Trends in Parasitology

CellPress OPEN ACCESS

Review

Evolution of the variant surface glycoprotein family in African trypanosomes

Sara Silva Pereira [1], Andrew P. Jackson [2], and Luísa M. Figueiredo [1,*]

An intriguing and remarkable feature of African trypanosomes is their antigenic variation system, mediated by the variant surface glycoprotein (VSG) family and fundamental to both immune evasion and disease epidemiology within host populations. Recent studies have revealed that the VSG repertoire has a complex evolutionary history. Sequence diversity, genomic organization, and expression patterns are species-specific, which may explain other variations in parasite virulence and disease pathology. Evidence also shows that we may be underestimating the extent to what VSGs are repurposed beyond their roles as variant antigens, establishing a need to examine VSG functionality more deeply. Here, we review sequence variation within the VSG gene family, and highlight the many opportunities to explore their likely diverse contributions to parasite survival.

Highlights

Research into variant antigen diversity has grown in recent years due to advances in sequencing and computational methods. We now have tools to assess the diversity and expression of variant surface glycoproteins (VSGs) systematically, and in depth, in multiple *Trypanosoma* species and in many strains.

Despite a common phenotype of antigenic variation, African trypanosome species have adopted distinct approaches to the generation of antigenic novelty, which affects the mechanism of antigenic variation per se and infection-persistence strategies.

Evidence for functional differentiation amongst the VSG repertoire points towards VSGs playing important roles in the host–trypanosome interaction beyond antigenic variation. Future research will dissect alternative functions of VSGs.

VSGs are variant antigens key to the survival of African trypanosomes

VSGs are major surface antigens of African trypanosomes, extracellular parasites of humans and animals. African trypanosomes alternate between a vector (typically a tsetse fly within Africa, but other blood-sucking insects beyond Africa) and a mammalian host (Box 1). African trypanosomes survive in the mammal using antigenic variation as a mechanism of immune evasion. VSGs are the variant antigens recognized by the host immune system during infection, and their modulation by the parasite can prevent long-lasting immunity (Box 2, and reviewed in [1]). African trypanosome genomes are furnished with several thousand VSG genes, perhaps 10–20% of total gene number, to provide the raw material for antigenic variation. But how has this pool of structural diversity evolved, and why are so many VSG genes required to evade immunity?

African trypanosomes are kinetoplastid parasites of the family Trypanosomatidae, which also includes the vector-borne parasites *Leishmania* spp. and *Trypanosoma cruzi* (Figure 1A). Although the genomes of these trypanosomatid lineages are quite similar, in terms of both **orthology** (see Glossary) and **synteny** [2], the molecular architecture of their cell surfaces, and the genes encoding cell-surface proteins, are mutually exclusive [3]. The cell surfaces of both Leishmaniinae and Stercorarian trypanosomes include several antigens encoded by multicopy gene families; by extension, we assume that the ancestral trypanosomatid also possessed such a heterogeneous cell surface. In contrast, African trypanosome cell surfaces are dominated by one family (VSG) and they have evolved a system of **monoallelic expression** to regulate VSG expression (reviewed in [4]). The transition from an ancestrally heterogeneous surface architecture to a cell coated with a single, highly abundant VSG defines the evolution of African trypanosomes. Increasingly, VSG genes are also implicated in essential functions besides antigenic variation. Here we discuss how VSGs have evolved, their sequence diversity within three African species of *Trypanosoma* (*T. brucei*, *T. congolense*, and *T. vivax*), and what this diversity could mean for protein function.

Throughout this review, we refer to the VSG gene family, which includes any gene sequence homologous to canonical *T. brucei* VSGs (i.e., established genes that encode major surface

[1]Instituto de Medicina Molecular – João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
[2]Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool L3 5RF, UK

*Correspondence:
lmf@medicina.ulisboa.pt (L.M. Figueiredo).

Box 1. The life cycle of African trypanosomes

Blood-sucking flies become vectors of African trypanosomiasis when they feed on an infected mammal. Biological transmission occurs only through tsetse flies and thus is restricted to their endemic region, sub-Saharan Africa. In the midgut of the tsetse fly, recently ingested bloodstream-form parasites differentiate into procyclic forms. Upon clonal expansion, these forms will travel anteriorly through the proventriculus towards the salivary glands (*T. brucei*) or mouthparts (*T. congolense* and *T. vivax*) of the fly, where they differentiate first into epimastigotes and then into metacyclic parasites. The latter are infective to mammals, so in the following bloodmeal they are injected into the skin of the animal with the fly's saliva. These parasites find their way to the bloodstream of the host and differentiate to bloodstream forms. Depending on the African trypanosome species, they can either remain in circulation (*T. congolense* and possibly *T. vivax*) or invade neighboring tissues (*T. brucei*). Exceptions to this cycle occur in *T. vivax, T. b. evansi*, and *T. b. equiperdum*. *T. vivax* does not have a procyclic form, differentiating directly from the bloodstream form to epimastigote in the fly's mouthparts [79]; *T. b. evansi* has lost its ability to differentiate into procyclic forms so it remains exclusively in its bloodstream form and is mechanically transmitted; and *T. b. equiperdum* has evolved to become sexually transmitted in horses [80]. VSGs are expressed by metacyclic and bloodstream parasites of all African trypanosome species.

