



OPEN

Comparative transcriptional analysis identifies genes associated with the attenuation of *Theileria parva* infected cells after long-term in vitro culture

Elisabeth M. D. L. van der Heijden¹✉, Lucas Lefevre², Anton Gossner², Thomas Tzelos^{2,4}, Timothy K. Connelley^{2,3,5} & Musa A. Hassan^{2,3,5}✉

Autologous administration of attenuated *Theileria parva*-infected cells induces immunity to *T. parva* in cattle. The mechanism of attenuation, however, is largely unknown. Here, we used RNA sequencing of pathogenic and attenuated *T. parva*-infected T-cells to elucidate the transcriptional changes underpinning attenuation. We observed differential expression of several host genes, including *TRAIL*, *PD-1*, *TGF- β* and granzymes that are known to regulate inflammation and proliferation of infected cells. Importantly, many genes linked with the attenuation of the related *T. annulata*-infected cells were not dysregulated in this study. Furthermore, known *T. parva* antigens were not dysregulated in attenuated relative to pathogenic cells, indicating that attenuation is not due to enhanced immunogenicity. Overall this study suggests that attenuation is driven by a decrease in proliferation and restoration of the inflammatory profile of *T. parva*-infected cells. Additionally, it provides a foundation for future mechanistic studies of the attenuation phenotype in *Theileria*-infected cells.

Theileria parva (*T. parva*) is a tick-transmitted obligate intracellular hemoprotozoan parasite in the order of the Piroplasmidae¹. Infection with *T. parva* causes East Coast Fever (ECF), a devastating disease of high morbidity and mortality, in cattle in sub-Saharan Africa². The life-cycle of *T. parva* is complex, relying on the bovine host and a tick vector, *Rhipicephalus appendiculatus* (*R. appendiculatus*)³. Infective *T. parva* sporozoites injected into the host by blood-feeding ticks rapidly invade host lymphocytes—predominantly T- and B-cells^{4,5}—and undergo schizogony to become multinucleate schizonts. This process is associated with a cancer-like transformation and uncontrolled proliferation of schizont-infected lymphocytes⁶, which give rise to the classical symptoms associated with ECF, such as generalized lymphadenopathy, respiratory distress, fever and anorexia⁷. The schizont-induced transformation effectively immortalizes the cell, and association of the parasite with the mitotic spindle in the cytosol of the host cell ensures concomitant infection of daughter cells with each division⁸. Unlike the transformation seen, for instance in cancerous cells⁴, *Theileria*-induced transformation is reversible by compounds such as buparvaquone.

Given the proliferative phenotype of *T. parva*-infected cells, they are readily maintained in in vitro culture by serial passaging. Autologous *T. parva*-infected T-cell-lines (TpM) have been shown to be highly pathogenic when administered to naïve animals early after establishment in vitro (within ~ 7 days), leading to the development of severe clinical symptoms of ECF, and death of the animal if left untreated^{9,10}. However, after long-term culture (≥ 42 days), in vitro established autologous TpM exhibit an attenuated phenotype, inducing only mild and transient clinical signs in naïve animals which subsequently recover and gain immunity to re-infection⁹. As a result, autologous TpM can be used as an alternative strategy to the current infection-and-treatment vaccination method^{11,12}. However, the molecular processes through which the attenuated phenotype is achieved

¹Division of Infectious Diseases and Immunology, Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands. ²Division of Immunology, The Roslin Institute, University of Edinburgh, Edinburgh, UK. ³Centre for Tropical Livestock Genetics and Health, Easter Bush Campus, Edinburgh, UK. ⁴Present address: Moredun Research Institute, Pentlands Science Park, Bush Loan, Penicuik EH26 0PZ, UK. ⁵These authors contributed equally: Timothy K. Connelley and Musa A. Hassan. ✉email: e.m.d.l.vanderheijden@uu.nl; Musa.Hassan@roslin.ed.ac.uk

are largely unknown. Furthermore, it is currently not known whether this phenomenon is driven by host- or parasite-derived factors. Elucidating the mechanisms underlying TpM attenuation would further our understanding of the host–pathogen interactions and help identify potential novel targets for vaccine development and therapeutics for ECF.

Although a significant amount of work studying the attenuation of *Theileria annulata*^{13–15} has been completed, very little equivalent work has been done for *T. parva*. The goal of this study, therefore, was to characterize the transcriptomic changes accompanying *T. parva*-infected bovine T-cells transitioning from pathogenic to attenuated phenotypes. To this end, the transcriptomes of naïve CD4⁺ T-cells (control) and in vitro-derived TpMs cultured for 7 and 69 days post-infection were obtained using high-throughput RNA sequencing and analysed. We found that in both day 7 (pathogenic) and day 69 (attenuated) TpM, there was up- and downregulation of cancer- and immunity-related host genes, respectively compared to naïve CD4⁺ T-cells. However, direct comparison of the day 69 cells to day 7 cells revealed three major changes associated with attenuation; a relative reactivation of a pro-inflammatory phenotype, curtailment of proliferation and a shift in metabolism.

Materials and methods

Ethics statement

All animal experiments were completed under a license from the UK government granted under the UK Animals (Scientific Procedures) Act 1986. All animal experiments were reviewed and approved by The Roslin Institute Animal Welfare and Ethics Review Board, and carried out according to relevant guidelines and regulations. All methods are reported in accordance with ARRIVE guidelines.

Study animals

Six 4–6 months old healthy Holstein–Friesian (HF) cattle housed at the University of Edinburgh Dryden Farm were used in this study. All animals were bled for the generation of *T. parva* infected cell-lines as described below. Animals were euthanized by a Schedule 1 approach in accordance with Home Office (UK government regulations). The agent used was pentobarbital sodium, administered intra-venously.

Isolation of CD4⁺ T-cells

Blood was collected by jugular venepuncture and peripheral blood mononuclear cells (PBMC) were isolated using standard density gradient centrifugation. PBMC were suspended in standard culture medium (SCM) consisting of RPMI-1640 (Thermo Fisher Scientific, Cheshire, UK), supplemented with 10% FBS, 50mM 2-mercaptoethanol (Merck, Gillingham, UK) and 1X penicillin–streptomycin–glutamine (Thermo Fisher Scientific), and MACS sorted using the CD4 specific antibodies listed in Table 1. Briefly, PBMC were stained with 0.5µg/mL of ILA-12 antibody at 4 °C for 15 min and washed in SCM by centrifugation at 4–8 °C 300 × g for 5 min. Cells were then resuspended in SCM and incubated with anti-IgG beads at 4 °C for 10 min, washed as above, and run through an LS-column (Miltenyi Biotec, Surrey, UK) to enrich for CD4⁺ cells. A minimum purity of 95%, checked as described below, was achieved for each sample.

Theileria parva infected CD4⁺ T-cell-lines

Theileria parva infected cell-lines (TpM) were established in vitro by exposing 2×10^7 CD4⁺ T-cells in 1mL of SCM to saturating levels of fresh *T. parva* Muguga sporozoites, as previously described¹⁶. At day 7 post-infection, cell aliquots from two TpM lines were separately harvested, sequentially washed once in SCM and PBS, with centrifugation after each wash at 400 × g for 5 min. Cell viability was assessed using Trypan Blue (Thermo Fisher Scientific) and cells resuspended in PBS at 2×10^4 /mL. Two animals were inoculated with 5mL of autologous TpM cell-lines in the left side of the neck, +– 10–15 cm dorsoanterior to the prescapular lymph node (PSLN). Animals were closely monitored for the development of pathology and euthanized within 1 week post-infection due to severe disease symptoms, which confirmed pathogenicity of the day 7 TpM cell-lines. The rest of the cell-lines were cultured in 75 cm² culture flasks by serial passaging (1:4) every 2–3 days until day 69.

Antibody/Stain	Specificity	Isotype	Source
<i>MACS sorting</i>			
ILA-12	CD4	IgG2a	In-house
<i>Immunofluorescence microscopy</i>			
ILS-40	PIM (<i>T. parva</i>)	IgG2a	In-house
IgGAM-FITC	Mouse IgG, IgA & IgM	Goat polyclonal	Merck
DAPI	Nuclei	n/a	Thermofisher Scientific
<i>Flowcytometry</i>			
CC8-PE	CD4	IgG2a	Thermofisher Scientific

Table 1. Antibodies and stains used for MACS sorting, immunofluorescence microscopy and flowcytometry.

Confirmation of host cell infection and phenotype

Confirmation of infection of each TpM cell-line was performed by indirect immunofluorescent staining using the antibodies in Table 1 and visualized under a microscope. Briefly, 5×10^4 cells were loaded onto slides by cytospin centrifugation and fixed in acetone before staining with $1 \mu\text{g}/\text{mL}$ ILS-40 primary antibody (specific for polymorphic immunodominant molecule (PIM)) and subsequently an anti-murine IgGAM-FITC secondary antibody, 4,6-diamidino-2-phenylindole (DAPI) counterstain. Slides were washed with PBS twice in between stains and rinsed in dH_2O before mounting and visualisation under a fluorescent microscope. The infection rate was determined as the percentage of infected cells (PIM^+ cells $\times 100\%/\text{DAPI}^+$ host cell nuclei).

In order to confirm that TpM cell-lines consisted of CD4^+ T-cells, phenotyping was performed with flow cytometry using the antibodies in Table 1. Briefly, an aliquot of each cell-line was harvested during passage, washed, and brought to 1×10^6 cells/mL in FACS buffer (PBS containing 0.5% BSA). $100 \mu\text{L}$ of cell suspension was stained with $10 \mu\text{L}$ of PE-conjugated CC8 antibody at 4°C for 15 min. The cells were washed with FACS buffer twice by pulse centrifuging to $600 \times g$ at 4°C and resuspended in FACS buffer. Data was acquired on the FACS Aria IIIu (BD Biosciences).

Characterization of the transcriptome of TpM cell-lines

At days 7 and 69 post-infection, total RNA was extracted from 3×10^6 cells of each TpM line using the miRNeasy mini kit (Qiagen, Manchester, UK) according to the manufacturer's protocol, with an additional DNase treatment step using the TURBO DNA-free kit (Ambion, Banchory, UK). RNA was also extracted from 3×10^6 uninfected CD4^+ T-cells (day 0, controls) using the same protocols. The quantity and integrity of the resulting total RNA samples were analysed using the Qubit BR RNA assay (Thermo Fisher Scientific) and high sensitivity RNA screentape on the TapeStation (Agilent, Manchester, UK), respectively, according to the manufacturers' protocols.

Library preparation and sequencing was carried out at the Wellcome Trust Clinical Research facility at the University of Edinburgh (Scotland, UK). Briefly, the libraries were separately prepared from 500 ng of each total-RNA sample using the NEBNext Ultra II Directional RNA Library Prep kit (New England BioLabs (NEB), Hitchin, UK #7760) and the Poly-A mRNA magnetic isolation module (NEB #E7490) according to the provided protocol. Following purification, the mRNA was fragmented using divalent cations under elevated temperature, primed with random hexamers and reverse transcribed into first strand cDNA. RNA templates were removed and a replacement strand synthesised incorporating dUTP in place of dTTP to generate ds cDNA. AMPure XP beads (Beckman Coulter, High Wycombe, UK #A63881) were then used to separate the ds cDNA from the second strand reaction mix, providing blunt-ended cDNA. Multiple indexing adapters were then ligated to the ends of the ds cDNA to prepare them for hybridisation onto a flow cell, before 11 cycles of PCR were used to selectively enrich those DNA fragments that had adapter molecules on both ends and amplify the amount of DNA in the library suitable for sequencing. After amplification libraries were purified using AMPure XP beads. Libraries were quantified by fluorometry using the Qubit dsDNA HS assay and assessed for quality and fragment size using the Agilent Bioanalyser with the DNA HS Kit (Agilent, #5067-4626). Libraries were multiplexed in an equimolar pool based on Qubit and Bioanalyser assay results and spiked with PhiX Control v3 (Illumina Inc, Cambridge, UK, #FC-110-3001). Sequencing was performed on the NextSeq 2000 platform (Illumina Inc, #20038897) using NextSeq 2000 P3 Reagents (100 Cycles) (Illumina Inc, #20040559).

