Fig 6. Inference of transcription factor binding motifs in the *T. gondii* genome. Analysis of the loci of open chromatin in the *T. gondii* genome reveals enrichment for several motifs (A). The purine (AG)-rich motif is the most abundantly represented at *T. gondii* promoters (B) followed by the known AP2 motif (GCATGCA), with the TATA-containing motif in 25.1% of promoters. Panel (C) shows the properties of the subset of genes with the AP2 motif at their promoters.

<https://doi.org/10.1371/journal.pone.0275226.g006>

found is the known apicomplexan AP2 TF binding site GCATGCA, first found in other apicomplexan parasites, *Plasmodium falciparum*, *Theileria*, and *Cryptosporidium* [40–42]. The third motif contains a strong TATA sequence, potentially representing the TATA box that has proven elusive in prior *T. gondii* studies [43].

We found the three motifs to be frequently co-located within the same gene promoters. In 1,349 genes all three of the motifs were found in the promoter region (Fig 6B). The candidate TATA box-containing motif is present at the promoters of 25.1% of *T. gondii* genes, comparable to the estimates of 20% of yeast promoters [44] and 27% of human promoters [45] having TATA boxes. When we analyzed the genes containing the better characterized AP2 motif, we found them to be significantly enriched for RNA and protein metabolic and biosynthetic processes (Fig 6C).

Potential *Toxoplasma gondii* transcription factor effects on host cell chromatin

Finally, we asked the exploratory question whether there was any evidence that TFs from *T. gondii* could be accessing the chromatin of the host cell and directly altering transcription, extending prior models of *T. gondii* effects on host chromatin [46–48]. We used the three major *T. gondii* TF motifs of Fig 6 and explored whether any of these were enriched at loci where chromatin becomes accessible during infection. We found 46 loci in the host genome to meet these criteria (S6 Table). In S7 Fig we show three examples, the first two (panels A-B) showing the AP2 motif at loci that open chromatin at intronic regions of the *LINC02703* and *PCMTD1* genes, while panel C shows the purine-rich motif at the bidirectional promoter of *TENM3-AS1/ENSG00000248266*, with chromatin opening at this locus associated with increased local transcription. While these findings could certainly represent a chance occurrence of overlap of the *T. gondii* TF motifs at loci that are undergoing chromatin remodelling during infection purely due to host TF activities, we speculate that it is possible that *T. gondii* TFs access the host nucleus and influence host chromatin states directly. The data presented define candidate loci for further exploration.

Discussion

This is the first study that looks simultaneously at the transcriptome and chromatin organization of both the host cell and an intracellular apicomplexan parasite during infection. Our goal was to gain insights into the interaction between the activity of the well-studied parasite *T. gondii* and the response of the infected human host cell, by studying gene expression and chromatin responses concurrently. We note that there have been elegant prior transcriptional [49, 50] and chromatin studies [33, 51, 52], and promoter mapping [35, 53] of apicomplexan organisms, and of host cell transcriptional responses [54] that show differences between cell types [55], a reminder that our results should be considered specific to infection of human fibroblasts. Our intent was to extend these studies to gain insights into the conflict being fought at the molecular level during infection by this very successful infectious agent.

The host gene expression changes are abundant, and mirror the results of several prior studies, including the induction of immune response genes [56], and the over-expression of matrix metalloproteases [57], while the induction of metabolic reprogramming is consistent with the auxotrophic nature of *T. gondii*, and the mitotic gene response reflective of prior observations of enhancement of cell cycle progression following *T. gondii* infection [58]. In order to understand how these responses are regulated, we were able to use TF binding motif information in the open chromatin regions of the host cell to infer the TF responses mediating these transcriptional changes. This analysis provided the insight that the NFκB response appeared to be predominantly involved in host cell immune responses, but JUN:FOS (AP-1) and RREB1 are likely to have additional properties of cell division, metabolic processes, and extracellular matrix gene regulation.

We were then able to add further step of inference by asking how these TFs could be altered in their functions with *T. gondii* infection. It is known that *T. gondii* exports into the host cell the GRA24 protein, which complexes with [59] and induces mitogen-activated protein (MAP) kinase p38α (P38 MAPK) [60]. The TFs that respond to P38 MAPK signaling include ATF1/ATF2, CEBP family members, SAP1, p53 and ETS1 [61], but their associated binding motifs were not prominent at loci gaining chromatin accessibility, making it unclear how GRA24 exerts transcriptional effects during infection. In Fig 7 we present a model for cell signaling pathways influencing JUN and RREB1, implicating the ERK1/2 and SAPK/JUNK MAPK signaling that are known to be induced by *T. gondii* infection [62, 63] in the induction of these