proteins of variant antigens) [5] (Figure 1B). Currently, all VSGs can be labeled by their subfamily (i.e., a-VSG, b-VSG, Fam1, Fam9, Fam13–16, Fam23–26) [6]. Each subfamily can be further subdivided into **phylotypes** (Figure 1B)**,** which include multiple, highly related VSGs that form robust clades in phylogenies and are present across all parasite strains (Figure 1C). Inclusion within the family is based on amino acid sequence similarity (i.e., protein primary structure). Therefore, VSGs include not only the originally established variant antigens but also genes that have acquired new functions and became nonvariant. The established variant antigens undergo antigenic variation and include most members of a-VSG, b-VSG, Fam13, Fam16, Fam23–25 (Figure 1C). Nonvariant VSGs have unsystematic names, sometimes associated with their predicted function or genomic locus. These include both single-copy [e.g., serum-resistance-associated (SRA), *T. b. gambiense*-specific glycoprotein (TgsGP), VSG^{sur}], and multicopy genes [e.g., VSG-related genes (VR), ESAG2, transferrin receptors (TfRs), procyclic-associated genes (PAGs), etc.] (Figure 1B). With the exception of *T. vivax*, the phylogenies of each subfamily suggest that the VSG repertoire is constantly producing novel, single-copy variant antigens (e.g., SRA), some of which may be repurposed and become nonvariant. As time progresses, these nonvariant VSGs may undergo gene duplication, giving rise to multicopy subfamilies (e.g., TfRs).

Interspecific variation in the genomic organization of VSGs

Surface proteins are key players in host–pathogen interactions and so are subject to strong selection pressures that result in rapid evolutionary change through mutation, recombination, and gene duplication. This translates into high gene copy-number, sequence diversity, and

Box 2. Trypanosome antigenic variation

Studies in *T. brucei* have shown that, in the bloodstream life stage, approximately ten million copies of VSG protein are produced to cover the entire surface of the parasite as a dense coat. The conformational epitopes of the NTD are exposed to the immune system, which recognizes them and elicits an antibody response [81]. At a single parasite growth peak we can detect a dominant parasite subpopulation expressing a superabundant VSG and many other subpopulations expressing alternative VSG variants, represented at much lower levels (less than 1%) [11,12]. Whilst parasites expressing the dominant VSG are cleared, the minority of cells expressing an alternative VSG can proliferate until a second antibody response is triggered. Theoretically, the cycle can continue until either the host immune system is exhausted or the limit for the parasite antigenic diversity is reached. Parasites can also spontaneously and autonomously switch the expressed VSG [82]. The temporal order of VSG switching is not arbitrary: genes harboring the telomeric expression sites are activated first, followed by those in the minichromosomes, then the VSGs in the subtelomeric tandem arrays, and finally the mosaics generated *de novo* during the course of infection [83–85]. Furthermore, the identity of switched VSGs is semipredictable [86,87], and when infections start with the same nonclonal inoculum, the pattern of VSG expression over time is reproducible, including the identity of the dominant VSGs [11,13,85,88]. This is also observed in *T. vivax* [14], but not in other antigenically variable pathogens, like the parasite *Plasmodium falciparum* [89], or the blood-borne bacterium *Borrelia* [90]. To date, *T. congolense* VSG expression patterns during mammalian host infections have not been described, but research is underway in several laboratories and should be elucidated in the near future.

Glossary

Coalescence time: the time to the most recent common ancestor for a defined set of descendants.
Flagellar pocket: an invagination of the plasma membrane that marks the beginning of the flagellum. It is involved in many cellular processes, including exocytosis, endocytosis, and cell division.
Gene conversion: the replacement of a DNA sequence (or part of it) by another one by homologous recombination.
Monoallelic expression: the mechanism of active transcription of a single allele of a gene. In trypanosomes, it happens by allelic exclusion.
Mosaicism: the combination of multiple ancestral sequence motifs within a novel coding sequence due to segmental gene conversion of multiple other genes or pseudogenes.
Negative selection: a selective process that removes deleterious mutations; negative selection against nonsynonymous mutations conserves the amino acid sequence and is indicated by a ratio of nonsynonymous to synonymous nucleotide substitutions less than one.
Neofunctionalization: an adaptive mutation process that results in a gene duplicate (paralog) acquiring a function different from that of its ancestral gene.
ω (dN/dS): the ratio of substitution rates at nonsynonymous and synonymous sites.
Orthology: a type of homology that occurs due to speciation, and not gene duplication.
Phylotype: a cladistic subdivision of VSG diversity intermediate between subfamily and a cluster of orthologous genes, commonly comprising several, closely related VSG genes and their in-paralogs.
Positive selection: a selective process that promotes advantageous mutations and leads to the emergence of new sequence variants with greater fitness; it is indicated by a ratio of nonsynonymous to synonymous nucleotide substitutions greater than one.
Subtelomeres: highly dynamic chromosomal regions immediately adjacent to the telomeres, often hemizygous.
Synteny: two genetic loci sharing the same chromosomal location (usually in different organisms).
Trypanolytic factor: a complex, present in the serum of humans and

