

PBMC transcriptomic signatures reflect *Trypanosoma cruzi* strain diversity and trained immunity in chronically infected macaques

Hans Desale, ... , Claudia Herrera, Eric Dumonteil

JCI Insight. 2025. <https://doi.org/10.1172/jci.insight.186003>.

Research In-Press Preview Infectious disease Microbiology

Chagas disease is a neglected tropical disease caused by *Trypanosoma cruzi* with clinical presentations ranging from asymptomatic to cardiac and/or gastrointestinal complications. The mechanisms of pathogenesis are still poorly understood, but *T. cruzi* strain diversity may be associated with disease progression. Therefore, we evaluated the transcriptomic response of PBMCs from macaques with natural chronic infections and tested for heterogeneity in their gene signatures. Remarkably, transcriptomic response to *T. cruzi* infection matched parasite strain profiles, indicating that parasite diversity is a key determinant of host response. While differences in adaptive immune responses were identified, more striking alterations of innate immune processes were detected. Thus, initial innate response to *T. cruzi* infection may be conditioned by parasite strain diversity, resulting in different profiles of trained immunity modulating subsequent adaptive responses, allowing parasite control or its persistence during the chronic phase. These results call for further characterization of the cross-talk between innate and adaptive immunity according to parasite diversity, and how altered trained immunity contributes to pathogenesis, as this may lead to better treatments and vaccines.

Find the latest version:

<https://jci.me/186003/pdf>



PBMC transcriptomic signatures reflect *Trypanosoma cruzi* strain diversity and trained immunity in chronically infected macaques

Hans Desale¹, Weihong Tu¹, Kelly Goff², Preston A. Marx^{1,2}, Claudia Herrera¹, Eric Dumonteil^{1*}

¹Department of Tropical Medicine and Infectious Disease, School of Public Health and Tropical Medicine, and Vector-Borne and Infectious Disease Research Center, Tulane University, New Orleans, Louisiana, USA

²Division of Microbiology, Tulane National Primate Research Center, Tulane University, Covington, Louisiana, USA.

***Corresponding Author:** Eric Dumonteil, Department of Tropical Medicine and Infectious Disease, School of Public Health and Tropical Medicine, and Vector-Borne and Infectious Disease Research Center, Tulane University, 1440 Canal St., New Orleans, LA, 70112, USA. E-mail: edumonte@tulane.edu.

Conflict of interest: The authors have declared that no conflict of interest exists

Abstract

Chagas disease is a neglected tropical disease caused by *Trypanosoma cruzi* with clinical presentations ranging from asymptomatic to cardiac and/or gastrointestinal complications. The mechanisms of pathogenesis are still poorly understood, but *T. cruzi* strain diversity may be associated with disease progression. Therefore, we evaluated the transcriptomic response of PBMCs from macaques with natural chronic infections and tested for heterogeneity in their gene signatures. Remarkably, transcriptomic response to *T. cruzi* infection matched parasite strain profiles, indicating that parasite diversity is a key determinant of host response. While differences in adaptive immune responses were identified, more striking alterations of innate immune processes were detected. Thus, initial innate response to *T. cruzi* infection may be conditioned by parasite strain diversity, resulting in different profiles of trained immunity modulating subsequent adaptive responses, allowing parasite control or its persistence during the chronic phase. These results call for further characterization of the cross-talk between innate and adaptive immunity according to parasite diversity, and how altered trained immunity contributes to pathogenesis, as this may lead to better treatments and vaccines.

Chagas disease is a neglected tropical disease with an estimated annual burden of \$627.46 million in health-care costs and 806,170 DALYs worldwide (1), caused by infections with the protozoan parasite *Trypanosoma cruzi*. Following infection, infected hosts undergo an acute phase which lasts a few weeks, characterized by flu-like non-specific signs and symptoms and high levels of circulating blood parasites. Patients then enter the chronic phase which is initially asymptomatic and often lasting many decades. About 20-40% will then progress to a symptomatic chronic phase with cardiac (e.g., cardiomyopathy) and/or gastrointestinal (e.g., megaesophagus and megacolon) complications of growing severity, leading to disability and death (2, 3). While these clinical manifestations are well described, the mechanisms of pathogenesis and disease progression are less understood, which impedes any reliable prognosis for patients. Indeed, there are no biomarkers allowing to identify patients who will remain asymptomatic, or those who will progress to cardiac or digestive disease (4), complicating patient care.

The immune response to *T. cruzi* infection has been extensively studied to understand the ability of parasites to evade host immune responses and persist over prolonged time, and identify immune mechanisms allowing parasite control. Current understanding suggests that the initial steps of *T. cruzi* infection and replication occur inside resident monocytes/macrophages at the parasite entry site, leading to the activation of innate mechanisms including type I interferon (IFN) signaling, and the recruitment of neutrophils and Natural Killer (NK) cells but these cells fail to be fully activated resulting in a delayed immune response favoring parasite establishment (5-7). A robust and effective B and T cell response is eventually induced, allowing for parasite control and the end of the acute phase, but the head start of the parasite is such that sterilizing immunity cannot be reached and low levels of parasite can persist for a prolonged time (8-10).

During the chronic phase, adaptive immunity is considered the most important mechanism contributing to parasite control through antigen-specific cytolytic T cells and

antibody production by B cells, and a balanced Th1/Th2/Th17 response (8, 11). Thus, a hyper-polarized Th1 response may lead to exacerbated inflammation and tissue damage, while immune exhaustion or a Th2 response may allow for parasite proliferation, both allowing for disease progression.

However, this paradigm has been predominantly built from the targeted analysis of specific cell populations and cytokines in knock-out mice or through flow-cytometry analyses, and the relative contribution of the multiple mechanisms to parasite control involved is unclear. Also, the contribution of innate immunity during the chronic phase of *T. cruzi* infection is often ignored, although a few studies suggest a role of innate mechanisms in the progression of chronic Chagasic cardiomyopathy (12-17). Thus, non-targeted approaches such as transcriptomic studies of patients with different clinical profiles have provided more integrated information, and suggested that NK/CD8⁺ T-cell cytotoxicity could play a key role in determining disease progression (18).

In addition, *T. cruzi* genetic diversity has also been thought for a long time to contribute to differences in disease progression and severity (19, 20), as major biological differences among strains can be observed both *in vitro* and *in vivo* (21, 22). However, the generalization of these findings has been challenging and the association between parasite discrete typing units (DTUs), the main genetic lineages structuring *T. cruzi* diversity, with disease progression has remained elusive (23, 24, 25).

Monitoring macaques with natural *T. cruzi* infection has shed new light on some of these aspects, as these animals recapitulate very well human responses to infection (26-28). Indeed, a clear association could be detected for the first time between chronic disease progression, defined by blood parasite levels and changes in electrocardiographic (ECG) recordings, and the diversity of infecting parasite strains (29), as a large proportion of natural infections include multiple strains. Thus, infections with a low diversity of strains were associated with increased blood parasite burden over time and changes in ECG profiles indicative of early cardiac disease

progression, while infections with a high diversity of strains were associated with a better control of blood parasites and the absence of significant changes in ECG profiles (29). These observations suggest that differences in host-parasite relationships according to parasite diversity may contribute to disease progression. As a first step to better characterize host responses, we evaluated here the transcriptomic response of unstimulated PBMCs from naturally-infected macaques and tested for potential heterogeneity in their gene signatures according to disease progression and parasite diversity. Such an integrative approach can provide a comprehensive view of some of the mechanisms underlying Chagas disease pathogenesis.

Results

Differential PBMCs gene signature from *T. cruzi* infected and uninfected macaques

We used PBMCs from 29 macaques, including 11 uninfected controls and 18 macaques with natural *T. cruzi* infection (Table 1) for a cross-sectional study of their unstimulated transcriptomic profile. These had been infected with *T. cruzi* for 3-7 years, as determined by the approximate time of seroconversion, and were on average 11 years old (range 4-19). We performed RNA sequencing to obtain about 20 million reads per sample that were mapped to the mmul10 macaque reference genome. Batch-adjusted, filtered read counts were used to call differentially expressed genes between groups. A clear clustering of the two macaque groups was observed through principal component analysis (PCA), indicating that chronic *T. cruzi* infection could explain part of the variation in gene expression among samples (Figure 1). Based on a mean expression fold change >1.5 between groups, there were 1,468 genes significantly upregulated, and 1,160 genes significantly downregulated between uninfected and

infected macaques (Figure 1B). Clustering of individual macaques according to *T. cruzi* infection status was further observed based on differentially expressed genes (Figure 1C).

To explore the function of the differentially expressed genes, up-regulated and down-regulated genes were queried for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways using over-representation analysis. Among infected macaques, up-regulated pathways included platelet activation, blood coagulation, integrin-mediated signaling pathway, defense response to virus, cell activation, and immune system processes (Figure 1D). Down-regulated pathways included multiple apoptotic signaling pathways/apoptotic processes, lymphocyte/leukocyte activation, immune system development and both intracellular and cell surface signaling pathways (Figure 1E). These are in general agreement with extensive alterations of immune cell activation and differentiation in response to the ongoing chronic infection. Gene Set Enrichment Analysis (GSEA) highlighted some similar pathways differentiating uninfected and infected macaques (Supplementary Figure 1), but only platelet activation and tumor necrosis factor (TNF) α pathways had some statistical support, possibly due to extensive heterogeneity in individual macaque responses to infection.