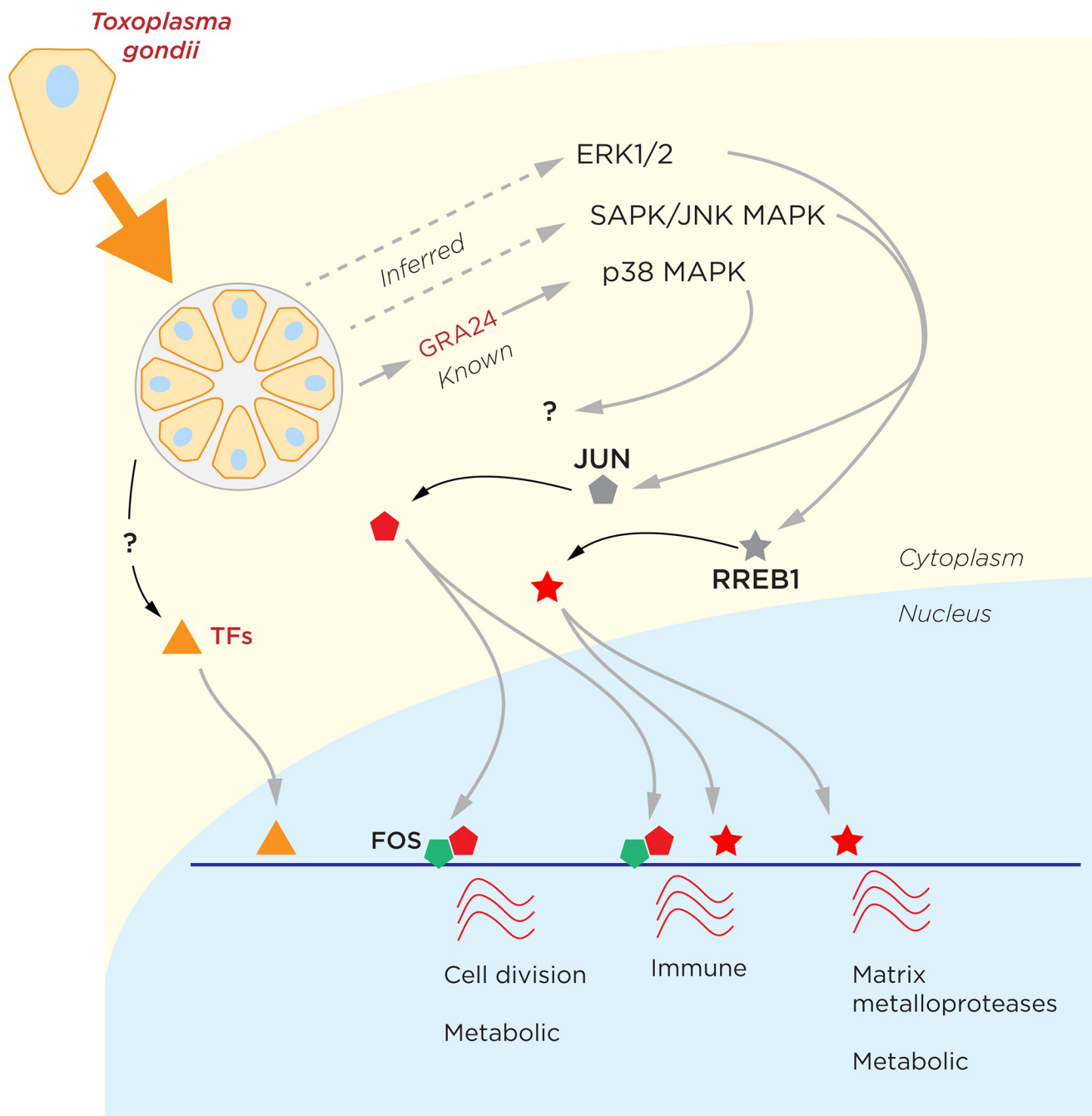


Fig 7. A model derived from genomic assay data for the host cellular response to *T. gondii* infection. By inferring the TFs mediating the host cell response, we can further predict the cell signaling pathways induced by *T. gondii* infection in human fibroblasts. How the known GRA24 induction of p38 MAPK signaling influences the transcriptional response remains uncertain, and we include the possibility that *T. gondii* TFs may contribute to host cell transcriptional dysregulation.

<https://doi.org/10.1371/journal.pone.0275226.g007>

TFs. These signaling pathways are also activated in *Leishmania* [64], *Plasmodium* and *Theileria* [65] infections, indicating how central these are in the host response to protozoal infections. The downstream response to the activation of these JUN and RREB1 TFs involves host responses likely to be favorable to the pathogen, and we also depict how there may also be the

possibility that *T. gondii* TFs directly influence host cell chromatin (S6 Fig). This Fig 7 model is undoubtedly incomplete but serves to illustrate how these genomic assay data can be used to infer a lot more about the pathogen-host interaction than just host transcriptional outcomes alone.

The data also reveal new insights into the *T. gondii* genome. We observe that stratification of tachyzoite gene expression by expression levels correlates with function. Mapping ATAC-seq peaks generally enriches the region close to the annotated transcription start site (TSS) of *T. gondii* genes, but we see numerous examples of open chromatin located at substantial distances from TSSs (Fig 4), providing a resource for improved TSS/promoter annotation in the *T. gondii* genome. We also find that there are ATAC-seq peaks more than 5 kb from any TSS, and show in Fig 5 how the combination of ATAC-seq and RNA-seq data allows new genes, antisense transcripts, intragenic alternative promoters and distal *cis*-regulatory elements to be predicted, a resource for enhancing the annotation of the *T. gondii* genome.

Motif analysis of the loci of open chromatin in *T. gondii* reassuringly rediscovered the known AP2 TF motif [40–42]. We also found enrichment of a purine-rich motif of uncertain significance present at a large proportion of TSS, a motif not previously identified in the few studies of *cis*-regulatory loci in *T. gondii*. The presence of the TATA motif within the TCGGCTTATATTC sequence raises the possibility that *T. gondii* does, in fact, use a TATA box [43].

In summary, the combined study of host and parasite transcription and its regulation offers insights into the conflict between the host and pathogen, with some evidence for host physiology being harnessed by the pathogen for its own purposes. The ability to infer through molecular genomic studies insights into cell signaling is relatively novel and typically an underutilized potential insight from molecular genomics studies, and is made possible by focusing on TFs as the mediators of molecular genomic changes.

Materials and methods

Cell and parasite cultures

Cell and parasite cultures were maintained as previously described [66]. Human Foreskin Fibroblasts (HFFs) were grown in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/mL penicillin, 100 µg/mL streptomycin and 2 mM L-glutamine and were maintained at 37°C with 5% CO₂. *T. gondii* Type I tachyzoites (RHΔHxΔKu80) were also maintained at 37°C with 5% CO₂ by infecting 25 cm² flasks containing confluent HFFs every 48 hours.

Host cell infections

The 25 cm² flasks containing freshly egressed *T. gondii* tachyzoites were then passaged through a 25-gauge needle three times to lyse out any remaining intracellular tachyzoites from their hosts, and the tachyzoites were spun down at 3,000 rpm for 8 minutes. 75 cm² flasks containing confluent HFFs were then infected with these tachyzoites using a multiplicity of infection (MOI) of 3, and 24 hours after the infection the percentage of infected cells per flask was calculated. Flasks counted as at least 80% infected at this time were harvested by scraping. A flask containing HFFs to be used as an uninfected control was also harvested in parallel by scraping. The cells were pelleted by centrifugation at 1,300 rpm for 5 minutes. The cells were harvested such that both RNA and DNA could be extracted from the same flask for each biological replicate, and a total of three biological replicates of both the control and infected flasks were harvested in total for each assay.

