Title

Tissue spaces are reservoirs of antigenic diversity for Trypanosoma brucei.

Journal

Nature: New biology, 636(8042)

Authors

Beaver, Alexander

Keneskhanova, Zhibek

Cosentino, Raúl

et al.

Publication Date

2024-12-01

DOI

10.1038/s41586-024-08151-z

Peer reviewed

Article

Tissue spaces are reservoirs of antigenic diversity for *Trypanosoma brucei*

https://doi.org/10.1038/s41586-024-08151-z

Received: 10 January 2023

Accepted: 3 October 2024

Published online: 30 October 2024

Open access

Check for updates

Alexander K. Beaver[1,2], Zhibek Keneskhanova[3,4], Raúl O. Cosentino[3,4], Brian L. Weiss[5], Erick O. Awuoche[5], Gretchen M. Smallenberger[5], Gracyn Y. Buenconsejo[2], Nathan P. Crilly[1,2], Jaclyn E. Smith[2], Jill M. C. Hakim[2], Bailin Zhang[2], Bryce Bobb[2], Filipa Rijo-Ferreira[6], Luisa M. Figueiredo[7], Serap Aksoy[5], T. Nicolai Siegel[3,4] & Monica R. Mugnier[2]✉

The protozoan parasite *Trypanosoma brucei* evades clearance by the host immune system through antigenic variation of its dense variant surface glycoprotein (VSG) coat, periodically 'switching' expression of the VSG using a large genomic repertoire of VSG-encoding genes[1–6]. Recent studies of antigenic variation in vivo have focused near exclusively on parasites in the bloodstream[6–8], but research has shown that many, if not most, parasites reside in the interstitial spaces of tissues[9–13]. We sought to explore the dynamics of antigenic variation in extravascular parasite populations using VSG-seq[7], a high-throughput sequencing approach for profiling VSGs expressed in populations of *T. brucei*. Here we show that tissues, not the blood, are the primary reservoir of antigenic diversity during both needle- and tsetse bite-initiated *T. brucei* infections, with more than 75% of VSGs found exclusively within extravascular spaces. We found that this increased diversity is correlated with slower parasite clearance in tissue spaces. Together, these data support a model in which the slower immune response in extravascular spaces provides more time to generate the antigenic diversity needed to maintain a chronic infection. Our findings reveal the important role that extravascular spaces can have in pathogen diversification.

Every pathogen must contend with the adaptive immune response of its host. *Trypanosoma brucei*, a protozoan parasite and causative agent of human and animal African Trypanosomiasis, has evolved a sophisticated strategy to evade this highly flexible and specific host response[14]. Transmitted by the bite of the tsetse fly, *T. brucei* lives extracellularly in the blood, lymph and interstitial tissue spaces of its mammalian host[15]. To escape clearance by a continuous onslaught of host antibodies, the parasite periodically 'switches' expression of its immunogenic variant surface glycoprotein (VSG) coat to new, antigenically distinct variants[1]. With a genomic repertoire of thousands of different VSG-encoding genes[2–5] and the ability to generate new VSGs through recombination[6,8,16], the parasite has an enormous capacity for altering its antigenic profile.

Studies examining *T. brucei* antigenic variation in vivo have focused nearly exclusively on parasites in the blood, revealing complex VSG expression dynamics[6–8]. However, it has recently become clear that many, if not most, *T. brucei* parasites inhabit extravascular spaces during both experimental and natural infections[9–13]. Although research has shown that tissue-resident parasites adapt to these environments[11], cause tissue-specific symptoms[15] and are associated with increased disease severity[13], it remains unclear why parasites invade tissue spaces and what role these populations might have in infection.

Several older studies suggested a role for tissue-resident parasites in antigenic variation. These studies found that brain- and lymph-resident parasite populations were antigenically distinct from those in the blood[17–20], and that antigenic types detectable in the lymphatic fluid could be detected in the blood at later time points during infection[19]. This led these researchers to propose that extravascular spaces might be a site for antigenic variation, with antigenic variants generated in extravascular spaces contributing to systemic immune evasion. However, a later study focusing on very early time points postinfection found no antigenic difference between *T. brucei* populations within the blood and several extravascular spaces, and the community consequently ruled out a role for extravascular parasites in antigenic variation[21,22]. These early investigations were limited by the methodology available at the time, which relied on the use of VSG-specific antisera to analyse parasite antigenic diversity. Modern high-throughput sequencing methods allow VSG expression to be measured accurately and in high resolution[7]. Given the mounting evidence that *T. brucei* parasites persist in and adapt to extravascular spaces, there is a clear need to reinvestigate the dynamics of VSG expression within tissue spaces.

Here, we use VSG-seq[7], a targeted RNA sequencing (RNA-seq) approach for profiling the VSGs expressed in *T. brucei* populations, to characterize the VSGs expressed by extravascular *T. brucei* parasites.

[1]Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD, USA. [2]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [3]Division of Experimental Parasitology, Faculty of Veterinary Medicine, Ludwig-Maximilians-Universität München, Munich, Germany. [4]Biomedical Center Munich, Division of Physiological Chemistry, Faculty of Medicine, Ludwig-Maximilians-Universität München, Munich, Germany. [5]Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA. [6]Division of Infectious Diseases and Vaccinology, Berkeley Public Health Molecular and Cell Biology Department, Berkeley, CA, USA. [7]Gulbenkian Institute for Molecular Medicine, Lisboa, Portugal. ✉e-mail: mmugnie1@jhu.edu

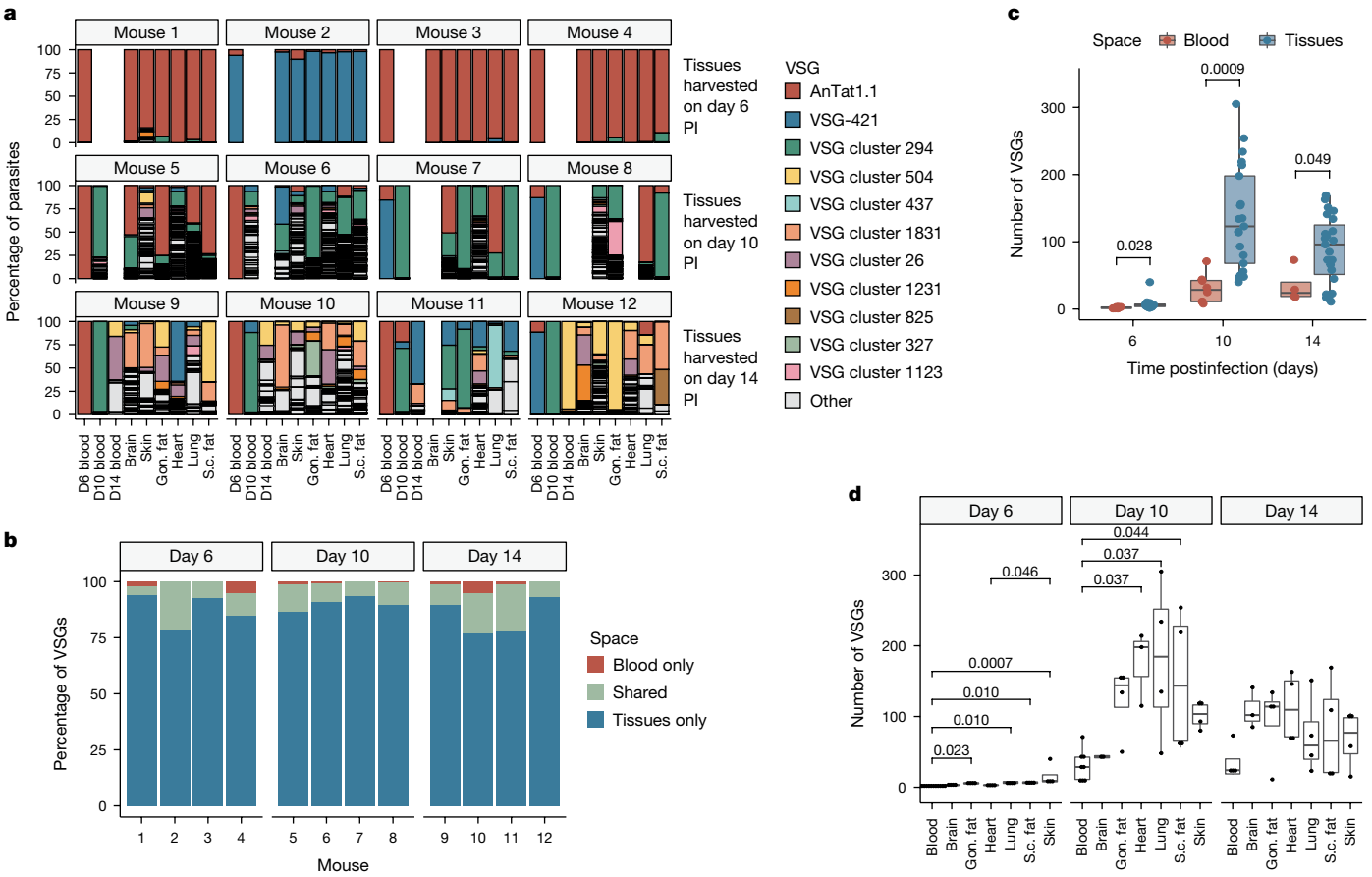


Fig. 1 | Extravascular parasites harbour most of the antigenic diversity in an infection. a, The percentage of parasites expressing each VSG within a space. The 11 VSGs with the highest overall expression are coloured, and all other VSGs are in grey as 'other'. **b**, Stacked bar graphs from each infected mouse representing the percentage of VSGs that were found exclusively within the blood (red), exclusively within tissue spaces (blue) or shared by both the blood and at least one tissue (green). **c**, Quantification of the number of VSGs found within the blood (red) or tissue spaces (blue) at each time point (Shapiro–Wilk normality test followed by a two-tailed Student's *t*-test Benjamini–Hochberg corrected). **d**, The number of VSGs in each tissue space (Shapiro–Wilk normality test followed by a two-tailed Dunnett's test). In **a**–**d**, $n = 12$ total mice with four biologically independent animals per time point over two independent experiments. In boxplots, boxes represent values between the first (25%) and third (75%) quartiles with a line at the median, and extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range. Gon. fat, gonadal fat; PI, postinfection; s.c. fat, subcutaneous fat.

Our results show that extravascular spaces are major reservoirs of antigenic diversity during *T. brucei* infection and that this parasite niche is central to the parasite's ability to continuously outmanoeuvre the immune system.

Most VSGs are in extravascular spaces

To investigate how tissue-resident parasites contribute to antigenic variation in vivo, we intravenously infected 12 mice, each with approximately 5 pleomorphic *T. brucei* EATRO1125 90-13 parasites[23]. We collected blood, then perfused mice with PBS glucose (Extended Data Fig. 1) and harvested the heart, lungs, gonadal fat, subcutaneous fat, brain and skin at 6, 10 and 14 days postinfection. For each sample, we extracted RNA and quantified *T. brucei* VSG expression using VSG-seq[7]. A single 'initiating' VSG (either AnTat1.1 or EATRO1125 VSG-421) dominated expression in both the blood and tissues on day 6 (Fig. 1a), in line with previous observations[24,25]. At later time points, VSG expression dynamics became more complex, with more VSGs expressed, a unique composition of VSGs in each tissue, and no single dominating variant (Fig. 1a). Although tissue-specific expression of variant surface proteins is a feature in other organisms that use antigenic variation[26–31], we found no evidence for tissue-specific VSGs or VSG sequence motifs (Extended Data Fig. 2). Instead, we observed an increase in antigenic diversity in extravascular spaces. The number of detectable VSGs in tissue spaces was, on average, two to four times higher than the blood (Fig. 1c). This did not correlate with parasite load (Extended Data Figs. 3 and 4a,b) and was not driven by any specific tissue (Fig. 1d). In addition, parasite differentiation to the non-dividing tsetse infective form via quorum sensing, which is marked by expression of the *PAD1* gene[32], did not correlate with VSG diversity in either the blood or tissues at the population level (Extended Data Fig. 4c). The overall contribution of tissue-resident parasites to antigenic diversity in a single infection was large: at any time, roughly 87% of expressed VSGs in any individual infection were found exclusively within extravascular spaces (Fig. 1b).

VSG-seq is a bulk measure of VSG expression. To be sure that the increased VSG diversity we observed in tissues was not due to the derepression of silent VSGs within individual cells, we performed single-cell RNA-seq using the SL-Smart-seq3xpress platform to quantify VSG expression in single cells from blood and tissue samples[33]. Because *T. brucei* can form new 'mosaic' VSGs through recombination[6,8,16,34], which are, by definition, absent from reference genomes, we initially used our VSG-seq pipeline to de novo assemble VSGs in each cell. This approach accounts for the possibility that expressed VSGs might be absent from the reference genome, either as a result of mosaic VSG formation or as a result of an incomplete genome assembly, potentially affecting quantification and/or read mapping. Only one VSG assembled

in most cells (94.2%; Extended Data Fig. 5a). We also mapped sequencing reads to the EATRO1125 genome, which could reveal more subtle signatures of derepression. By this analysis, there was no obvious difference between the blood and the tissues in the number of expressed VSGs in each cell (Extended Data Fig. 5b) or in the relative expression of the most abundant VSG in each cell (Extended Data Fig. 5c).

To estimate the number of cells maintaining monogenic expression, we defined a cell as maintaining monogenic expression if 80% of unique molecular identifiers (UMIs) mapping to VSGs mapped to a single VSG. Although mapping to the genome suggested that most cells (62.7%) maintained monogenic expression based on this threshold, the proportion of cells estimated to maintain monogenic expression was lower than estimated by de novo assembly. Comparison of the two analyses showed that of those cells expressing more than one VSG by genome alignment, 85% were found to express only one VSG by mapping to the de novo assembled VSGs and using the same 80% threshold for defining monogenic expression (Supplementary Data 1–3). Further investigation revealed that in most of these cases (95.3%) the alignment to many genomic VSGs was an artefact, where the assembled VSG was not well represented within the annotated sequences of the EATRO1125 genome or there were several VSGs with high similarity to the assembled VSG, leading to inaccurate VSG expression quantification (Extended Data Fig. 5d). We estimate that 93.4% of cells were probably expressing only one VSG, with no bias for multigenic VSG expression in tissue spaces (Extended Data Fig. 5e, shades of green, and Extended Data Table 1). In 5.2% of cells, reads mapped to several EATRO1125 VSGs, but no VSG could be assembled (grey). Although it is impossible to distinguish between multi- and monogenic VSG expression in this set of cells, the proportion of cells in this category did not differ between blood and tissues. The few cells that seem to express more than one VSG (1.05% of cells in the blood and 1.03% of cells in the tissues) could indicate sorting doublets or could represent cells mid-switch (Extended Data Fig. 5e, red and purple). Overall, these data indicate that VSG monogenic expression is maintained by most cells in both extravascular spaces and the blood and that the increased VSG diversity we observe in tissue spaces is unlikely to be due to the specific derepression of silent VSGs in tissue populations.