Immune profile of *T. cruzi* infected macaques

To better characterize the immune profile of *T. cruzi* infected macaques, we focused on the expression profile of cytokine, chemokines, and immune marker genes, in an integrated manner. Interleukin (*IL*)15 and *IFNG* genes were upregulated in infected macaques, while transforming growth factor β 1 (*TGFB1*), *IL23*, and *IL10* genes were downregulated, and other cytokines appeared unaffected (Figure 2A). Infection was also associated with the marked upregulation of CXC motif chemokine ligand 10 (*CXCL10*) and CX3C motif chemokine receptor 1 (*CX3CR1*) genes, associated with CD8⁺ T cell response and the recruitment of NK cells, respectively, while C-C motif chemokine ligand 22 (*CCL22*), CXC motif chemokine ligand 8

(*CXCL8*), CXC motif chemokine receptor 4 (*CXCR4*) and *CXCR5* genes, which are involved in the chemotaxis and migration of T cells and neutrophils, were downregulated (Figure 2B). Analysis of additional markers further indicated a mixture of activation and inhibition of several immune cell populations (Figure 2C). These included antigen presenting cells, with the upregulation of *CD1C* and TAP binding protein (*TAPBP*) genes in infected macaques, while *CD83* was down-regulated, and *CD40* remained expressed at the same level. There was also a mixture of up-regulated anti-inflammatory marker genes (*CD84*) together with downregulated markers (*CD9*, *CD93*, cytotoxic T-lymphocyte associated protein 4 (*CTLA4*)). B cell marker *CD79B* was up-regulated, while no difference was detected in *CD4* or *CD8* expression. Lymphocyte activation marker CD40 ligand (*CD40LG*) gene was up-regulated, while *CD69* was down-regulated. T cell exhaustion/senescence marker *CD160* was not differentially expressed, while killer cell lectin-like receptor G1 (*KLRG1*) was upregulated in PBMCs from infected macaques, and co-stimulatory receptor *CD28* gene was downregulated. Interferon regulatory factor 4 (*IFRA4*) and T-cell immunoglobulin and mucin-domain containing 3 (*TIM3*), markers for Th2 and Th1 cells, respectively, were down-regulated, and Treg marker forkhead box P3 (*FOXP3*) tended to be down-regulated as well.

Next, multivariate analysis was used to identify immune markers most associated with chronic infection. Between group analysis (BGA) of gene expression revealed strong separation of chronically infected and uninfected macaques along the main BGA discrimination axis (Figure 3A). By overlaying coordinate data of the samples over the clustering of gene expression, genes found at the ends of the main BGA discrimination axis represents genes most associated with groups (Figure 3B). The top 20 genes associated with chronic *T. cruzi* infection in macaques were identified along the main BGA discrimination axis (Figure 3C), most of which had also been identified as differentially expressed by DESeq2 analysis. Infection was associated with the upregulation of pro-inflammatory genes *CXCL10*, *IFNG*, and *IL15*, alongside several NK and T cell activation markers such as *CD244*, *CD40LG* and *CD226*. The *CD1C* marker of dendritic

cells was upregulated, as well as *CD59*, which mediates complement-mediated cell lysis and lymphocyte signaling. On the other hand, there was a downregulation of the pro-inflammatory cytokine genes *IL1B*, and of the anti-inflammatory *TGFB1*, alongside chemotactic molecules genes *CXCL8* and *CXCL16*. Both inflammation marker *CD93*, and anti-inflammation marker *CD9* were downregulated. There was also a downregulation of activation markers *CD70*, *CD82* and *CD83*, suggesting altered Treg function and/or incomplete activation of T and B cells, and co-stimulatory molecule *CD28* gene was also downregulated. Together, these data suggested that chronically infected macaques present a mixture of immune activation to target *T. cruzi* parasites, together with immune suppression to prevent uncontrolled tissue inflammation and damage, with the involvement of both innate and adaptive immune mechanisms, with important heterogeneity in individual responses.

Differential PMBC gene signature from controller and progressor macaques

Because there is extensive variability in disease progression and severity in *T. cruzi* infected hosts, we next tested for differences in gene expression profiles among macaques with progressive or controlled infection, based on changes in their parasite burden (Table 1). PCA analysis of gene expression profile revealed a differential clustering of progressor, controller and uninfected macaques, respectively (Figure 4A), with both infected groups presenting a large number of differentially expressed genes with uninfected macaques, as well as 28 significantly up-regulated and 74 down-regulated genes between progressor and controller macaques (Figure 4B). These data were in agreement with a different host response in progressor and controller macaques. Focusing on genes involved in the immune response, we detected a set of genes significantly up- and down-regulated in both groups of infected macaques, including the up-regulation of *IL15*, *CXCL10*, and *CD244* and the down regulation of *IL10*, *TGFB1*, or *CD28* genes (Figure 4C). On the other hand, some genes were specifically modulated in PBMCs from controller and progressor macaques, respectively. For example, *IFNG*, *IL6* and *IL24* genes were

upregulated in controller macaques, together with B cell marker *CD79B* gene as well as anti-inflammatory marker *CD109*, while *CCL22* gene involved in the chemotaxis and migration of T cells was downregulated, together with activation marker *CD83* (Figure 4C), suggesting an active adaptative response with a controlled inflammation. Conversely, PBMCs from progressor macaques presented an upregulation of *CD4*, as well as of *CD59*, which inhibits complement-mediated cell lysis to protect cells from complement damage, and *CD101*, which has been proposed as a marker of pathological responses to dysbiosis, and down-regulation of *IFNL1*, which plays a role in viral and bacterial infections, as well as the pro-inflammatory cytokine gene *IL1B*. Macrophage anti-inflammatory markers *CD9* and *CD163*, and lymphocyte activation marker *CD69* were also down-regulated in progressor macaques (Figure 4C), suggesting a more pro-inflammatory response. For an integrated analysis of the immune response, we again performed a multivariate analysis of the gene expression profile among macaque groups (Figure 4D), which confirmed the unique gene signatures of PBMCs from controller and progressor macaques.

Differential PBMC gene signatures reflect *T. cruzi* strain diversity

To further understand the heterogeneity of host responses to *T. cruzi* infection, we next performed an unsupervised clustering of infected macaques, based on their PBMC differential gene expression profiles. This analysis indicated the presence of three clusters of unique gene signatures among infected macaques (Figure 5A). Clusters A and B included mostly macaques that had been classified as progressors, although they presented clear differences in their gene expression profiles, while cluster C comprised controller macaques. One macaque (HN75), initially classified as controller, rather clustered with progressor macaques from Cluster B. Remarkably, this clustering of macaques based on their transcriptomic response to *T. cruzi* infection matched well their parasite strain profiles. Indeed, macaques from cluster A were predominantly infected by TcI parasites, while those from cluster B were predominantly infected

with TcIV parasites, and those from cluster C were infected with the largest diversity of strains that included mixtures of TcI, TcII, TcIV, TcV or TcVI in variable proportions (Figure 5A and B). Linear discriminant analysis (LDA) of parasite strain composition was also able to significantly discriminate the three clusters (PERMANOVA, $P=0.01$), and over 77% of macaques could be correctly reclassified into their corresponding group of gene expression profile based solely on *T. cruzi* parasite composition data (Figure 5C). These results strongly suggested that infecting *T. cruzi* strain composition is a major determinant of host response and disease progression.

We further assessed differences in immune profile among the three clusters of infected macaques. Analysis of cytokine and immune marker gene expression indicated clear differences in the immune response of each infected macaque group (Figure 6A and B). Thus, the PBMCs response of progressor macaques from cluster A was characterized by an increased expression of *IFNG*, *IL15* and *IL17* cytokine genes, together with the up-regulation of *CD163* associated with alternatively activated macrophages, and *CD1A* associated with dendritic cells. At the same time, the up-regulation of *CD300A* and *CD300E* and the down-regulation of *CD82* genes suggested the inhibition of effector functions and T cell responses. Furthermore, *CD177* and *CD33* were up-regulated, in agreement with poor pathogen control as observed in SARS-Cov2 and HIV infections (30, 31).

PBMCs of progressor macaques from cluster B presented a down-regulated expression of *IFNG*, and even though the up-regulation of *CD40LG* suggested T cell activation, this may be dampened by the up-regulation of *CD5* and *CD6* genes, which mediate the control of aberrant immune responses and inhibit T cells, respectively (Figure 6A and B). Furthermore, the down-regulation of *CD93* suggested a low immune infiltration.

Lastly, PBMCs from controller macaques from cluster C also presented a high level of *IFNG* expression, together with the up-regulation of *IL15*, and *IL17* cytokine genes, but the up-regulation of *CD300LG* suggested higher lymphocyte migration, while the down-regulation of *CD177* was in agreement with improved pathogen control (Figure 6A and B).

These differences in immune response profiles among the three groups of *T. cruzi* infected macaques were also largely supported by a BGA of the gene expression profiles (Figure 6C). Thus, progressor macaques from cluster A were characterized by *CD163*, *CD177*, *CD36*, *CD33* and *CXCL16* gene expression, which associate with a rather pro-inflammatory response, which those from cluster B were associated with the expression of *CD5*, *CD59*, *CD247*, and *TGFB1* genes suggesting more regulatory processes. On the other hand, the gene profile from controller macaques from cluster C included genes such as *IFNG*, *CD226*, *CD79B*, *CD1C* and *CXCL8*, suggesting a broad activation of B and T cells as well as from innate cells, and the expression of *CD109*, *CD40* genes also suggested some regulatory processes (Figure 6C).