DNA extraction

Genomic DNA (gDNA) was extracted from each sample as follows: 500 μ L 10% SDS and 10 μ L RNase A (10 mg/mL) were added to each sample followed by an incubation at 37°C for 1 hour. 50 μ L of proteinase K (20 mg/mL) was then added and the samples were incubated at 55°C overnight. An equal volume of saturated phenol was then added to each tube, and the samples were mixed slowly at room temperature for 15 minutes. The samples were then spun down at 3,000 rpm for 10 minutes and the supernatant was collected. A total of 3 phenol extractions were performed, followed by 2 chloroform extractions. The samples were then pipetted into a dialysis bag and further purified at 4°C through three changes of a 20x NaCl, sodium citrate (SSC) buffer over the course of 24 hours. Dehydration was then performed by covering each sample in a pile of polyethylene glycol (PEG) for ~1 hour. The PEG was rinsed from each dialysis bag and the gDNA extracted from each sample was collected.

RNA extraction

RNA was extracted using TRIzol reagent (Invitrogen Cat. # 15596018) according to the manufacturer's instructions as follows. One milliliter of TRIzol was added to each sample and incubated at room temperature for 5 minutes. A volume of 200 μ L of chloroform was added to each sample and the tubes were shaken vigorously for 15 seconds and then incubated for 2–3 minutes at room temperature. The samples were then centrifuged at 12,000 rpm for 15 minutes at 4°C, and the aqueous phase of each sample were collected. Following this, 0.5 mL of 100% isopropanol was added to the aqueous phase, and the samples were incubated for 10 minutes at room temperature, followed by centrifugation at 12,000 g for 10 minutes at 4°C. The supernatant was removed from each tube and the pellet was washed in 1 mL of 75% ethanol and centrifuged at 7,500 g for 5 minutes at 4°C. The supernatant was again removed and the pellet was air dried for 5 minutes. The RNA pellet was then resuspended in RNase-free water and incubated at 55°C for 10 minutes. Following the RNA extraction, the samples were then treated with DNase to remove contaminating gDNA, and further purified using the miRNeasy Mini Kit (QIAGEN cat# 217004). The Bioanalyzer was used to assess total RNA integrity prior to library preparation, and only samples with an RNA integrity number (RIN) greater than 8 was used for further downstream library preparation.

Directional RNA-seq library preparation

Directional RNA-seq libraries were prepared as previously described by our lab [67]. One microgram of RNA was DNase-treated and rRNA-depleted (Ribozero rRNA Removal Kit, Epicentre). RNA was used as a template for SuperScript III first-strand cDNA synthesis (ThermoFisher Scientific, SuperScript III kit cat# 18080–044), using oligo-dT as well as random hexamers. Actinomycin D was added to the reaction to prevent any possible amplification from contaminating gDNA. During second-strand synthesis, a (dUTP/dATP/dCTP/dGTP) mix was used to create directional libraries. Before library preparation, cDNA samples were Covaris-fragmented to 300 bp fragments. The samples were then end-filled, 3' terminal A extended and ligated with TruSeq-indexed Illumina adapters. Uracil-DNA-glycosylase (UDG) treatment preceded the PCR reaction to amplify exclusively the originally oriented transcripts. Libraries were amplified using P5 and P7 Illumina primers. Prior to sequencing the libraries were gel-extracted for size selection, primer-dimers were removed and the quality was assessed on the Bioanalyzer. The six libraries were multiplexed on a single lane of the Illumina Hi-Seq 2500 platform twice to obtain 100 bp single-end reads. The reads combined from the two lanes of the sequencer for each sample were used for further downstream analysis, resulting in a mean of 25 million reads per sample (S7 Table).

ATAC-seq library preparation

The assay for transposase accessible chromatin (ATAC)-seq libraries were prepared as previously described [29]. A total of three biological replicates each of uninfected and *T. gondii* infected HFFs were harvested at 24 hours post-infection. To prepare the nuclei, we spun 75,000 cells at 500 g for 5 minutes. Each sample was washed using 50 μ L of cold 1x PBS. Samples were then centrifuged at 500 g for 5 minutes. Cells were lysed in cold lysis buffer (10 mM Tris- HCL, pH 7.4, 10 mM NaCl, 3 mM MgCl₂ and 0.1% IGEPAL CA-630). Immediately following lysis, nuclei were spun at 500 g for 10 minutes at 4°C. The pellet was then resuspended in the transposase reaction mix (25 μ L 2x TD buffer, 2.5 μ L transposase and 22.5 μ L nuclease-free water).

The transposition reaction was carried out for 30 minutes at 37°C and the samples were then purified using a Zymo DNA Clean and Concentrator purification kit. Following the purification, we amplified the libraries with the following PCR conditions: 72°C for 5 minutes; 98°C for 30 seconds; and a total of 10 cycles of 98°C for 10 seconds; 63°C for 30 seconds; 1 minute for 72°C. Libraries were purified using Agencourt AMPure beads. Libraries were further amplified using the PCR conditions: 98°C for 30 seconds; and a total of 14 cycles of 98°C for 10 seconds; 63°C for 30 seconds and 72°C for 1 minute. The six ATAC-seq samples were multiplexed on a single lane of the Illumina Hi-Seq 2500 to obtain 100 bp paired-end reads, which resulted in a mean of 53 million reads per sample (S8 Table).

RNA-seq read processing, alignment, and quantification

101 bp single-end reads were trimmed to remove low quality base calls and Illumina universal adapters using Trim Galore! (version 0.6.5) with default parameters and then assessed using fastQC (version 0.11.4) and multiqc (version 1.10.1) [68]. Trimmed reads were aligned to a combined human (GRCh38), human decoy (GRCh38), and *T. gondii* (ME49) genome using Hisat2 (version 2.0.4) [69] with default parameters. Reads were then split into human and *T. gondii* fractions using samtools (version 1.9) [70]. Aligned reads were then assessed for each organism separately using RSeQC (version 2.6.4) [71] and Qualimap (version 2.2.1) [72]. Aligned reads were then quantified using HTseq (version 0.12.3) [73] in union mode using default parameters. Alignment tracks were generated for forward and reverse strand reads using deepTools (version 3.1.0) [74] normalized using scaling factors calculated using the RLE method in DESeq2 (version 1.32) [30], and viewed in IGV (version 2.9.4) [75].