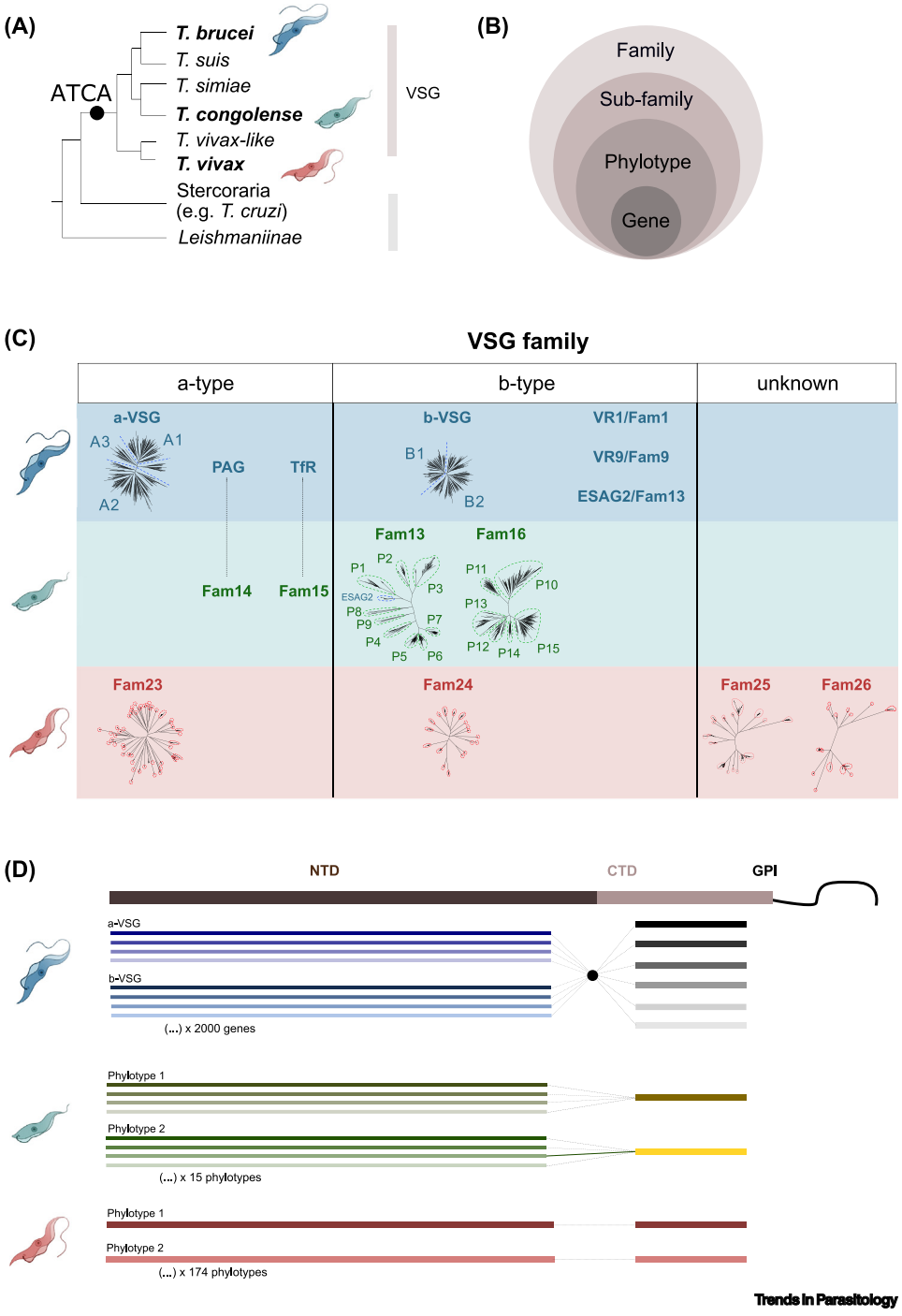

Figure 1. The variant surface glycoprotein (VSG) repertoires of African trypanosomes. (A) Schematic phylogeny based on small-subunit ribosomal RNA genes [97], showing African trypanosomes in the context of other pathogenic trypanosomatids, and the presence/absence of a VSG coat. Species in bold are discussed in this review. ATCA, African trypanosome common ancestor. (B) Current nomenclature used to describe VSGs. Sequences in African trypanosome species homologous to canonical *T. brucei* VSGs comprise the VSG family. Within each species we can identify subfamilies that can be split into phylotypes, which are robust clades with common C-terminal domain (CTD) sequences, each of which

(Figure legend continued at the bottom of the next page.)

some nonhuman primates, that lyses trypanosomes. In its smaller form (TLF-1), it is a high-density lipoprotein, apolipoprotein L1, and hemoglobin-binding haptoglobin-related protein, whilst its larger form (TLF-2) also includes an associated IgM molecule.

VSG expression sites: specialized structures near the telomeres of chromosomes from where VSGs and VSG-associated genes are expressed by RNA polymerase I.

unstable genomic loci [7]. Consequently, studying the VSG repertoire is challenging, and VSGs are often overlooked especially in large-scale analyses. However, in the last decade, annotation of VSG repertoires in reference genome sequences [8–10], the development of sensitive methods of expression analysis (like VSG- or trypanosome-specific cDNA amplification prior to sequencing [11–14]), and of VSG profiling methods from clusters of orthologs or protein motifs [15], have produced a systematic understanding of VSG diversity and its dynamics.

The canonical *T. brucei* VSG consists of an N-terminal domain (NTD) of 300–400 amino acids and a smaller C-terminal domain (CTD) of <100 amino acids [16], glycosylphosphatidylinositol (GPI)-anchored to the plasma membrane [17] (Figure 1D). The NTD contains antigenic epitopes and is highly variable [18]. Homology among NTDs largely relies on conservation of cysteine, glycine, and tryptophan residues [19,20], patterns of which separate *T. brucei* VSGs into two subfamilies: a- and b-VSGs [8,19,20] (Figure 1C). Phylogenetic analyses of VSG in the *T. brucei* 927 reference genome further separated NTDs into five subgroups (N1–N5) [8,19,21]. N4 contains all b-VSGs, while N1–3 and N5 (formerly c-type VSG) encompass a-VSGs [8,21]. Cross *et al*. [10] analyzed the VSG repertoire of another model strain, *T. brucei* 427, and chose to include N5 within N3, but otherwise proposed very similar systematics, comprising three a-type clades (A1–A3, corresponding to N1–N3) and two b-type clades (B1, B2, corresponding to N4). The CTD, which is inaccessible to antibodies [22], is far less variable and has been subdivided into six subgroups based on nucleotide sequences (C1–C6) [8,19,21,23]. A single NTD can form a functional VSG with CTDs of different types [23]. Additionally, the recent resolution of the crystal structure of the MiTat1.3/VSG3 NTD has shown a considerable, and perhaps unexpected, level of topological diversity between a- and b-VSGs [24], even though they still share a common CTD.