GSEA further identified multiple pathways that were differentially associated with each cluster of macaques. Again, gene signatures from progressor macaques from Cluster A presented the most pro-inflammatory responses, predominantly involving innate immune processes mediated by antigen presenting cells and monocytes, and a down regulated platelet activation pathway, while pathways associated with B and T cell adaptive responses were marginally affected compared to controller macaques from Cluster C (Figure 7). This response also included exacerbated TLR and inflammation pathways, as well as upregulated blood coagulation pathway compared to Cluster C macaques (Supplementary Figure 2).

Progressor macaques from Cluster B also tended to present a pro-inflammatory innate component, although not as marked as Cluster A, and an up-regulated platelet activation pathway (Figure 7 and supplementary Figure 3), with some upregulation of B cell markers compared to Cluster A (Supplementary Figure 3). On the other hand, gene expression from pathways associated with B and T cells was not significantly different between Cluster B and C macaques, except for some upregulation of mitotic CD4 cells in Cluster C macaques (Supplementary Figure 4). Thus, controller macaques from Cluster C presented the most anti-inflammatory response and a controlled innate response compared to Clusters A and B.

Finally, we assessed the potential of some of the expressed genes to serve as biomarkers of host immune response profile and disease progression. Several potential biomarkers were specifically up- or down regulated in PBMCs of the different macaque groups (Figure 8A). The combination of only eight genes, including Latent Transforming Growth Factor Beta Binding Protein 1 (*LTBP1*), apolipoprotein D (*APOD*), TIMP metalloproteinase inhibitor 1 (*TIMP1*), oncostatin M (*OSM*), transforming growth factor-beta 1 (*TGFB1*), 2'-5'-oligoadenylate synthetase 2 (*OAS2*), Galectin 3 Binding Protein (*LGALS3BP*) and nucleobindin 2 (*NUCB2*) genes was sufficient to correctly discriminate among the three clusters of infected macaques, as indicated by LDA analysis (PERMANOVA, $P=0.002$), and 25/29 (86.2%) of macaques could be correctly reclassified based on the signature of these markers (Figure 8B).

Discussion

Host responses to *T. cruzi* infection can lead to a broad range of clinical manifestations, ranging from an asymptomatic status to severe cardiac and digestive disease (3, 32). Many studies have focused on specific components of the immune response to identify cell populations, cytokines and immune processes involved, but integration of this knowledge has been challenging, and the relative contribution of the multiple immune components for parasite control and disease progression remains unclear (33, 34). Here we used RNA-sequencing of PBMCs from naturally infected macaques to investigate the transcriptomic response of these cells in an integrated perspective.

As expected, we found that chronic *T. cruzi* infection was associated with major alterations of PBMC gene expression profiles, characterized by a mixture of pro-inflammatory immune activation, together with immune suppression which may reflect control of hyperpolarized inflammation and/or immune exhaustion associated with parasite persistence and insufficient control. This is in general agreement with a broad adaptive immune response observed in the chronic phase of infection (9, 11). Our integrated approach also points to the

important contribution of innate immune processes such as platelet activation, as well as hemostasis and blood coagulation. Indeed, hypercoagulability and thrombosis have been involved in the chronic phase, independently of the development of Chagasic cardiomyopathy (13, 35-37). Altered platelet activation has also been described in mice experimentally infected with *T. cruzi* (38), as well as in patients (14). In addition to their role in hemostasis and thrombosis, platelets can also play an immunomodulatory role in infection through multiple mechanisms (39, 40). While the role of innate responses in the early steps of *T. cruzi* infection have been described, the continued involvement of innate processes after years of infection is more striking (5, 17, 41).

Remarkably, our results also shed light on some of the key mechanisms underlying the diversity of host responses to *T. cruzi* infection, which has been a major hurdle in understanding Chagas disease progression. Indeed, major differences in the transcriptomic response of host PMBCs were detected according to whether these hosts were able to control or not parasite multiplication. Furthermore, while controller macaques were a fairly homogenous group based on their gene expression profile, progressor macaques presented additional heterogeneity in their profiles, with two main gene signatures, suggesting that there are multiple ways to have an inadequate immune response against *T. cruzi* (Clusters A and B of progressor macaques) but likely less options to correctly control the parasite (Cluster C of controller macaques). Strikingly, these differences in host responses could be explained in large part by the composition of infecting strains of *T. cruzi*, with progressor macaques from Cluster A infected predominantly with TcI, those from Cluster B with TcIV, while controller macaques from Cluster C infected with a high diversity of strains including TcI, TcII, TcIV, TcV or TcVI in variable proportions. This expands previous work hinting at differences in cytokine profiles in patients infected with TcI, TcII or mixtures of TcI and TcII (34), and strongly support the hypothesis that disease progression may be predicted based on parasite diversity as proposed before (29). Importantly, this implies that such infections with multiple parasite strains occur at once rather than

sequentially, which seems supported by observations in triatomine vectors (42, 43) and the low likelihood of sequential infections (44).

Analysis of the gene expression profile heterogeneity among *T. cruzi* infected macaques further revealed key features of the immune response during the chronic phase in an integrated manner. Thus, macaques predominantly infected with TcI had the most pro-inflammatory responses compared to controller macaques infected with a high parasite diversity. Strikingly, many of the enriched pathways in these progressor macaques were associated with the innate immune response, with alterations in platelet, monocyte, myeloid cells and NK cell activation, and TLR pathways, rather than with the adaptive immune response, with limited enrichment in B or T cell associated pathways. On the other hand, controller macaques displayed a strong adaptative response but in a less pro-inflammatory environment. These results are in agreement with a comparable study in Chagas disease patients, as these displayed blood gene signatures from several pathways associated with innate immunity such as TLR signaling and monocyte and NK cell responses that were associated with different levels of severity of chronic cardiomyopathy, although no information was available on *T. cruzi* strains infecting these patients (18). Thus, while adaptive immunity is clearly needed (9, 11) and can play a role in Chagas disease progression as observed before (45) our results suggest that innate immunity may be a predominant contributor to pathogenesis, even though it is often overlooked in chronic infections. More specifically, it is likely that trained immunity can be responsible for differences in the inflammatory environment. Indeed, trained immunity involves the long-term reprogramming of immune cells including monocytes, macrophages and/or natural killer cells, through epigenetic modifications (46). This training/programming of the innate cells may occur during the acute phase of infection, and vary according to infecting *T. cruzi* strain diversity, to condition the subsequent inflammatory environment and the adaptive response during the chronic phase (47), and thus mediate pathogenesis. Trained immunity has been implicated in host response to pathogens such as *Mycobacterium* (48), and can also mediate non-specific

effects of vaccines against non-target pathogens, such as the effect of BCG vaccination in modulating clinical and immune responses to *Plasmodium falciparum* infection (49), through the reprogramming of neutrophils (50). Remarkably, memory-like NK cells generated in response to vaccination of mice with a highly attenuated *T. cruzi* strain have recently been found to be critical for protection against a secondary infectious challenge, strengthening the role of trained innate immunity in long term parasite control (51).

Finally, we also identified a simple biomarker signature that can discriminate among host transcriptomic response profiles, and that also reflects *T. cruzi* strain diversity and disease progression. As such, it may have important clinical relevance to identify patients most at risk of progressing to chronic cardiomyopathy (52). Previous studies identified several clinical parameters as predictors of *T. cruzi* PCR positivity in patients (53), and clinical parameters from cardiac echography together with brain natriuretic peptide levels but not those of TGF β 1 have been associated with Chagas disease cardiovascular events (54). However, TGF β serum levels increased in experimentally infected mice and correlated well with cardiac fibrosis and echographic alterations (55). Further validation of these biomarkers may lead to new tools for prognosis of Chagas disease.

A limitation of our study is that transcriptomic responses were only assessed at a single time-point, and no longitudinal assessment of individual responses was performed. Such studies would be key to document transcriptomic changes as disease progresses, but may be difficult to perform given the slow progression of clinical disease. The limited extent of cardiac disease in our cohort of macaques is also another limitation, as none of the animals presented severe cardiomyopathy, which may be associated with additional transcriptomic changes. Indeed, all appeared to have normal ECGs, with only statistical differences detectable through multivariate analysis of ECG parameters (29), suggesting very early stages of cardiac alterations. Nonetheless, stratification of progressor/controller macaques based on changes in blood parasite burden remained helpful to assess the heterogeneity of individual responses to

infection. Nonetheless, we still detected some inconsistencies among individual macaque transcriptomic responses, likely due to additional factors that may further modulate host response and the progression of the infection, including host genetics, host status at the time of infection, infecting parasite dose, etc...), but these could not be taken into account.

In conclusion, our integrative study demonstrates that *T. cruzi* parasite diversity is a major determinant of host response to natural infection observed during the chronic phase in macaques, leading to differences in parasite control and disease progression. While differences in adaptive immune responses can be observed according to *T. cruzi* diversity, more striking alterations of innate immune processes can be detected suggesting that trained immunity is a major component of Chagas disease pathogenesis. A likely hypothesis is that the initial innate response to *T. cruzi* infection is conditioned by parasite strain diversity, resulting in different profiles of trained immunity that modulate the adaptive immune response, ultimately shaping parasite control or its persistence during the chronic phase. These results call for further characterization of the cross-talk between innate and adaptive immunity and how altered trained immunity contributes to Chagas disease pathogenesis, as this may lead to better treatments and more effective vaccines.