Human and *Toxoplasma gondii* differential expression analysis

Differential expression analysis was performed using DESeq2 and gene set enrichment analysis performed using clusterProfiler (version 4.0.5) [76]. A gene ontology annotation for *T. gondii* ME49 was downloaded from ToxoDB. For *T. gondii*, only the infected replicates were analyzed.

ATAC-seq read processing, alignment, and quantification

101 bp paired-end reads were trimmed to remove low quality base calls and Illumina universal adapters using Trim Galore! (version 0.6.5) with default parameters and then assessed using fastQC (version 0.11.4) and multiqc (version 1.10.1). Trimmed reads were aligned to a combined human (GRCh38), human decoy (GRCh38), and *T. gondii* (ME49) genome using bwa mem (version 0.7.17) [77] with default parameters. Samtools (version 1.9) [70], Picard tools (version 1.92), and deeptools (version 3.1.0) [74] were used for removing mitochondrial reads, PCR optical duplicates, low quality reads, and shifting reads to account for Tn5 transposase

offset. ATACSeqQC (version 1.16) [78] was used to assess aligned reads. Narrow peaks were called using MACS2 (version 2.1.0) [79] using default parameters for each replicate and reproducible peaks across all replicates in the same condition were called using ChIP-R (version 1.1.0) [39]. Reproducible peaks from each condition were combined into non-redundant peaks, blacklisted regions were filtered, and peaks were quantified using Rsubread (version 2.6.4) [80] and GenomicAlignments (version 1.28.0) [81]. Alignment tracks were generated using deepTools, normalized using scaling factors calculated using the RLE method in DESeq2, and viewed in IGV.

Human and *Toxoplasma gondii* differential accessibility analysis

Differential accessibility analysis was performed using DESeq2 and gene set enrichment analysis was performed on annotated peaks using clusterProfiler. A gene ontology annotation for *T. gondii* ME49 was downloaded from ToxoDB (toxodb.org). For *T. gondii*, only the infected replicates were analyzed.

Human peak annotation and motif enrichment

Differentially accessible peaks were annotated using ChipPeakAnno (version 3.26.4) [82] with parameters adjusted to output all overlapping features within 5 kb upstream or downstream of the nearest transcription start site. For *de novo* motif discovery, 100 bp fragments were extracted from the middle of the top 1,000 differentially accessible upregulated or downregulated peaks and input into MEME-ChIP (version 5.4.1) [83] with default parameters. chromVAR (version 1.14.0) [84] was used with the JASPAR2020 vertebrate core motif database [85] to identify variable motifs between conditions, testing all differentially accessible peaks.

Toxoplasma gondii peak annotation and motif enrichment

Reproducible peaks in the infected condition were annotated using ChipPeakAnno with parameters adjusted to output all overlapping features within 5 kb upstream or downstream of the nearest transcription start site. For *de novo* motif discovery, 100 bp fragments were extracted from the middle of all reproducible peaks in the infected condition and input into MEME-ChIP with default parameters. To detect potential *T. gondii*-specific transcription factors acting on the human genome, *de novo* motifs were searched against the human differentially accessible peak sequences using FIMO (version 5.4.1) [86].

Supporting information

S1 Fig. The gene set enrichment analysis-derived Gene Ontology (GO) terms characterizing the properties of the genes increasing expression with *T. gondii* infection. (TIF)

S2 Fig. A PCA plot of ATAC-seq of human host cell genes, using the 5,000 most variable peaks across all replicates. The analysis shows that one infected sample (I1) clusters with the uninfected samples (U1-3). We interpreted this to indicate that the infection in this replicate was poor, and excluded this I1 sample from further analyses. (TIF)

S3 Fig. The location of ATAC-seq peaks (left) and differentially-accessible regions (DARs, right) relative to annotated transcription start sites (TSS) in the human genome. Whereas ATAC-seq peaks are strongly enriched at TSS, only 51 (9.6%) of DARs are located within 5 kb

of annotated transcription start sites (yellow shading).
(TIF)

S4 Fig. A heat map illustrating the enrichment of transcription factor binding site motifs from a JASPAR analysis of the loci where chromatin becomes more accessible with infection. We see the RELA, RREB1 and FOSB::JUNB motifs now accompanied by many other transcription factor motifs, including many NFκB family members and AP-1 targets. These are all listed in [S3 Table](#).

(TIF)

S5 Fig. In (A) we show the number of RNA-seq reads from *T. gondii* to be consistently greater than 2 million in all replicates. In (B) we illustrate how most genes in the *T. gondii* genome are represented consistently across replicates, also showing this quantitatively in the UpSet plot of panel (C).

(TIF)

S6 Fig. The expected nucleosomal periodicity pattern of chromatin is revealed from the insert size plots of panel (A), representing the reads aligning to the *T. gondii* nuclear genome, whereas in (B) we see no evidence for such nucleosomal organization in reads mapping to the *T. gondii* apicoplast DNA.

(TIF)

S7 Fig. We show examples of loci where the motifs enriched at *T. gondii* loci of open chromatin are also found in the host genome at sites that open chromatin during *T. gondii* infection, representing candidate loci where *T. gondii* TFs may be acting directly on host chromatin.

(TIF)

S1 Table. The list of host genes changing expression with *T. gondii* infection.

(XLSX)

S2 Table. Gene set enrichment analysis-derived Gene Ontology (GO) terms characterizing genes increasing expression with *T. gondii* infection.

(XLSX)

S3 Table. The results of our analysis of loci of open chromatin in infected cells using the vertebrate core motif database JASPAR2020, the data used for [S4 Fig](#).

(XLSX)

S4 Table. The detailed expression data for the ME49 reference genome annotated genes. We present a raw counts file with data for each replicate, and the normalized counts file on which our analyses were based.

(XLSX)

S5 Table. Gene ontology (GO) annotations of *T. gondii* ME49 genes from ToxoDB (toxodb.org) applied to genes in expression level categories of high, medium and low.