Bioinformatic analysis pipeline

Raw reads were subjected to quality control using FastQC (v 0.11.9)¹⁷ and subsequently trimmed of Illumina adapters and poor quality reads (using the default cut-off Phred score of 20) in TrimGalore (v 0.6.4)¹⁸, with the following parameters: --illumina --length 35 --paired. Next, the reads were pseudoaligned to Ensembl release 108¹⁹ bosTau9 (ARS-UCD1.2)²⁰ and PiroplasmDB release 60²¹ *T. parva* Muguga (GCA_000165365.1, Jun 11, 2020)²² transcripts respectively, using kallisto (v 0.46.1)²³ with default parameters and 10 bootstraps. Transcript-level abundances were imported into R²⁴ and summarized into gene-level counts using tximport (v 1.24.0)²⁵. Differential expression analysis was carried out using DESeq2 (v 1.36.0)²⁶ with log-fold shrinkage (apeglm²⁷). The Database for Annotation, Visualization and Integrated Discovery (DAVID)²⁸ was used to explore enriched gene ontology terms and clustering of KEGG^{29–31} pathways based on the lists of differentially expressed bovine genes. Next, gene set enrichment analysis (GSEA) was performed using GSEA-P (v 4.3.2)^{32,33} with "Hallmark" gene sets³⁴ from the Molecular Signature Database³⁵ on the Log₂ fold change (L2FC) pre-ranked *B. taurus* differentially expressed gene lists. Finally, leading edge (LE) analysis was performed on the significantly enriched gene sets to determine which genes contributed most to the enrichment of (multiple) gene sets. The R packages ggplot2³⁶ and ComplexHeatmap³⁷, and Graphpad Prism v9.3.1 and BioRender.com were used for visualization of the results.

Statistics and reproducibility

The level of statistical significance was set at false discovery rate (FDR) ≤ 0.05 and L2FC ≥ 2 as determined by DESeq2, or FDR ≤ 0.05 as determined by GSEA-P.

Results and discussion

Establishing a system to capture the transcriptomic profiles of bovine CD4^+ T-cells during pathogenic and attenuated phases of *Theileria parva* infection in vitro

T. parva-transformed CD4^+ T-cell-lines (TpM) from 6 healthy HF calves were established ex vivo by infection of purified CD4^+ T-cells derived from PBMC with saturating levels of *T. parva* (Muguga strain) and maintained in vitro by passaging every 2–3 days for up to 69 days post-infection. To capture the changes in the host and parasite transcriptomes that occur in the early and late stages of *T. parva* infection, high throughput RNA sequencing (RNA-seq) was performed on uninfected CD4^+ T-cells (day 0—controls), and TpM aliquots taken on

day 7 (pathogenic) and day 69 (attenuated) post-infection. Two of the six TpM lines could not be maintained in culture for the intended 69-day period. Flow cytometry and immunohistochemistry analyses performed at each time-point confirmed that $\geq 95\%$ of the cells were CD4⁺ T-cells and that on days 7 and 69, $\geq 95\%$ of the cells were infected with *T. parva* (i.e. contained *T. parva* schizonts). To verify that day 7 cultured TpM were pathogenic, 1×10^5 cells were administered to the autologous naïve animals, who subsequently developed acute and severe classical symptoms of ECF³⁸ and were euthanized within 1 week of infection ($n=2$, data not shown). From each sample, the number of reads that passed quality-control was between 31 M and 46 M, and the percentage of reads mapping uniquely to the bovine reference genome was between 70 and 87%. One TpM collected on day 69 was excluded because the RNA-seq reads showed poor alignment to both the host and parasite transcriptomes. The final dataset thus, comprised 6 ($n=6$), 6 ($n=6$), and 3 ($n=3$) biological replicates at day 0, 7 and 69 post-infection, respectively.

Transcriptional changes in bovine CD4⁺ T-cells induced by infection with *T. parva*

To study the transcriptional changes induced in T-cells following infection with *T. parva*, differential gene expression analyses of day 7 and 69 TpMs, relative to uninfected (day 0) controls was performed. Downstream analysis was restricted to 13,956 bovine genes that had at least 10 uniquely mapped RNA-seq reads in at least three biological replicates (the minimum number of biological replicates at each time-point). Most genes (13,350, 95.6%) were expressed at all time-points, with only 19, 7 and 3 genes uniquely expressed at day 0, 7 and 69, respectively, and 577 genes (4.1%) expressed at 2 out of the 3 time-points (Fig. 1a). Exploratory principal component analysis (PCA) showed that infection status, captured by PC1, was the major source of transcriptional heterogeneity among the samples (Fig. 1b).

In total, 5,713 unique genes were differentially expressed (\log_2 fold change ≥ 2 and FDR ≤ 0.05) in at least one time-point relative to the controls (Supplementary File 1). Despite the distinct and well-characterised phenotypic changes evident in TpM (i.e. transformation), compared to uninfected CD4⁺ T-cells, a majority of genes ($\sim 70\%$) were not significantly dysregulated at either day 7 or day 69. In comparison to day 0 uninfected cells, a larger proportion of genes were downregulated than upregulated in TpM, with 1275 and 2230 genes down-regulated, and 970 and 1837 genes up-regulated at day 7 and 69, respectively (Fig. 1c). This suggests infection with *T. parva* predominantly was associated with an inhibition rather than activation of specific host gene expression and biological processes.

Day 7 vs. Day 0

Preliminary observation of the most differentially expressed genes at day 7 revealed that 31 out of the 100 most downregulated DEGs, including *CIQA-C*, *IL1B*, *IL7R*, *CTLA4* and *CXCL3*, were associated with immune and inflammatory responses (Fig. 1d and Supplementary File 1), suggesting that early in infection the parasite induces an anti-inflammatory profile. Subversion of the local inflammatory milieu during the early phase of infection may contribute to the dysregulation of the primary host T-cell response that has been reported previously³⁹. This in vivo dysregulation observed in primary *T. parva* infection results in the generation of a non-protective and aberrant T-cell response, which has been proposed to facilitate the parasite reaching the critical mass required for its survival and subsequent pathogenesis. Conversely, the top 10 most upregulated genes at day 7 included a number of genes associated with cell proliferation and cancer, including *FKBP10*⁴⁰, *STC2*⁴¹ and *AK4*⁴². Of note, some of the genes associated with the TGF- β pathway—known to be dysregulated in leukocytes transformed by the related parasite *T. annulata*¹³—such as matrix metalloproteinase 9 (*MMP9*)^{14,15}, *MMP2*, *SMAD1* and *GDF10*^{43,44}, were also amongst the 100 most upregulated genes. However, the up-/down-regulation of other genes identified in a recent study of the transcriptome of *T. annulata*-infected cells (TaA)¹⁴ was not reproduced in our data (e.g. *IL21R*, *A2M* and *IGSF9B*; Supplementary File 1), suggesting there are both similarities and differences in the host genes that are manipulated by *T. parva* and *T. annulata* parasites to effect transformation of their host cells.

Day 69 vs. Day 0

Relative to day 0, more genes were differentially expressed at day 69 than at day 7; 4,067 (29%) vs 2,245 (16%), respectively (Fig. 1c)—indicating that attenuation was not predominantly associated with a reversion of the transcriptome towards a naïve CD4⁺ T-cell phenotype (Fig. 1c). However, the majority ($\pm 83\%$) of day 7 top 100 differentially expressed genes (DEGs) are also differentially expressed at day 69 (Fig. 1e; top panel—comparing the central and right panes), and the majority ($\pm 65\%$) of the top 100 DEGs at day 69 exhibit the same ‘direction’ of differential expression at day 7 (Fig. 1e; bottom panel—comparing the central and right panes). This suggests that there is a ‘core’ of transcriptional changes observed in TpM that are shared at day 7 and day 69 when compared to uninfected cells. As a consequence, similar to day 7, many top 100 DEG at day 69 were associated with either immunity (16/100) and cancer-like phenotypes (27/100), or both. For example, *ZFP36L2*, a known suppressor of regulatory T-cells (T_{regs})⁴⁵, and *PNRC1*, a putative tumor suppressor⁴⁶, were downregulated at day 69 whilst *PYCR1*, which has been shown to be important in sustaining tumors⁴⁷ was upregulated. Of note, the majority of the genes cited as examples above showed similar up/down-regulation at both day 7 and day 69 (Fig. 1 and Supplementary File 1), reflecting the broad similarities in the transcriptional changes observed at these two time-points. However, there were also clear subsets of genes that showed disparate patterns of regulation at day 7 and 69 when compared to day 0; indicating that although there were similarities in the transcriptomic profiles of the pathogenic (day 7) and attenuated (day 69) cells, there were also differences (Fig. 1e; top and bottom panels—comparing the central and right panes).

To provide a more holistic contextualisation of the transcriptional differences seen at day 7 and day 69, relative to day 0, functional enrichment analyses using DAVID²⁸ and GSEA^{32,33} was performed. At FDR ≤ 0.05 , several gene sets, including E2F targets, MYC targets (V1 and V2), glycolysis, and G2M checkpoint, as well as gene