Common VSGs appear in tissues first

The high antigenic diversity observed in tissues could serve to maintain a chronic infection. If antigenic variation occurs relatively rarely in the bloodstream, then parasites from extravascular spaces might serve as a source of new, antigenically distinct, VSGs. In line with this, tissue spaces contain more 'unique VSGs', those VSGs that are expressed exclusively within one space in an infection, than the blood (Fig. 2a,b). To examine the potential for tissue-resident VSGs to contribute to antigenic variation systemically, we identified VSGs only expressed in tissues on day 6 postinfection and analysed whether they later appeared within the blood. Most (74%) of these VSGs were expressed within the blood on day 10 or 14. Analysis of individual VSGs revealed that rare VSGs expressed at low levels exclusively in tissue spaces also have the potential to become ubiquitously expressed within a host (Fig. 2c). In addition to a model in which tissue spaces provide new VSGs to re-seed the blood, it is possible that tissue-resident parasites undergo antigenic variation before blood-resident populations. Therefore, these data could be explained by either a trigger within the tissue environment that induces parasite switching or a differential selective pressure imposed by the tissue environment.

Parasite clearance is delayed in tissues

Indeed, our data suggest that the environment within tissue spaces is distinct from the blood, with parasite clearance occurring at different rates in each space. Although parasites expressing the initiating VSG were cleared from the blood by day 10 postinfection, they were not cleared from tissues until at least day 14 (Fig. 3a). This suggests that VSG-specific parasite clearance from extravascular spaces is delayed, but not abolished, compared to the blood. VSG-seq is a measure of VSG expression at the transcript level, however. To confirm this observation at the protein level, we performed flow cytometry on *T. brucei* cells from the blood, lungs and gonadal fat, using the tdTomato-expressing 'triple reporter' *T. brucei* EATRO1125 AnTat1.1E cell line[35] (Fig. 3b,c and Extended Data Table 2). The flow cytometry analysis showed a detectable AnTat1.1-expressing tissue parasite population at day 13 postinfection, a time point at which AnTat1.1^{+} parasites were undetectable, or nearly undetectable, in the blood. Similar to the increase in antigenic diversity we observed in every tissue space, this delay in clearance, observed at both the RNA and protein levels, was not tissue-specific. Thus, the immune mechanisms influencing extravascular parasite clearance seem to be general features of extravascular spaces.

High diversity after infection by tsetse

A benefit of starting infections with a small intravenous inoculum is that it creates convenient and reproducible infections. In nature, however, a tsetse fly bite introduces thousands of parasites into the skin, each expressing a single metacyclic VSG (mVSG) from a limited repertoire[36]. To ensure that our observations held true in this more complex context, we repeated our infections using a more natural tsetse bite infection model. Infections were initiated in five mice by tsetse bite using flies infected with RUMP 503 *T. brucei* parasites[37]. We used VSG-seq to quantify VSG expression in the blood on day 5 postinfection and the blood and tissues on day 14 postinfection. In line with our previous observations, we found that in tsetse-initiated infections roughly 80% of VSGs were exclusively expressed within extravascular spaces (Fig. 4a) and tissue populations harboured more VSGs than the blood (Fig. 4b). This demonstrates that extravascular spaces are the primary reservoir of antigenic diversity, even when infections are initiated by fly bite.

Parasite populations were more diverse at early time points in tsetse infections than intravenous infections, probably due to the larger and more heterogeneous inoculum delivered by the fly. To analyse VSG-specific parasite clearance in tissues, we quantified expression of the most abundantly expressed VSG in the blood of each mouse on day 5 as well as the known mVSG repertoire from this parasite strain, which should represent the repertoire of VSGs expressed at the start of an infection. In both cases, we found that on day 14, tissue spaces still contained parasites expressing the most abundant VSG and/or mVSGs, whereas these VSGs were expressed by no or very few parasites in the blood (Fig. 4c,d). This suggests that in tsetse bite-initiated infections, as we observed in intravenous infections, tissue parasite populations are cleared at a slower rate than parasites in the blood.

Delayed clearance increases diversity

The increased antigenic diversity in extravascular spaces could be explained by the distinct clearance dynamics we observe, as prolonged survival in tissues could provide more time for parasites to switch. To test whether there was a link between parasite survival and increased diversity, we sought to interrupt parasite clearance in tissues. Because parasite clearance in the blood coincides with the appearance of anti-VSG IgM, between days 8 and 10, and parasite clearance in the tissues coincides with the anti-VSG IgG response, between days 10 and 14 (refs. 1,12,38–41), we proposed that clearance in tissues is dependent on the anti-VSG IgG response. Thus, the loss of IgG might abrogate the clearance of tissue-resident parasites. To test this hypothesis, we infected activation-induced cytidine deaminase (AID) knockout (AID Cre) mice[42], which only produce IgM antibodies (Extended Data Fig. 6), and analysed blood and tissues from days 6 and 14 postinfection by VSG-seq. As expected, clearance of the initiating VSG was severely

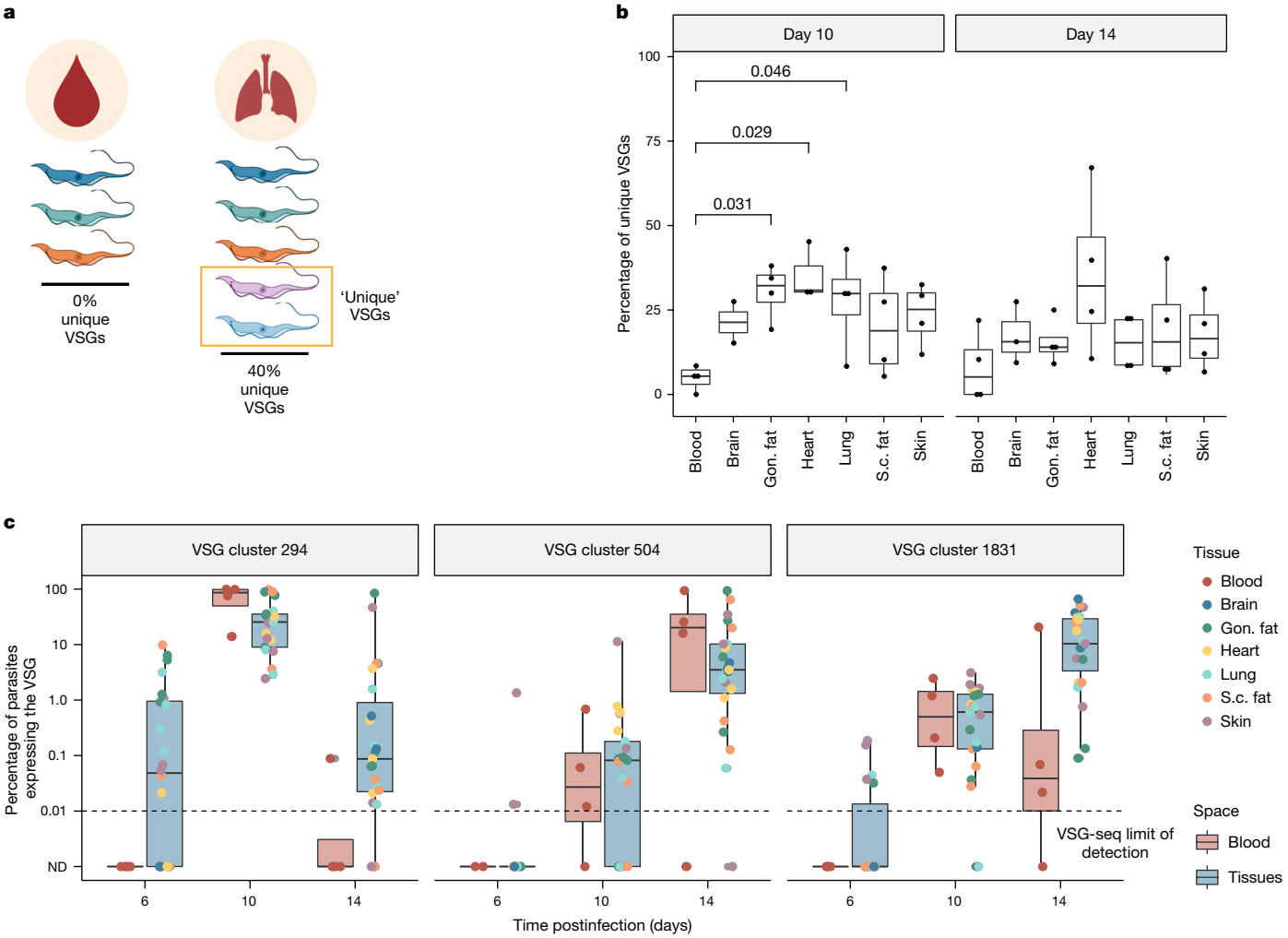


Fig. 2 | Tissue-resident parasites express a unique repertoire of VSGs during infection. a, We define 'unique' VSGs as those VSGs solely found within a specific space in a mouse. **b**, The percentage of VSGs that were unique to one space within a mouse (Shapiro–Wilk normality test followed by a two-tailed Dunnett's test). Day 6 samples were excluded from this analysis because few VSGs are expressed at this point. $n = 4$ biologically independent animals per time point over two independent experiments (total of eight mice). **c**, The expression of three representative VSGs (cluster 294, 504 and 1831) within blood and tissue samples on days 6, 10, and 14. $n = 12$ total mice with four biologically independent animals per time point over two independent experiments. ND indicates that the VSG was not detected. In boxplots, boxes represent values between the first (25%) and third (75%) quartiles with a line at the median, and extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range. Illustration in **a** created using BioRender (https://biorender.com).

delayed in tissues on day 14 in AID$^{-/-}$ mice, suggesting that IgG is important, if not critical, for the clearance of extravascular parasites. This could be explained by the fact that IgM, a bulky pentamer, does not diffuse efficiently into tissue spaces[43,44], whereas IgG, a monomer, readily diffuses. We also observed a defect in the clearance of blood-resident parasites in AID$^{-/-}$ mice compared to the blood of wild-type mice (Fig. 5a). In both the blood and tissues of AID$^{-/-}$ mice, where clearance was delayed, more VSGs were detected on day 14 postinfection compared to wild-type (Fig. 5b), showing a direct relationship between the timing of parasite clearance and VSG diversity. Regardless of their local environment (intra- or extravascular), longer-lived parasite populations generated more diverse sets of VSGs.

Discussion

The idea that antigenic variation might occur outside the bloodstream, with extravascular populations contributing to immune evasion within the bloodstream, is not a new one. Here, using modern high-resolution techniques, we provide evidence for this long-standing hypothesis. In both needle- and tsetse bite-initiated infections, we find that extravascular spaces are the primary reservoir of VSGs, accounting for most antigenic diversity in any individual infection. The number of VSGs we detected in the blood matches previous estimates[6–8], whereas the diversity in tissues is two to four times higher. Our data indicate that this is at least partially due to slower clearance dynamics in extravascular spaces and highlight the role that tissue spaces can have in pathogen diversification.

Although extravascular parasite populations were highly antigenically diverse, we saw no evidence of tissue-specific VSG expression. Because parasites invade tissues efficiently before much VSG switching has occurred, it seems unlikely that any specific VSG is required for tissue invasion. Whether VSGs could influence parasite fitness in specific host spaces is less clear. We measured VSG expression up to day 14 postinfection, at which point tissue-resident populations are just beginning to diverge from one another and the blood. It is therefore possible that, as these populations further evolve, there may be a selection for VSGs better adapted to certain tissue spaces.

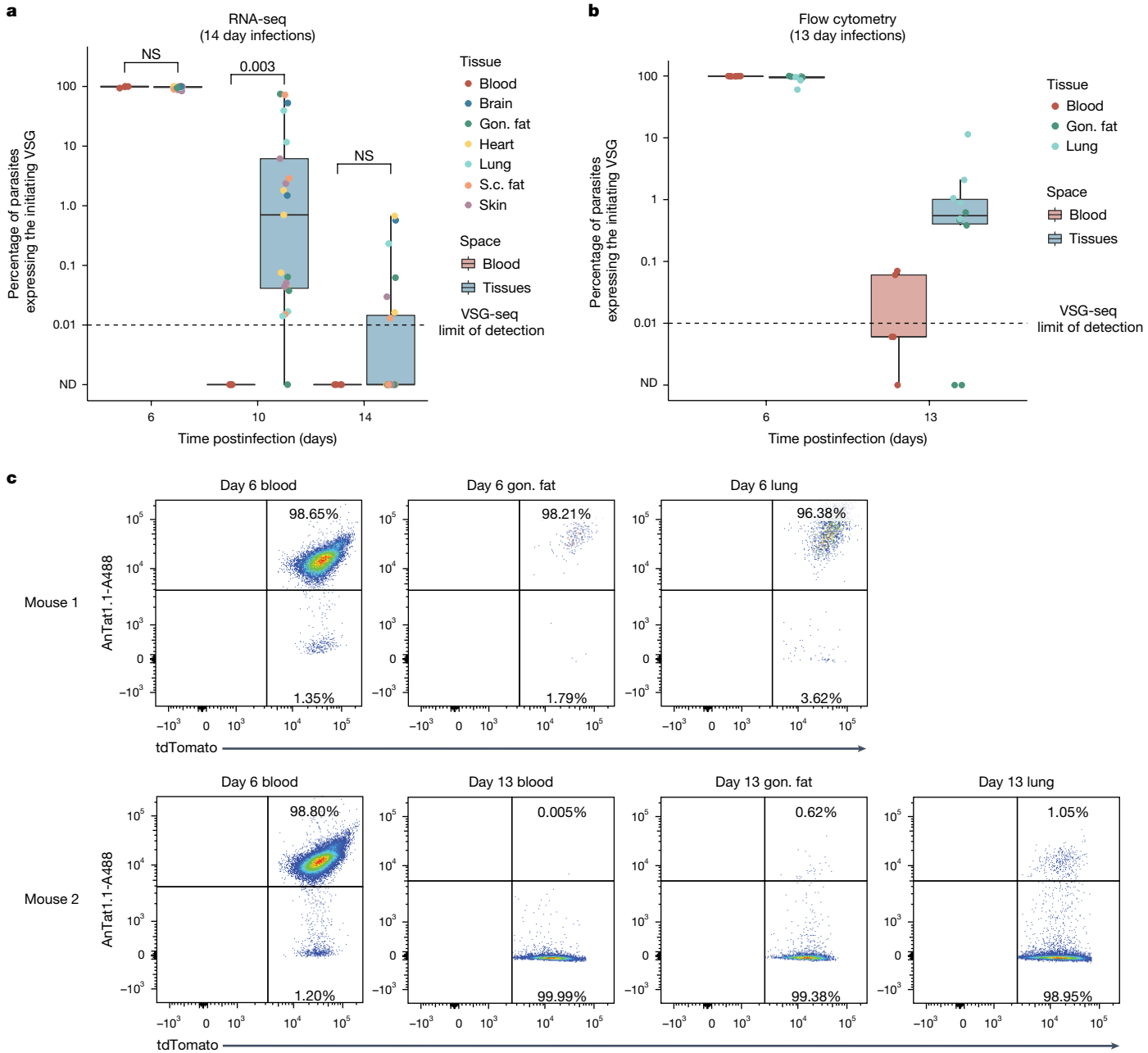


Fig. 3 | VSG-specific parasite clearance is slower in tissues than in the blood. **a**, The percentage of parasites expressing the initiating VSG (AnTat1.1 or VSG-421) at days 6, 10 and 14 postinfection. Tissue samples were grouped together (blue) and compared to blood samples (red) (two-tailed Wilcoxon test). $n = 12$ total mice with four biologically independent animals per time point over two independent experiments. **b**, Quantification of the number of parasites that were tdTomato positive and stained positive for AnTat1.1 by flow cytometry ($n = 10$ total mice with five biologically independent mice per time point examined over one independent experiment). The horizontal dotted line represents the limit of detection for VSG-seq. **c**, Representative flow cytometry plots from tissues collected from mice infected with chimeric triple-marker parasites that express tdTomato constitutively in their cytoplasm. Parasites were stained with anti-AnTat1.1 antibody. In boxplots, boxes represent values between the first (25%) and third (75%) quartiles with a line at the median, and extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range. NS, not significant.