Besides the variant antigen subfamilies, there are others homologous to a-VSGs (i.e., PAG and TfR genes, encoded by ESAG6 and ESGA7) and b-VSGs (i.e., ESAG2, VR genes) (Figure 1C). VR genes include both ancient and more recent lineages that have ceased to be variant antigens; some appear to be derivations of the canonical *T. brucei* b-VSG (with familiar CTD but lacking a complete NTD, e.g., Tb927.2.2060), others lack the hypervariable region of the NTD and the characteristic CTD, closely resembling *T. congolense* VSGs, (e.g., Tb927.3.2540), while yet others lack any CTD at all (e.g., Tb927.3.1500).

The VSGs of *T. brucei* subspecies (i.e., *T. b. evansi*, *T. b. gambiense*, *T. b. rhodesiense*) share the genetic diversity of *T. b. brucei* VSGs, described above [25–27]. Orthology is maintained between strains, and strain-specific clades are small, indicative of a common origin and similar evolutionary forces affecting the repertoires. *T. brucei* VSGs are found in three main loci: the small chromosomes, the **subtelomeres** of megabase chromosomes, and the telomeric **VSG expression sites** [28–30].

T. congolense and *T. vivax* have VSG repertoires of comparable size to that of *T. brucei*, and mostly homologous to either a- or b-VSGs. However, sequence comparison shows that they

may contain several genes. (C) The *Trypanosoma brucei* VSG repertoire contains two subfamilies that contain both variant and nonvariant antigens (a-VSG and b-VSG) and five subfamilies of exclusively nonvariant antigens (PAG, TfR, VR1/Fam1, VR9/Fam9, ESAG2/Fam13). The subfamilies a-VSG and b-VSG contain 1500–2500 genes and have been classified based on N-terminal (NTD) and CTD types [only NTD types are labeled (A1–3 and B1,2)] [10]. In *Trypanosoma congolense*, we have identified two subfamilies (Fam13 and Fam16) of b-type VSGs that divide into 15 phylotypes (P1–15) and ~1500 genes [9,31]. *Trypanosoma vivax* VSGs divide into four subfamilies (Fam23–26, inclusive) and 174 phylotypes [9,15]. For each species, the current systematic subdivisions are annotated onto maximum likelihood phylogenies of major subfamilies. Phylotypes are encircled by broken lines. (D) On top, the structure of a VSG, comprising an NTD, a CTD, and a glycosylphosphatidylinositol (GPI) anchor. Below, the constraints (or lack thereof) for VSG recombination within each species. Broken lines represent potential for recombination between NTD and CTD. Abbreviations: PAGs, procyclin-associated genes; TfR, transferrin receptor; VR-, VSG-related.

are mutually exclusive in content [9]. There are no orthologous VSG genes shared by multiple trypanosome species; each repertoire has diverged to become species-specific in its phylogenetic composition.

Thus far, comparative analyses have shown that *T. congolense* variant antigens are composed exclusively of b-type VSGs, albeit with older and more diverse evolutionary origins than *T. brucei* b-VSG [9]. The repertoire divides into two subfamilies, Fam13 (which also includes *T. brucei* ESAG2-like genes) and Fam16 (Figure 1C). Among these two distinct ancestral lineages, 15 conserved phylotypes were identified. Each phylotype has a distinct CTD sequence lacking homology with other phylotypes [31], in contrast to the homologous CTD shared by all canonical *T. brucei* VSGs. Recombination between *T. congolense* VSG genes with distinct CTDs is rare [14], suggesting that CTD heterogeneity constrains the generation of antigenic diversity (Figure 1D). Although it lacks a-type variant antigens, the *T. congolense* genome does contain two a-type VSG subfamilies (Fam14 and Fam15) that are not used in antigenic variation but are instead homologous to *T. brucei* PAG and TfR genes [6,9] (Figure 1C). Like *T. brucei*, *T. congolense* VSG genes are also spread throughout the subtelomeres, small chromosomes, and telomere-proximal sites [32].

T. vivax VSGs are divided into four subfamilies (Fam23–26 [6]) (Figure 1C) and 174 phylotypes [14] (Figure 1D). Fam23 and Fam24 are homologous to a- and b-type VSGs, respectively, but neither subfamily has a conserved CTD as in *T. brucei* [9] (Figure 1D). Fam25 and Fam26 are *T. vivax*-specific and share only ~20% amino acid identity to other VSG subfamilies. Fam25 contains genes that have been detected as the dominant VSG transcripts during bloodstream infections, and therefore is assumed to encode functional variant antigens [14,33], but the same does not apply to Fam26, so their function remains ambiguous (Figure 1C). Sequence divergence between phylotypes is remarkably high, (only 30.1% amino acid sequence identity [9]), but variation within phylotypes is low (73.8% amino acid conservation across the species on average [14]), such that genes within the same phylotype may not encode serologically distinct antigens. This indicates that the *T. vivax* VSG repertoire is smaller than those of *T. brucei* and *T. congolense* [14] and might perhaps encode fewer immunologically distinct antigens. *T. vivax* VSGs seem to be exclusively located in the subtelomeres and but not immediately adjacent to the telomeric repeats, more data is required to confirm this.