(XLSX)

S6 Table. The candidate transcription factor binding site motifs from [Fig 6](#) were used to identify loci in the host human genome where these motifs are present and chromatin becomes accessible during infection, defining candidate loci where *T. gondii* transcription factors may be directly acting on the host human genome. We list the 46 candidate loci and the associated motifs.

(XLSX)

S7 Table. RNA-seq sequencing data metrics for the human and *T. gondii* genomes.
(XLSX)

S8 Table. ATAC-seq sequencing data metrics for the human and *T. gondii* genomes.
(XLSX)

Acknowledgments

The authors thank Drs. Zofia Wicik, Sanchari Bhattacharyya and Amit Verma for their input into this project, and Dr. Kami Kim (University of South Florida) and Dr. Inessa Gendlina (Albert Einstein College of Medicine) for generating and characterizing the *in vitro* *Toxoplasma* infections. Einstein core facilities involved were the High-Performance Computing Core, the Epigenomics Shared Facility, and the Genomics Core Facility, with support from the Albert Einstein Cancer Center (P30CA013330) and the Center for Epigenomics.

Author Contributions

Conceptualization: John M. Greally.

Data curation: Ronald Cutler, N. Ari Wijetunga, Matthew McKnight Croken, Andrew D. Johnston, Yu Kong.

Formal analysis: Netha Ulahannan, Ronald Cutler, Reanna Doña-Termine, N. Ari Wijetunga, Yu Kong, Masako Suzuki.

Funding acquisition: John M. Greally.

Investigation: Netha Ulahannan, Ronald Cutler, Claudia A. Simões-Pires, Matthew McKnight Croken, Andrew D. Johnston, Shahina B. Maqbool, John M. Greally.

Methodology: Ronald Cutler, John M. Greally.

Project administration: John M. Greally.

Resources: John M. Greally.

Software: Ronald Cutler, Reanna Doña-Termine, Andrew D. Johnston, Yu Kong.

Supervision: Claudia A. Simões-Pires, Masako Suzuki, John M. Greally.

Writing – original draft: Netha Ulahannan, N. Ari Wijetunga, Matthew McKnight Croken, Yu Kong, Shahina B. Maqbool, Masako Suzuki, John M. Greally.

Writing – review & editing: Ronald Cutler, Reanna Doña-Termine, Andrew D. Johnston, John M. Greally.

References

1. Bahia D, Satoskar AR, Dussurget O. Editorial: Cell Signaling in Host-Pathogen Interactions: The Host Point of View. *Front Immunol.* 2018; 9: 221. <https://doi.org/10.3389/fimmu.2018.00221> PMID: 29520263
2. Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admetlla A, Pattini L, Nielsen R. Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet.* 2011; 7: e1002355. <https://doi.org/10.1371/journal.pgen.1002355> PMID: 22072984
3. Tanowitz HB, Weiss LM. Toxoplasmosis: A comprehensive clinical guide:toxoplasmosis: A comprehensive clinical guide. *Clin Infect Dis.* 2002; 34: 1664–1665. <https://doi.org/10.1086/340624>
4. Laliberté J, Carruthers VB. Host cell manipulation by the human pathogen *Toxoplasma gondii*. *Cell Mol Life Sci.* 2008; 65: 1900–1915. <https://doi.org/10.1007/s00018-008-7556-x> PMID: 18327664