It has previously been shown that in the bloodstream alone *T. brucei* expresses more VSGs than seem to be required for immune evasion[7]. Here we find that the expressed VSG diversity within host tissues is even greater. Although on the surface it could seem to be disadvantageous for *T. brucei* to use so many different antigens this quickly, this striking diversity could be important for the parasite. In natural infections, particularly in wild animals in which pre-existing anti-VSG immunity is more likely to exist, a high switch rate may be required to ensure some parasites successfully evade the host's existing antibody repertoire. Moreover, infections in the wild can last for months to years[14,45]. During these long infections, the large reservoirs of VSGs found in tissues may be essential for the maintenance of a chronic infection.

Indeed, our data support the idea that the large reservoirs of antigenic diversity in extravascular spaces contribute to systemic infection when parasites re-enter the blood after switching: rare VSGs expressed exclusively in tissues at early time points are expressed in the blood and other spaces later. This is also in line with another recent study, which showed that blood-resident parasites are largely non-replicative, indicating that tissue-resident parasites may be required to re-seed the blood[46]. There is another intriguing explanation for this observation,

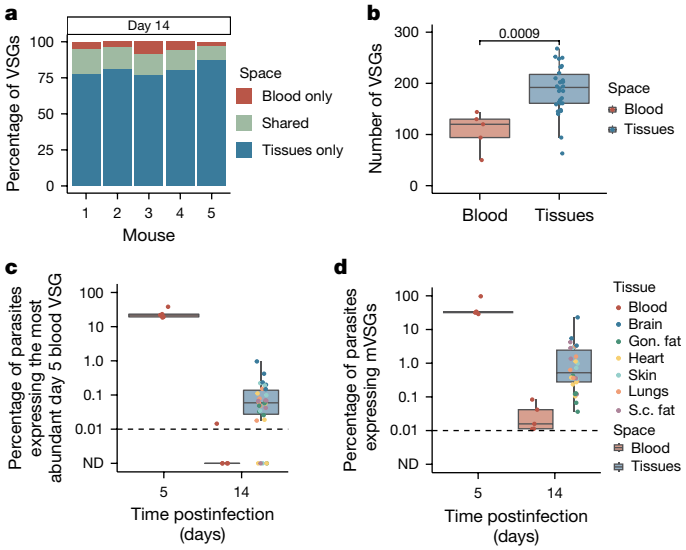


Fig. 4 | Tsetse bite-initiated infections show increased antigenic diversity and delayed immune clearance in extravascular spaces. Data from five mice infected with RUMP 503 parasites from a tsetse fly bite. **a**, Bar graphs representing the percentage of VSGs in each mouse that were found exclusively within the blood (red), exclusively within tissue spaces (blue) or shared by both the blood and at least one tissue (green). **b**, Quantification of the number of VSGs found within the blood (red) or tissue spaces (blue) on day 14 postinfection (Shapiro–Wilk normality test followed by a two-tailed Student's *t*-test Benjamini–Hochberg corrected). **c**, The percentage of parasites within each mouse expressing the most abundant VSG from the day 5 blood. **d**, The percentage of parasites expressing one of the mVSGs found to be expressed by RUMP 503 parasites in the salivary gland of a tsetse fly. In boxplots, boxes represent values between the first (25%) and third (75%) quartiles with a line at the median, and extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range. For **a**–**d**, $n = 5$ biologically independent mice examined over one independent experiment.

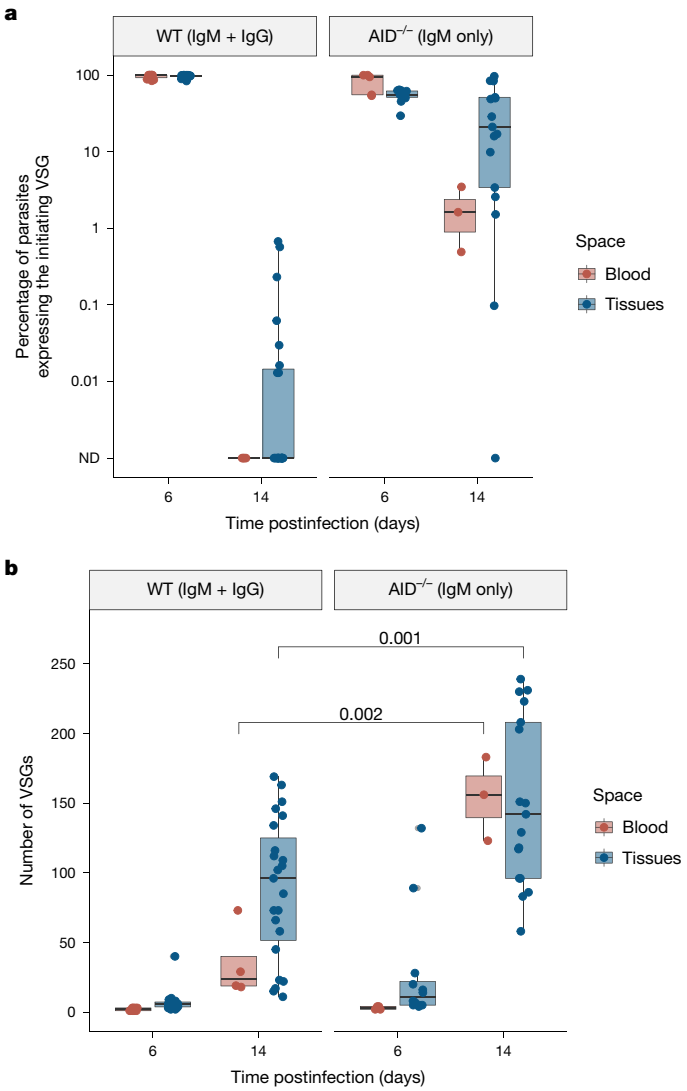


Fig. 5 | Delayed parasite clearance correlates with an increase in VSG diversity. a, The percentage of parasites expressing the initiating VSG (AnTat1.1 or VSG-421) in both wild-type (WT) and AID$^{-/-}$ mice. **b**, The number of VSGs expressed within the blood (red) and tissues (blue) of WT and AID$^{-/-}$ mice (Shapiro–Wilk normality test followed by a two-tailed Student's *t*-test). For WT mice, $n = 12$ total biologically independent mice with four mice representing harvested tissues for each time point. For AID$^{-/-}$ experiments, $n = 5$ total biologically independent mice with tissues from two mice represented on day 6 and tissues from three mice represented on day 14. Day 6 blood represents samples from all five mice. In boxplots, boxes represent values between the first (25%) and third (75%) quartiles with a line at the median, and extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range.

however. Switching in *T. brucei* is known to be semipredictable[7,47], and it is possible that tissue-resident parasites simply switch earlier, or more frequently, than those in the blood. In this case, the same VSGs would arise independently in every population without any parasite movement between spaces. Whereas the high vascular permeability observed after the initial stages of infection indicates that parasites probably move back and forth between the vasculature and extravascular spaces, the fact that tissue-resident populations contain many unique VSGs suggests that blood re-entry may represent a bottleneck for the parasite. An increased rate of switching in tissues could be explained by a higher proportion of dividing slender form parasites in these spaces, as has been observed in the adipose tissue[13], but we found no correlation at the population level between *PAD1* expression and increased VSG diversity. Notably, increased diversity is still observed when the overall parasite load in tissues is lower than the blood (Extended Data Fig. 3c). It is therefore exciting to speculate that some aspect of the extravascular environment supplies a molecular or physical stimulus that promotes VSG switching.

Whereas the timing of VSG switching could be the result of an environmental trigger, our data indicate that the higher antigenic diversity in extravascular spaces compared to the blood can also be explained, at least in part, by the dynamics of the immune response to *T. brucei* in each space. In tissue spaces, we observed slower VSG-specific clearance of parasites than in the blood, potentially providing these populations more time to undergo antigenic variation. Furthermore, newly switched parasites are still vulnerable to immune clearance by antibodies against their previous VSG for roughly 29 h, so moderate delays in immune clearance could allow more switched parasites to survive[48]. The direct relationship we observed between the timing of parasite clearance and antigenic diversity in AID$^{-/-}$ mice supports this model, with even small delays in clearance showing great effects on parasite VSG diversity in both the tissues and the blood.

Our results show that the production of IgG has a key role in clearing parasites from tissues, which we propose could be related to the ready diffusion of this molecule within extravascular spaces facilitating parasite clearance. It is important to note, however, that whereas the difference in timing between the anti-VSG IgM and IgG responses inspired the hypothesis that IgG might be important for parasite clearance in tissues, our study does not prove that this difference in timing explains the delayed clearance in tissue spaces. The local immune

response to *T. brucei* is complex, with many mechanisms likely to have a role in parasite detection and clearance; our results indicate that IgG is an important player in this response. The increased VSG diversity we observe is certainly multifactorial and could be influenced by parasite factors, such as metabolism, division, motility and antibody internalization, and host factors, such as extravascular environmental stresses, the local immune response and vascular permeability. More research will be required to fully understand the complex nature of the *T. brucei* host–pathogen interaction within the extravascular niche.

Altogether, our results outline a model in which *T. brucei* parasites 'hide' in extravascular spaces to generate new antigenic variants capable of exiting tissues and aiding in systemic immune evasion. Coupled with other recent studies[9–13,46,49], this suggests a framework for the progression and pathogenesis of *T. brucei* infections where, instead of being the primary parasite reservoir, the blood may represent a transient population that is regularly reseeded by extravascular parasites. The vasculature, then, might act as a highway system for movement between the tissue spaces and for eventual transmission back into the tsetse fly.

Interfering with the egress from or establishment within tissue spaces might be a strategy for treating *T. brucei* infections or other infections with pathogens that rely on the distinct features of the extravascular environment. In line with this, one recent study showed that partial inhibition of *T. brucei* tissue invasion using antibodies against P- and E-selectins results in prolonged survival in mice[13]. This also fits with data from another group that found immotile *T. brucei* parasites, probably unable to invade tissues, were no longer infectious[50]. It is possible that without the proper establishment of parasite tissue reservoirs, overall antigenic diversity is lowered, limiting the parasite's capacity for immune evasion and leading to a decrease in parasite burden. Defining the dynamics and variation of parasites both within and between spaces, as well as the unique host environment within each tissue space, will be central to understanding how *T. brucei* consistently avoids immune clearance and harnessing this mechanism for disease control.

More broadly, these data demonstrate how different environmental and immune pressures within a host can influence pathogen diversification. The production of genetic heterogeneity within an infection is important for many pathogen virulence processes, including establishing and maintaining infection, facilitating immune evasion, generating drug resistance and adapting to different host environments. The extravascular environment has a unique role in promoting pathogen evolution, and *T. brucei* serves as a valuable model for understanding this aspect of the host–pathogen interface.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-08151-z.