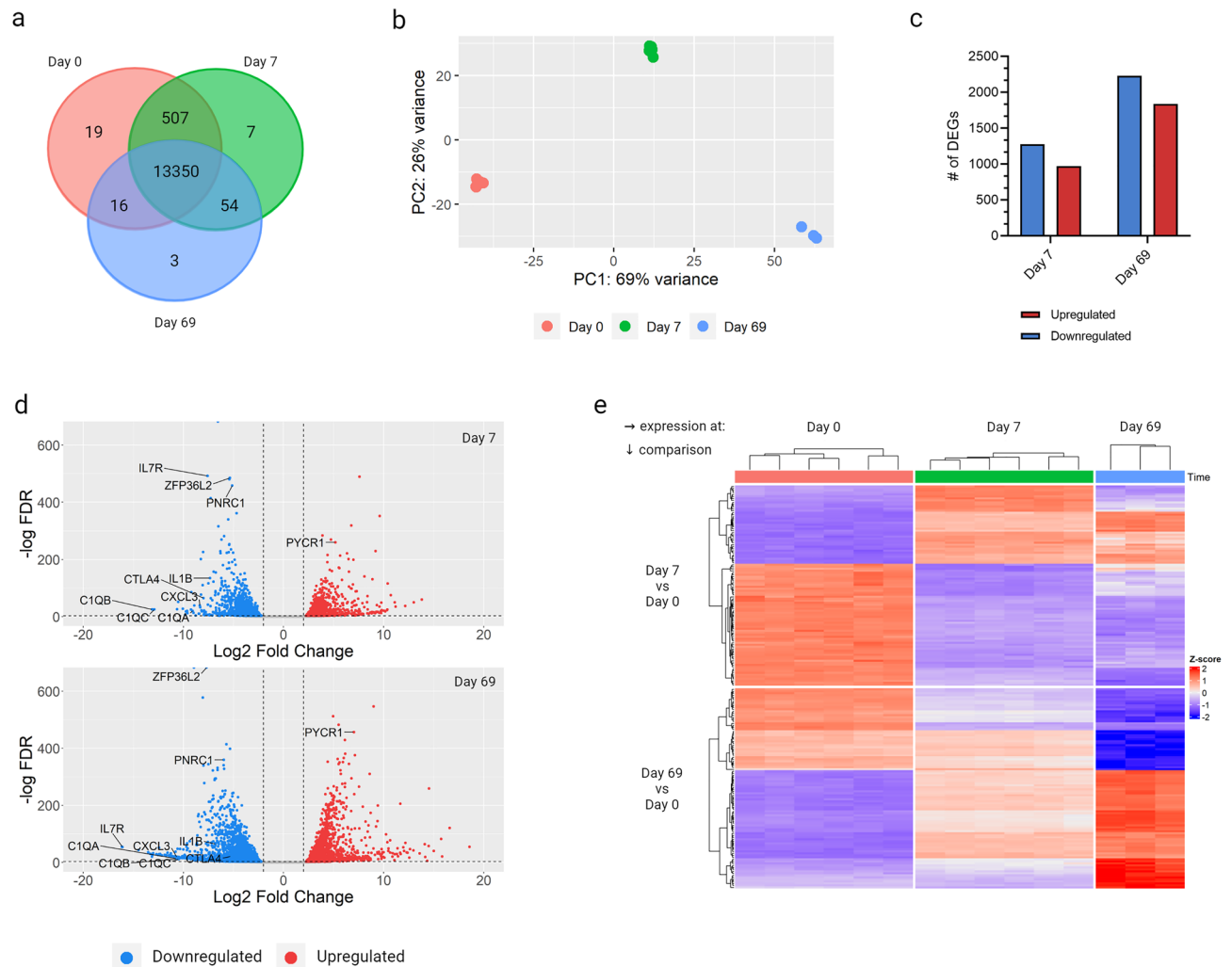


Figure 1. Expression of *Bos taurus* genes in different states of *Theileria parva* infection of bovine CD4⁺ T-cells. **(a)** The total number of expressed genes at day 0 (red), day 7 (green) and day 69 (blue). **(b)** Principal component analysis plot showing the first and second principle components; based on variance stabilized transformed counts of all expressed genes. **(c)** The number of significantly upregulated (red) or downregulated (blue) genes at day 7 and day 69, relative to day 0, as determined by DE analysis using DESeq2 (FDR ≤ 0.05 and L2FC $|\leq 2|$). **(d)** Distribution of all significantly DEGs according to the L2FC and $-\log(\text{adjusted } p\text{-value})$ at day 7 and day 69, relative to day 0. **(e)** Expression of the top 100 significantly DEGs at day 7 and day 69, relative to day 0, displayed as z-score transformed counts. Created with ComplexHeatmap v2.12.1 and ggplot2 v3.4.2 in R Studio v2023.03.1, Graphpad Prism v9.3.1 and BioRender.com.

sets related to metabolic events were positively enriched at both day 7 and day 69 (Fig. 2, Supplementary File 2, and Supplementary Figs. 1 and 2). These results are consistent with previous studies which have demonstrated that manipulation of pathways such as E2F signalling and c-Myc activation by *Theileria* are critical to the transformation of *Theileria*-infected cells^{48,49}. E2F transcription factors are crucial in cell-cycle progression and an elevation in expression of these factors has been linked to malignancy and hyper-proliferation^{50,51}; subversion of this pathway by *T. parva* has been demonstrated to stimulate cell-proliferation of the transformed cells⁴⁹. The *c-myc* gene is a well-studied proto-onco gene associated with a wide variety of malignancies, and is associated with enhanced proliferation⁵²; *Theileria*-transformed cells have been shown to be critically dependent on c-Myc for survival through inhibition of caspase-dependent apoptosis⁴⁸. Enrichment of these gene sets at both day 7 and 69, and downregulation of GO terms including negative regulation of cell proliferation and growth (data not shown) provide the transcriptomic basis for the strongly proliferative phenotype that both pathogenic and attenuated TpM exhibit. Another gene set that is enriched in both pathogenic and attenuated cells is the glycolysis pathway—presumably reflecting the heightened metabolic demands associated with rapidly proliferating cells. Conversely, several gene sets associated with the host immune response, such as TNF α signalling via NF- κ B, inflammatory response, and IL-6/JAK-STAT3 signalling, had a negative enrichment score at both day 7 and day 69 (Fig. 2 and Supplementary File 2), showing there are elements of the anti-inflammatory profile consistent between pathogenic and attenuated cells.

However, there were several gene sets that showed different regulation at day 7 and day 69. Notable amongst these was TGF- β signalling which was negatively enriched at day 69, but not day 7. Previous work in *T. annulata*

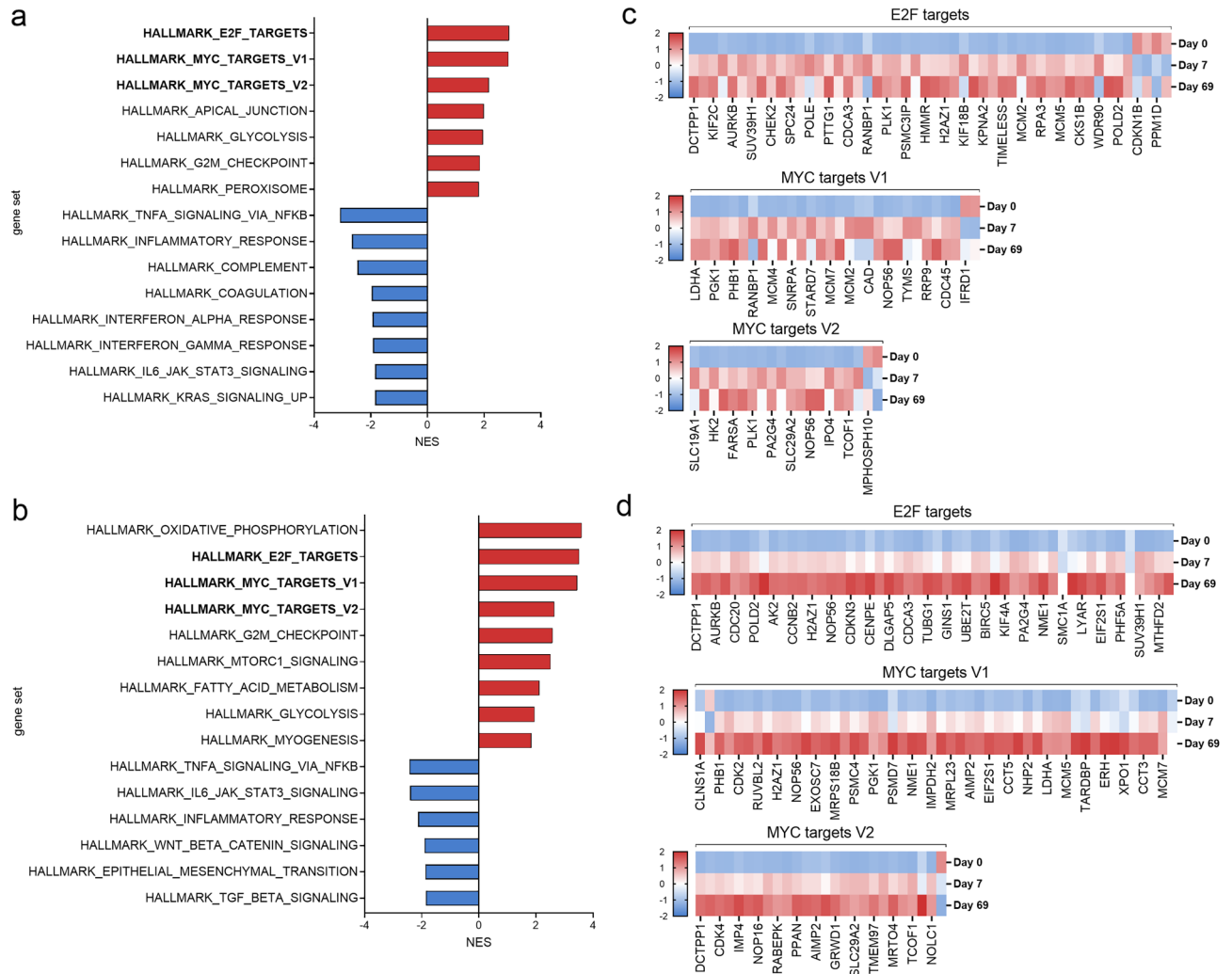


Figure 2. Gene set enrichment analysis of differentially expressed genes in pathogenic and attenuated cells. Normalized enrichment scores of gene sets that were either positively (red) or negatively (blue) significantly enriched at (a) day 7, or (b) day 69, relative to day 0. Expression of genes in selected significantly enriched gene sets at (c) day 7, or (d) day 69, relative to day 0. Expression is shown in z-scores of variance transformed counts. Only the top 50 genes of E2F targets at day 7 and day 69 and MYC targets V1 at day 69 are shown. Created with Graphpad Prism v.9.3.1.

has shown that TGF- β expression is inversely correlated with pathogenicity of transformed cells¹³ and that TGF- β signalling involves a number of other factors that are also associated with the pathogenicity of TaA (e.g. *Grb2*⁵³—which in our data also is downregulated at day 69 but not day 7—Supplementary File 1), suggesting that this mechanism may contribute to the attenuation of both *T. parva* and *T. annulata*-infected cells.

Overall, the functional enrichment analyses show that at days 7 and 69, the transcriptomic profile of *T. parva*-infected cells have many shared features including those associated with anti-inflammation and enhanced proliferation (with a commensurate increase in metabolic activity). The transcriptomic profiles are consistent with the known phenotypes of *T. parva*-transformed cells (i.e. rapid proliferation and immunomodulation) and demonstrates the general validity of utilising RNA-sequencing to analyse the transcriptional profiles underlying the phenotype of *T. parva*-induced T-cell transformation. Despite the commonality in the transcriptomes of pathogenic and attenuated *T. parva* cells there were obvious differences—to further examine this the transcriptomes of the day 7 and day 69 cells were compared directly.