Materials and Method

Sex as a biological variable

Infected macaques of both sexes were included in the study as indicated in Table 1, and similar findings are reported for both sexes.

Naturally infected macaques

Chagasic macaques were part of a cohort of naturally infected rhesus macaques (*Macaca mulatta*) held at the Tulane National Primate Research Center with confirmed *T. cruzi*

infection (29). Eighteen animals were included in this study. These had been infected with *T. cruzi* for 3-7 years, as determined by the approximate time of seroconversion, and were on average 11 years old (range 4-19) (Table 1). An additional 11 uninfected healthy control macaques were included for comparison. Whole blood was collected with EDTA-K2, and ECG recording were performed during routine veterinary care of the animals and represent an opportunistic sample. A cross-sectional study of unstimulated PBMCs transcriptome was performed at a single time-point.

Disease progression was assessed by measuring blood parasite burden by qPCR as described before (29), every 6-8 months over a 24-48 months period. Changes in blood parasite burden over time were calculated and macaques were classified as controller when presenting a decreasing blood parasite burden over time, or as progressor when presenting an increase in blood parasite burden (Table 1). *T. cruzi* parasites infecting the macaques were genotyped by next-generation sequencing of the mini-exon marker as before (29), to identify parasite DTUs present and their relative proportion. Cardiac function was also monitored by ECG recordings and analysis of the main wave patterns (29). Specific abnormalities/arrhythmias were not observed, and all ECGs were normal and no macaque presented with severe cardiac disease. However, most macaques in the progressor group presented significant changes in ECG profile over time suggesting early conduction defects, while most in the controller group had no significant change over time (Table 1). Brain natriuretic peptide (BNP) plasma concentration was measured by ELISA (RayBiotech Inc.) and tended to be elevated in progressor macaques although this did not reach statistical significance (Kuskal-Wallis, $P=0.22$, Supplementary Figure 5).

RNA Extraction from PBMCs and sequencing

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood centrifugation with a Ficoll-Paque™ (Millipore-Sigma, Inc) solution, washed with RPMI and PBS

and cryopreserved in Cryostor™ (Biolife Solutions, Inc) solution until use. Frozen PBMCs were thawed and assessed for viability through counting in a hemocytometer with trypan blue dye staining. Thawed unstimulated PBMCs with at least 90% viability were used for RNA extraction. RNA was extracted using the PerfectPure RNA Cultured Cell Kit™ (5 Prime, Inc.) as per manufacturer's instructions and RNA integrity was assessed with an Agilent Bioanalyzer and all samples had a RIN>8. About 100-200 ng of RNA per sample was used for library preparation and sequencing on an Illumina MiSeq platform.

Differentially Expressed Gene Calling

Data analysis and visualization was performed using R version 4.2.1 (56). Reads were filtered for quality and mapped to the rhesus macaque mmul10 reference genome (RefSeq assembly accession: GCF_003339765.1) using the STAR (Spliced Transcripts Alignment to a Reference) (57) package under default parameters. Quality mapped read counts were batch adjusted using the R package ComBat-seq (58). Read counts were normalized to account for differences in sequencing depth among, filtered for genes with low counts or outliers, and differentially expressed genes called using DESeq2 (59) under the apeglm (60) shrinkage estimator to reduce noise. Significant genes were called at a significance alpha of 0.05 adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method and a fold change of 1.5. Principle component analysis (PCA) was performed through the DESeq2 package to visualize the overall gene expression profile among individual macaques.

Pathway Analysis

Pathway analysis of up and down regulated DEGs in infected macaques were performed using both the clusterProfiler (61) R package and the ShinyGO (62) webapp to query the Gene Ontology (GO) (63, 64) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (65)

databases using over-representation analysis under a false-discovery rate cutoff < 0.05 . Two query algorithms and two pathway databases were used to provide increased confidence in relevance of identified pathway; and only pathways shared between the two search algorithm results were retained. Functional pathways were visualized using the Pathview (66) R package. Additional pathway enrichment analysis was performed with Gene Set Enrichment Analysis (GSEA) 4.3.2.(67). Annotated gene sets from the Hallmark molecular signature database (68) and the blood transcriptome module (BTM) (69) were used for these analyses.

To assess potential differences in gene expression profile among infected macaques, these were initially stratified as progressor and controllers, as defined above. We assessed for difference in the overall gene expression profile using PCA before focusing on the expression of immune cell markers and cytokines/chemokines among these infected macaque groups. We also assessed the expression of genes that may be used as biomarkers of parasite control by identifying genes differentially expressed between controller and progressor macaques, which protein products may be secreted and measured in blood/serum samples. Linear discriminant analysis (LDA) was performed to assess the ability of a set of biomarkers able to identify controller and progressor macaques. Significance of the LDA was assessed by one-way permutation ANOVA (PERMANOVA).

Next, we tested for further differentiation of host responses and performed an unsupervised hierarchical clustering of infected macaques based on a complete linkage analysis of their differential gene expression profiles, which calculates the maximum pairwise dissimilarity between elements in two clusters when building the clusters. The clustering was performed on a subset of 85 genes presenting a significant difference in expression level among infected macaques ($P < 0.05$ FDR corrected, > 1.5 fold change). Infected macaques were then stratified according to these new clusters to assess their immune and infection profiles. We also tested for the composition of parasite strains infecting macaques as a predictor for these new clusters, using LDA of parasite strain composition, and statistical significance was evaluated by

PERMANOVA. The reclassification matrix was also calculated to estimate the accuracy of the LDA.

Multivariate analysis of gene expression profile was also performed using the MADE4 package (70) which is specifically built to perform supervised dimension reduction and between group analysis (BGA) using gene read count data. BGA is a supervised PCA method seeking to ordinate groups rather than individual samples (70).

Statistical analysis

Changes in blood parasite burden were compared between progressor and controller macaques with two-tailed Student t test. Differentially expressed genes were called at a significance alpha of 0.05 adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) method and a fold change of 1.5 in DESeq2 (59) under the apegglm (60) shrinkage estimator to reduce noise. Principle component analysis (PCA) was performed through the DESeq2 package. Pathways associated with differentially expressed genes were identified using over-representation analysis under a false-discovery rate cutoff < 0.05 . Between group multivariate analysis (BGA) of differential gene expression was performed in MADE4 (70). Linear discriminant analysis (LDA) of parasite DTU composition and of biomarkers gene expression level was performed in PAST4 and statistical significance was assessed by one-way permutation ANOVA (PERMANOVA).

Study approval

The study was approved by the Tulane Institutional Animal Care & Use Committee.

Data availability

All RNA sequences have been deposited in NCBI SRA database and are available under BioProject #PRJNA1139163, Biosamples #SAMN42764022 to SAMN42764047, and

BioProject #PRJNA1010169, Biosamples #SAMN37182435, SAMN37182439, and SAMN37182443. Supporting data values are available in a supplementary file.

Acknowledgements

This work was supported by grants R21AI175523 and R01AI162907 from the National Institute of Allergy and Infectious Diseases to ED. We also acknowledge the support from the National Center for Research Resources and the Office of Research Infrastructure Programs (ORIP) of the NIH through grant P51 OD011104, U42 OD024282 and U42 OD010568 to the Tulane National Primate Research Center.

Author contribution

E.D. designed the study and acquired funding with support from C.H. and P.A.M.; H.D. contributed to the design of discrete experiments and analyses; H.D., W.T. and K.G. conducted the experiments and analyzed data with technical support and conceptual advice from E.D. and C.H.; H.D. and E.D. prepared the manuscript with reviewing and editing input from all authors.

Table 1: Macaque demographic data

ID	Age (years)	Sex	Months Infected	Parasite DTU	Change in parasite burden ^a	Change in ECG ^b	Disease progression
HA67	14	Female	35	Tcl, TcIV	-0.6	No	Controller
HN75	14	Female	59	Tcl, TcIV	-1.0	ND	Controller
IN54	12	Female	82	Tcl, TcII, TcIV, TcV	-1.4	No	Controller
IP64 ^c	12	Female	89	Tcl, TcIV, TcV, TcVI	-0.8	No	Controller
JC08	11	Female	63	Tcl, TcII, TcIV, TcV, TcVI	-1.3	No	Controller
JM46	10	Male	77	Tcl, TcII, TcIV, TcVI	-0.7	ND	Controller
JN64	10	Female	72	Tcl, TcIV, TcV, TcVI	-1.1	No	Controller
KP37	8	Male	77	Tcl, TcIV	-0.6	No	Controller
LL64	5	Female	57	Tcl, TcIV	-1.9	No	Controller
ED57 ^c	19	Female	75	Tcl, TcIV	1.5	Yes	Progressor
EP36	18	Female	85	Tcl, TcIV	0.6	Yes	Progressor
GI52 ^c	15	Female	75	Tcl, TcIV	1.6	No	Progressor
JL71	11	Male	82	Tcl, TcIV	1.5	Yes	Progressor
JT42	10	Female	81	Tcl, TcII, TcIV, TcVI	2.0	Yes	Progressor
KA90	9	Female	81	Tcl, TcII, TcIV	0.9	Yes	Progressor
KC05	9	Female	81	Tcl, TcIV	0.9	Yes	Progressor
KL57	6	Female	81	Tcl, TcII, TcIV	2.7	ND	Progressor
MD12	4	Female	34	Tcl, TcII, TcV	0.2	Yes	Progressor
ID59	15	Male	0	-	-	-	Uninfected
HD69	14	Female	0	-	-	-	Uninfected
JN58	10	Female	0	-	-	-	Uninfected
KC78	9	Male	0	-	-	-	Uninfected
KK63	9	Male	0	-	-	-	Uninfected
KR83	9	Male	0	-	-	-	Uninfected
KL72	7	Male	0	-	-	-	Uninfected
LB83	6	Male	0	-	-	-	Uninfected
LD53	6	Male	0	-	-	-	Uninfected
ME13	6	Male	0	-	-	-	Uninfected
AG23	2	Female	0	-	-	-	Uninfected

^a Change in parasite burden was measured by qPCR in blood samples and is expressed in Log₁₀ parasites/ml change during a 12-24 months period. Negative values indicate a decrease, and positive values an increase in blood parasite burden over time. There was a significant difference between Progressor and Controller macaques (two-tailed Student $t=8.1$, $P<0.0001$).