1. Magez, S. et al. The role of B-cells and IgM antibodies in parasitemia, anemia, and VSG switching in *Trypanosoma brucei*-infected mice. *PLoS Pathog.* **4**, e1000122 (2008).
2. Cross, G. A. M., Kim, H. S. & Wickstead, B. Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.* **195**, 59–73 (2014).
3. Müller, L. S. M. et al. Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**, 121–125 (2018).
4. Hertz-Fowler, C. et al. Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS ONE* **3**, e3527 (2008).
5. Cosentino, R. O., Brink, B. G. & Nicolai Siegel, T. Allele-specific assembly of a eukaryotic genome corrects apparent frameshifts and reveals a lack of nonsense-mediated mRNA decay. *NAR Genom. Bioinform.* **3**, lqab082 (2021).
6. Hall, J. P. J., Wang, H. & David Barry, J. Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. *PLoS Pathog.* **9**, e1003502 (2013).
7. Mugnier, M. R., Cross, G. A. M. & Papavasiliou, F. N. The in vivo dynamics of antigenic variation in *Trypanosoma brucei*. *Science* **347**, 1470–1473 (2015).
8. Jayaraman, S. et al. Application of long read sequencing to determine expressed antigen diversity in *Trypanosoma brucei* infections. *PLoS Negl. Trop. Dis.* **13**, e0007262 (2019).
9. Capewell, P. et al. The skin is a significant but overlooked anatomical reservoir for vector-borne African trypanosomes. *eLife* **5**, e17716 (2016).
10. Camara, M. et al. Extravascular dermal trypanosomes in suspected and confirmed cases of gambiense human African trypanosomiasis. *Clin. Infect. Dis.* **73**, 12–20 (2021).
11. Trindade, S. et al. *Trypanosoma brucei* parasites occupy and functionally adapt to the adipose tissue in mice. *Cell Host Microbe* https://doi.org/10.1016/j.chom.2016.05.002 (2016).
12. Carvalho, T. et al. *Trypanosoma brucei* triggers a marked immune response in male reproductive organs. *PLoS Negl. Trop. Dis.* https://doi.org/10.1371/journal.pntd.0006690 (2018).
13. De Niz, M. et al. Organotypic endothelial adhesion molecules are key for *Trypanosoma brucei* tropism and virulence. *Cell Rep* **36**, 109741 (2021).
14. *Control and Surveillance of Human African Trypanosomiasis: Report of a WHO Expert Committee*. WHO Technical Report Series (WHO, 2013).
15. Crilly, N. P. & Mugnier, M. R. Thinking outside the blood: perspectives on tissue-resident *Trypanosoma brucei*. *PLoS Pathog.* **17**, e1009866 (2021).
16. Kamper, S. M. & Barbet, A. F. Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **53**, 33–44 (1992).
17. Seed, J. R. & Effron, H. G. Simultaneous presence of different antigenic populations of *Trypanosoma brucei* gambiense in *Microtus montanus*. *Parasitology* **66**, 269–278 (1973).
18. Seed, J. R., Edwards, R. & Sechelski, J. The ecology of antigenic variation. *J. Protozool.* **31**, 48–53 (1984).
19. Barry, J. D. & Emery, D. L. Parasite development and host responses during the establishment of *Trypanosoma brucei* infection transmitted by tsetse fly. *Parasitology* **88**, 67–84 (1984).
20. Tanner, M., Jenni, L., Hecker, H. & Brun, R. Characterization of *Trypanosoma brucei* isolated from lymph nodes of rats. *Parasitology* **80**, 383–391 (1980).
21. Vickerman, K. Trypanosome sociology and antigenic variation. *Parasitology* **99**, S37–S47 (1989).
22. Barry, J. D. & Turner, C. M. R. The dynamics of antigenic variation and growth of African trypanosomes. *Parasitol. Today* **7**, 207–211 (1991).
23. Engstler, M. & Boshart, M. Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. *Genes Dev.* **18**, 2798–2811 (2004).
24. Turner, C. M., Hunter, C. A., Barry, J. D. & Vickerman, K. Similarity in variable antigen type composition of *Trypanosoma brucei* Rhodesiense populations in different sites within the mouse host. *Trans. R. Soc. Trop. Med. Hyg.* **80**, 824–830 (1986).
25. Turner, C. M. R. & Barry, J. D. High frequency of antigenic variation in *Trypanosoma brucei* rhodesiense infections. *Parasitology* **99**, 67–75 (1989).
26. Salanti, A. et al. Evidence for the involvement of VAR2CSA in pregnancy-associated malaria. *J. Exp. Med.* **200**, 1197–1203 (2004).
27. Duffy, P. E. & Fried, M. *Plasmodium falciparum* adhesion in the placenta. *Curr. Opin. Microbiol.* **6**, 371–376 (2003).
28. Jonsson, A. -B., Ilver, D., Falk, P., Pepose, J. & Normark, S. Sequence changes in the pilus subunit lead to tropism variation of *Neisseria gonorrhoeae* to human tissue. *Mol. Microbiol.* **13**, 403–416 (1994).
29. Nassif, X. et al. Antigenic variation of pilin regulates adhesion of *Neisseria meningitidis* to human epithelial cells. *Mol. Microbiol.* **8**, 719–725 (1993).
30. Rudel, T., van Putten, J. P. M., Gibbs, C. P., Haas, R. & Meyer, T. F. Interaction of two variable proteins (PilE and PilC) required for pilus-mediated adherence of *Neisseria gonorrhoeae* to human epithelial cells. *Mol. Microbiol.* **6**, 3439–3450 (1992).
31. Virji, M. & Heckels, J. E. The role of common and type-specific pilus antigenic domains in adhesion and virulence of gonococci for human epithelial cells. *J. Gen. Microbiol.* **130**, 1089–1095 (1984).
32. Dean, S., Marchetti, R., Kirk, K. & Matthews, K. R. A surface transporter family conveys the trypanosome differentiation signal. *Nature* **459**, 213–217 (2009).
33. McWilliam, K. R. et al. High-resolution scRNA-seq reveals genomic determinants of antigen expression hierarchy in African Trypanosomes. Preprint at *bioRxiv* https://doi.org/10.1101/2024.03.22.586247 (2024).
34. Smith, J. E. et al. DNA damage drives antigen diversification through mosaic VSG formation in *Trypanosoma brucei*. Preprint at *bioRxiv* https://doi.org/10.1101/2024.03.22.582209 (2024).
35. Calvo-Alvarez, E., Cren-Travaillé, C., Crouzols, A. & Rotureau, B. A new chimeric triple reporter fusion protein as a tool for in vitro and in vivo multimodal imaging to monitor the development of African trypanosomes and Leishmania parasites. *Infect. Genet. Evol.* **63**, 391–403 (2018).
36. Hutchinson, S. et al. The establishment of variant surface glycoprotein monoallelic expression revealed by single-cell RNA-seq of *Trypanosoma brucei* in the tsetse fly salivary glands. *PLoS Pathog.* **17**, e1009904 (2021).
37. Savage, A. F. et al. Transcript expression analysis of putative *Trypanosoma brucei* GPI-anchored surface proteins during development in the tsetse and mammalian hosts. *PLoS Negl. Trop. Dis.* **6**, 1708 (2012).
38. Schopf, L. R., Filutowicz, H., Bi, X. J. & Mansfield, J. M. Interleukin-4-dependent immunoglobulin G1 isotype switch in the presence of a polarized antigen-specific Th1-cell response to the trypanosome variant surface glycoprotein. *Infect. Immun.* **66**, 451 (1998).
39. Liu, G. et al. Distinct contributions of $CD4^+$ and $CD8^+$ T cells to pathogenesis of *trypanosoma brucei* infection in the context of gamma interferon and interleukin-10. *Infect. Immun.* **83**, 2785–2795 (2015).
40. Reinitz, D. M. & Mansfield, J. M. T-cell-independent and T-cell-dependent B-cell responses to exposed variant surface glycoprotein epitopes in trypanosome-infected mice. *Infect. Immun.* **58**, 2337–2342 (1990).

41. Radwanska, M. et al. Comparative analysis of antibody responses against HSP60, invariant surface glycoprotein 70, and variant surface glycoprotein reveals a complex antigen-specific pattern of immunoglobulin isotype switching during infection by *Trypanosoma brucei*. *Infect. Immun.* **68**, 848–860 (2000).
42. Robbiani, D. F. et al. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. *Cell* **135**, 1028–1038 (2008).
43. Hector, R. F., Collins, M. S. & Pennington, J. E. Treatment of experimental *Pseudomonas aeruginosa* pneumonia with a human IgM monoclonal antibody. *J. Infect. Dis.* **160**, 483–489 (1989).
44. Barth, W. F., Wochner, R. D., Waldmann, T. A. & Fahey, J. L. Metabolism of human gamma macroglobulins. *J. Clin. Invest.* **43**, 1036 (1964).
45. Mehlitz, D. & Molyneux, D. H. The elimination of *Trypanosoma brucei* gambiense? Challenges of reservoir hosts and transmission cycles: expect the unexpected. *Parasite Epidemiol. Control* **6**, e00113 (2019).
46. Larcombe, S. D., Briggs, E. M., Savill, N., Szoor, B. & Matthews, K. The developmental hierarchy and scarcity of replicative slender trypanosomes in blood challenges their role in infection maintenance. *Proc. Natl Acad. Sci. USA* **120**, e2306848120 (2023).
47. Morrison, L. J., Majiwa, P., Read, A. F. & Barry, J. D. Probabilistic order in antigenic variation of *Trypanosoma brucei*. *Int. J. Parasitol.* **35**, 961–972 (2005).
48. Pinger, J., Chowdhury, S. & Papavasiliou, F. N. Variant surface glycoprotein density defines an immune evasion threshold for African trypanosomes undergoing antigenic variation. *Nat. Commun.* **8**, 828 (2017).
49. Trindade, S. et al. Slow growing behavior in African trypanosomes during adipose tissue colonization. *Nat. Commun.* **13**, 7548 (2022).
50. Shimogawa, M. M. et al. Parasite motility is critical for virulence of African trypanosomes. *Sci. Rep.* **8**, 9122 (2018).

Methods

Intravenous mouse infections and sample collection

Female C57Bl/6J (wild-type, strain no. 000664 Jackson Laboratory) or B6.129P2-*Aicda*$^{tm1(cre)Mnz}$/J (AID$^{-/-}$, strain no. 007770 Jackson Laboratory)[42] between 7 and 10 weeks old were housed at 68–76 °C (target 72 °C) with 30–70% relative humidity (target 42%) under a 14.5 h:9.5 h light:dark photoperiod. Mice were infected by intravenous tail vein injection with roughly five pleiomorphic EATRO1125 AnTat1.1E 90-13 *T. brucei* parasites[23]. Blood parasitaemia was counted by tail bleed every 2 days starting on day 4 postinfection by haemocytometer with a limit of detection of 2.22×10^5 parasites per millilitre. Blood (25 µl) was collected by a submandibular bleed on days 6, 10 and 14 postinfection and placed into TRIzol LS. For wild-type mice, four mice were anaesthetized and perfused at days 6, 10 and 14 postinfection. Infected AID$^{-/-}$ mice were anaesthetized and perfused at days 6 (two mice) and 13 (three mice) postinfection. Sample sizes were chosen based on preliminary data and for experiment feasibility. For all experiments, mice were assigned to each collection time point randomly from the full set of mice; mice were not housed by sample group. Samples were not blinded at any point. Mice were perfused with 50 ml of PBS glucose (0.055 M D-glucose) with heparin. After perfusion, tissues were dissected and placed immediately into 1 ml of RNA Later. The heart, lungs, gonadal fat, subcutaneous fat, brain and skin (ear) were collected.

For flow cytometry, immunofluorescence experiments and single-cell sorting experiments, 7–10-week-old female C57Bl/6J mice were infected by intravenous tail vein injection with roughly five AnTat1.1E chimeric triple reporter *T. brucei* parasites that express tdTomato[35]. Blood was collected by a submandibular bleed at designated time points. For flow cytometry, mice were anaesthetized and perfused on days 6 and 13 postinfection as discussed above and the gonadal fat and lungs were harvested. For immunofluorescence experiments, perfused tissues were collected at day 13 postinfection. For single-cell sorting, blood and tissues were only collected on day 14 postinfection. All animal studies were approved by the Johns Hopkins Animal Care and Use Committee (protocol no. MO22H163).

VSG-seq sample and library preparation

RNA was isolated from blood samples stored in TRIzol LS (ThermoFisher, catalogue no. 10296010) by phenol–chloroform extraction. Tissue samples were weighed and homogenized in TRIzol, and then RNA was isolated by phenol–chloroform extraction. RNA from each sample was DNase treated using Turbo DNase and cleaned up with Mag-Bind TotalPure NGS beads (Omega Bio-Tek, M1378-00). First-strand complementary DNA (cDNA) synthesis was performed using SuperScript III Reverse Transcriptase and a primer that binds to the conserved VSG 14-mer in the 3′ untranslated region (5′-GTGTTAAAATATATC-3′). Products were cleaned up using Mag-Bind TotalPure NGS beads (Omega Bio-Tek, M1378-01). Next, a VSG-specific PCR with Phusion polymerase (ThermoFisher, F530L) was performed using primers for the spliced-leader (5′-ACAGTTTCTGTACTATATTG-3′) and SP6-VSG 14-mer sequences (5′-GATTTAGGTGACACTATAGTGTTAAAATATATC-3′) for 25 cycles. VSG-PCR products were cleaned up using Mag-Bind TotalPure NGS beads and quantified using the QuBit HS DNA kit (Life Technologies). Finally, sequencing libraries were prepared with the Nextera XT DNA Sample Prep Kit (Illumina) using the manufacturer's guidelines, and libraries were sequenced with 100 bp single-end reads on an Illumina HiSeq 2500.

Tissue-load and *PAD1* qPCRs

First-strand synthesis was performed with SuperScript III Reverse Transcriptase (ThermoFisher Scientific, 18080051) and random hexamer primers on tissue RNA samples. Quantitative PCR (qPCR) was performed in triplicate using SYBR Green qPCR Master Mix (Invitrogen, 4309155). *TbZFP3* primers were used to estimate parasite load in tissue samples (forward 5′-CAGGGGAAACGCAAAACTAA-3′; reverse 5′-TGTCACCCCAACTGCATTCT-3′). Cycle threshold (Ct) values were averaged between the triplicates and parasite loads per mg of tissue were estimated using a standard curve of values from RNA isolated from known numbers of cultured parasites (standard curves can be found in Extended Data Fig. 3d).

For *PAD1* expression quantification, RNA extraction, first-strand synthesis and qPCR were performed following the same methods as above. *PAD1* expression was quantified by normalizing to *TbZFP3* as a control gene (same primers as above) (*PAD1* primers; forward 5′-CAGCGGC GATTATTGCATTGG-3′; reverse 5′-AGGAAGAAGGTTCCTTTGGTC-3′). Ct values were averaged between the triplicates and samples were compared using the delta-CT between *PAD1* and *TbZFP3*.

VSG-seq analysis

Analysis of sequencing results was performed following the method we reported previously[7], with two changes: no mismatches were allowed for bowtie alignments and each sample was analysed (assembly, alignment and quantification) separately. The data were analysed using the VSG-seq pipeline available on GitHub (https://github.com/mugnierlab/VSGSeqPipeline, commit 226a8a1b3cc050391c6e62f9ababc3594177d0dd). The following software and versions were used in the pipeline: Trinity[51,52] (v.2.8.5), Bowtie[53] (v.1.2.3), Biopython[54] (v.1.72), Blast[55,56] (v.2.9), Bedtools[57] (v.2.29.2), cd-hit[58,59] (v.4.8.1), trim-galore[60] (v.0.6.4) and samtools[61] (v.1.9). To compare expressed VSG sets between samples, all assembled VSGs were clustered using CD-HIT-EST[58,59] (v.4.8.1). VSGs with more than 98% identity to one another were conservatively treated as one VSG. VSGs were then identified by their Cluster number for further analysis. Samples that had less than 100,000 successfully aligning reads to VSGs were excluded from further analysis. Four samples, three brain and one heart, were discarded because fewer than 100,000 reads aligned to VSG (Extended Data Fig. 3a). Downstream analysis of expression data and generation of figures was performed in R v.4.3.1.