Attenuation is associated with a mitigation of the anti-inflammatory profile, curtailment of proliferation and a metabolic shift

To track the changes in gene regulation as the cells transitioned from naïve CD4⁺ T-cells to pathogenic and then attenuated TpM, a Sankey plot showing genes up-, down- and non-differentially regulated at day 7 and day 69 was generated (Fig. 3a). The plot highlighted that the majority of genes that were differentially expressed at day 7 ($n = 2245$), remained regulated in the same direction at day 69 ($n = 1208$, 54%) whilst only a negligible proportion of genes ($n = 15$, 0.01%) inverted direction, indicating a substantial continuity in the transcriptomics of pathogenic and attenuated cells. While a high number of up-/down-regulated genes at day 7 become non-differentially

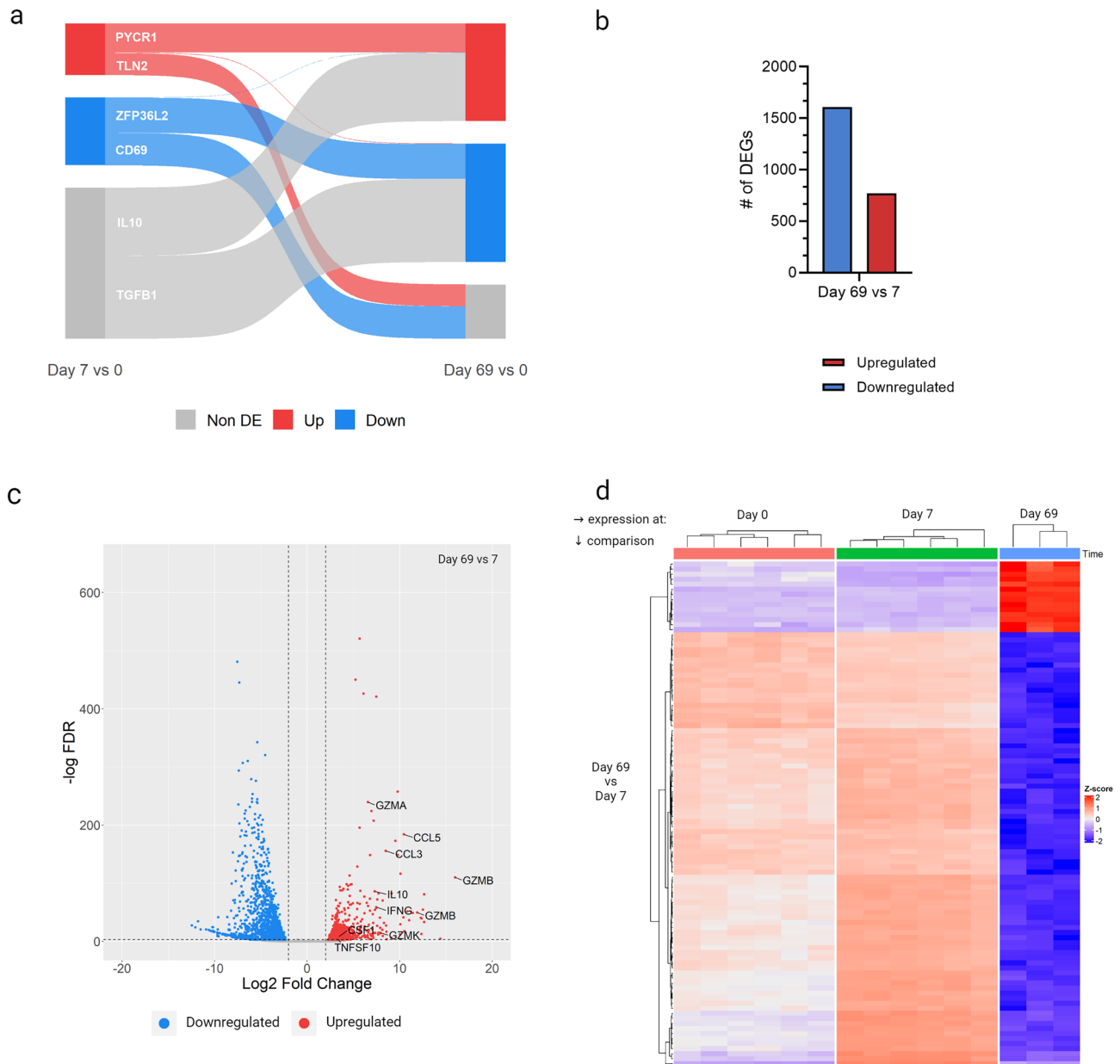


Figure 3. Direct comparison of attenuated to pathogenic *T. parva* infected cells. **(a)** Changes in directionality of genes significantly differentially expressed at either day 7 or day 69, relative to day 0, as determined by DE analysis using DESeq2 ($FDR \leq 0.05$ and $L2FC \leq 2$). **(b)** The number of significantly upregulated (red) or downregulated (blue) genes at day 69 compared to day 7. **(c)** Distribution of all significantly DEGs according to the L2FC and $-\log(\text{adjusted } p\text{-value})$ at day 69 compared to day 7. **(d)** Expression of the top 100 significantly DEGs at day 69 compared to day 7, displayed as z-score transformed counts. Created with ggscankey v0.0.9999, ComplexHeatmap v2.12.1 and ggplot v3.4.2 in R Studio v2023.03.1 and BioRender.com. Blue = downregulated; red = up-regulated; grey = non differentially expressed (DE).

regulated at day 69 ($n = 1022$, 46%), a greater number of genes ($n = 2844$) that were non-differentially regulated at day 7 became up-/down-regulated at day 69 re-iterating that attenuation is associated with a greater perturbation of the transcriptome than pathogenicity and that may be associated with disruption of more cellular pathways and processes.

In direct day 69 vs. day 7 comparison, in total 2378 genes were differentially expressed; of which 770 and 1608 were up- and down-regulated, respectively, in attenuated (day 69), compared to pathogenic (day 7) cells (Fig. 3b). Amongst the genes most upregulated in day 69 cells were a suite of genes associated with immune/inflammatory responses, including a number of chemokine ligands (*CCL1*, *CCL3* and *CCL5*), interferon- γ (*IFNG*), *IL-10*, and a number of granzymes (granzyme A, B and K) (Fig. 3c,d). The upregulation of *IL-10* at day 69, but not day 7 was unexpected as upregulation of this cytokine has previously been reported to be universal in TpM, irrespective of lymphocyte lineage, and had been proposed to be associated with the dysregulation of the primary T-cell response

discussed above³⁹. However, our data suggests that the kinetics of the induction of IL-10 is not consistent with this hypothesis, as IL-10 expression is upregulated in attenuated but not pathogenic TpM. The upregulation of the granzyme genes in day 69 TpM cells is also of particular note as in a recent study by Rchiad et al.¹⁴, granzyme A (*GZMA*) was one of only 4 regulated genes consistently associated with an attenuated phenotype in the TaA. Additionally, the reduction in *GZMA* expression in the attenuated TaA was associated with reduced TGF- β signalling, which as commented above was also identified as a downregulated pathway at day 69 but not day 7 TpM, when compared to naïve, uninfected CD4⁺ T-cells. This adds further support to at least a component of the attenuation mechanisms being shared between *T. parva* and *T. annulata* transformed cells.

Functional annotation enrichment analysis and clustering using GSEA corroborated the positive enrichment of several gene sets related to immune processes (interferon- α (IFNA) response, IFNG response, allograft rejection, IL-2/STAT5 signalling and inflammatory response) in day 69 TpM when considering the global transcriptome (Fig. 4 and Supplementary Fig. 3). The only other enriched gene sets were oxidative phosphorylation (metabolism) and cell proliferation pathways (E2F targets and MYC targets V1). Metheni et al.⁵⁴, proposed that attenuated TaA transition from the characteristic Warburg glycolysis to oxidative phosphorylation as cell proliferation slows, which our data suggests may also be happening in TpM (Fig. 4 and Supplementary File 2), drawing another parallel in the attenuation process between the two parasites. In their study, however, HIF1A was found to be up-regulated in both pathogenic and attenuated TaA as compared to negative controls⁵⁵, whereas we found HIF1A only to be up-regulated when directly comparing attenuated to pathogenic cells. The enrichment of E2F targets and MYC targets V1 in the GSEA appeared counter-intuitive as, based on previous work by Morrison et al.⁵⁶, and informal observations during general establishment and culture of TpM (unpublished), attenuated TpM were anticipated to have a reduced, rather than enhanced, proliferative phenotype. As GSEA doesn't capture the directionality or interaction of genes, functional annotation clustering using DAVID was used to further explore and confirm the observations made with GSEA. The analysis at day 69 confirmed the upregulation of functional clusters including oxidative phosphorylation (and metabolic pathways) and immune functions (such as NK cell mediated cytotoxicity and chemokine signalling) but down-regulation of clusters that include cancer-like phenotypes and VEGF, ErbB, HIF-1 and Ras signalling pathways, which are closely associated with cell growth and proliferation^{57–60} (Table 2 and Supplementary File 3). Notably, an additional functional cluster found to be down-regulated at day 69 in this analysis included the PD-1 checkpoint pathway, a crucial immune checkpoint in cancer⁶¹, suggesting that attenuation of this pathway may also be associated with a decreased capacity of attenuated TpMs to inhibit, and thus evade elimination by T-cells.

Leading edge analysis was carried out to determine the key genes that may be driving attenuation. A number of genes formed part of the core enrichment of at least 3 of the aforementioned enriched gene sets (Table 3, Fig. 5 and Supplementary File 4). Of note, colony stimulating factor 1 (*CSF1*) emerged as the gene that contributed to the enrichment of the highest number of gene sets (allograft rejection, IFN- α response, inflammatory response and IL-2/STAT5 signalling). *CSF1* drives monocytes to produce a plethora of cytokines and as such plays a key role in innate immunity and inflammation⁶², suggesting *CSF1* has a role in promoting the pro-inflammatory

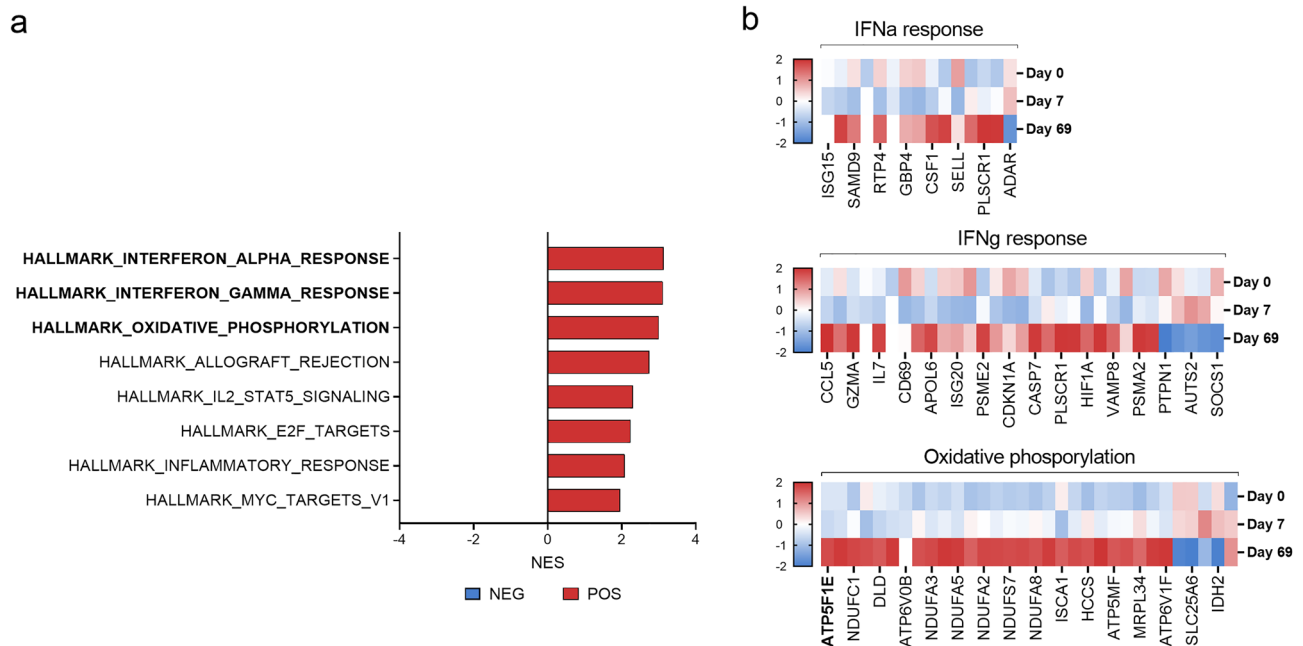


Figure 4. Gene set enrichment analysis of differentially expressed genes in attenuated vs pathogenic cells. (a) Normalized enrichment scores of gene sets that were either positively (red) or negatively (blue) significantly enriched at day 69 compared to day 7. (b) Expression of genes in selected significantly enriched gene sets day 69 compared to day 7. Expression is shown in z-scores of variance transformed counts. Created with Graphpad Prism v.9.3.1.