5. Gissot M, Choi S-W, Thompson RF, Grealley JM, Kim K. *Toxoplasma gondii* and *Cryptosporidium parvum* lack detectable DNA cytosine methylation. *Eukaryotic Cell*. 2008; 7: 537–540. <https://doi.org/10.1128/EC.00448-07> PMID: 18178772
6. Gissot M, Kelly KA, Ajioka JW, Grealley JM, Kim K. Epigenomic modifications predict active promoters and gene structure in *Toxoplasma gondii*. *PLoS Pathog*. 2007; 3: e77. <https://doi.org/10.1371/journal.ppat.0030077> PMID: 17559302
7. Espindula E, Sperb ER, Bach E, Passaglia LMP. The combined analysis as the best strategy for Dual RNA-Seq mapping. *Genet Mol Biol*. 2019; 42: e20190215. <https://doi.org/10.1590/1678-4685-GMB-2019-0215> PMID: 32442239
8. Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res*. 2003; 92: 827–839. <https://doi.org/10.1161/01.RES.0000070112.80711.3D> PMID: 12730128
9. Elkington PTG, O’Kane CM, Friedland JS. The paradox of matrix metalloproteinases in infectious disease. *Clin Exp Immunol*. 2005; 142: 12–20. <https://doi.org/10.1111/j.1365-2249.2005.02840.x> PMID: 16178851
10. Lu C-Y, Lai S-C. Matrix metalloproteinase-2 and -9 lead to fibronectin degradation in astroglia infected with *Toxoplasma gondii*. *Acta Trop*. 2013; 125: 320–329. <https://doi.org/10.1016/j.actatropica.2012.11.002> PMID: 23201304
11. Ólafsson EB, Varas-Godoy M, Barragan A. *Toxoplasma gondii* infection shifts dendritic cells into an amoeboid rapid migration mode encompassing podosome dissolution, secretion of TIMP-1, and reduced proteolysis of extracellular matrix. *Cell Microbiol*. 2018;20. <https://doi.org/10.1111/cmi.12808> PMID: 29119662
12. Lei Y, Takahama Y. XCL1 and XCR1 in the immune system. *Microbes Infect*. 2012; 14: 262–267. <https://doi.org/10.1016/j.micinf.2011.10.003> PMID: 22100876
13. Iwasaki A, Medzhitov R. Control of adaptive immunity by the innate immune system. *Nat Immunol*. 2015; 16: 343–353. <https://doi.org/10.1038/ni.3123> PMID: 25789684
14. Kroczeck RA, Henn V. The Role of XCR1 and its Ligand XCL1 in Antigen Cross-Presentation by Murine and Human Dendritic Cells. *Front Immunol*. 2012; 3: 14. <https://doi.org/10.3389/fimmu.2012.00014> PMID: 22566900
15. Dev A, Iyer S, Razani B, Cheng G. NF- κ B and innate immunity. *Curr Top Microbiol Immunol*. 2011; 349: 115–143. https://doi.org/10.1007/82_2010_102 PMID: 20848362
16. Jones SA. Directing transition from innate to acquired immunity: defining a role for IL-6. *J Immunol*. 2005; 175: 3463–3468. <https://doi.org/10.4049/jimmunol.175.6.3463> PMID: 16148087
17. McCarl C-A, Khalil S, Ma J, Oh-hora M, Yamashita M, Roether J, et al. Store-operated Ca²⁺ entry through ORAI1 is critical for T cell-mediated autoimmunity and allograft rejection. *J Immunol*. 2010; 185: 5845–5858. <https://doi.org/10.4049/jimmunol.1001796> PMID: 20956344
18. Linge HM, Collin M, Nordenfelt P, Mörgelin M, Malmsten M, Egesten A. The human CXC chemokine granulocyte chemotactic protein 2 (GCP-2)/CXCL6 possesses membrane-disrupting properties and is antibacterial. *Antimicrob Agents Chemother*. 2008; 52: 2599–2607. <https://doi.org/10.1128/AAC.00028-08> PMID: 18443119
19. Zare-Bidaki M, Assar S, Hakimi H, Abdollahi SH, Nosratabadi R, Kennedy D, et al. TGF- β in Toxoplasmosis: Friend or foe? *Cytokine*. 2016; 86: 29–35. <https://doi.org/10.1016/j.cyto.2016.07.002> PMID: 27449809
20. Meylan F, Richard AC, Siegel RM. TL1A and DR3, a TNF family ligand-receptor pair that promotes lymphocyte costimulation, mucosal hyperplasia, and autoimmune inflammation. *Immunol Rev*. 2011; 244: 188–196. <https://doi.org/10.1111/j.1600-065X.2011.01068.x> PMID: 22017439
21. Brown EJ, Baltimore D. Essential and dispensable roles of ATR in cell cycle arrest and genome maintenance. *Genes Dev*. 2003; 17: 615–628. <https://doi.org/10.1101/gad.1067403> PMID: 12629044
22. Vigneswara V, Ahmed Z. The Role of Caspase-2 in Regulating Cell Fate. *Cells*. 2020;9. <https://doi.org/10.3390/cells9051259> PMID: 32438737
23. Lee K-J, Lin Y-F, Chou H-Y, Yajima H, Fattah KR, Lee S-C, et al. Involvement of DNA-dependent protein kinase in normal cell cycle progression through mitosis. *J Biol Chem*. 2011; 286: 12796–12802. <https://doi.org/10.1074/jbc.M110.212969> PMID: 21330363
24. Poirier S, Mayer G, Benjannet S, Bergeron E, Marcinkiewicz J, Nassoury N, et al. The proprotein convertase PCSK9 induces the degradation of low density lipoprotein receptor (LDLR) and its closest family members VLDLR and ApoER2. *J Biol Chem*. 2008; 283: 2363–2372. <https://doi.org/10.1074/jbc.M708098200> PMID: 18039658

25. Cases S, Stone SJ, Zhou P, Yen E, Tow B, Lardizabal KD, et al. Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. *J Biol Chem*. 2001; 276: 38870–38876. <https://doi.org/10.1074/jbc.M106219200> PMID: 11481335
26. Geng S, Zhou S, Glowacki J. Effects of 25-hydroxyvitamin D(3) on proliferation and osteoblast differentiation of human marrow stromal cells require CYP27B1/1 α -hydroxylase. *J Bone Miner Res*. 2011; 26: 1145–1153. <https://doi.org/10.1002/jbmr.298> PMID: 21542014
27. Menendez JA, Lupu R. Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis. *Nat Rev Cancer*. 2007; 7: 763–777. <https://doi.org/10.1038/nrc2222> PMID: 17882277
28. Blume M, Seeber F. Metabolic interactions between *Toxoplasma gondii* and its host. [version 1; peer review: 2 approved]. *F1000Res*. 2018;7. <https://doi.org/10.12688/f1000research.16021.1> PMID: 30467519
29. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013; 10: 1213–1218. <https://doi.org/10.1038/nmeth.2688> PMID: 24097267
30. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15: 550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
31. Ma W, Noble WS, Bailey TL. Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat Protoc*. 2014; 9: 1428–1450. <https://doi.org/10.1038/nprot.2014.083> PMID: 24853928
32. Deng Y-N, Xia Z, Zhang P, Ejaz S, Liang S. Transcription Factor RREB1: from Target Genes towards Biological Functions. *Int J Biol Sci*. 2020; 16: 1463–1473. <https://doi.org/10.7150/ijbs.40834> PMID: 32210733
33. Nardelli SC, Silmon de Monerri NC, Vanagas L, Wang X, Tampaki Z, Sullivan WJ, et al. Genome-wide localization of histone variants in *Toxoplasma gondii* implicates variant exchange in stage-specific gene expression. *BMC Genomics*. 2022; 23: 128. <https://doi.org/10.1186/s12864-022-08338-6> PMID: 35164683
34. Brooks CF, Francia ME, Gissot M, Croken MM, Kim K, Striepen B. *Toxoplasma gondii* sequesters centromeres to a specific nuclear region throughout the cell cycle. *Proc Natl Acad Sci USA*. 2011; 108: 3767–3772. <https://doi.org/10.1073/pnas.1006741108> PMID: 21321216
35. Markus BM, Waldman BS, Lorenzi HA, Lourido S. High-Resolution Mapping of Transcription Initiation in the Asexual Stages of *Toxoplasma gondii*. *Front Cell Infect Microbiol*. 2020; 10: 617998. <https://doi.org/10.3389/fcimb.2020.617998> PMID: 33553008
36. Farhat DC, Swale C, Dard C, Cannella D, Ortet P, Barakat M, et al. A MORC-driven transcriptional switch controls *Toxoplasma* developmental trajectories and sexual commitment. *Nat Microbiol*. 2020; 5: 570–583. <https://doi.org/10.1038/s41564-020-0674-4> PMID: 32094587
37. Sindikubwabo F, Ding S, Hussain T, Ortet P, Barakat M, Baumgarten S, et al. Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites. *Elife*. 2017;6. <https://doi.org/10.7554/eLife.29391> PMID: 29101771
38. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res*. 2012; 22: 1813–1831. <https://doi.org/10.1101/gr.136184.111> PMID: 22955991
39. Newell R, Pienaar R, Balderson B, Piper M, Essebier A, Bodén M. ChIP-R: Assembling reproducible sets of ChIP-seq and ATAC-seq peaks from multiple replicates. *Genomics*. 2021; 113: 1855–1866. <https://doi.org/10.1016/j.ygeno.2021.04.026> PMID: 33878366
40. De Silva EK, Gehrke AR, Olszewski K, León I, Chahal JS, Bulyk ML, et al. Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc Natl Acad Sci USA*. 2008; 105: 8393–8398. <https://doi.org/10.1073/pnas.0801993105> PMID: 18541913
41. Balaji S, Babu MM, Iyer LM, Aravind L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res*. 2005; 33: 3994–4006. <https://doi.org/10.1093/nar/gki709> PMID: 16040597
42. Boeva V. Analysis of genomic sequence motifs for deciphering transcription factor binding and transcriptional regulation in eukaryotic cells. *Front Genet*. 2016; 7: 24. <https://doi.org/10.3389/fgene.2016.00024> PMID: 26941778
43. Yamagishi J, Wakaguri H, Ueno A, Goo Y-K, Tolba M, Igarashi M, et al. High-resolution characterization of *Toxoplasma gondii* transcriptome with a massive parallel sequencing method. *DNA Res*. 2010; 17: 233–243. <https://doi.org/10.1093/dnares/dsq013> PMID: 20522451
44. Basehoar AD, Zanton SJ, Pugh BF. Identification and distinct regulation of yeast TATA box-containing genes. *Cell*. 2004; 116: 699–709. [https://doi.org/10.1016/s0092-8674\(04\)00205-3](https://doi.org/10.1016/s0092-8674(04)00205-3) PMID: 15006352