Analysis of VSG sequence motifs

To identify whether there were tissue-specific VSG sequence motifs, the similarity of N-terminal sequences from all assembled VSGs were compared. N-terminal sequences were identified using a HMMEr scan[62] (v.3.1b2) against a database curated by Cross et al.[2,63]. All N termini were compared in an all versus all BLAST[55,56] (v.2.9) using default parameters. All VSG pairwise comparisons with an e-value higher than 1×10^{-3} were considered sufficiently similar to one another for further analysis. VSGs that were found in a given tissue were binned into that tissue group, and the distribution of the BLAST bitscores in a given compartment was compared against the total population of similar VSGs.

Flow cytometry

Once mice were perfused, tissues were dissected and washed with Hank's balanced salt solution (ThermoFisher Scientific, 14175095). Tissue samples were minced and placed in DMEM (ThermoFisher Scientific, 11995065) containing either 1 mg ml^{-1} collagenase type 1 (ThermoFisher Scientific, 17100017) for adipose fat or 2 mg ml^{-1} collagenase type 2 (ThermoFisher Scientific, 17101015) for lung samples. Hearts were dissociated using 2 mg ml^{-1} collagenase type 2, 50 U ml^{-1} DNase I and 20 U ml^{-1} hyaluronidase. These were then incubated in a 37 °C water bath for 1 h and briefly vortexed every 10 min. Next, samples were passed through a 70 µM filter and centrifuged at 2,600*g* for 8 min at 4 °C, and the cell pellet was taken for antibody staining.

Blood samples were collected by submandibular bleed and red blood cells were depleted by magnetic-activated cell sorting with anti-Ter-119 MicroBeads (Miltenyi Biotech, 130-049-901) following the manufacturer's protocol. Cells were pelleted and washed with HMI-9 media.

All samples, both blood and tissues, were stained with Zombie Aqua dye at 1:100 in PBS and washed with PBS following the manufacturer's protocol (BioLegend, 423101). Samples were then stained for 10 min at 4 °C with a rabbit anti-AnTat1.1 polyclonal antibody diluted 1:15,000 in HMI-9 media and washed once with HMI-9 (antibody courtesy of J. Bangs[64]). Then, secondary antibody staining was performed while shaking for 10 min at 4 °C with Anti-Rabbit IgG (H+L), F(ab')2 Fragment conjugated to Alexa Fluor 488 fluorescent dye (Cell Signaling Technology, 4412S). Finally, samples were washed with cold PBS and resuspended in PBS for flow cytometry analysis. Samples were run on a Beckton Dickenson A3 Symphony flow cytometer and analysis was performed using FlowJo (v.10.6.1) (see Extended Data Fig. 7a for the gating strategy).

Immunofluorescence

Mice infected with AnTat1.1E chimeric triple reporter *T. brucei* parasites that express tdTomato were euthanized and perfused as previously described at days 6 and 13 postinfection. Lung, heart and gonadal fat were collected and fixed in 4% paraformaldehyde in PBS for 12 h at 4 °C. Postfixation, tissues were frozen, embedded in Optimal Cutting Temperature Compound (Tissue-Tek) and cut by cryostat microtome into 10 µm sections.

The following antibodies were applied to sections: rat antimouse CD31 (PECAM-1) (Santa Cruz Biotechnology no. sc-18916, 1:200) with goat antirat Fluor 488 (Cell Signaling Technology, no. 4416, 1:1,000). Coverslips were mounted using ProLong Gold (Life technologies). Tissues were imaged with ×4, ×10 and ×20 objectives using a Nikon Eclipse 90i fluorescence microscope (Nikon) and X-Cite 120 fluorescent lamp (Excelitas) with an ORCA-ER digital CCD camera (Hammamatsu) and ImageJ v.1.53 image analysis software. Image collection and analysis followed published guidelines for rigour and reproducibility[65].

Serum antibody ELISA quantification

Blood (25 µl) was collected by submandibular bleed on days 0, 6, 10 and 14 postinfection from mice infected with roughly five pleiomorphic EATRO1125 AnTat1.1E 90-13 *T. brucei* parasites. Serum was isolated using serum separator tubes (BD Microtainer SST tubes, 365967). IgM and IgG were quantified by enzyme-linked immunosorbent assay (ELISA) using ThermoFisher IgM and IgG kits following manufacturer protocols (IgG catalogue no. 88-50400-88, IgM catalogue no. 88-50470-88).

Serum sample flow cytometry on *T. brucei*

Blood was collected from two mice infected with AnTat1.1E chimeric triple reporter *T. brucei* parasites, which initially express the VSG AnTat1.1, by cheek bleed. Serum was isolated by spinning blood samples at 10,000*g* for 5 min and pipetting off the top serum layer. EATRO1125 AnTat1.1E 90-13 *T. brucei* parasites[23] expressing the VSG AnTat1.1 and Monomorphic Single Marker Lister427 VSG221 TetR T7RNAP bloodstream form (NR42011; Lot 61775530)[66], which express VSG-2, were used for flow cytometry. Next, 10^6 parasites were stained in duplicate while shaking for 10 min at 4 °C using mouse serum diluted 1:100 in PBS. As a positive control, a rabbit anti-AnTat1.1 polyclonal antibody diluted 1:15,000 in PBS was used following the same staining procedure (antibody courtesy of J. Bangs[64]). Then, secondary antibody staining was performed while shaking for 10 min at 4 °C with Antimouse IgG (H+L), F(ab')2 Fragment conjugated to Alexa Fluor 647 fluorescent dye (Cell Signaling Technology, 4410S) or Anti-Rabbit IgG (H+L), F(ab')2 Fragment conjugated to Alexa Fluor 647 fluorescent dye (Cell Signaling Technology, 4414S). Finally, samples were washed with cold PBS and resuspended in PBS for flow cytometry analysis. Samples were run on an Attune Nxt flow cytometer (Invitrogen) and analysis was performed using FlowJo (v.10.6.1).

Single-cell sorting and RNA-seq library preparation

Blood and tissue samples (heart, gonadal fat and lung) from two mice were collected after perfusion as described above. For sorting parasites, the same tissue dissociation and flow cytometry was performed as described above, with the exception of viability staining, which was performed using propidium iodide instead of Zombie Aqua. Samples were kept on ice as much as possible through this process. Single live, tdTomato$^+$ *T. brucei* cells were sorted into chilled 384-well plates for SL-Smart-seq3xpress library preparation containing lysis buffer and an RNA spike-in control using a Beckman Coulter MoFlo XDP cell sorter (see Extended Data Fig. 7b for the gating strategy). For each blood and tissue sample, a single plate of parasites were sorted for a total of 370 cells per sample.

The SL-Smart-seq3xpress library preparation approach was followed as described in ref. 33. In brief, single cells were lysed by incubation at 72 °C for 10 min in 0.3 µl of lysis buffer. Reverse transcription was done by adding 0.1 µl of reverse transcription mix to each well and incubation at 42 °C for 90 min, followed by ten cycles of 50 °C for 2 min and 42 °C for 2 min, with a final incubation at 85 °C for 5 min. Preamplification was done by adding 0.6 µl of a mix containing primers annealing to the Spliced-Leader sequence and to a conserved sequenced added by the oligodT primer during reverse transcription, with the following cycling conditions: 95 °C for 1 min, 16 cycles of: 98 °C for 10 s, 65 °C for 30 s, 68 °C for 4 min and finally 72 °C for 10 min. Following, the amplified cDNA was diluted by adding 9 µl of water per well. Next, 1 µl of each well was transferred to a new plate, 1 µl of tagmentation mix was added and the plate incubated at 55 °C for 10 min for tagmentation. The reaction was stopped by adding 0.5 µl of 0.2% SDS to each well and incubating for 5 min. The final index PCR was done by adding 1 µl of specific index primer combinations to each well and 1.5 µl of PCR mix. The following cycling conditions were used: 72 °C for 3 min, 95 °C for 30 s, 14 cycles of: 95 °C for 10 s, 55 °C for 30 s, 72 °C for 1 min; followed by 72 °C for 5 min. Single-cell libraries were then pooled and purified using AMPure XP beads at a ratio of 1:0.7. Libraries were run on a 4% non-denaturing PAGE gel and purified according to standard polyacrylamide gel purification protocols. Purified libraries from several plates were pooled and sequenced on a NextSeq 1000 sequencing platform to produce paired-end reads of 101 nucleotides (nt) (cDNA) and 19 nt (TAG + UMI read), and 8 nt for the index reads.

Single-cell RNA-seq primary processing and VSG de novo assembly

The primary processing of the sequencing data was as described in ref. 33. In brief, the two reads containing the indexes (8 nt each) and the one containing the TAG + UMI (19 nt) were concatenated into a 35 nt read. Artefact reads containing the TAG sequence (or its reverse complement) in the cDNA reads were filtered out with Cutadapt[67] (v.4.3).

For analysis of derepression by de novo assembly of VSGs, the filtered reads were sorted into individual files for each cell and these read files were run through our VSG-seq analysis pipeline individually using the same parameters as described above for bulk VSG-Seq. Using this approach, VSG open reading frames (ORFs) were assembled and quantified for each cell individually. VSG ORFs were then clustered among all single cells (VSG clusters were not related to previous VSG clusters from bulk VSG-seq and cannot be compared to the previous analysis based on cluster names).

For analysis of derepression by alignment to the genome, filtered reads were mapped with STAR[68] (v.2.7.10a) to a hybrid fasta file combining the *T. brucei* EATRO1125 strain genome assembly (v.67, downloaded from TriTrypDB[69]) and the set of ten sequences used as RNA spike-in. The count matrix obtained was then corrected with the index hopping filtering pipeline scSwitchFilter, v.1.0.0 (https://github.com/colomemaria/scSwitchFilter). Only cells with at least 500 genes detected, 1,000 gene UMI transcript counts, 30 spike-in UMI counts and ten VSG UMI counts were used for downstream analyses. A total of 1,216 total cells fit these criteria out of 2,960 total cells sequenced. For each tissue and blood sample, a single plate (370 cells) was sequenced. For assessment

of potential derepression (Extended Data Fig. 5b,e), VSGs with more than one UMI count were considered expressed in a cell. Cells were considered to be monogenically expressing a VSG (Extended Data Fig. 5e) if the VSG represented more than or equal to 80% of VSG UMI counts. Alignment and assembly data for each cell can be found in Supplementary Data 1–3.

For evaluating read coverage of VSGs and VSG clusters, Bowtie indexes were created for each reference sequence then reads from a single cell were aligned using Bowtie[53] (v.1.2.3). Read coverage was calculated using deepTools[70] (v.3.5.5) to convert BAM alignment files to bigWig coverage tracks. Coverage was then visualized using the ggcoverage[71] package in R (https://github.com/showteeth/ggcoverage, v.1.2.0).

Tsetse fly husbandry and fly bite-inoculated mouse infections

Glossina morsitans morsitans were maintained in the Yale School of Public Health insectary at 25 °C with 65–70% relative humidity under a 12 h:12 h light:dark photoperiod. All flies received defibrinated sheep blood (Lampire Biologicals) every 48 h through an artificial membrane feeding system[72]. Newly eclosed adult female flies were administered per os an initial blood meal containing 1×10^6 per millilitre of *Trypanosoma brucei brucei* (strain RUMP 503; previously expanded in rats) and cysteine (10 μM; to increase the infection prevalence[73]). After this single parasite challenge, flies were maintained on normal blood every other day.

Thirty-five days postchallenge (the time it takes *T. b. brucei* to complete their developmental cycle within the tsetse fly and become infectious to a new vertebrate host), 6–8-week-old female C57Bl/6J mice were exposed to the bite of individual, trypanosome challenged flies 72 h after the flies had taken their last blood meal. Following the consumption of mouse blood, individual flies were microscopically dissected to confirm that their salivary glands were infected with vertebrate-infectious metacyclic stage *T. b. brucei* (if not, another fly was allowed to feed on the mouse until it was confirmed that an infectious fly had taken a blood meal). Five mice were infected using this method. All experiments using mice were performed in strict accordance with the Yale University Institutional Animal Care and Use Committee policies (Protocol 2014–07266 renewed on March 2023).

Once mice were infected, blood parasitaemia was counted by tail bleed every 2 days starting on day 4 postinfection by haemocytometer with a limit of detection of 2.22×10^5 parasites per millilitre. Blood (25 μl) was collected by a submandibular bleed on days 6, 10 and 14 postinfection and placed into TRIzol LS. Five mice were anaesthetized and perfused at day 14 postinfection. Mice were perfused with 50 ml of PBS glucose (0.055 M D-glucose) with heparin. After perfusion, tissues were dissected and placed immediately into 1 ml of RNA Later. The heart, lungs, gonadal fat, subcutaneous fat, brain and skin (ear) were collected. Sequencing libraries were prepared and analysed following the methodology described above.

To quantify the mVSG repertoire of RUMP 503, we also collected a pool of tsetse saliva containing RUMP 503 *T. brucei* parasites. This sample was stored in TRIzol LS, RNA was extracted and the sample was prepared for sequencing as described above. The VSG-seq pipeline was used to quantify mVSG expression in the sample.

Statistics and figures

Normality was tested for all Students *t*-tests and Dunnet's tests analyses and can be found in the code on the repository at https://github.com/mugnierlab/Beaver2022. Nearly all VSG diversity measurements were found to be normally distributed, except for some samples with low VSG counts or sample numbers. We thus assumed normality for all VSG diversity measurements. All reported *P* values have been corrected for multiple comparisons using the Benjamini–Hochberg procedure. For all figures with boxplots, the box represents the first (25%) and third (75%) quartiles with a line at the median. Extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range from either end of the box.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data for generating the analysis and figures in this paper are available at GitHub (https://github.com/mugnierlab/Beaver2022/) and Zenodo (https://doi.org/10.5281/zenodo.13684001)[74]. The EATRO1125 genome used for single-cell RNA-seq analysis is available from TriTrypDB (version 67, available at https://w1.tritrypdb.org/common/downloads/release-67/TbruceiEATRO1125/fasta/data/TriTrypDB-67_TbruceiEATRO1125_Genome.fasta). Raw sequencing data are available in National Center for Biotechnology Information (NCBI) Sequence Read Archive under accession number PRJNA858046. Source data are provided with this paper.

Code availability

All code used to perform the analysis and generate figures in this paper is available at GitHub https://github.com/mugnierlab/Beaver2022 and Zenodo (https://doi.org/10.5281/zenodo.13684001)[75].

51. Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
52. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
53. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
54. Cock, P. J. A. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
55. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinf.* **10**, 421 (2009).
56. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
57. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
58. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
59. Li, W. & Godzik, A. CD-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
60. Krueger, F. et al. FelixKrueger/TrimGalore: v.0.6.4 - add default decompression path. Zenodo https://doi.org/10.5281/zenodo.5127898 (2023).
61. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
62. Eddy, S. R. Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
63. So, J. et al. VSGs expressed during natural *T. b. gambiense* infection exhibit extensive sequence divergence and a subspecies-specific bias towards type B N-terminal domains. *mBio* **13**, e02553-22 (2022).
64. Gruszynski, A. E., DeMaster, A., Hooper, N. M. & Bangs, J. D. Surface coat remodeling during differentiation of *Trypanosoma brucei*. *J. Biol. Chem.* **278**, 24665–24672 (2003).
65. Lee, J.-Y. & Kitaoka, M. A beginner's guide to rigor and reproducibility in fluorescence imaging experiments. *Mol. Biol. Cell* https://doi.org/10.1091/mbc.E17-05-0276 (2018).
66. Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. M. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* **99**, 89–101 (1999).
67. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
68. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
69. Shanmugasundram, A. et al. TriTrypDB: an integrated functional genomics resource for kinetoplastida. *PLoS Negl. Trop. Dis.* **17**, e0011058 (2023).
70. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
71. Song, Y. & Wang, J. ggcoverage: an R package to visualize and annotate genome coverage for various NGS data. *BMC Bioinf.* **24**, 309 (2023).
72. Moloo, S. K. An artificial feeding technique for Glossina. *Parasitology* **63**, 507–512 (1971).
73. MacLeod, E. T., Maudlin, I., Darby, A. C. & Welburn, S. C. Antioxidants promote establishment of trypanosome infections in tsetse. *Parasitology* **134**, 827–831 (2007).
74. Beaver, A. mugnierlab/Beaver2022: Release for publication. Zenodo https://doi.org/10.5281/zenodo.13684001 (2024).
75. Barnett, S. A. The skin and hair of mice living at a low environmental temperature. *Q. J. Exp. Physiol. Cogn. Med. Sci.* **44**, 35–42 (1959).

Acknowledgements We thank J. Bangs for supplying us with the anti-AnTat1.1 antibody and B. Rotureau for providing us with the chimeric triple reporter cell line. We thank T. Nilles and W. Allak at the Becton Dickinson Immunology and Flow Cytometry core at Johns Hopkins Bloomberg School of Public Health for training, support and technical assistance using the BD FACS Symphony A3. We also thank H. Zhang and J. Margolick from the Flow Cytometry Cell Sorting Core Facility at Bloomberg School of Public Health, Johns Hopkins University for doing fluorescence-activated cell sorting. The facility was supported by grant nos. CFAR: 5P30AI094189-04 (Chaisson), 1S10OD016315-01 and 1S10RR13777001. Data analysis was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), which is supported by the National Science Foundation (NSF) grant number OAC 1920103. A.K.B. was supported by grant no. NIH T32AI007417. N.P.C. was supported by grant no. NIH T32 OD011089. F.R.-F. is supported by grant no. NIH NIGMS 1K99GM132557-01 and is an Investigator in the Chan Zuckerberg Biohub. M.R.M., A.K.B., J.E.S., N.P.C., J.M.C.H., B.B. and B.Z. were supported by Office of the Director, NIH (grant no. DP5OD023065). In addition, this study was supported by grant no. NIAID R01AI158805 awarded to S.A., and the German Research Foundation (grant no. SI 1610/2-2) and an European Research Council Consolidator Grant (grant no. SwitchDecoding 101044320) awarded to T.N.S. Z.K. was supported by a MSCA ITN Cell2Cell fellowship.

Author contributions Conceptualization: A.K.B., M.R.M., L.M.F., F.R.-F., T.N.S., R.O.C., S.A. and B.L.W. Methodology: A.K.B., R.O.C., Z.K., B.L.W., E.O.A., G.M.S., N.P.C., G.Y.B., J.M.C.H., J.E.S., B.Z. and B.B. Investigation: A.K.B., R.O.C., Z.K., B.L.W., N.P.C., G.Y.B., J.M.C.H. and J.E.S. Visualization: A.K.B., N.P.C. and J.M.C.H. Funding acquisition: M.R.M., L.M.F., T.N.S. and S.A. Project administration: M.R.M. Supervision: M.R.M. Writing—original draft: A.K.B. and M.R.M. Writing—review and editing: A.K.B., J.E.S., G.Y.B., B.Z., M.R.M., F.R.-F., L.M.F., T.N.S. and R.O.C.

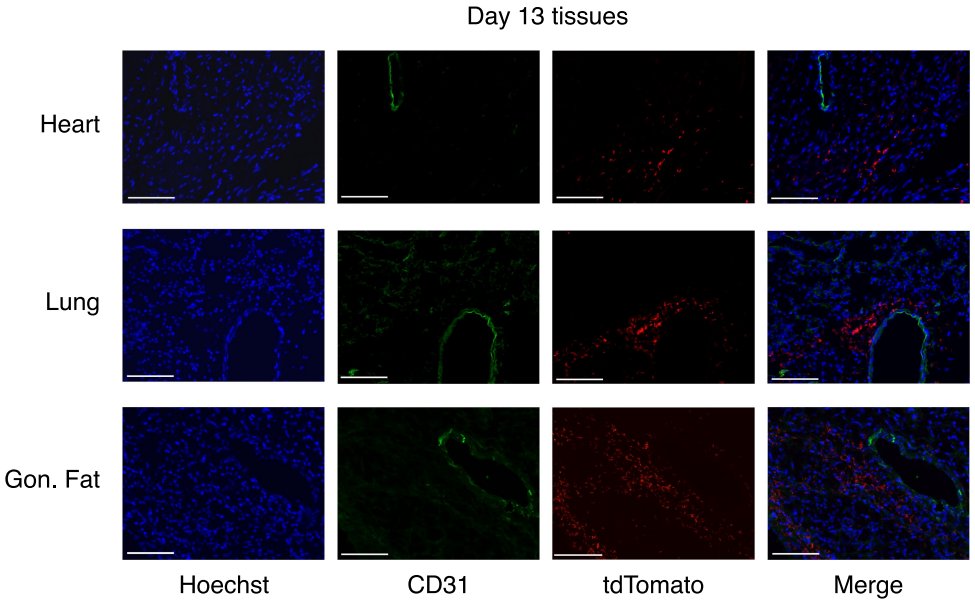
Competing interests The authors declare no competing interests.

Additional information

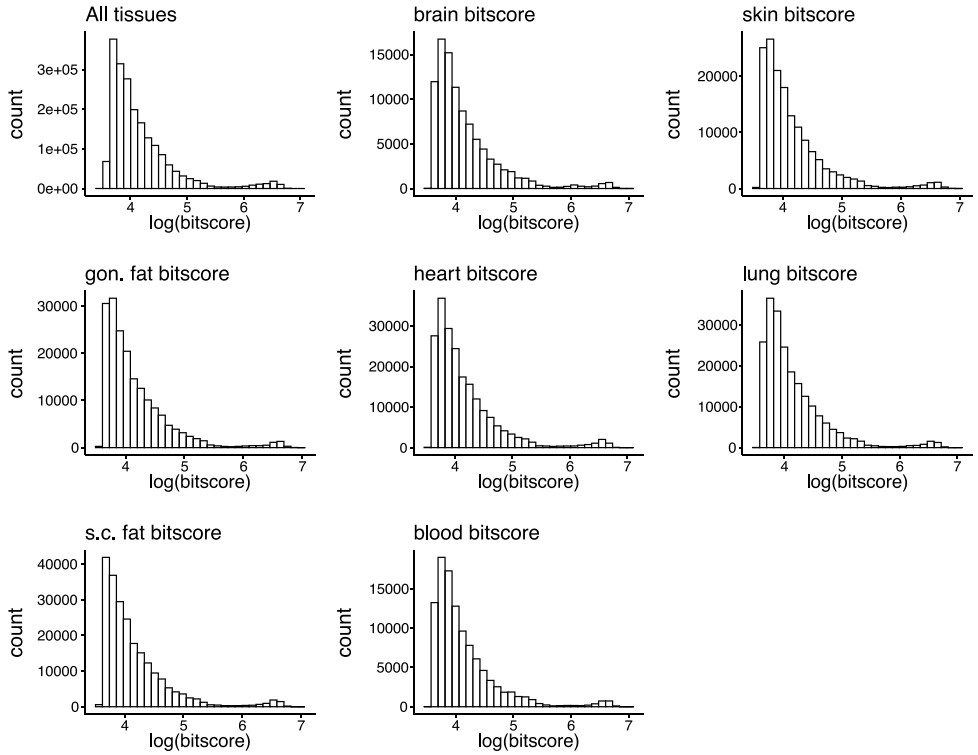
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08151-z.

Correspondence and requests for materials should be addressed to Monica R. Mugnier.

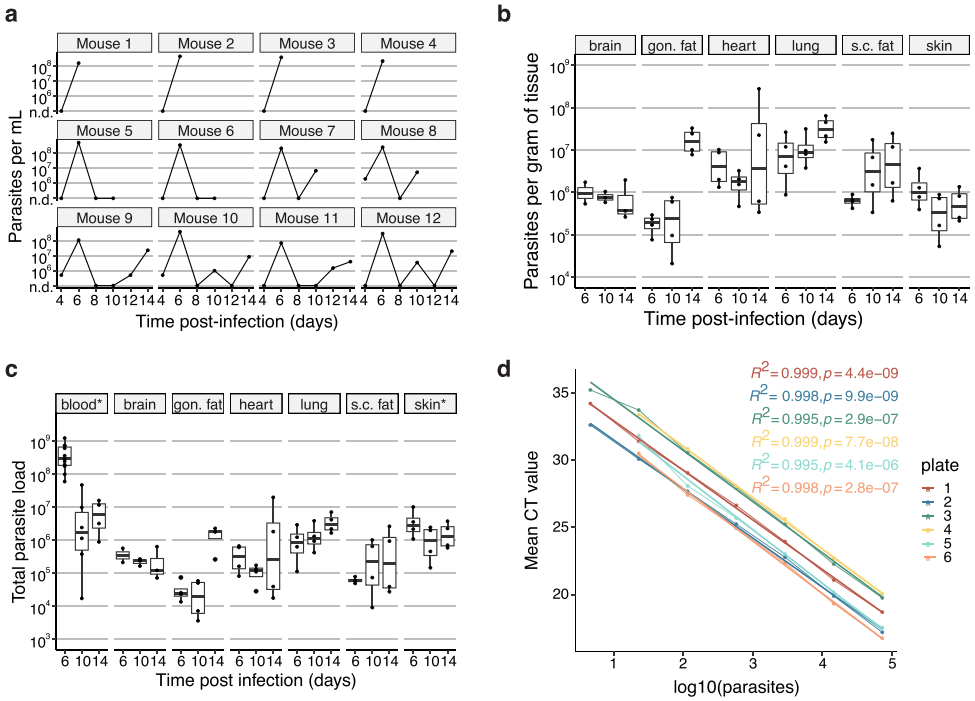
Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.



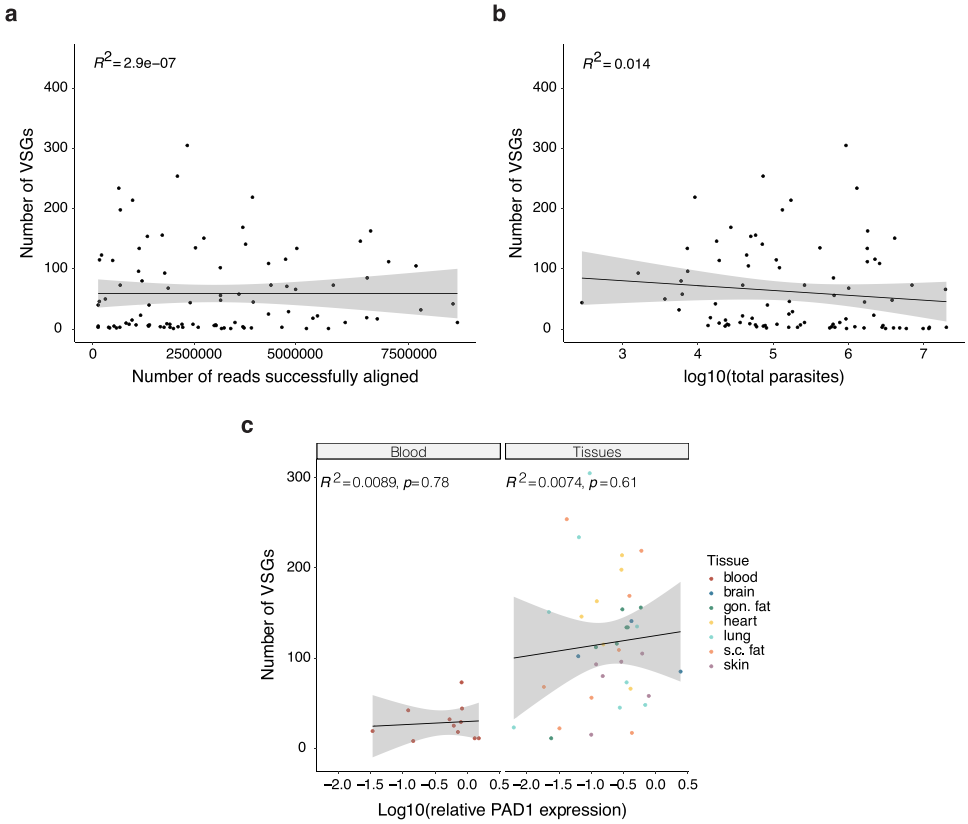
Extended Data Fig. 1 | *T. brucei* parasites are found in interstitial tissue spaces. Representative immunofluorescence images of tdTomato expressing parasites (red) from cross sections of perfused tissues stained with hoechst (blue) and anti-CD31 antibody (green). TdTomato-positive parasites localized separately from CD31 lined spaces, showing that parasites are extravascular. n = 3 biologically independent mice examined over three independent experiments. The scale bars represent 100 microns.