Cluster	Direction	Score	Majority of terms related to (number/total)
1	Down	2.95	Cancers/cell proliferation (7/9)
2	Down	2.07	T-cell differentiation / regulation (2/3)
3	Down	1.98	Cancers/cell proliferation (21/71)
			Signalling (pathways) (19/71)
4	Down	1.75	Endocrinological processes (11/15)
5	Down	1.72	Cancers/cell proliferation (2/3)
1	Up	6.60	Cell metabolism (3/13)
			Degenerative diseases (6/13)
2	Up	2.29	(Auto-) Immunity (5/6)
3	Up	2.02	Signalling (pathways) (3/3)

Table 2. Summarized functional annotation clustering results. This table includes only those clusters that were significantly enriched based on their enrichment score (the geometric mean of the negative log-adjusted p -values for each term in the cluster). An enrichment score of ≥ 1.3 was considered significant. The final column provides insight into the type of process the enriched terms in each cluster relate to most. See Supplementary File 3 for complete results. (<https://david.ncifcrf.gov/>; accessed on 25 April 2023).

Symbol	Description	Function	Number of gene sets	Gene sets (Hallmark)
<i>CSF1</i>	Colony stimulating factor 1	Regulation of survival, proliferation and differentiation of hematopoietic precursor cells and release of pro-inflammatory chemokines	4	Ar, IaR, Ir, ISs
<i>CKDN1A</i>	Cyclin dependent kinase inhibitor 1A	A regulator of cell cycle progression at G1	3	E, I γ R, Ir
<i>TNFSF10</i>	TNF Superfamily Member 10	Induces apoptosis in transformed and tumor cells	3	I γ R, Ir, ISs
<i>BST2</i>	Bone Marrow Stromal Cell Antigen 2	<i>Undetermined</i>	3	IaR, I γ R, Ir
<i>SELL</i>	Selectin L	A cell surface adhesion molecule	3	IaR, Ir, ISs
<i>IL7</i>	Interleukin 7	A cytokine important for B- and T-cell development	3	Ar, I γ R, IaR
<i>RTP4</i>	Receptor Transporter Protein 4	Involved in defense response to virus	3	IaR, I γ R, Ir
<i>IL10</i>	Interleukin 10	Cytokine with pleiotropic effects in immunoregulation and inflammation	3	Ar, Ir, ISs
<i>GBP4</i>	Guanylate Binding Protein 4	Induced by interferon, hydrolyses GTP to GDP and GMP	3	IaR, I γ R, ISs
<i>CCL5</i>	C-C Motif Chemokine Ligand 5	Chemoattractant for blood monocytes, T helper cells and eosinophils	3	Ar, I γ R, Ir

Table 3. Leading edge analysis of the significantly enriched gene sets. Leading edge analysis was performed only on those gene sets that were significantly enriched as determined by the pre-ranked GSEA. Only genes that formed part of the leading edge in ≥ 3 gene sets are listed in this table. The order of genes is descending according to number of gene sets they enriched. Description and function of genes derived from GeneCards—the human gene database (<https://www.genecards.org/>)^{69,70}. Gene sets = each gene set in which the gene formed part of the core enrichment. Gene sets: Ar = allograft rejection; E = E2F targets; IaR = IFN α response; I γ R = IFN γ response; Ir = inflammatory response; ISs = IL2/STAT5 signalling.

profile of attenuated TpM (Fig. 5a). However, previous work has shown GM-CSF (closely related to CSF1) to be expressed in TpM and part of an autocrine loop that can contribute to TpM proliferation⁶³; such evidence points to the potentially contradictory roles that families of pleiotropic cytokines may play in determining the phenotypes of TpM. Similarly, interpretation of how IL10, a cytokine with multiple functions may influence TpM attenuation is complex due to evidence that in addition to its well-known anti-inflammatory roles, it can potentiate anti-cancer CD8⁺ T-cell responses^{64–66} and so may either enhance and/or diminish the immune response against day 69 TpM and contribute to or mitigate the attenuation phenotype (Fig. 5b). Cyclin-dependent kinase inhibitor 1A (*CKDN1A*—also known as p21), was among the leading edge of 5 enriched gene sets in our data. p21 is a known regulator of cell cycle progression, and has previously been described as having tumour suppressive properties, inducing anti-proliferative signals and senescence⁶⁷ (Fig. 5c). Therefore, upregulation of *CKDN1A* in attenuated cells may contribute to a mitigation of the proliferative state. However, similar to *CSF1* and *IL-10*, p21 has also been reported to have contradictory effects, with proto-oncogenic activity also having been observed⁶⁷. The TNF Superfamily Member 10 (*TNFSF10*—also known as tumour necrosis factor-related apoptosis-inducing ligand (TRAIL)), which was enriched in 4 of the top gene sets, has been shown to induce apoptotic cell death in cancerous and transformed cells⁶⁸ and so, unlike the other genes discussed here, there is an unambiguous potential to contribute to the attenuation of TpM at day 69 (Fig. 5d).

Prolonged in vitro culture induces alterations in *T. parva* gene expression

To explore if the transition to attenuation of the TpM was associated with changes in *T. parva* transcriptome, we compared the gene expression patterns observed at day 7 and day 69. After quality control, an average of 1.4 (3.3%) and 3.1M (10.0%) reads uniquely aligned to the *T. parva* reference transcriptome in pathogenic (day

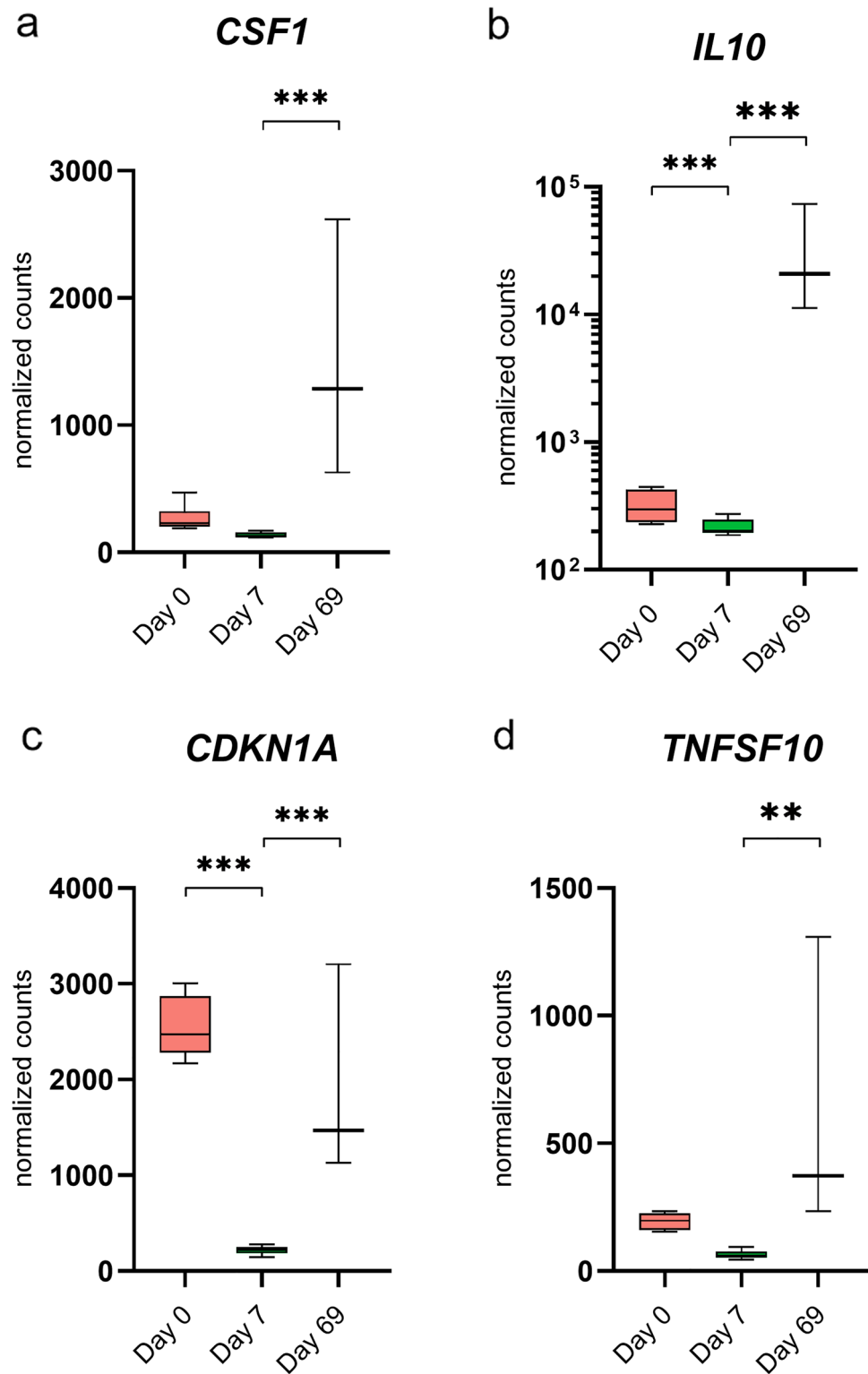


Figure 5. Expression of specific host genes identified in leading edge analysis. Box and whisker plots depicting the min–max values of normalized counts of *Bos taurus* genes (a) *CSF1*; (b) *IL10*; (c) *CDKN1A*; and (d) *TNFSF10*; expressed at day 0, day 7 and day 69. ** p -value ≤ 0.01 ; *** p -value ≤ 0.001 . Normalized counts of *IL10* are depicted on a log-scale Y-axis. Created with Graphpad Prism v.9.3.1.

7) and attenuated (day 69) TpM, respectively. Of the 4129 defined *T. parva* genes²², 3625 genes (> 87%) were represented in the transcriptome dataset (with a minimum of 10 unique mapped reads per gene). Most of the expressed genes overlapped between day 7 and day 69, the only exception being 6 genes uniquely expressed in pathogenic cells (Fig. 6a and Supplementary File 5). As with analysis of the bovine genes, PC analysis of the *T. parva* transcriptomes revealed clustering of samples according to days post-infection (Fig. 6b). Of the 3625 genes analysed only 118 (3%) genes were significantly differentially expressed (FDR ≤ 0.05 and L2FC of ≥ 2). The majority (90/118) of DEGs were upregulated in attenuated cells, whilst only 28 DEGs were downregulated (Fig. 6c). A hierarchical clustering heatmap of all DEGs highlighted that the DEGs were remarkably consistent across all samples at the same time-point (Fig. 6e).