^b The presence or absence of changes in ECG recordings parameters over time (12-24 months period) was evaluated. ND: not done. ^c Macaques of Chinese ancestry. All others are of Indian ancestry.

|

References

1. Lee BY, et al. Global economic burden of Chagas disease: a computational simulation model. *Lancet Infect Dis.* 2013;13(4):342-8.
2. Rassi A, Jr., et al. Chagas' heart disease. *Clin Cardiol.* 2000;23(12):883-9.
3. Rassi A, Jr., et al. Chagas disease. *Lancet.* 2010;375(9723):1388-402.
4. Pinazo MJ, et al. Biomarkers of therapeutic responses in chronic Chagas disease: state of the art and future perspectives. *Mem Inst Oswaldo Cruz.* 2015;110(3):422-32.
5. Padilla AM, et al. Delayed Activation of T Cells at the Site of Infection Facilitates the Establishment of *Trypanosoma cruzi* in Both Naive and Immune Hosts. *mSphere.* 2023;8(1):e0060122.
6. Padilla AM, et al. Insufficient TLR activation contributes to the slow development of CD8+ T cell responses in *Trypanosoma cruzi* infection. *J Immunol.* 2009;183(2):1245-52.
7. Chessler AD, et al. *Trypanosoma cruzi* triggers an early type I IFN response in vivo at the site of intradermal infection. *J Immunol.* 2009;182(4):2288-96.
8. Perez-Mazliah D, et al. Host-parasite dynamics in Chagas disease from systemic to hyper-local scales. *Parasite immunol.* 2021;43(2):e12786.
9. Tarleton RL. CD8+ T cells in *Trypanosoma cruzi* infection. *Sem Immunopathol.* 2015;37(3):233-8.
10. Tarleton RL. The role of T-cell subpopulations in experimental Chagas' disease. *Res Immunol.* 1991;142:130-3.
11. Jones KM, et al. Preclinical advances and the immunophysiology of a new therapeutic Chagas disease vaccine. *Exp Rev Vaccines.* 2022;21(9):1185-203.
12. Medeiros NI, et al. Evidence of Different IL-1beta Activation Pathways in Innate Immune Cells From Indeterminate and Cardiac Patients With Chronic Chagas Disease. *Front Immunol.* 2019;10:800.
13. Choudhuri S, and Garg NJ. Platelets, Macrophages, and Thromboinflammation in Chagas Disease. *J Inflamm Res.* 2022;15:5689-706.
14. Pengue C, et al. Impaired frequencies and function of platelets and tissue remodeling in chronic Chagas disease. *PLoS one.* 2019;14(6):e0218260.
15. Fernandes M, et al. *Trypanosoma cruzi*: course of infection in platelets-depleted mice. *Rev Inst Med Trop Sao Paulo.* 1992;34(1):9-13.
16. Gomes JAS, et al. Systems Biology Reveals Relevant Gaps in Fc-gammaR Expression, Impaired Regulatory Cytokine Microenvironment Interfaced With Anti-*Trypanosoma cruzi* IgG Reactivity in Cardiac Chagas Disease Patients. *Front Microbiol.* 2018;9:1608.
17. Sathler-Avelar R, et al. Innate immunity and regulatory T-cells in human Chagas disease: what must be understood? *Mem Inst Oswaldo Cruz.* 2009;104 Suppl 1:246-51.
18. Ferreira LR, et al. Blood Gene Signatures of Chagas Cardiomyopathy With or Without Ventricular Dysfunction. *J Infect Dis.* 2017;215(3):387-95.
19. Tibayrenc M, and Ayala FJ. Isozyme variability in *Trypanosoma cruzi*, the agent of Chagas' disease: genetical, taxonomical, and epidemiological significance. *Evolution.* 1988;42:277-92.
20. Brisse S, et al. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. *Int J Parasitol.* 2000;30:35-44.
21. Zingales B, et al. A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. *Mem Inst Oswaldo Cruz.* 2009;104(7):1051-4.

22. Zingales B, et al. The revised *Trypanosoma cruzi* subspecific nomenclature: rationale, epidemiological relevance and research applications. *Infect, Genet Evol.* 2012;12(2):240-53.
23. Tavares de Oliveira M, et al. Quantification of parasite burden of *Trypanosoma cruzi* and identification of Discrete Typing Units (DTUs) in blood samples of Latin American immigrants residing in Barcelona, Spain. *PLoS Negl Trop Dis.* 2020;14(6):e0008311.
24. Nielebock MAP, et al. Association between *Trypanosoma cruzi* DTU TcII and chronic Chagas disease clinical presentation and outcome in an urban cohort in Brazil. *PLoS one.* 2020;15(12):e0243008.
25. Apt W, et al. *Trypanosoma cruzi* burden, genotypes, and clinical evaluation of Chilean patients with chronic Chagas cardiopathy. *Parasitol Res.* 2015;114(8):3007-18.
26. Sathler-Avelar R, et al. Phenotypic and Functional Signatures of Peripheral Blood and Spleen Compartments of Cynomolgus Macaques Infected With *T. cruzi*: Associations With Cardiac Histopathological Characteristics. *Front Cell Infect Microbiol.* 2021;11:701930.
27. Vitelli-Avelar DM, et al. Cynomolgus macaques naturally infected with *Trypanosoma cruzi*-I exhibit an overall mixed pro-inflammatory/modulated cytokine signature characteristic of human Chagas disease. *PLoS Negl Trop Dis.* 2017;11(2):e0005233.
28. Sathler-Avelar R, et al. Phenotypic Features of Circulating Leukocytes from Non-human Primates Naturally Infected with *Trypanosoma cruzi* Resemble the Major Immunological Findings Observed in Human Chagas Disease. *PLoS Negl Trop Dis.* 2016;10(1):e0004302.
29. Dumonteil E, et al. Intra-host *Trypanosoma cruzi* strain dynamics shape disease progression: the missing link in Chagas disease pathogenesis. *Microbiol Spectr.* 2023;11(5):e0423622.
30. Levy Y, et al. CD177, a specific marker of neutrophil activation, is associated with coronavirus disease 2019 severity and death. *iScience.* 2021;24(7):102711.
31. Duran-Castells C, et al. Plasma proteomic profiling identifies CD33 as a marker of HIV control in natural infection and after therapeutic vaccination. *EBioMedicine.* 2023;95:104732.
32. Bern C. Chagas' Disease. *The New England journal of medicine.* 2015;373(5):456-66.
33. Lewis MD, et al. Host and parasite genetics shape a link between *Trypanosoma cruzi* infection dynamics and chronic cardiomyopathy. *Cel Microbiol.* 2016;18(10):1429-43.
34. Poveda C, et al. Cytokine profiling in Chagas disease: towards understanding the association with infecting *Trypanosoma cruzi* discrete typing units (a BENEFIT TRIAL sub-study). *PLoS one.* 2014;9(3):e91154.
35. Samuel J, et al. Cardiac thrombosis and thromboembolism in chronic Chagas' heart disease. *Am J Cardiol.* 1983;52(1):147-51.
36. Herrera RN, et al. Inflammatory and prothrombotic activation with conserved endothelial function in patients with chronic, asymptomatic Chagas disease. *Clin Appl Thromb Hemost.* 2011;17(5):502-7.
37. Echeverria LE, et al. Coagulation disorders in Chagas disease: A pathophysiological systematic review and meta-analysis. *Thromb Res.* 2021;201:73-83.
38. Tanowitz HB, et al. Enhanced platelet adherence and aggregation in Chagas' disease: a potential pathogenic mechanism for cardiomyopathy. *Am J Trop Med Hyg.* 1990;43(3):274-81.
39. Portier I, and Campbell RA. Role of Platelets in Detection and Regulation of Infection. *Arterioscler Thromb Vasc Biol.* 2021;41(1):70-8.
40. Middleton E, and Rondina MT. Platelets in infectious disease. *Hematology Am Soc Hematol Educ Program.* 2016;2016(1):256-61.