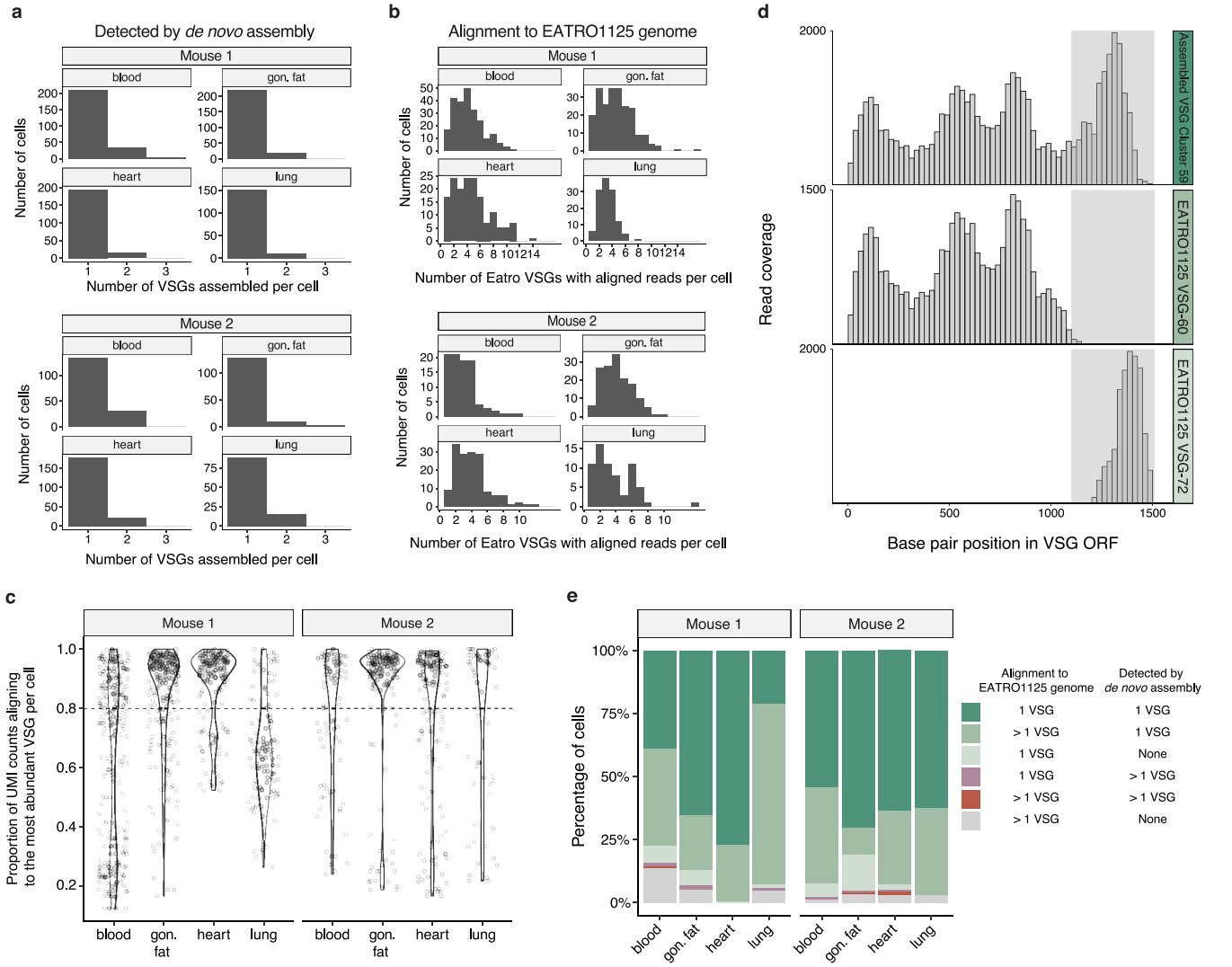
Extended Data Fig. 2 | Comparison of VSG sequences expressed by parasites in the blood and tissues. The similarity of VSGs detected in each tissue measured by bitscore, a sequence similarity metric normalized to the database size, allows for comparing tissue compartments with different numbers of total VSGs expressed in each tissue. No statistical significance was found.



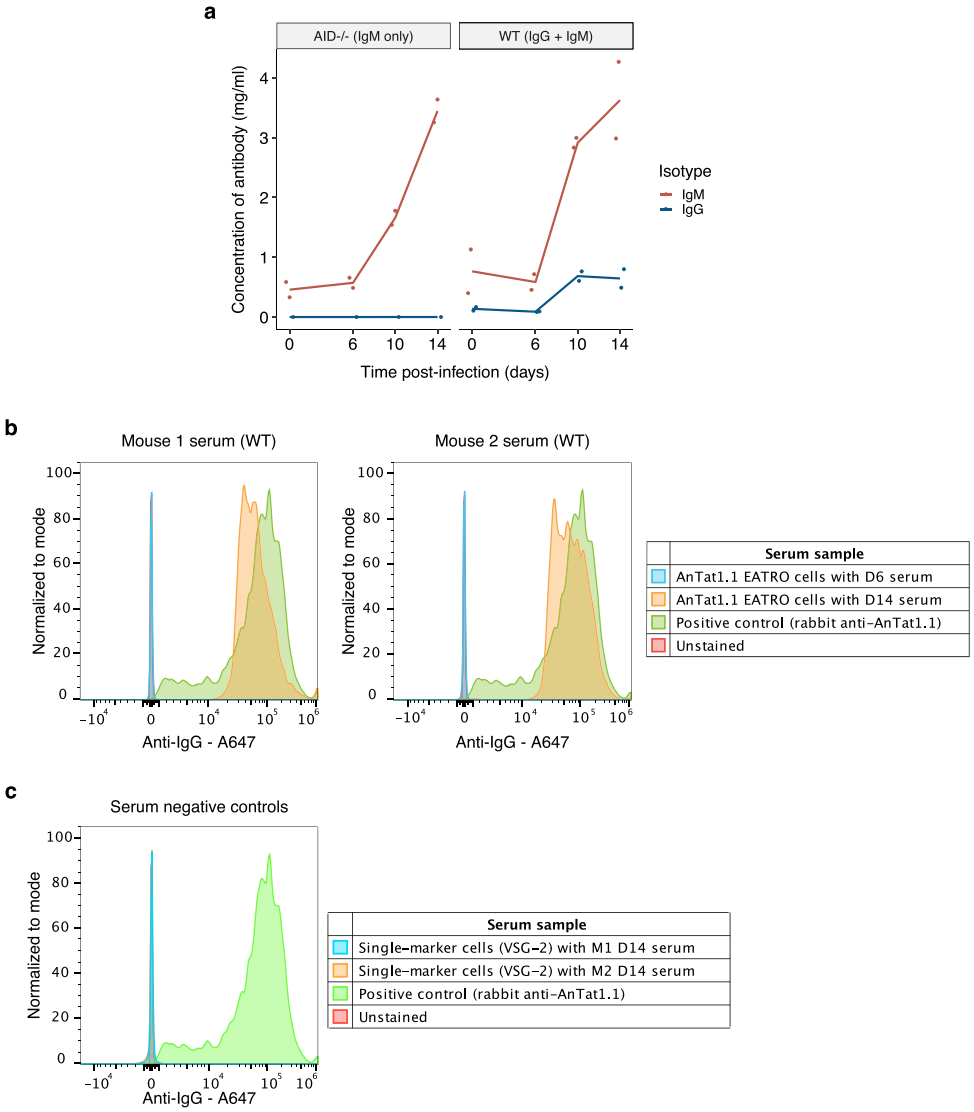
Extended Data Fig. 3 | Parasite load in blood and tissues. (**a**) Parasitemia of 12 mice infected with AnTat1.1E *T. brucei* counted from tail blood by hemocytometer ("n.d." = not detectable, limit of detection of 2.22×10^5 parasites/mL). (**b**) Estimated parasite load per gram of tissue using QPCR. tbZFP3 was used as the control and RNA from known quantities of parasites was used to make standard curves. (**c**) The approximate total number of parasites represented in each organ. This was calculated using the estimated number of parasites from QPCR and the recorded organ mass. For the blood and skin, it was assumed that each mouse had 1.5 mL of blood and 2.73 grams of skin[75] e point over 2 independent experiments. (**d**) The qPCR standard curves used for each plate of samples. These were used to estimate the number of parasites represented in each tissue sample based on RNA from known parasite concentrations (cultured parasites counted using a hemocytometer). A linear regression model was used to examine correlation between standards, no adjustments for multiple comparisons were made. In boxplots, boxes represent values between the first (25%) and third (75%) quartiles with a line at the median, and extending lines represent the maximum and minimum values not including outliers that are further than 1.5 times the interquartile range.



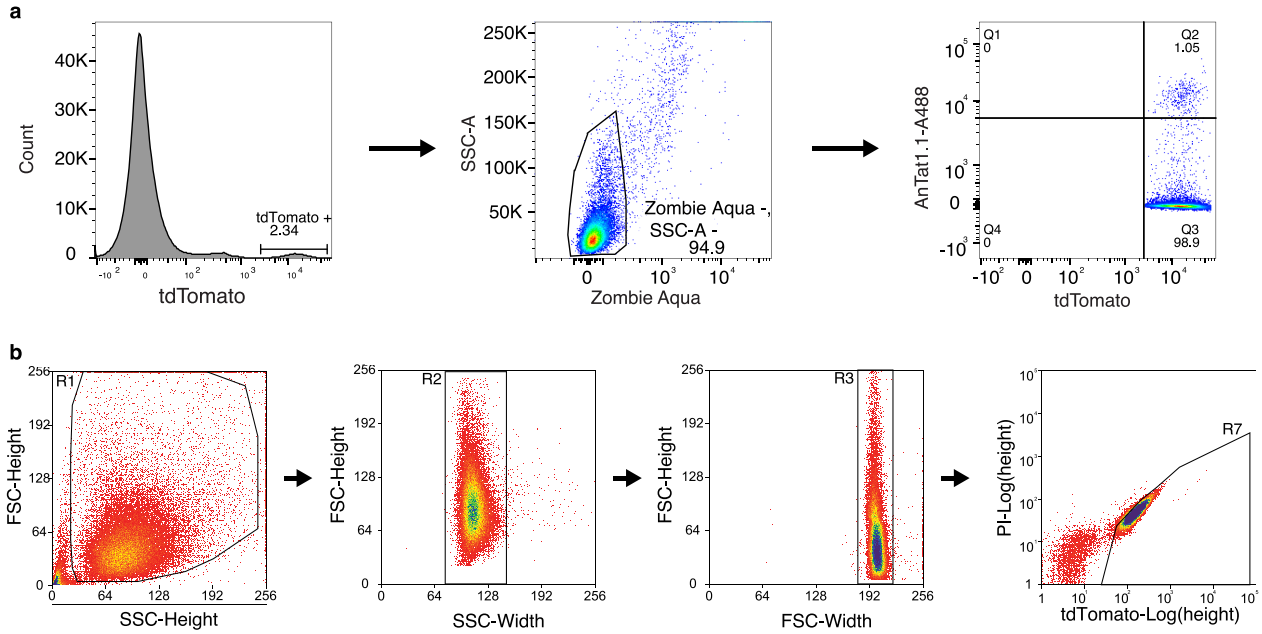
Extended Data Fig. 4 | Correlation of VSG counts with read mapping, parasite load, and PAD1 expression. (**a**) A comparison of the number of reads successfully aligned in a sample and the number of VSGs observed. (**b**) A comparison of the total number of parasites and the number of VSGs found in each sample. (**c**) The correlation between PAD1 expression relative to the housekeeping gene tbZFP3 and the number of VSGs expressed at the population level for each sample. Error bands in the above panels represent the 95% confidence interval for predictions from a linear model. Linear regression models were used to calculate R^2 and P values where noted without adjustments for multiple comparisons.



Extended Data Fig. 5 | Single-cell sequencing analysis of VSG expression. (**a**) The number of VSG open reading frames (ORFs) detected per cell by de novo assembly with the VSG-Seq pipeline. In this figure, all cells sequenced were evaluated for VSG ORF assembly without any filtering (1377 total cells had a de novo assembled VSG). (**b**) The number of EATRO1125 VSGs detected per single cell. Cells were evaluated for VSG expression only if they had at least 500 genes detected, 1000 gene UMI counts, 30 spike-in UMI counts, and 10 VSG UMI counts (1216 cells fit these criteria out of 2960 total sequenced cells). Only VSGs with >1 UMI count were considered for quantification of VSG expression. These filtering parameters were also applied to genomic mapping analyses in panels **c** and **e**. (**c**) Alignment to the EATRO1125 genome assembly was used to quantify the fraction of VSG UMI counts coming from the most abundant VSG gene within each cell. The dashed line represents the 0.8 fraction of total VSG UMI counts (80%) threshold set to define monogenic expression with one dominant VSG in a cell. (**d**) Representative histograms of coverage for reads from one cell. Read coverage is shown for the de novo assembled VSG cluster 59 and two genomic VSG ORFs, VSG-60 and VSG-72, which represent a common scenario that creates ambiguous VSG expression if reads are only aligned to the EATRO1125 genome. (**e**) The VSG expression classification for each cell using both alignment to the EATRO1125 genome and de novo assembly. In alignment and assembly, if a VSG represented 80% or more of the VSG UMI counts in a cell (for genomic mapping) or 80% of the population (for VSG-seq analysis), that cell was considered to be expressing only the dominant VSG.



Extended Data Fig. 6 | General and VSG-specficic antibody levels during infection. (**a**) Quantification of serum IgM and IgG concentrations in infected AID$^{-/-}$ and wildtype (WT) mice by ELISA (n = 2 biologically independent animals per group over one independent experiment). (**b**) AnTat1.1 EATRO cells (expressing VSG AnTat1.1) stained with serum, followed by an anti-IgG secondary antibody that cross-reacts with IgM and other isotypes, from day 6 and day 14 of two mice infected with AnTat1.1 Triple-marker cells. A polyclonal rabbit anti-AnTat1.1 antibody on cells known to be expressing AnTat1.1 was used as a positive control. (**c**) Single-marker cells (expressing VSG-2) stained with day 14 serum from two mice infected with AnTat1.1 Triple-marker cells. These were used as negative controls to show that serum from infected mice specifically binds AnTat1.1 and no other VSGs.



Extended Data Fig. 7 | Flow cytometry and cell sorting gating strategies. (**a**) Example gating strategy for flow cytometry experiments from Fig. 3. Samples were first gated for tdTomato-positive cells, which represent *T. brucei* parasites expressing tdTomato in their cytoplasm. Then live cells were gated based on live/dead Zombie Aqua™ staining. Finally, quadrants were placed around AnTat1.1-A488 positive and negative cells. (**b**) Gating strategy for single-cell sorting into 384-well plates for the SL-Smart-seq3xpress platform. This example is from the blood of mouse 2. Singlet parasites were selected and then tdTomato positive, PI (propidium iodide) negative *T. brucei* cells were sorted into single wells.