Most of the characterised *T. parva* antigens (including Tp1-10, Tp12-36, p32 and p67)⁷¹⁻⁷³ were non-differentially expressed at day 7 and day 69—the only exceptions were Tp3 (TpMuguga_01g00868) and p32 (TpMuguga_01g01056), which were both upregulated at day 69 (Figs. 6d, 7a and b, and Supplementary File 6). The upregulation of p32, which is a merozoite surface antigen⁷⁴, at day 69 fits in with the biological life-cycle of the parasite (i.e. upregulation of a merozoite antigen would be anticipated to occur late, rather than early, in the schizont phase), as does the absence of p67 expression at either time-point, which has previously been considered a sporozoites-specific antigen⁷⁵. New evidence by Tonui, et al.⁷⁶ is now challenging this view, who did find low levels of expression of p67 in the schizont, which is in conflict with our findings. The finding that most antigens

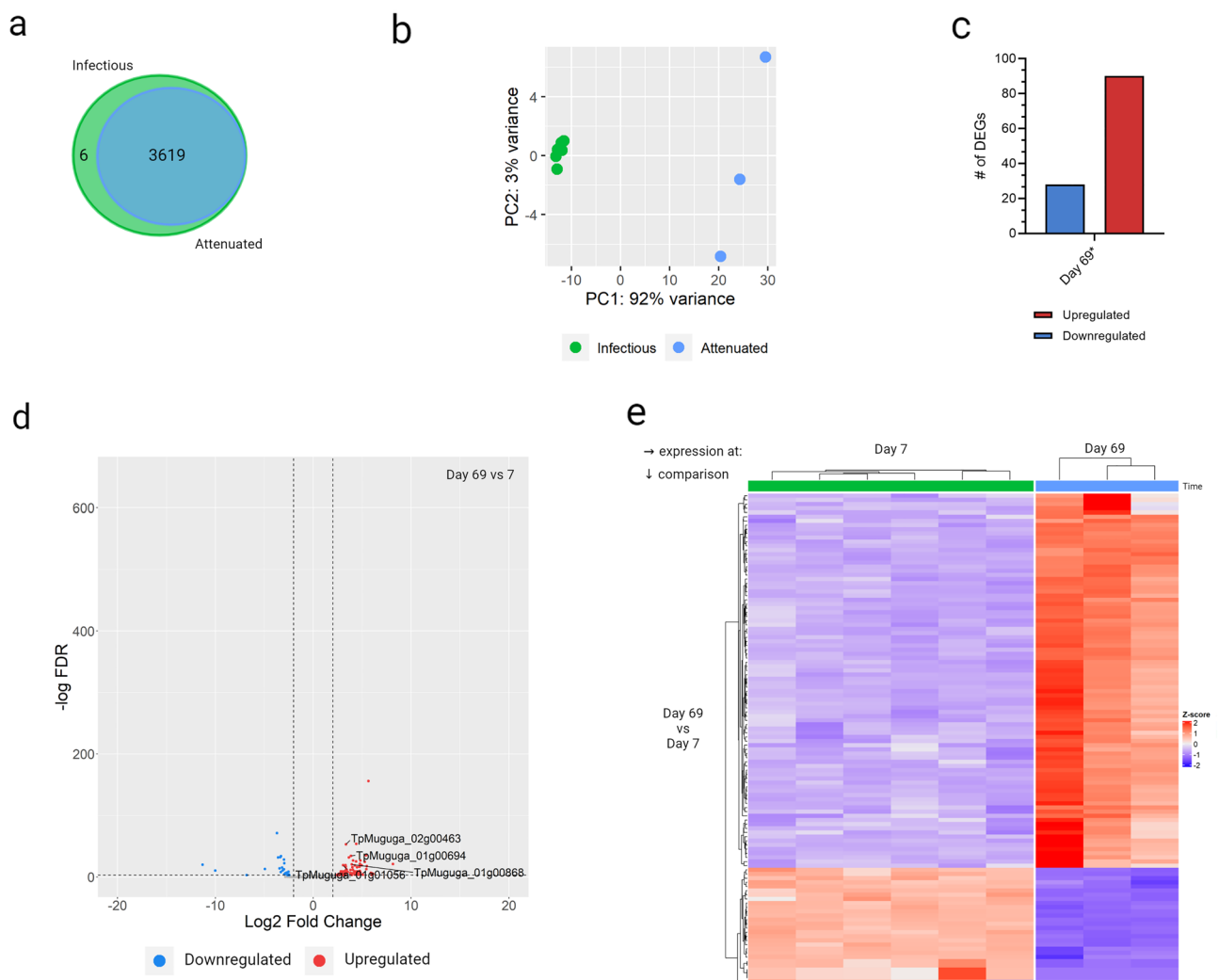


Figure 6. Expression of *Theileria parva* genes in different states of *T. parva* infection of bovine CD4⁺ T-cells. **(a)** The total number of expressed genes at day 7 (green) and day 69 (blue). **(b)** Principal component analysis plot showing the first and second principle components; based on variance stabilized transformed counts of all expressed genes. **(c)** The number of significantly upregulated (red) or downregulated (blue) genes day 69 compared to day 7, as determined by DE analysis using DESeq2 (FDR ≤ 0.05 and L2FC ≤ 2). **(d)** Distribution of all significantly DEGs according to the L2FC and $-\log(\text{adjusted p-value})$ at day 69 compared to day 7. **(e)** Expression of the all significantly DEGs at day 69 compared to day 7, displayed as z-score transformed counts. Created with ComplexHeatmap v2.12.1 and ggplot2 v3.4.2 in R Studio v2023.03.1, Graphpad Prism v9.3.1 and BioRender.com.

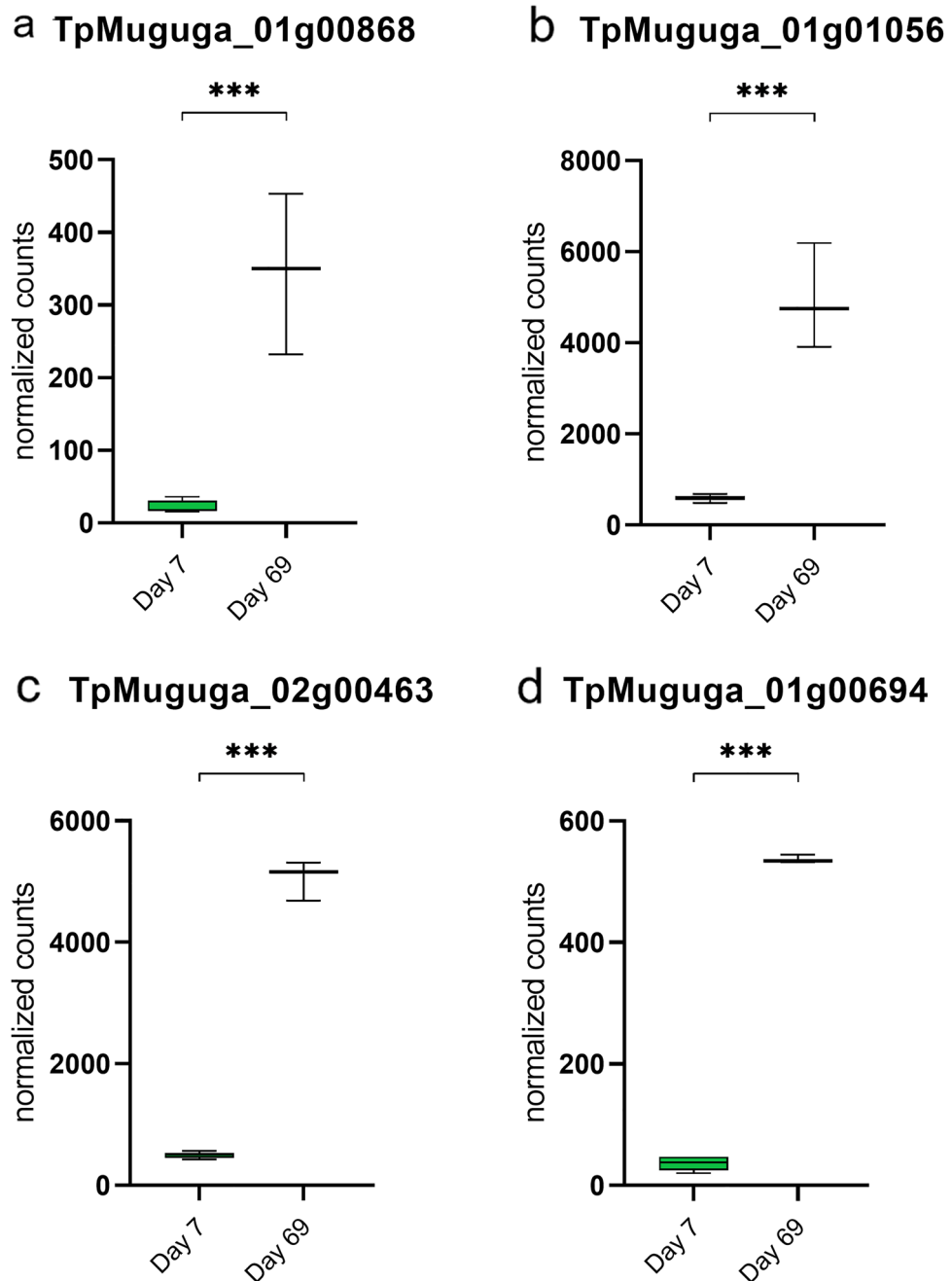


Figure 7. Expression of specific parasite genes in significantly enriched pathways or gene sets in attenuated cells. Box and whisker plots depicting the min–max values of normalized counts of *Theileria parva* genes (a) TpMuguga_01g00868; (b) TpMuguga_01g01056; (c) TpMuguga_02g00463; and (d) TpMuguga_01g00694; expressed at day 7 and day 69. ***p-value ≤ 0.001 . Created with Graphpad Prism v.9.3.1.

were expressed at equivalent levels at day 7 and day 69 suggests that differential expression of antigens (and so potential exposure to *T. parva*-specific B-cell or T-cell responses) was not evidently a mechanism contributing to attenuation of the TpM.

Of the 118 differentially regulated genes there were two major protein groupings. The first of these were genes encoding ribosomal proteins (n = 9), all of which were downregulated at day 69; notably 3/6 of the *T. parva* genes that were expressed at day 7 but not at day 69 also encoded ribosomal proteins (Supplementary File 5). Therefore, ribosomal downregulation, probably reflecting a relative decrease in parasite protein production in attenuated cells, was a general feature of day 69 cells. The second group was membrane proteins (n = 22, inner membrane complex proteins/(putative) integral membrane proteins) of which all but one were upregulated at day 69; this probably reflects the increase in the size of the schizonts in infected cells later in the infection, the consequent increase in the surface area of the schizonts, and so a need for greater quantities of membrane proteins. It is

likely that the relative downregulation of ribosomes and upregulation of membrane proteins seen at day 69 are consequences of attenuation rather than changes that drive the attenuation phenotype.