Overall, despite their apparently conserved antigenic variation phenotype, recent analyses show that VSG repertoires in *T. brucei*, *T. congolense*, and *T. vivax* are substantially different in terms of sequence variation, lineage composition, and genomic organization. These differences may affect the mechanism of antigenic variation. For example, the CTD common to both a-VSG and b-VSG in *T. brucei* is currently thought to be essential for VSG switching (Box 3) [9,34]. Yet, this conserved CTD is unique to *T. brucei*; other species have a variety of CTD structures. Since the a-VSG and b-VSG lineages are much older than *T. brucei* itself (see below), it was concluded that harmonization of *T. brucei* CTDs is a recent adaptation for VSG transposition [9]. It follows that *T. congolense* and *T. vivax* must have different mechanisms for antigenic switching to some extent, and these remain unknown, although *T. congolense* has telomeric domains at least superficially reminiscent of *T. brucei* VSG expression sites [32]. We should remember that these VSG repertoires are defined by sequence homology, and description of a gene as a VSG does not guarantee that it functions as a variant antigen. Hence, interspecies variation in VSG repertoire may represent functional variation, which we explore below through several examples of VSG **neofunctionalization**.

The origins of VSGs

Comparison of VSG repertoires allows us to understand how old these genes are relative to the species themselves, and how the different VSG lineages might have emerged or disappeared

Box 3. The mechanism of trypanosome antigenic switching

T. brucei VSG transcription is restricted to specialized telomere-proximal structures in megabase chromosomes. In the bloodstream life stage of the parasite, VSGs are expressed from bloodstream expression sites, which comprise a promoter, a 70 bp repeat, the VSG, several non-VSG expression-site associated genes (i.e., ESAGs), and sporadic retrotransposons [28–30]. During the metacyclic life stage, VSGs are expressed from metacyclic expression sites, which are shorter, lack the 70 bp repeat, and have a specific promoter and a monocistronic structure [39,67].

VSG switching occurs at 10^2–10^3 switches per cell per population [91] and is intimately connected to the VSG expression site. It can occur through transcriptional switching, via the silencing of the active VSG expression site and subsequent activation of a different one [92,93], but it most commonly happens through recombinatorial switching. The latter involves the replacement of the VSG in the active expression site by an alternative through gene conversion [40,94,95].

T. congolense bloodstream VSG expression is also telomeric [96], but expression sites have not been sequenced yet. A recent study provided a candidate for the expression site; VSGs may be transcribed from a widespread genomic locus near the telomeres of small and megabase chromosomes, that is, topologically similar to *T. brucei* VSG expression sites, but with clear differences in gene composition and conserved noncoding sequences [32]. We do not know how VSG switching operates in *T. congolense,* but the lack of the 70 bp repeat in the candidate expression sites, and the lack of an identifiable recombination breakpoint in the VSG CTD, suggest that antigenic switching involves different players from those of *T. brucei*.

In *T. vivax*, a specialized expression locus has not yet been identified, so the mechanisms of antigenic switching remain uncertain. However, the lack of evidence for VSG mosaic expression in *T. vivax* suggest that the mechanism of antigenic switching described for *T. brucei* cannot be operating in *T. vivax* [14]. This raises the possibility that a different mode of survival in the mammalian host, particularly in terms of the ability to withstand the VSG-directed antibody immune response and of persisting within the mammalian host, is in place.

throughout evolution. In fact, Jackson *et al*. [9] showed that a-type VSGs (i.e., a-VSGs, Fam23, Fam14/PAG, and Fam15/TfRs) shared a single, common ancestor after splitting from b-type genes (i.e., b-VSGs, Fam13, Fam16, Fam 24, ESAG2, VR9/Fam9, and VR1/Fam1). This indicates that the principal VSG subfamilies originated prior to the speciation of modern African trypanosomes [9]. Interestingly, the division of *T. congolense* VSG genes into Fam13 and Fam16 also likely arose before speciation because their closest relatives are ESAG2 and VR9/Fam9 genes in *T. brucei*, respectively, rather than each other. Conversely, it is suggested that all *T. brucei* canonical b-VSG genes are derived from a single ancestral lineage, with most of the remaining lineages lost or neofunctionalized, that is, they are derived from a diversity bottleneck [9].

However, this bottleneck in diversity does not mean that VSG 'population size' was constricted at any point in the evolution of *T. brucei*. Rather, we suggest that it occurs through the constant replacement of ancestral genes by derived sequences through **gene conversion** until the repertoire consists purely of new genes. Divergence in this manner was proposed for the concerted evolution of VSG within species and the lack of orthology between species [9]. Despite the recent derivation of the repertoire, it is important to realize that antigenic diversity in *T. brucei* is not reduced relative to other species since there is a seemingly unlimited capacity for mosaicism among both a-VSGs and b-VSGs. Indeed, alongside the unique acquisition of a common CTD, their single, recent origin may have promoted this mosaicism since recombination frequency increases proportionally with sequence similarity between templates [35]. Hence, we speculate that the recent diversity bottleneck affecting *T. brucei* VSGs may be part of a general adaptation for greater antigenic diversity in this species through **mosaicism**, which is discussed in the next section.

As *T. vivax* occupies the most basal branch in the African trypanosome phylogeny, its repertoire offers an important insight into that of African trypanosome common ancestor (ATCA). It is unclear whether *T. vivax*-specific VSG subfamilies (i.e., Fam25,26) are derived from ancestral

lineages lost in *T. congolense* and *T. brucei*, or otherwise derived from a-type and b-type VSG lineages independently (and again, it is not confirmed that all of these encode variant antigens). However, based on the *T. vivax* repertoire, we can say that the ATCA VSG repertoire contained multiple, distinct lineages of a-type and b-type VSGs, and perhaps also additional lineages now extant in *T. vivax* only. The multiple VSG lineages have subsequently experienced different evolutionary pressures in each species to produce the repertoires we see today: some lineages were lost (e.g., a-VSG in *T. congolense*), others diversified rapidly (b-VSG in *T. brucei*), whilst some acquired invariant roles (e.g., TfR in the ancestor of *T. brucei* and *T. congolense*).