41. Fernandez-Villegas A, et al. The innate immune response status correlates with a divergent clinical course in congenital Chagas disease of twins born in a non-endemic country. *Acta Trop.* 2014;140:84-90.
42. Murillo-Solano C, et al. Diversity of *Trypanosoma cruzi* parasites infecting *Triatoma dimidiata* in Central Veracruz, Mexico, and their One Health ecological interactions. *Inf Genet Evol.* 2021;95:105050.
43. Dumonteil E, et al. Interactions among *Triatoma sanguisuga* blood feeding sources, gut microbiota and *Trypanosoma cruzi* diversity in southern Louisiana. *Mol Ecol.* 2020;29(19):3747-61.
44. Nouvellet P, et al. The improbable transmission of *Trypanosoma cruzi* to human: The missing link in the dynamics and control of Chagas disease. *PLoS Negl Trop Dis.* 2013;7(11):e2505.
45. Almeida GG, et al. Chagasic cardiomyopathy is marked by a unique signature of activated CD4(+) T cells. *J Transl Med.* 2022;20(1):551.
46. Ochando J, et al. Trained immunity - basic concepts and contributions to immunopathology. *Nat Rev Nephrol.* 2023;19(1):23-37.
47. Murphy DM, et al. The Effects of Trained Innate Immunity on T Cell Responses; Clinical Implications and Knowledge Gaps for Future Research. *Front Immunol.* 2021;12:706583.
48. Joosten SA, et al. Mycobacterial growth inhibition is associated with trained innate immunity. *J Clin Invest.* 2018;128(5):1837-51.
49. Walk J, et al. Outcomes of controlled human malaria infection after BCG vaccination. *Nat Commun.* 2019;10(1):874.
50. Moorlag S, et al. BCG Vaccination Induces Long-Term Functional Reprogramming of Human Neutrophils. *Cell Rep.* 2020;33(7):108387.
51. Horta AL, et al. Memory-like NK Cells Are a Critical Component of Vaccine-Induced Immunity to *Trypanosoma cruzi* Infection. *J Immunol.* 2024;212(4):617-31.
52. Pinazo MJ, et al. Biological markers for evaluating therapeutic efficacy in Chagas disease, a systematic review. *Expert Review Anti-inf Ther.* 2014;12(4):479-96.
53. de Lima ACB, et al. Predictors of *Trypanosoma cruzi* PCR positivity in patients with chronic Chagas disease. *Mem Inst Oswaldo Cruz.* 2023;118:e230115.
54. Mendes VG, et al. Biomarkers and Echocardiographic Predictors of Cardiovascular Outcome in Patients With Chronic Chagas Disease. *J Am Heart Assoc.* 2023;12(12):e028810.
55. Hoffman KA, et al. Improved Biomarker and Imaging Analysis for Characterizing Progressive Cardiac Fibrosis in a Mouse Model of Chronic Chagasic Cardiomyopathy. *J Am Heart Assoc.* 2019;8(22):e013365.
56. R Core Team. R: A language and environment for statistical computing. <https://www.R-project.org/>.
57. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.
58. Zhang Y, et al. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform.* 2020;2(3):lqaa078.
59. Love MI, et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
60. Zhu A, et al. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics.* 2019;35(12):2084-92.
61. Yu G, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284-7.
62. Ge SX, et al. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics.* 2020;36(8):2628-9.

63. Ashburner M, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25(1):25-9.
64. Gene Ontology C. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* 2021;49(D1):D325-D34.
65. Kanehisa M, et al. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* 2021;49(D1):D545-D51.
66. Luo W, and Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics.* 2013;29(14):1830-1.
67. Subramanian A, et al. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics.* 2007;23(23):3251-3.
68. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Nat Acad Sci USA.* 2005;102(43):15545-50.
69. Li S, et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. *Nat Immunol.* 2014;15(2):195-204.
70. Culhane AC, et al. MADE4: an R package for multivariate analysis of gene expression data. *Bioinformatics.* 2005;21(11):2789-90.

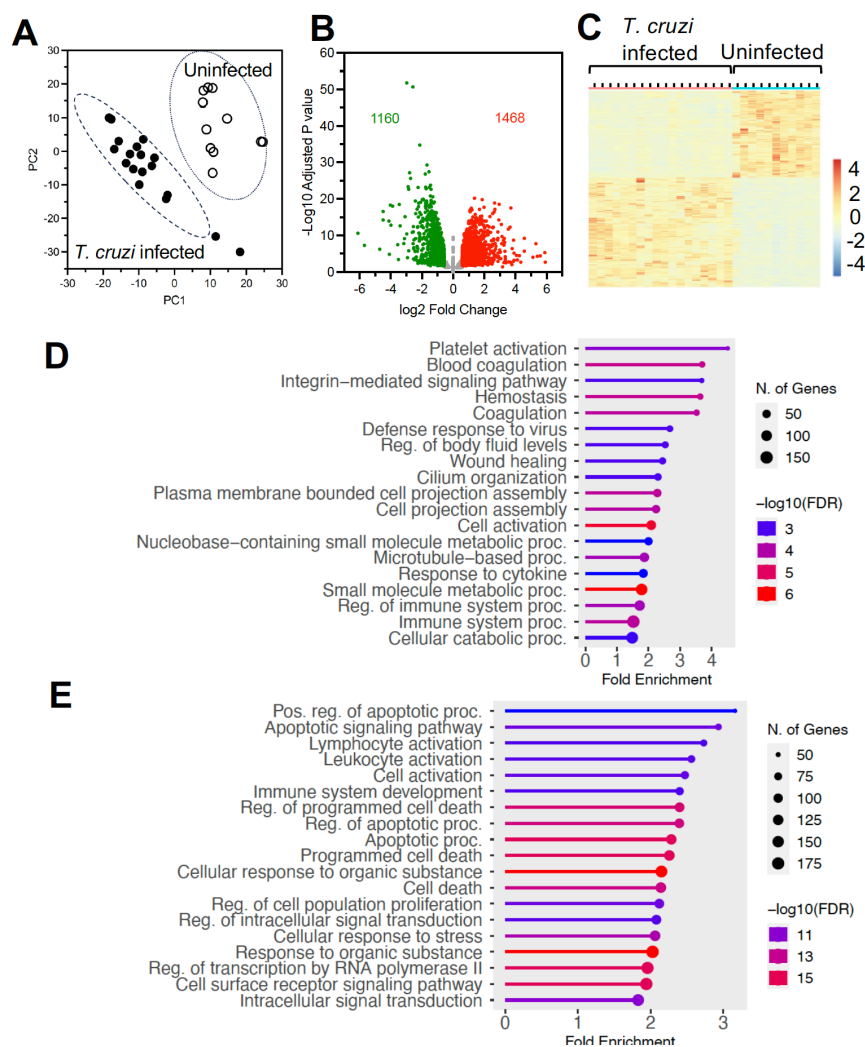


Figure 1. Differences in gene expression profile between *T. cruzi* infected and uninfected macaques

(A) Principal Component Analysis (PCA) plot of PBMC gene expression profiles. Normalized adjusted read counts were used as the inputs for PCA to visualize differences in gene expression profiles among samples. **(B)** Volcano plot of differential gene expression between uninfected and infected macaques. A total of 2,628 genes were differentially expressed (adjusted $P < 0.05$, fold change > 1.5) in PBMCs from *T. cruzi* infected and uninfected macaques, with 1,468 up-regulated (Red) and 1,160 down-regulated genes (Green). **(C)** Heat map of differentially expressed genes, which differences in expression level is color-coded as indicated in the color scale (\log_{10} fold change). Note the clustering of up- and downregulated genes in PBMCs from infected and uninfected macaques. **(D)** Lollipop charts of significantly upregulated pathways in infected macaques. **(E)** Lollipop charts of significantly downregulated pathways in infected macaques.

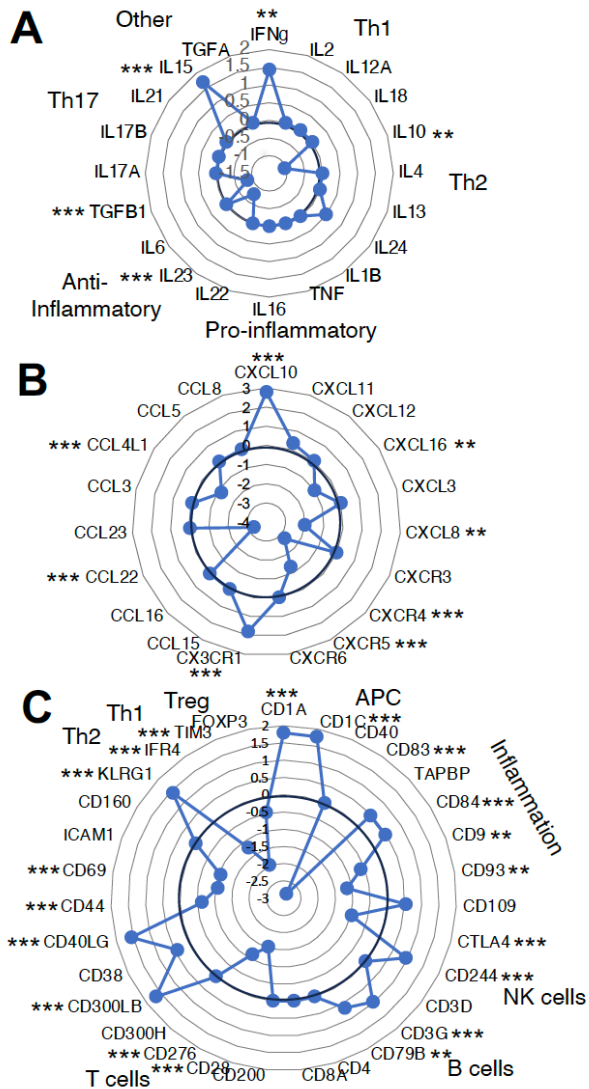


Figure 2. Cytokines, chemokines and immune cell marker expression levels

Radar plots of the expression level for the indicated genes, expressed as the Log₂ fold difference between PBMCs from *T. cruzi*-infected and uninfected macaques, are shown. The thicker line at 0 indicate no difference in expression level. **(A)** Cytokines. **(B)** Chemokines and chemokine receptors. **(C)** Immune cell markers. ** and *** indicate significant differences in expression level between infected and uninfected macaques (adjusted $P < 0.01$ and $P < 0.001$, respectively).