Of the remaining 87 differentially expressed genes only approximately half ($n = 43$) encoded annotated proteins, none of which could be easily linked to functional activities that may be related to attenuation. However, upregulation of *RCD1* (TpMuguga_02g00463) and a histidine protein (TpMuguga_01g00694) (Figs. 6d, 7c and d), both of which are involved in the control of reactive oxygen species (ROS) in plants^{77,78}, may be linked to the oxidative stress and subsequent altered metabolic state of the cells as observed in the host transcriptome. Due to the small number of annotated differentially expressed *T. parva* genes, functional analysis using the Gene Ontology Enrichment tool on PiroplasmaDB (VeuPathDB²¹) did not provide additional insight (data not shown). The remaining 44 differentially regulated ‘hypothetical proteins’ (38 upregulated and 6 downregulated at day 69), for which there is obviously no implied functional information available, may therefore be of particular interest, as it is feasible to suppose that the proteins contributing to attenuation may be encoded by genes within this set of uncharacterised genes. In order to further investigate possible functions of these proteins, the search tool on PiroplasmDB (VeuPathDB²¹) was used to identify possible orthologs in *T. annulata*. Only 9 of the unannotated *T. parva* genes were annotated in *T. annulata*, but their functions did not provide additional clarity on their relevance in the present study (Supplementary File 7). To provide clarity on this, additional work applying novel approaches to provide some characterisation of the structure, orthology and/or function of these genes, would be required^{79,80}.

Limitations of the study

At the commencement of this study there was no transcriptional data profiling attenuated vs pathogenic *T. parva*-infected cells, and the rationale of this study was thus to fill this knowledge gap. Due to the fragility of early passage *T. parva* cells, meaning they can't be reliably cryopreserved, it was not possible to retrospectively perform phenotypic analyses, based on the transcriptome data, of the cells at day 7. However, the data from this study provides a strong platform to select and integrate relevant phenotypic assays in future iterations of the experiments, which are currently being planned. Thus, whilst phenotypic confirmation of the traits implied by the transcriptional profiles was not possible within this study, they will form core components of future work.

As a consequence of the SARS-CoV-2 pandemic the animals involved in this study were culled prior to day 69 of the experiment and so it was not possible to administer the autologous cell-lines in vivo at the day 69 timepoint. Although this precluded our ability to empirically confirm that the cells were attenuated at this time, we are confident that the cells would have had an attenuated phenotype. This is based on additional information from Prof. W.I. Morrison that *T. parva*-infected cell-lines are reliably attenuated within 42 days of passage (personal communication—which extends the observations made in published data⁹) and our own experience where in vitro established cell-lines are used as a means of immunizing animals without the need for concurrent treatment. We are aware that this timeline is distinct from that observed in *T. annulata*, where a longer period of in vitro culture is required to guarantee attenuation.

Conclusion

Our data indicates that attenuation of *T. parva*-infected cells is associated with a decrease in proliferation and a restoration of an inflammatory profile, as well as a change in the metabolism of TpM. These transcriptomic changes broadly fit with observations made on the phenotype of TpM and can be rationalised as contributing to attenuation (i.e. it is intuitive that a reduction in proliferation and a mitigation of the anti-inflammatory profile may facilitate attenuation). Perhaps surprisingly, attenuation was associated with an increase in the perturbation of the transcriptomic landscape of the host cell, with many more genes differentially regulated at day 69 than at day 7, suggesting attenuation is a progressive process with a greater divergence from, rather than a partial reversion to, the transcriptome of uninfected T-cells.

Some of the differentially regulated host genes identified in this study (e.g. *TRAIL*, *PD-1*, *TGF- β* , granzymes), are particularly attractive candidates for further exploration, as they have also been identified in attenuated *T. annulata*-infected cell-lines. The finding that only a small fraction of the *T. parva* genes were differentially regulated between the pathogenic and attenuated TpM is encouraging and provides a restricted number of genes that are candidates for further exploration. Which of these host and pathogen genes (or pathways) drive attenuation requires further functional and transcriptomic studies to provide better resolution. Fortunately, the *T. parva* model provides a number of opportunities for further comparative transcriptomic analyses. For example comparison of B-cell TpM (which are effectively attenuated very early after in vitro infection) with T-cell TpM (and even T-cell subsets). The fact that both *T. parva* and *T. annulata* can infect B-cells offers opportunities for direct cross-species comparisons; notably, some of the transcriptomic changes seen in attenuated TpM are similar to those observed in TaA (e.g. granzyme upregulation in attenuated cells), whilst others appear to diverge between the species. Thus, although transcriptional changes cannot be extrapolated across these two parasites (*T. parva* and *T. annulata*) a better understanding of the commonalities and differences could provide novel molecular insights into the attenuation process. Comparison of transcriptomic differences of *T. parva*-infected cells derived from animals of the recently identified *T. parva* tolerant family lineage⁸¹ from those derived from susceptible animals may be interesting to see if tolerance is mediated through similar gene regulation as that observed with attenuation (notably tolerance has been associated with a reduced proliferative capacity of TpM⁸²). Combining greater understanding of transcriptomic changes associated with attenuation and novel approaches to genetic manipulation will offer new insights on the host/pathogen interaction and lead to new methods that can be exploited to control these important livestock pathogens.

Data availability

The raw and processed data are available through GEO accession number GSE245180.

Received: 1 February 2024; Accepted: 8 April 2024

Published online: 18 April 2024

References

- Dobbelaere, D. & Heussler, V. Transformation of leukocytes by *Theileria parva* and *T. annulata*. *Annu. Rev. Microbiol.* **53**, 1–42. <https://doi.org/10.1146/annurev.micro.53.1.1> (1999).
- Neitz, W. O. Theileriosis, gonderioses and cytauxzoonoses: A review. *Onderstepoort. J. Vet. Res.* **27**, 275–326 (1957).
- Nene, V. *et al.* The biology of *Theileria parva* and control of East Coast fever—Current status and future trends. *Ticks Tick Borne Dis.* **7**, 549–564. <https://doi.org/10.1016/j.ttbdis.2016.02.001> (2016).
- Dobbelaere, D. A. & Rottenberg, S. *Theileria*-induced leukocyte transformation. *Curr. Opin. Microbiol.* **6**, 377–382. [https://doi.org/10.1016/s1369-5274\(03\)00085-7](https://doi.org/10.1016/s1369-5274(03)00085-7) (2003).
- Emery, D. L., MacHugh, N. D. & Morrison, W. I. *Theileria parva* (Muguga) infects bovine T-lymphocytes in vivo and induces coexpression of BoT4 and BoT8. *Parasite Immunol.* **10**, 379–391. <https://doi.org/10.1111/j.1365-3024.1988.tb00228.x> (1988).
- Dobbelaere, D. A., Fernandez, P. C. & Heussler, V. T. *Theileria parva*: Taking control of host cell proliferation and survival mechanisms. *Cell. Microbiol.* **2**, 91–99. <https://doi.org/10.1046/j.1462-5822.2000.00045.x> (2000).
- Fry, L. M. *et al.* East coast fever caused by *Theileria parva* is characterized by macrophage activation associated with vasculitis and respiratory failure. *PLoS One* **11**, e0156004. <https://doi.org/10.1371/journal.pone.0156004> (2016).
- Hulliger, L., Wilde, K. H., Brown, C. G. & Turner, L. Mode of multiplication of theileria in cultures of bovine lymphocytic cells. *Nature* **203**, 728–730. <https://doi.org/10.1038/203728a0> (1964).
- Morrison, I., MacHugh, N. & Lalor, P. A. Pathogenicity of *Theileria parva* is influenced by the host cell type infected by the parasite. *Infect. Immun.* **64**, 557–562 (1996).
- Buscher, G., Morrison, W. I. & Nelson, R. T. Titration in cattle of infectivity and immunogenicity of autologous cell lines infected with *Theileria parva*. *Vet. Parasitol.* **15**, 29–38. [https://doi.org/10.1016/0304-4017\(84\)90107-9](https://doi.org/10.1016/0304-4017(84)90107-9) (1984).
- Brown, C. G. D. in *Advances in the Control of Theileriosis: Proceedings of an International Conference held at the International Laboratory for Research on Animal Diseases in Nairobi, 9–13th February, 1981* A. D. Irvin, M. P. Cunningham, & A. S. Young (eds) 104–119 (Springer Netherlands, 1981).
- Nene, V. & Morrison, W. I. Approaches to vaccination against *Theileria parva* and *Theileria annulata*. *Parasite Immunol.* **38**, 724–734. <https://doi.org/10.1111/pim.12388> (2016).
- Chaussepied, M. *et al.* TGF- β 2 induction regulates invasiveness of *Theileria*-transformed leukocytes and disease susceptibility. *PLoS Pathog.* **6**, e1001197. <https://doi.org/10.1371/journal.ppat.1001197> (2010).
- Rchiad, Z. *et al.* Novel tumour suppressor roles for GZMA and RASGRP1 in *Theileria annulata*-transformed macrophages and human B lymphoma cells. *Cell Microbiol.* **22**, e13255. <https://doi.org/10.1111/cmi.13255> (2020).
- Adamson, R. & Hall, R. A role for matrix metalloproteinases in the pathology and attenuation of *Theileria annulata* infections. *Parasitol. Today* **13**, 390–393. [https://doi.org/10.1016/s0169-4758\(97\)01127-7](https://doi.org/10.1016/s0169-4758(97)01127-7) (1997).
- Rocchi, M. S., Ballingall, K. T., MacHugh, N. D. & McKeever, D. J. The kinetics of *Theileria parva* infection and lymphocyte transformation in vitro. *Int. J. Parasitol.* **36**, 771–778. <https://doi.org/10.1016/j.ijpara.2006.02.017> (2006).
- FastQC: A Quality Control Tool for High Throughput Sequence Data (2010).
- Trim Galore! A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. (2012).
- Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res.* **50**, D988–D995. <https://doi.org/10.1093/nar/gkab1049> (2022).
- Rosen, B. D. *et al.* De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience* <https://doi.org/10.1093/gigascience/giaa021> (2020).
- Amos, B. *et al.* VEuPathDB: The eukaryotic pathogen, vector and host bioinformatics resource center. *Nucleic Acids Res.* **50**, D898–D911. <https://doi.org/10.1093/nar/gkab929> (2022).
- Tretina, K. *et al.* Re-annotation of the *Theileria parva* genome refines 53% of the proteome and uncovers essential components of N-glycosylation, a conserved pathway in many organisms. *BMC Genom.* **21**, 279. <https://doi.org/10.1186/s12864-020-6683-0> (2020).
- Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527. <https://doi.org/10.1038/nbt.3519> (2016).
- R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2021).
- Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521. <https://doi.org/10.12688/f1000research.7563.2> (2015).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
- Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: Removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092. <https://doi.org/10.1093/bioinformatics/bty895> (2019).
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57. <https://doi.org/10.1038/nprot.2008.211> (2009).
- Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
- Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**, 1947–1951. <https://doi.org/10.1002/pro.3715> (2019).
- Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res.* **51**, D587–D592. <https://doi.org/10.1093/nar/gkac963> (2023).
- Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550. <https://doi.org/10.1073/pnas.0506580102> (2005).
- Subramanian, A., Kuehn, H., Gould, J., Tamayo, P. & Mesirov, J. P. GSEA-P: A desktop application for Gene Set Enrichment Analysis. *Bioinformatics* **23**, 3251–3253. <https://doi.org/10.1093/bioinformatics/btm369> (2007).
- Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425. <https://doi.org/10.1016/j.cels.2015.12.004> (2015).
- Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740. <https://doi.org/10.1093/bioinformatics/btr260> (2011).
- Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag, 2009).
- Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313> (2016).
- Irvin, A. D. & Mwamachi, D. M. Clinical and diagnostic features of East Coast fever (*Theileria parva*) infection of cattle. *Vet. Rec.* **113**, 192–198. <https://doi.org/10.1136/vr.113.9.192> (1983).