These various hypotheses of the origins of VSG subfamilies stem from comparative analyses, which are always limited by our sampling effort. Genome sequences for additional African trypanosomes will test these ideas, in particular, the genomes of the *T. vivax*-like species (i.e., *T. vivax* sister lineage [36]) would provide another perspective on the apparent greater phylogenetic diversity in the ATCA; while genomes for *T. suis* clades (the sister lineage to *T. brucei* [37]) would directly examine the hypothesis that *T. brucei* antigenic diversity has increased through diversity bottleneck and mechanistic adaptations such as the common CTD.

Forces driving diversification of the VSG repertoire

We might expect genes involved in host–parasite interactions, such as VSGs, to be influenced by natural selection. Evolutionary biologists measure the direction of selection using the ratio of nonsynonymous to synonymous nucleotide substitutions [**ω (dN/dS)**], which equals 1 under neutral conditions. More nonsynonymous than synonymous substitutions indicates **positive selection** of new gene variants and molecular adaptation. Given the parasites' need to diversify their antigenic repertoire to continually undergo antigenic variation, one could predict that VSG sequence evolution would be under positive selection. This is not the case. While most genes in *T. brucei* are under **negative selection** (i.e., ω of non-VSGs falls below 0.3 [26]), this effect is much weaker for VSG genes (i.e., ω of *T. brucei* VSGs ranges between 0.31 and 0.7 [26]), thus contributing to the removal of around half of the nonsynonymous substitutions [10]. Therefore, the unusual diversity of VSGs must be driven by an alternative force.

In fact, we know that antigenic novelty in *T. brucei* is driven predominantly by frequent gene duplication coupled with gene conversion, often partial, resulting in sequence mosaicism [8]. Gene conversion in *T. brucei* is facilitated by two conserved flanking sequences that favor sequence exchange (the 70 bp repeat upstream of the VSG gene [34,38,39] and the common CTD downstream [40]) which remain *in situ* after gene conversion and result in a recombination breakpoint between the NTD and the CTD (Figure 1D) [8,40]. Mosaic VSGs arise from the conversion of segments from multiple VSG gene or pseudogene donors until a functional gene is obtained [8,11,13]. Mosaicism is key to antigenic variation, especially during chronic infection [8,11–13,41], as it generates novel antigens that promote immune exhaustion. For instance, Jayaraman *et al*. [13] showed that between 27 and 187 VSGs were expressed simultaneously in the first 12 days of mouse infection, also that mosaic VSGs appear within 3 days postinfection and could be assembled from donor sequences with only 40% nucleotide identity. However, the combination of gene duplication and gene conversion also leaves a legacy of redundant VSG copies, which we hypothesize provides a reservoir of structural potential for the evolution of novel functions. To support this, we discuss the alternative roles for the VSG family below.

To different degrees, conservation of primary structure among the *T. congolense* and *T. vivax* VSGs seems to impose limits to genetic exchange, resulting in lower VSG recombination rates, which goes some way to explaining why the phylogenies from each species in Figure 1C have different shapes [9]. In *T. congolense*, VSGs recombine almost exclusively with genes of the

same phylotype, effectively limiting the potential for the generation of diversity [14] (Figure 1D). We speculate that negative selection operates to maintain distinct phylotypes, preventing major fluctuations in gene numbers [31]. This indicates functional differentiation within the repertoire but the precise reason for the strict conservation of *T. congolense* phylotypes remains unknown.

The scenario is more extreme in *T. vivax*. The generally low sequence similarity among phylotypes indicated that genetic exchange between them would be difficult [9] (Figure 1D). Indeed, a recent population genomic study showed that recombination between VSG phylotypes is negligible [14]. **Coalescence times** of VSG phylotypes are uniformly large, suggesting that they were created a long time ago and broadly contemporaneously, rather than continuously and recently as required if gene conversion were happening. For a given *T. vivax* VSG, complete orthologs (including CTD) may be found in distantly related strains and across the whole species, contrary to what is seen in *T. brucei* [42]. This raises searching questions about the mechanism of antigenic variation. Given that the *T. vivax* VSG repertoire does not appear to diversify [14], how does this parasite establish long-term infections? While *T. brucei* (and probably *T. congolense*) do so through the continual generation of serologically distinct mosaic VSGs, *T. vivax* perhaps has an alternative mechanism for persistence, and understanding this may bring new opportunities for preventing or treating disease.

VSG neofunctionalization and the evolution of host–parasite interactions

When parasites respond to a new selective pressure, they usually repurpose a redundant gene duplicate whose original function remains assured by paralogs. The mechanisms that rapidly recombine VSG sequences bring about antigenic switching and immune evasion. Yet, we speculate that they are also an engine for structural innovation, ensuring that raw material is constantly available for new gene functions. As such, VSG neofunctionalization drives host–parasite coevolution and is a pertinent subject when considering many aspects of African trypanosome biology. VSGs play other functions besides antigenic variation, namely immune modulation, host range (through resistance to human serum), nutrient scavenging (by mediating iron transport), and drug resistance (Figure 2).

Immune modulation

VSGs directly interact with the immune system and modulate the host's cellular responses by various modes [43]. For example, they induce a proinflammatory response via IFN-γ production by macrophages [44–47], increased MHC-II presentation, and activation of $CD4^+$ T cells [48], which results in inflammatory anemia and other symptoms, and they can specifically inhibit activation of the alternative pathway of the complement system to prevent trypanosome lysis (reviewed in [43]).