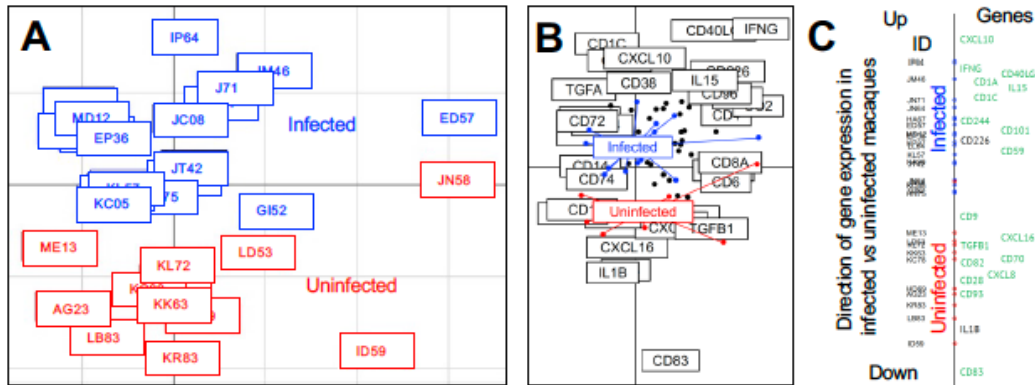


Figure 3. Multivariate analysis of PBMC gene expression between *T. cruzi* infected and uninfected macaques.

(A) BGA of infected and uninfected macaques, indicating clustering of individual macaques. Infected macaques are indicated in blue, uninfected in red. **(B)** Overlay of macaque clustering and gene clustering from BGA analysis. Colored dots and lines represent individual macaques, infected in blue, uninfected in red. Black dots and boxed gene names represent individual genes, with genes found at the ends of the main BGA discrimination axis representing genes most associated with macaque groups. **(C)** Top 20 genes which expression profile is most associated with each macaque group. Genes in the top part of the graph are upregulated, while those at the bottom are downregulated in infected macaques vs uninfected macaques. A significant difference in gene expression level was detected by DESeq2 analysis for the genes labeled in green.

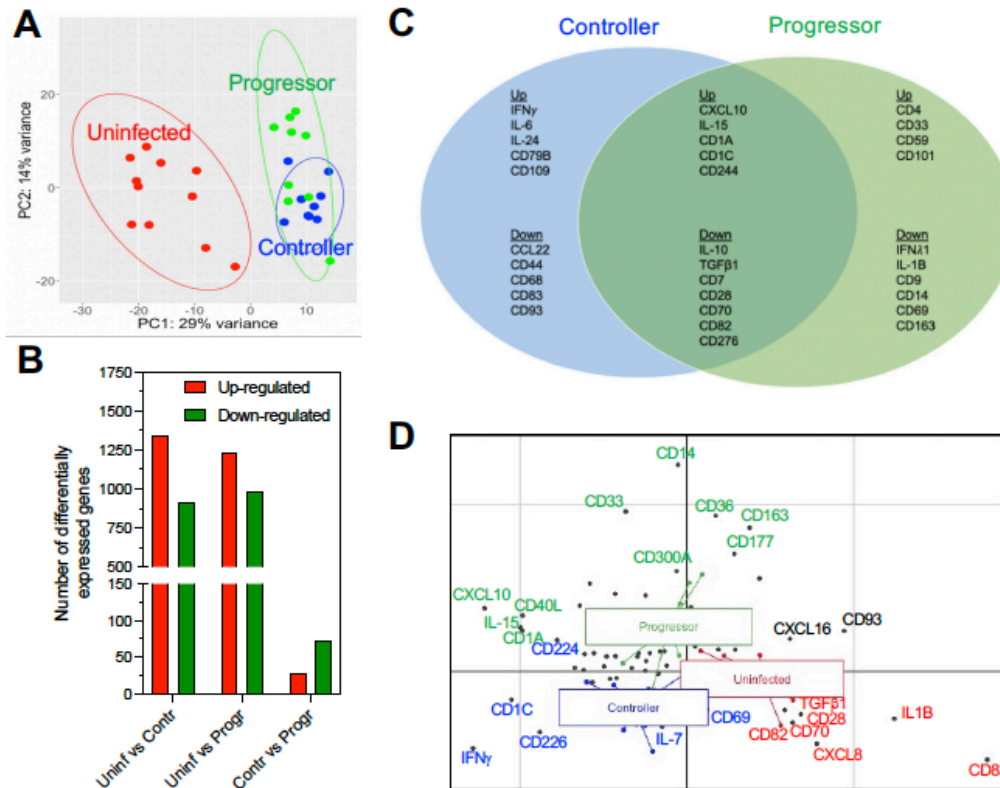


Figure 4. PBMCs gene signature associated with *T. cruzi* parasite control

(A) PCA analysis of gene expression profile of PBMCs from uninfected, progressor and controller macaques. (B) Number of statistically significant (adjusted $P < 0.05$) differentially expressed genes of PBMCs from uninfected, progressor and controller macaques. (C) Venn diagram of statistically significant (adjusted $P < 0.05$) differentially expressed genes associated with the immune response of PBMCs from controller and progressor macaques. Shared and unique up- and down-regulated genes from the two groups are indicated. (D) BGA multivariate analysis of the gene expression signatures differentiating the three groups of macaques. Colored dots and lines represent individual macaques, colored as controller (blue), progressor (green) and uninfected (red). Black dots represent individual genes, with gene labels colored based on which macaque group they are associated with.

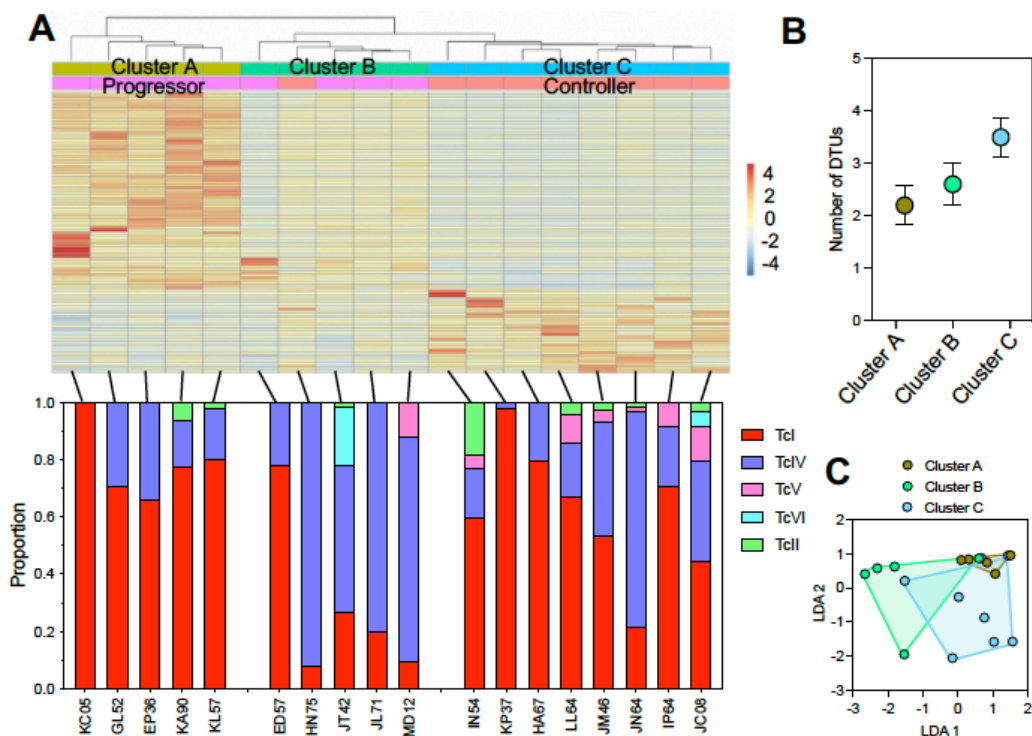


Figure 5. Unsupervised clustering of infected macaques based on PBMC gene signatures
(A) Heat map represents expression level of differentially expressed genes, color coded as indicated (Log₂ fold change). Hierarchical clustering of infected macaques resulted in three clusters, labeled as A, B and C, that showed some overlap with progressor/controller classification: cluster C comprised only controller macaques, while cluster A and B included progressor macaques and one controller macaque that was assigned to cluster B. Bottom panel shows *T. cruzi* parasite DTU composition among individual macaques from the respective gene expression clusters. **(B)** Average number of DTU infecting individual macaques from the respective clusters. Data presented as mean \pm SEM. **(C)** LDA of macaques based on the DTU composition of infecting parasite strains, indicating significant clustering corresponding to that based on PBMC gene signatures (PERMANOVA, $P=0.012$), with up to 77% of individual macaques correctly classified to their respective cluster.