39. Houston, E. F. *et al.* Infection of cattle with *Theileria parva* induces an early CD8 T cell response lacking appropriate effector function. *Int. J. Parasitol.* **38**, 1693–1704. <https://doi.org/10.1016/j.ijpara.2008.05.014> (2008).
40. Wang, R. G. *et al.* FKBP10 functioned as a cancer-promoting factor mediates cell proliferation, invasion, and migration via regulating PI3K signaling pathway in stomach adenocarcinoma. *Kaohsiung J. Med. Sci.* **36**, 311–317. <https://doi.org/10.1002/kjm2.12174> (2020).
41. Qie, S. & Sang, N. Stanniocalcin 2 (STC2): A universal tumour biomarker and a potential therapeutical target. *J. Exp. Clin. Cancer Res.* **41**, 161. <https://doi.org/10.1186/s13046-022-02370-w> (2022).
42. Klepinin, A. *et al.* Adenylate kinase and metabolic signaling in cancer cells. *Front. Oncol.* **10**, 660. <https://doi.org/10.3389/fonc.2020.00660> (2020).
43. Baster, Z., Li, L., Rajfur, Z. & Huang, C. Talin2 mediates secretion and trafficking of matrix metalloproteinase 9 during invadopodium formation. *Biochim. Biophys. Acta Mol. Cell Res.* **1867**, 118693. <https://doi.org/10.1016/j.bbamcr.2020.118693> (2020).
44. Quintanilla, M., del Castillo, G., Kocić, J. & Santibáñez, J. F. 1–37 (Nova Science Publishers, 2012).
45. Makita, S. *et al.* RNA-binding protein ZFP36L2 downregulates helios expression and suppresses the function of regulatory T cells. *Front Immunol.* **11**, 1291. <https://doi.org/10.3389/fimmu.2020.01291> (2020).
46. Gaviraghi, M. *et al.* Tumor suppressor PNR1 blocks rRNA maturation by recruiting the decapping complex to the nucleolus. *EMBO J.* <https://doi.org/10.15252/embj.201899179> (2018).
47. Westbrook, R. L. *et al.* Proline synthesis through PYCR1 is required to support cancer cell proliferation and survival in oxygen-limiting conditions. *Cell Rep.* **38**, 110320. <https://doi.org/10.1016/j.celrep.2022.110320> (2022).
48. Dessauge, F. *et al.* c-Myc activation by *Theileria* parasites promotes survival of infected B-lymphocytes. *Oncogene* **24**, 1075–1083. <https://doi.org/10.1038/sj.onc.1208314> (2005).
49. Tretina, K. *et al.* *Theileria* parasites subvert E2F signaling to stimulate leukocyte proliferation. *Sci. Rep.* **10**, 3982. <https://doi.org/10.1038/s41598-020-60939-x> (2020).
50. Ren, B. *et al.* E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev.* **16**, 245–256. <https://doi.org/10.1101/gad.949802> (2002).
51. Segeren, H. A. *et al.* Excessive E2F transcription in single cancer cells precludes transient cell-cycle exit after DNA damage. *Cell Rep.* **33**, 108449. <https://doi.org/10.1016/j.celrep.2020.108449> (2020).
52. Dang, C. V. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol. Cell. Biol.* **19**, 1–11. <https://doi.org/10.1128/MCB.19.1.1> (1999).
53. Haidar, M. *et al.* TGF- β 2 induces Grb2 to recruit PI3-K to TGF-RII that activates JNK/AP-1 signaling and augments invasiveness of *Theileria*-transformed macrophages. *Sci. Rep.* **5**, 15688. <https://doi.org/10.1038/srep15688> (2015).
54. Metheni, M., Lombes, A., Bouillaud, F., Batteux, F. & Langsley, G. HIF-1 α induction, proliferation and glycolysis of *Theileria*-infected leukocytes. *Cell. Microbiol.* **17**, 467–472. <https://doi.org/10.1111/cmi.12421> (2015).
55. Metheni, M. *et al.* The level of H(2)O(2) type oxidative stress regulates virulence of *Theileria*-transformed leukocytes. *Cell. Microbiol.* **16**, 269–279. <https://doi.org/10.1111/cmi.12218> (2014).
56. Morrison, W. I. *et al.* *Theileria parva*: Kinetics of infection in the lymphoid system of cattle. *Exp. Parasitol.* **52**, 248–260. [https://doi.org/10.1016/0014-4894\(81\)90080-1](https://doi.org/10.1016/0014-4894(81)90080-1) (1981).
57. Masoud, G. N. & Li, W. HIF-1 α pathway: Role, regulation and intervention for cancer therapy. *Acta Pharm. Sin. B* **5**, 378–389. <https://doi.org/10.1016/j.apsb.2015.05.007> (2015).
58. Goel, H. L. & Mercurio, A. M. VEGF targets the tumour cell. *Nat. Rev. Cancer* **13**, 871–882. <https://doi.org/10.1038/nrc3627> (2013).
59. Wang, Z. ErbB receptors and cancer. *Methods Mol. Biol.* **1652**, 3–35. https://doi.org/10.1007/978-1-4939-7219-7_1 (2017).
60. Gimple, R. C. & Wang, X. RAS: Striking at the core of the oncogenic circuitry. *Front. Oncol.* **9**, 965. <https://doi.org/10.3389/fonc.2019.00965> (2019).
61. Han, Y., Liu, D. & Li, L. PD-1/PD-L1 pathway: Current researches in cancer. *Am. J. Cancer Res.* **10**, 727–742 (2020).
62. Chitu, V. & Stanley, E. R. Colony-stimulating factor-1 in immunity and inflammation. *Curr. Opin. Immunol.* **18**, 39–48. <https://doi.org/10.1016/j.coi.2005.11.006> (2006).
63. Baumgartner, M. *et al.* Constitutive PI3-K activity is essential for proliferation, but not survival, of *Theileria parva*-transformed B cells. *Cell. Microbiol.* **2**, 329–339. <https://doi.org/10.1046/j.1462-5822.2000.00062.x> (2000).
64. Guo, Y. *et al.* Metabolic reprogramming of terminally exhausted CD8(+) T cells by IL-10 enhances anti-tumor immunity. *Nat. Immunol.* **22**, 746–756. <https://doi.org/10.1038/s41590-021-00940-2> (2021).
65. Saxton, R. A. *et al.* Structure-based decoupling of the pro- and anti-inflammatory functions of interleukin-10. *Science* **371**, 6535. <https://doi.org/10.1126/science.abc8433> (2021).
66. Oft, M. Immune regulation and cytotoxic T cell activation of IL-10 agonists—Preclinical and clinical experience. *Semin. Immunol.* **44**, 101325. <https://doi.org/10.1016/j.smim.2019.101325> (2019).
67. Abbas, T. & Dutta, A. p21 in cancer: Intricate networks and multiple activities. *Nat. Rev. Cancer* **9**, 400–414. <https://doi.org/10.1038/nrc2657> (2009).
68. Manzo, F. *et al.* TNF-related apoptosis-inducing ligand: Signalling of a “smart” molecule. *Int. J. Biochem. Cell Biol.* **41**, 460–466. <https://doi.org/10.1016/j.biocel.2007.12.012> (2009).
69. GeneCards – the human gene database. <https://www.genecards.org/>.
70. Stelzer, G. *et al.* The GeneCards suite: From gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinform.* <https://doi.org/10.1002/cpb.5> (2016).
71. Graham, S. P. *et al.* *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3286–3291. <https://doi.org/10.1073/pnas.0511273103> (2006).
72. Morrison, W. I., Connelley, T., Hemmink, J. D. & MacHugh, N. D. Understanding the basis of parasite strain-restricted immunity to *Theileria parva*. *Annu. Rev. Anim. Biosci.* **3**, 397–418. <https://doi.org/10.1146/annurev-animal-022513-114152> (2015).
73. Morrison, W. I. *et al.* CD4 T cell responses to *Theileria parva* in immune cattle recognize a diverse set of parasite antigens presented on the surface of infected lymphoblasts. *J. Immunol.* **207**, 1965–1977. <https://doi.org/10.4049/jimmunol.2100331> (2021).
74. Skilton, R. A. *et al.* A 32 kDa surface antigen of *Theileria parva*: Characterization and immunization studies. *Parasitology* **120**(Pt 6), 553–564. <https://doi.org/10.1017/s0031182099005934> (2000).
75. Dobbelaere, D. A., Webster, P., Leitch, B. L., Voigt, W. P. & Irvin, A. D. *Theileria parva*: Expression of a sporozoite surface coat antigen. *Exp. Parasitol.* **60**, 90–100. [https://doi.org/10.1016/s0014-4894\(85\)80026-6](https://doi.org/10.1016/s0014-4894(85)80026-6) (1985).
76. Tonui, T. *et al.* Transcriptomics reveal potential vaccine antigens and a drastic increase of upregulated genes during *Theileria parva* development from arthropod to bovine infective stages. *PLoS One* **13**, e0204047. <https://doi.org/10.1371/journal.pone.0204047> (2018).
77. Hiltscher, H. *et al.* The radical induced cell death protein 1 (RCD1) supports transcriptional activation of genes for chloroplast antioxidant enzymes. *Front. Plant Sci.* **5**, 475. <https://doi.org/10.3389/fpls.2014.00475> (2014).
78. Hayashi, S. *et al.* Control of reactive oxygen species (ROS) production through histidine kinases in *Aspergillus nidulans* under different growth conditions. *FEBS Open Bio* **4**, 90–95. <https://doi.org/10.1016/j.fob.2014.01.003> (2014).
79. van Schaick, G., Haselberg, R., Somsen, G. W., Wuhrer, M. & Dominguez-Vega, E. Studying protein structure and function by native separation-mass spectrometry. *Nat. Rev. Chem.* **6**, 215–231. <https://doi.org/10.1038/s41570-021-00353-7> (2022).
80. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589. <https://doi.org/10.1038/s41586-021-03819-2> (2021).

81. Wragg, D. *et al.* A locus conferring tolerance to Theileria infection in African cattle. *PLoS Genet.* **18**, e1010099. <https://doi.org/10.1371/journal.pgen.1010099> (2022).
82. Latre de Late, P. *et al.* Inherited tolerance in cattle to the apicomplexan protozoan theileria parva is associated with decreased proliferation of parasite-infected lymphocytes. *Front. Cell. Infect. Microbiol.* **11**, 751671. <https://doi.org/10.3389/fcimb.2021.751671> (2021).

Author contributions

T.C. and M.H. provided research ideas and designed experiments. E.H., L.L., A.G. and T.T. performed the experiments. E.H. performed bioinformatic analyses. E.H., T.C. and M.H. wrote the manuscript. E.H. prepared figures and supplementary data. All authors reviewed and approved the manuscript prior to submission.

Funding

This research was funded in part by the Bill & Melinda Gates Foundation and with UK aid from the UK Foreign, Commonwealth and Development Office (Grant Agreement OPP1127286) under the auspices of the Centre for Tropical Livestock Genetics and Health (CTLGH), established jointly by the University of Edinburgh, SRUC (Scotland's Rural College), and the International Livestock Research Institute, and strategic funding to the Roslin Institute by the Biotechnology and Biological Sciences Research Council (BBS/E/D/20002173). The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation nor the UK Government.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-59197-y>.

Correspondence and requests for materials should be addressed to E.M.D.L.v.d. or M.A.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024