45. Trinklein ND, Aldred SJF, Saldanha AJ, Myers RM. Identification and functional analysis of human transcriptional promoters. *Genome Res.* 2003; 13: 308–312. <https://doi.org/10.1101/gr.794803> PMID: 12566409
46. Hakimi M-A, Olias P, Sibley LD. Toxoplasma effectors targeting host signaling and transcription. *Clin Microbiol Rev.* 2017; 30: 615–645. <https://doi.org/10.1128/CMR.00005-17> PMID: 28404792
47. Gay G, Braun L, Brenier-Pinchart M-P, Vollaie J, Josserand V, Bertini R-L, et al. Toxoplasma gondii TgIST co-opts host chromatin repressors dampening STAT1-dependent gene regulation and IFN- γ -mediated host defenses. *J Exp Med.* 2016; 213: 1779–1798. <https://doi.org/10.1084/jem.20160340> PMID: 27503074
48. Olias P, Etheridge RD, Zhang Y, Holtzman MJ, Sibley LD. Toxoplasma Effector Recruits the Mi-2/NuRD Complex to Repress STAT1 Transcription and Block IFN- γ -Dependent Gene Expression. *Cell Host Microbe.* 2016; 20: 72–82. <https://doi.org/10.1016/j.chom.2016.06.006> PMID: 27414498
49. Lee VV, Judd LM, Jex AR, Holt KE, Tonkin CJ, Ralph SA. Direct Nanopore Sequencing of mRNA Reveals Landscape of Transcript Isoforms in Apicomplexan Parasites. *mSystems.* 2021;6. <https://doi.org/10.1128/mSystems.01081-20> PMID: 33688018
50. Xue Y, Theisen TC, Rastogi S, Ferrel A, Quake SR, Boothroyd JC. A single-parasite transcriptional atlas of Toxoplasma Gondii reveals novel control of antigen expression. *Elife.* 2020;9. <https://doi.org/10.7554/eLife.54129> PMID: 32065584
51. Bunnik EM, Venkat A, Shao J, McGovern KE, Batugedara G, Worth D, et al. Comparative 3D genome organization in apicomplexan parasites. *Proc Natl Acad Sci USA.* 2019; 116: 3183–3192. <https://doi.org/10.1073/pnas.1810815116> PMID: 30723152
52. Toenhake CG, Fraschka SA-K, Vijayabaskar MS, Westhead DR, van Heeringen SJ, Bártfai R. Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying Plasmodium falciparum Blood-Stage Development. *Cell Host Microbe.* 2018; 23: 557–569.e9. <https://doi.org/10.1016/j.chom.2018.03.007> PMID: 29649445
53. Chappell L, Ross P, Orchard L, Russell TJ, Otto TD, Berriman M, et al. Refining the transcriptome of the human malaria parasite Plasmodium falciparum using amplification-free RNA-seq. *BMC Genomics.* 2020; 21: 395. <https://doi.org/10.1186/s12864-020-06787-5> PMID: 32513207
54. Albuquerque SS, Carret C, Grosso AR, Tarun AS, Peng X, Kappe SHI, et al. Host cell transcriptional profiling during malaria liver stage infection reveals a coordinated and sequential set of biological events. *BMC Genomics.* 2009; 10: 270. <https://doi.org/10.1186/1471-2164-10-270> PMID: 19534804
55. Swierzy IJ, Händel U, Kaever A, Jarek M, Scharfe M, Schlüter D, et al. Divergent co-transcriptomes of different host cells infected with Toxoplasma gondii reveal cell type-specific host-parasite interactions. *Sci Rep.* 2017; 7: 7229. <https://doi.org/10.1038/s41598-017-07838-w> PMID: 28775382
56. Sasai M, Yamamoto M. Innate, adaptive, and cell-autonomous immunity against Toxoplasma gondii infection. *Exp Mol Med.* 2019; 51: 1–10. <https://doi.org/10.1038/s12276-019-0353-9> PMID: 31827072
57. Geurts N, Opdenakker G, Van den Steen PE. Matrix metalloproteinases as therapeutic targets in protozoan parasitic infections. *Pharmacol Ther.* 2012; 133: 257–279. <https://doi.org/10.1016/j.pharmthera.2011.11.008> PMID: 22138604
58. Velásquez ZD, Conejeros I, Larrazabal C, Kerner K, Hermosilla C, Taubert A. Toxoplasma gondii-induced host cellular cell cycle dysregulation is linked to chromosome missegregation and cytokinesis failure in primary endothelial host cells. *Sci Rep.* 2019; 9: 12496. <https://doi.org/10.1038/s41598-019-48961-0> PMID: 31467333
59. Braun L, Brenier-Pinchart M-P, Yogavel M, Curt-Varesano A, Curt-Bertini R-L, Hussain T, et al. A Toxoplasma dense granule protein, GRA24, modulates the early immune response to infection by promoting a direct and sustained host p38 MAPK activation. *J Exp Med.* 2013; 210: 2071–2086. <https://doi.org/10.1084/jem.20130103> PMID: 24043761
60. Pellegrini E, Palencia A, Braun L, Kapp U, Bougdour A, Belrhali H, et al. Structural basis for the subversion of MAP kinase signaling by an intrinsically disordered parasite secreted agonist. *Structure.* 2017; 25: 16–26. <https://doi.org/10.1016/j.str.2016.10.011> PMID: 27889209
61. Yang Y, Kim SC, Yu T, Yi Y-S, Rhee MH, Sung G-H, et al. Functional roles of p38 mitogen-activated protein kinase in macrophage-mediated inflammatory responses. *Mediators Inflamm.* 2014; 2014: 352371. <https://doi.org/10.1155/2014/352371> PMID: 24771982
62. Valère A, Garnotel R, Villena I, Guenounou M, Pinon JM, Aubert D. Activation of the cellular mitogen-activated protein kinase pathways ERK, P38 and JNK during Toxoplasma gondii invasion. *Parasite.* 2003; 10: 59–64. <https://doi.org/10.1051/parasite/2003101p59> PMID: 12669350
63. Kim L, Butcher BA, Denkers EY. Toxoplasma gondii interferes with lipopolysaccharide-induced mitogen-activated protein kinase activation by mechanisms distinct from endotoxin tolerance. *J Immunol.* 2004; 172: 3003–3010. <https://doi.org/10.4049/jimmunol.172.5.3003> PMID: 14978104