Extended Data Table 1 | Summary of single cell sequencing results. *T. brucei* single cells were sorted from blood and tissues on day 14 post-infection from 2 mice and sequenced using the SL-Smart-seq3xpress platform

		Number of single cells									
Mouse	*Tissue*	*Total cells sequenced*	*Cells passing QC cutoffs*	*Average successfully aligning UMI counts per cell*	*Cells with a de novo assembled VSG ORF*	*1 VSG by alignment and 1 VSG by de novo assembly*	*>1 VSG by alignment and 1 VSG by de novo assembly*	*1 VSG by alignment and no VSG by de novo assembly*	*1 VSG by alignment and >1 VSG by de novo assembly*	*>1 VSG by alignment and >1 VSG by de novo assembly*	*>1 VSG by alignment and no VSG by de novo assembly*
Mouse 1	blood	370	238	5736	227	93	91	16	4	1	33
	gon. fat	370	227	4275	228	148	50	13	4	0	12
	heart	370	162	5375	201	125	36	1	0	0	0
	lung	370	123	2952	156	26	88	2	1	0	6
Mouse 2	blood	370	92	3572	148	50	35	5	1	0	1
	gon. fat	370	151	3997	133	106	16	22	1	1	5
	heart	370	156	4699	188	99	46	3	1	2	5
	lung	370	67	5097	96	42	23	0	0	0	2
Totals		2960	1216	4462	1377	689	385	62	12	4	64

For de novo assembly of VSGs, all cells were considered without filters, but for all other columns only cells that had at least 500 genes detected, 1000 gene UMI counts, 30 spike-in UMI counts, and 10 VSG UMI counts were considered. Additionally, only VSGs that had >1 UMI counts when aligning to the EATRO1125 genome were considered to be detected.

Extended Data Table 2 | Summary table of the flow cytometry counts of tdTomato-positive parasites from infections with triple-marker *T. brucei* parasites stained with an anti-AnTat1.1 antibody

Biological experiment	*Mouse*	*Tissue*	*Day*	*Count of AnTat1.1 + cells*	*Count of AnTat1.1 - cells*	*Total number of cells*	*Percent of cells AnTat1.1+*
1	1	blood	6	20394	486	20880	97.67%
1	1	gon. fat	6	411	32	443	92.78%
1	1	lung	6	817	49	866	94.34%
1	2	blood	6	42969	1458	44427	96.72%
1	2	blood	13	1	19922	19923	0.01%
1	2	gon. fat	13	0	250	250	0.00%
1	2	lung	13	573	4460	5033	11.38%
2	3	blood	6	19737	270	20007	98.65%
2	3	gon. fat	6	220	4	224	98.21%
2	3	lung	6	1464	55	1519	96.38%
2	4	blood	6	39373	592	39965	98.52%
2	4	gon. fat	6	379	10	389	97.43%
2	4	lung	6	676	30	706	95.75%
2	5	blood	6	49406	600	50006	98.80%
2	5	blood	13	1	19885	19886	0.005%
2	5	gon. fat	13	31	4982	5013	0.62%
2	5	lung	13	318	29889	30207	1.05%
2	6	blood	6	10317	31	10348	99.70%
2	6	blood	13	0	1001	1001	0.00%
2	6	gon. fat	13	3	781	784	0.38%
2	6	lung	13	265	29733	29998	0.88%
3	7	blood	6	65135	249	65384	99.62%
3	7	gon. fat	6	69	3	72	95.83%
3	7	lung	6	275	182	457	60.18%
3	8	blood	6	49239	632	49871	98.73%
3	8	gon. fat	6	71	0	71	100.00%
3	8	lung	6	770	137	907	84.90%
3	9	blood	6	49758	147	49905	99.71%
3	9	blood	13	3	5042	5045	0.06%
3	9	gon. fat	13	0	1537	1537	0.00%
3	9	lung	13	403	18944	19347	2.08%
3	10	blood	6	49286	642	49928	98.71%
3	10	blood	13	14	20204	20218	0.07%
3	10	gon. fat	13	33	7034	7067	0.47%
3	10	lung	13	187	38096	38283	0.49%

Parasites were collected from the blood and dissociated tissues of 10 mice on either day 6 or 13 post-infection and stained with an anti-AnTat1.1 polyclonal antibody. Here we report the raw number of live tdTomato-positive parasites that were Antat1.1 positive and negative in each sample. These data were used to generate Fig. 3b.

nature portfolio

Corresponding author(s): Monica R Mugnier

Last updated by author(s): Jul 18, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For VSG-Seq, Sequencing libraries were sequenced with 100bp single-end reads on an Illumina HiSeq 2500. For single cell sequencing, samples were sequenced on a NextSeq 1000 sequencing platform to produce paired-end reads of 101nt (cDNA) and 19nt (TAG+UMI read), and 8nt for the index reads.
Data analysis	Sequencing data was analyzed using the VSG-Seq pipeline available on github (github: https://github.com/mugnierlab/VSGSeqPipeline). The following software and versions were used in the pipeline: Trinity (2.8.5), Biopython (1.72), Blast (2.9), Bedtools (2.29.2), cd-hit (4.8.1), trim-galore (0.6.4), bowtie(v1.2.3) and samtools (1.9). Data analysis was done in R (version 4.3.1). All code to generate figures and statistics can be found on github at https://github.com/mugnierlab/Beaver2022 (doi: 10.5281/zenodo.13684001). Analysis to investigate the existence of tissue-specific VSGs used HMMer (v3.1b2) Single cell data were analyzed using STAR (version 2.7.10a), Cutadapt (v4.3), and scSwitchFilter", v1.0.0 (https://github.com/colomemaria/scSwitchFilter. Expression matrices were analyzed in R (version 4.3.1). VSG coverage analysis was performed using Flow cytometry data was analyzed using bowtie (v1.2.3) deepTools (v3.5.5) and ggcoverage(v1.2.0) FlowJo (version 10.6.1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data for generating the analysis and figures in this paper are available at https://github.com/mugnierlab/Beaver2022 (doi:10.5281/zenodo.13684001). The EATRO1125 genome used for scRNA-seq analysis is available from TriTrypDB (version 67, available at https://w1.tritrypdb.org/common/downloads/release-67/TbruceiEATRO1125/fasta/data/TriTrypDB-67_TbruceiEATRO1125_Genome.fasta). Raw sequencing data are available in National Center for Biotechnology Information (NCBI) Sequence Read Archive under accession number PRJNA858046.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were chosen based on preliminary data and for experiment feasibility. We chose four mice per time-point and collected tissues that have been shown to contain extravascular T. brucei. Two mice were infected and used for single-cell sequencing. For flow cytometry experiments, 5 mice were infected. For tsetse-initiated infections 5 mice were used as well. In single cell sequencing experiments, blood and tissue samples from 2 mice were collected and sequenced.
Data exclusions	For VSG-Seq, we excluded four samples (3 brain and 1 heart samples) because fewer than 100,000 reads aligned to VSG. For single-cell sequencing data, only cells which had 500 genes detected, 1000 gene UMI transcript counts, 30 spike-in UMI counts, and 10 VSG UMI counts were used for downstream analyses.
Replication	VSG-seq experiments in witdtype mice were performed using 4 independent biological replicates per time point over two independent experiments. VSG-seq experiments in AID-/- were performed using 2 (day 6) or 3 (day 14) biological replicates in one independent experiment. The overall result in wildtype mice was also replicated by repeating this experiment using a differnet parasite strain and a different infection route. This tsetse replication experiment was performed only once. The scRNA-seq experiment was performed twice, with 2 biological replicates per experiment. The two experiments showed similar results. Flow cytometry experiments were performed using 5 biologically independent replicate mice per timepoint in one experiment. The ELISA experiment and flow cytometric analysis of anti-VSG antibody were performed using 2 biologically independent animals examined over one independent experiment. The immunohistochemistry experiments were performed on 3 biologically independent animals over one experiment and representative images are shown.
Randomization	For all experiments, mice were assigned to each collection timepoint randomly, using a random number generator; mice were not housed by group with the exception for AID-/- and WT groups where each strain was in their own cages. Collected samples were prepared in batches based on tissue type. For VSG-seq experiments, all samples from an individual mouse were sequenced on the same flow cell. Single-cell sequencing plates were prepared in two batches and sequenced between two different runs.

Blinding	Blinding samples was not necessary for this study as the code used to analyze data was the same regardless of sample type. All measures were quantitative such that knowledge of sample groups could not influence the analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	We used a custom rabbit anti-AnTat1.1 polyclonal antibody courtesy of Jay Bangs. In addition, the following commercially available antibodies were used: Zombie Aqua™(BioLegend, 423101), propidium iodide (BD, BDB556463), Anti-mouse IgG (H+L), F(ab')2 Fragment conjugated to Alexa Fluor® 647 fluorescent dye (Cell Signalling Technology, 4410S) or Anti-Rabbit IgG (H+L), F(ab')2 Fragment conjugated to Alexa Fluor® 647 fluorescent dye (Cell Signaling Technology, 4414S) , Alexa Fluor® 488 fluorescent dye (Cell Signaling Technology, 4412S), thermoFisher IgM and IgG kits following manufacturer protocols (IgG cat# 88-50400-88, IgM cat# 88-50470-88), and rat anti-mouse CD31 (PECAM-1) (Santa cruz biotechnology, cat# sc-18916) with goat anti-rat Fluor 488 (Cell signaling technology cat #4416).
Validation	We validated our anti-AnTat1.1 polyclonal antibody by flow cytometry on in vitro T. brucei cells expressing AnTat1.1 and T. brucei cells not expressing AnTat1.1 as a negative control. Live dead Zombie Aqua and propidium iodide staining was validated using live T. brucie cells and heat killed T. brucei cells.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	*EATRO 1125 AnTat1.1E 90-13 T. brucei (gifted from Keith Mathews, Ref: Engstler, M. & Boshart, M. Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in Trypanosoma brucei. Genes Dev. 18, 2798 (2004)). *AnTat1.1E chimeric triple reporter T. brucei (from Brice Rotureau; Ref: Calvo-Alvarez, E., Cren-Travaillé, C., Crouzols, A. & Rotureau, B. A new chimeric triple reporter fusion protein as a tool for in vitro and in vivo multimodal imaging to monitor the development of African trypanosomes and Leishmania parasites. Infect. Genet. Evol. 63, 391–403 (2018)) . *Monomorphic Single Marker Lister427 VSG221 TetR T7RNAP bloodstream form (NR42011; LOT: 61775530) Wirtz, E., Leal, S., Ochatt, C. & Cross, G. A. M. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Mol. Biochem. Parasitol. 99, 89–101 (1999). *Trypanosoma brucei brucei strain RUMP 503 supplied by Serap Aksoy.
Authentication	EATRO 1125 AnTat1.1E 90-13 and Lister427 T. brucei were authenticated by PCR of drug resistence markers and growth in drug selection. AnTat1.1E chimeric triple reporter T. brucei cells were validated by the growth in drug selection and by flow cytometry showing AnTat1.1 staining and tdTomato fluorescence.
Mycoplasma contamination	Cell lines were not tested for Mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	C57Bl/6J (WT, strain# 000664 Jackson Laboratory) or B6.129P2-Aicdatm1(cre)Mnz/J (AID-/-, strain# 007770 Jackson Laboratory) between 7-10 weeks old were used. Mice were infected by intravenous injection in the tail-vein or by tsetse-bite from flies with a confirmed salivary gland infection. 30-35 day old adult Glossina morsitans morsitans that had been fed a blood meal containing

	infectious bloodstream form parasites at 2 days old were used to initiate infections.
Wild animals	No wild animals were used in this study.
Reporting on sex	Only female mice were infected and analyzed. This is standard in the field and some of our preliminary data suggested that sex does not have a major influence on antigenic variation dynamics in vivo. Financial constraints also limited our ability to evaluate sex-specific effects. Female flies were used for tsetse bite infections. This is because female flies live longer; it takes at least 30 days before transmissible salivary gland infections can be presumed, and most male flies do not survive this long.
Field-collected samples	No field-collected sample were used in this study.
Ethics oversight	All animal studies were approved by the Johns Hopkins Animal Care and Use Committee (protocol # MO22H163) and by Yale University Institutional Animal Care and Use Committee policies (Protocol 2014–07266 renewed on March 2023).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Tissues were dissected and washed with HBSS (Hanks balanced salt solution, ThermoFisher Scientific 14175095). Tissue samples were minced and placed in DMEM (ThermoFisher Scientific, 11995065) containing either 1 mg/mL collagenase type 1 (ThermoFisher Scientific, 17100017) for adipose fat or 2 mg/mL collagenase type 2 (ThermoFisher Scientific, 17101015) for lung samples. Hearts were dissociated using 2 mg/mL collagenase type 2, 50U/mL DNase I, and 20U/mL Hyaluronidase. These were then incubated in a 37°C water bath for 1 hour and briefly vortexed every 10 minutes. Next, samples were passed through a 70µM filter and centrifuged at 2600 x g for 8 mins at 4 C, and the cell pellet was taken for antibody staining. Blood samples were collected by submandibular bleed and red blood cells were depleted by magnetic-activated cell sorting (MACS) with anti-Ter-119 MicroBeads (Miltenyi Biotech, 130-049-901) following the manufacturer's protocol. Cells were pelleted and washed with HMI-9 media. All samples, both blood and tissues, were stained with Zombie Aqua™ dye at 1:100 in PBS or with propidium iodide and washed with PBS following the manufacturer's protocol (BioLegend, 423101). Samples were then stained for 10 minutes at 4°C with a rabbit anti-AnTat1.1 polyclonal antibody diluted 1:15,000 in HMI-9 media and washed once with HMI-9 (antibody courtesy of Jay Bangs). Then, secondary antibody staining was performed while shaking for 10 minutes at 4°C. Finally, samples were washed with cold PBS and resuspended in PBS for flow cytometry analysis.
Instrument	Becton Dickenson A3 Symphony flow cytometer or for sorting, a Beckman Coulter MoFlo™ XDP cell sorter.
Software	FlowJo (version 10.6.1)
Cell population abundance	For single-cell sequencing, single cells were sorted into individual wells of a 384-well plate. Sample purity was not assessed because a T. brucei-specific library prep was performed, such that only T. brucei RNA would be sequenced in each library regardless of contamination.

Gating strategy

For flow cytometry experiments, we first gated on tdTomato positive cells, which represent T. brucei cells. We then gated on live cells based on live/dead controls. Finally, we gated AnTat1.1-A488 positive and negative cells. For single-cell sorting, Singlet parasites were selected and then tdTomato positive, PI (propidium iodide) negative T. brucei cells were sorted into single wells of a 384-well plate containing containing lysis buffer and an RNA spike-in control.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.