Resistance to human serum

Two strains of *T. brucei*, (*T. b. rhodesiense* and *T. b. gambiense*), have evolved to infect humans by independently repurposing a VSG to bypass the effects of the **trypanolytic factors** (TLF-1 and 2). *T. b. rhodesiense* resistance to human serum is conferred by the VSG *SRA* gene [49,50] and is cotranscribed with the expressed VSG in the bloodstream VSG expression site [51]. The *SRA* gene has an internal deletion of 126 amino acids compared to its canonical a-VSG paralog in *T. b. brucei* [52]. This deletion spans the primary sequence that usually forms the surface-exposed antigenic loops, and includes two conserved cysteine residues, so that the traditional disulfide bridging pattern is compromised [52]. Therefore, the external faces of the two alpha-helices are exposed, which may serve as a binding site. Unlike canonical VSGs, SRA is not a surface protein, localizing instead to the endosomal network [53,54] where it binds to apolipoprotein L-I (APOL-1), targeting it for degradation [55–57].

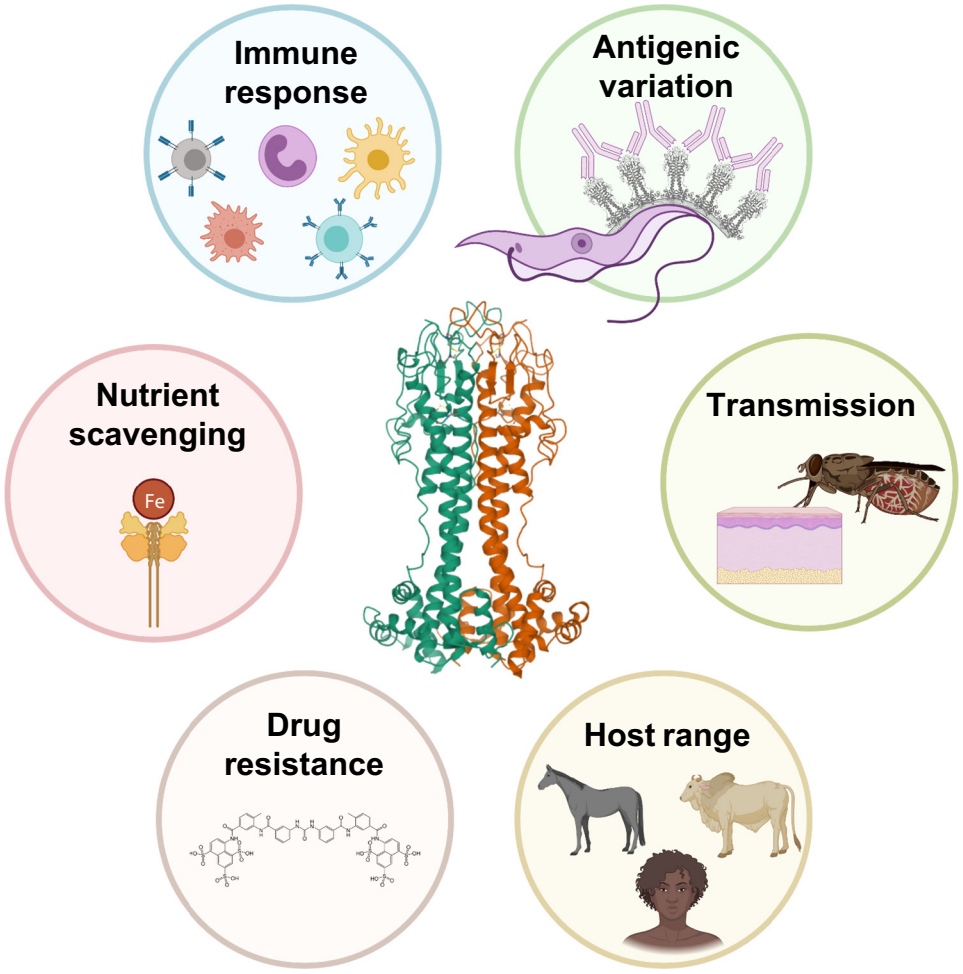

Figure 2. The different roles of variant and nonvariant variant surface glycoprotein (VSG) genes. Besides antigenic variation and its impact on the host's antibody response, VSGs affect transmission through metacyclic VSG expression; host range, because they can confer human infectivity to some *T. brucei* strains; resistance to the drug suramin; nutrient scavenging through the mediation of transferrin binding for iron import; and immune modulation, as they are known to trigger a proinflammatory response whilst suppressing the activation of complement. Figure created with BioRender.com.

In contrast, *T. b. gambiense* neutralizes TLFs via a multifactorial mechanism that includes the TgsGP [58,59], a GPI-anchored b-VSG protein lacking the CTD that localizes to the **flagellar pocket** [58,60]. Unlike the *SRA* gene, *TgsGP* is transcribed by pol-II and is not contained within the VSG expression site, but rather in the subtelomere of chromosome 21 [60]. Rather than directly interacting with APOL-1, TgsGP induces membrane stiffening, rendering APOL-1-mediated cell lysis inefficient. Haptoglobin-hemoglobin receptor inactivation due to an amino acid substitution (L210S), and increased acidity in the endosomes, also contribute to the resistance mechanism by preventing TLF-1 internalization and accelerating APOL-1 degradation, respectively [59]. Both SRA and TgsGP show how host range evolves by VSG neofunctionalization.

Iron transport

T. brucei and *T. congolense* have a repurposed a-type VSG lineage that encodes TfRs (Fam15), essential iron-binding proteins expressed at the flagellar pocket [61]. The *TfR* gene subfamilies of

T. brucei (*n* = 23) and *T. congolense* (*n* = 45) are sister lineages derived from an a-type VSG in the common ancestor of these two species. This follows because *T. vivax* lacks the *TfR* gene subfamily but has a-type VSGs, making the TfR lineage younger [9]. In *T. brucei*, they are typically arranged in tandem pairs (ESAG6 and 7) within the bloodstream VSG expression sites. In *T. congolense*, *TfR* genes are mostly distributed throughout the subtelomeres, although a few copies have been found associated with possible *T. congolense* VSG expression sites [32].