64. Becker S, Jaffe CL. Effect of protein kinase inhibitors on the growth, morphology, and infectivity of *Leishmania* promastigotes. *Parasitol Res.* 1997; 83: 273–280. <https://doi.org/10.1007/s004360050246> PMID: 9089725
65. Shaw MK. *Theileria parva* sporozoite entry into bovine lymphocytes involves both parasite and host cell signal transduction processes. *Exp Parasitol.* 1996; 84: 344–354. <https://doi.org/10.1006/expr.1996.0123> PMID: 8948324
66. Nardelli SC, Che F-Y, Silmon de Monerri NC, Xiao H, Nieves E, Madrid-Aliste C, et al. The histone code of *Toxoplasma gondii* comprises conserved and unique posttranslational modifications. *MBio.* 2013; 4: e00922–13. <https://doi.org/10.1128/mBio.00922-13> PMID: 24327343
67. Ramos M-P, Wijetunga NA, McLellan AS, Suzuki M, Greally JM. DNA demethylation by 5-aza-2'-deoxycytidine is imprinted, targeted to euchromatin, and has limited transcriptional consequences. *Epigenetics Chromatin.* 2015; 8: 11. <https://doi.org/10.1186/s13072-015-0004-x> PMID: 25806086
68. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* 2016; 32: 3047–3048. <https://doi.org/10.1093/bioinformatics/btw354> PMID: 27312411
69. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015; 12: 357–360. <https://doi.org/10.1038/nmeth.3317> PMID: 25751142
70. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009; 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
71. Wang L, Wang S, Li W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics.* 2012; 28: 2184–2185. <https://doi.org/10.1093/bioinformatics/bts356> PMID: 22743226
72. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics.* 2016; 32: 292–294. <https://doi.org/10.1093/bioinformatics/btv566> PMID: 26428292
73. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015; 31: 166–169. <https://doi.org/10.1093/bioinformatics/btu638> PMID: 25260700
74. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 2014; 42: W187–91. <https://doi.org/10.1093/nar/gku365> PMID: 24799436
75. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011; 29: 24–26. <https://doi.org/10.1038/nbt.1754> PMID: 21221095
76. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (N Y).* 2021; 2: 100141. <https://doi.org/10.1016/j.xinn.2021.100141> PMID: 34557778
77. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25: 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324> PMID: 19451168
78. Ou J, Liu H, Yu J, Kelliher MA, Castilla LH, Lawson ND, et al. ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics.* 2018; 19: 169. <https://doi.org/10.1186/s12864-018-4559-3> PMID: 29490630
79. Feng J, Liu T, Qin B, Zhang Y, Liu XS. Identifying ChIP-seq enrichment using MACS. *Nat Protoc.* 2012; 7: 1728–1740. <https://doi.org/10.1038/nprot.2012.101> PMID: 22936215
80. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 2019; 47: e47. <https://doi.org/10.1093/nar/gkz114> PMID: 30783653
81. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. *PLoS Comput Biol.* 2013; 9: e1003118. <https://doi.org/10.1371/journal.pcbi.1003118> PMID: 23950696
82. Zhu LJ, Gazin C, Lawson ND, Pagès H, Lin SM, Lapointe DS, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics.* 2010; 11: 237. <https://doi.org/10.1186/1471-2105-11-237> PMID: 20459804
83. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics.* 2011; 27: 1696–1697. <https://doi.org/10.1093/bioinformatics/btr189> PMID: 21486936
84. Schep AN, Wu B, Buenrostro JD, Greenleaf WJ. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nat Methods.* 2017; 14: 975–978. <https://doi.org/10.1038/nmeth.4401> PMID: 28825706

85. Fornes O, Castro-Mondragon JA, Khan A, van der Lee R, Zhang X, Richmond PA, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 2020; 48: D87–D92. <https://doi.org/10.1093/nar/gkz1001> PMID: 31701148
86. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics.* 2011; 27: 1017–1018. <https://doi.org/10.1093/bioinformatics/btr064> PMID: 21330290