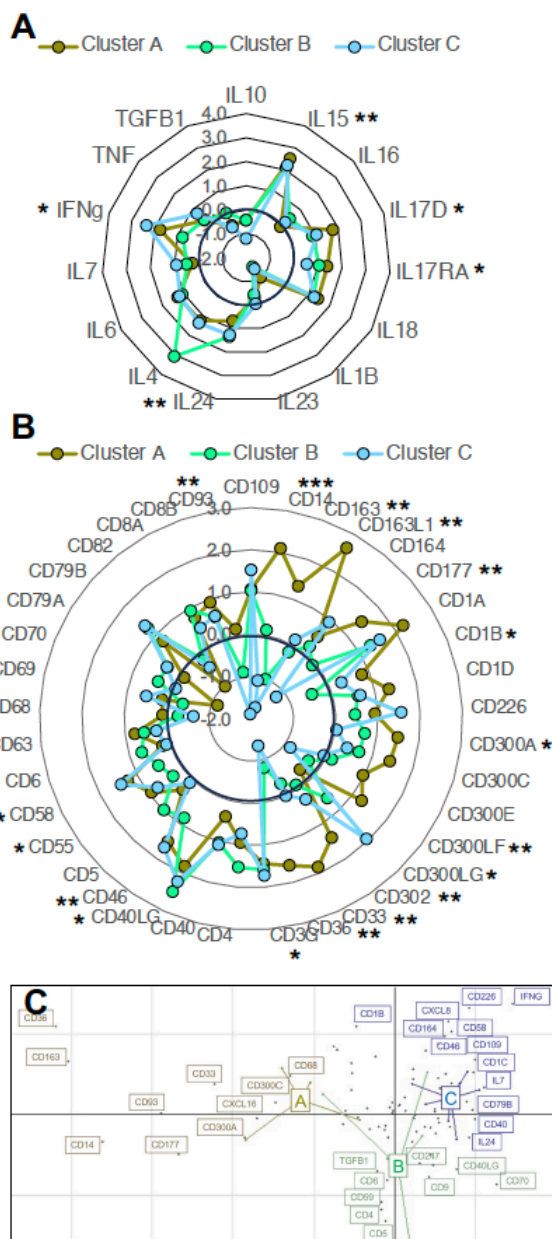


Figure 6. Immune profile associated with *T. cruzi* strain infecting macaques.

(A) Cytokine and **(B)** Immune marker differential expression levels among clusters A, B and C of infected macaques. Radar plots indicate differential expression of the indicated genes, as Log₂ fold change. Thicker black circle corresponds to no change. *, ** and *** indicate significant differences in expression level between infected and uninfected macaques ($P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively). **(C)** BGA multivariate analysis of PBMC gene signatures among the three clusters of macaques. Colored dots and lines represent individual macaques, colored as cluster A (brown), cluster B (green) and cluster C (blue). Black dots represent individual genes, with gene labels colored based on which macaque group they are associated with.

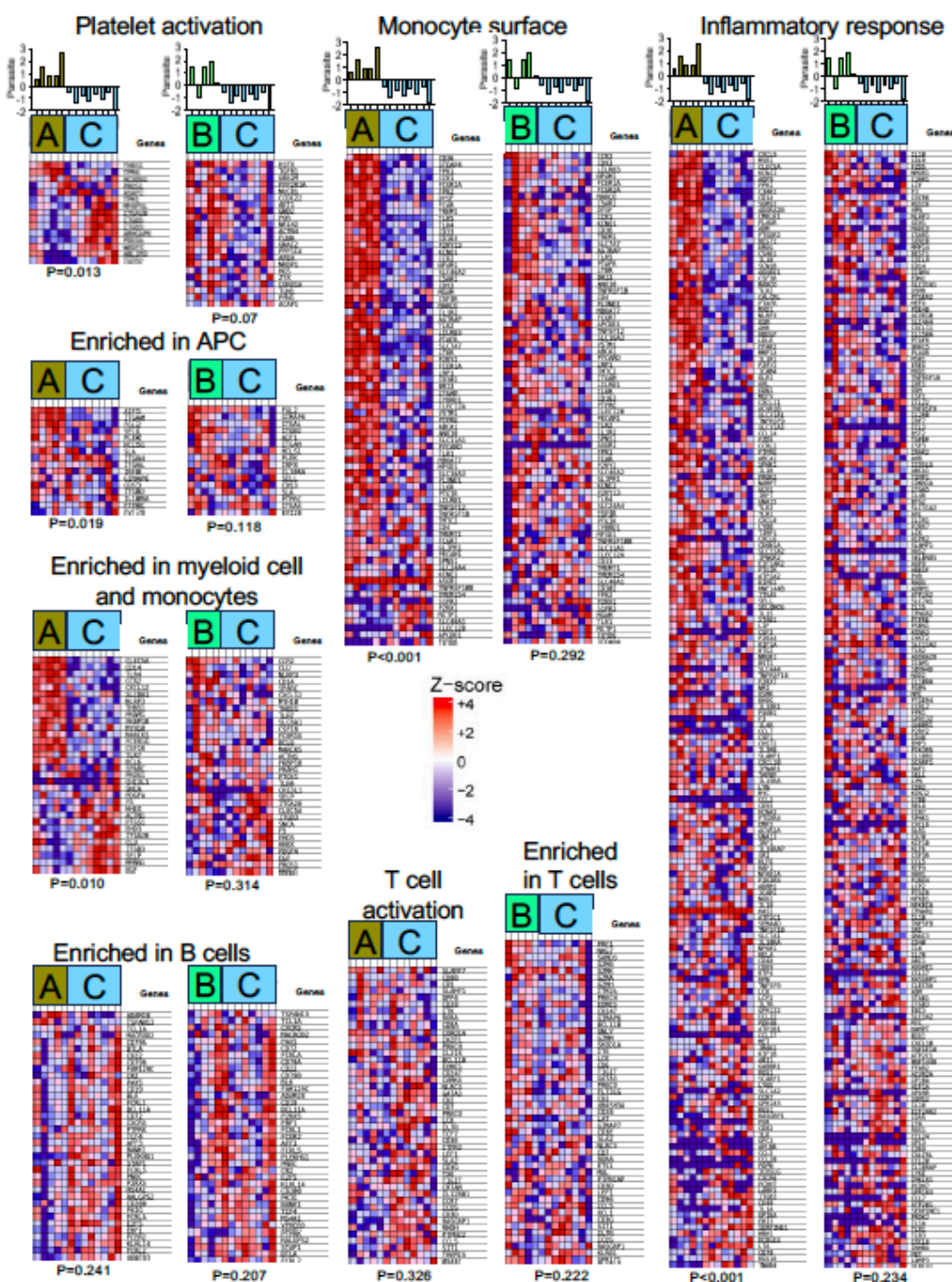


Figure 7. GSEA pathways associated with macaque PBMC responses.

The indicated pathways from the BTM and Hallmark modules were analyzed among progressor macaques from Clusters A and B, compared to controller macaques from Cluster C. Each column represent an individual macaque. The top bar graphs represent changes in blood parasite burden overtime, expressed in Log₂ values, with positive values indicating an increase, and negative values a decrease in parasite burden. The statistical significance of differences in gene expression profile between the indicated macaque clusters is indicated at the bottom of each heatmap.

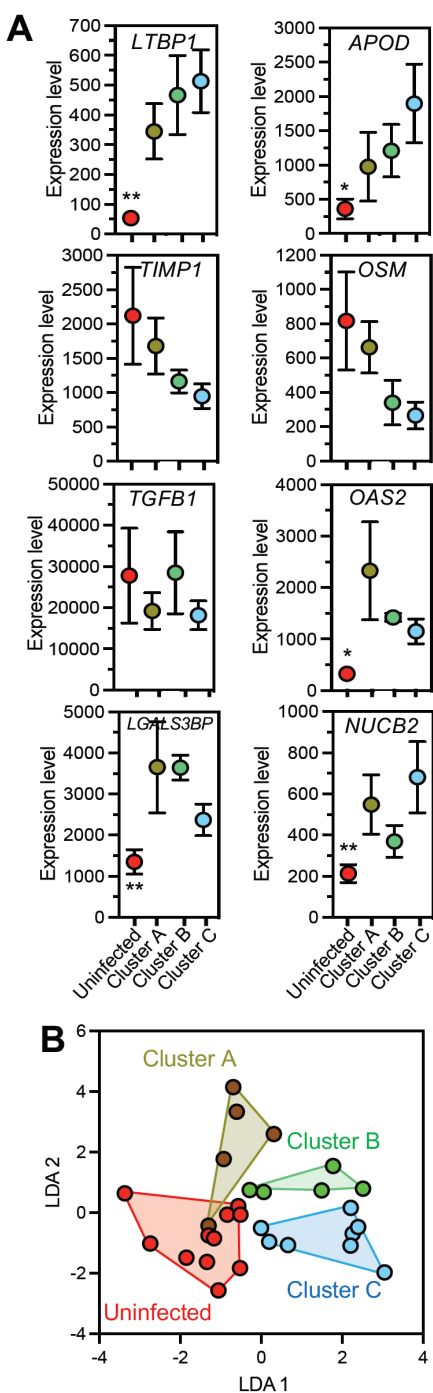


Figure 8. Biomarkers for differential host response to *T. cruzi* infection.

(A) Expression level of the indicated genes as potential biomarkers of the host response/parasite strain composition expressed as mean normalized counts \pm SEM. * and ** indicate significant differences among groups (ANOVA, $P < 0.05$ and $P < 0.01$, respectively). **(B)** LDA of candidate biomarkers indicating a significant discrimination among macaque groups based on the signature of eight biomarkers (PERMANOVA, $P = 0.007$), with 25/29 (86.2%) of individuals correctly classified in their respective group.