Drug resistance

In contrast to the long evolutionary timescale of the preceding examples, evolution of drug resistance occurred in the last century due to the strong selective pressure imposed by drug treatment. Suramin is a first-line drug for first-stage acute human sleeping sickness, but resistance has become a major issue in clinics. *T. b. rhodesiense* and *T. b. brucei* resistance to suramin is caused by the expression of a divergent VSG (VSG^{sur}) [62,63]. Although the exact mechanism of resistance is unclear, drug internalization via the endocytic lysosomal pathway seems to be reduced [63]. Recent resolution of the tertiary structure of VSG^{sur} by X-ray crystallography revealed that VSG^{sur} binds to suramin, reducing the concentration of bioavailable drug entering the trypanosome cell [64]. Furthermore, newly produced VSG^{sur} has been hypothesized to bind the imported suramin inside the endosome, and to be then trafficked to the cell surface, effectively removing the drug from cells [64].

Besides these examples, structural variation among VSGs might reflect roles in other biological processes, not yet identified.

Transmission

Whilst bloodstream-form parasite populations express a superabundant variant antigen, *T. brucei* and *T. congolense* metacyclic parasites express a pool of coabundant metacyclic VSGs [65–69]. In *T. brucei* and *T. congolense*, metacyclic VSG genes are different from one fly

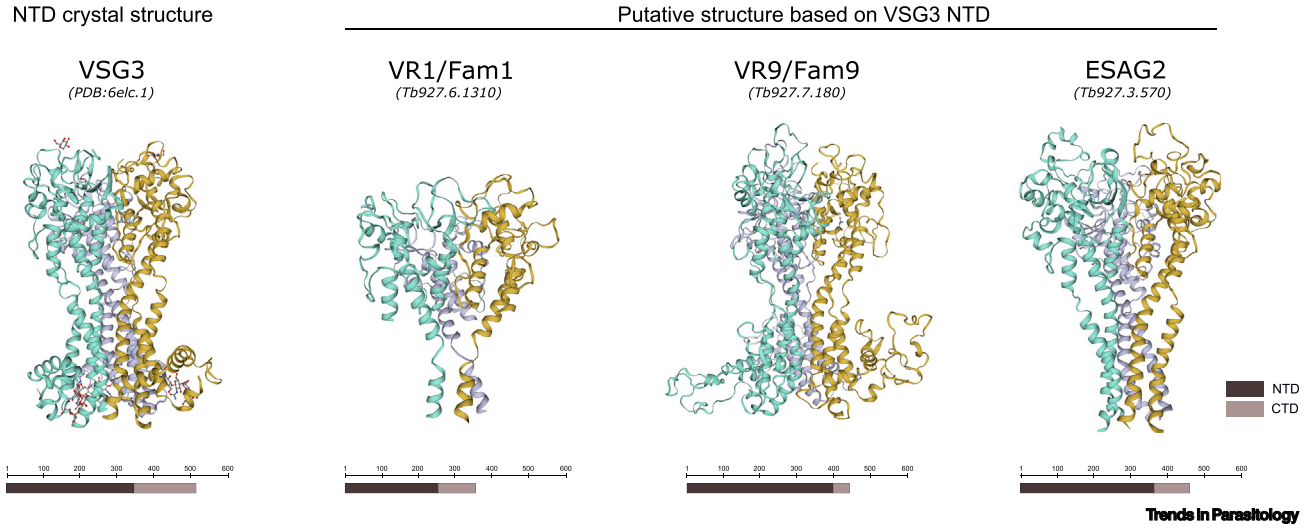

Figure 3. Putative 3D structures of the N-terminal domains of b-type nonvariant variant surface glycoprotein (VSG) genes. The solved structure of the VSG3 N-terminal domain (NTD) [24] was used as a template to predict the tertiary structures of three b-type nonvariant VSG genes by homology-modeling, using SWISS-MODEL [98] at the ExPASy server [99]. VSG3 is spatially arranged as a homotrimer, so each color represents a different monomer/chain. Tb927.6.1310 is a representative of the VR1/Fam1 subfamily, which has been shown to encode a membrane protein restricted to the flagellar pocket and endosomes of bloodstream forms [9,61]. Tb927.3.1500 is a member of the VR9/Fam9 subfamily, which is expressed on the bloodstream-form cell surface [61]. Tb927.3.570 encodes ESAG2 (Fam13), which is constitutively expressed across the whole cell body membrane of bloodstream forms [61]. Below each 3D structure we show a comparison of NTD and C-terminal domain (CTD) length of each protein. To determine the CTD and NTD of each gene, we chose an arbitrary boundary at the conserved tryptophan position (residue number 342 for VSG3). Scale represents amino acid sequence length.

generation to another. However, we showed recently that *T. congolense* VSG phylotype (P8), the smallest phylotype in the *T. congolense* genome, is consistently expressed by metacyclic parasites and maintained across transmission cycles [31]. Specifically, TcIL3000_0_09520 (TcIL3000.A.H_000381200) was significantly enriched within metacyclic transcriptomes, relative to its low copy-number in the genome, and was abundant in three metacyclic-stage VSG studies [31,70,71]). The role of this specific gene remains unknown but it is also weakly expressed by bloodstream-form parasites, in addition to the active VSG (unpublished datasets and [72]). It may be constitutively expressed and no longer encoding a variant antigen, much like ESAG2 in *T. brucei* [61], perhaps functioning as a receptor for host-derived factors or nutrients.

Tissue tropism

Antigenically variable pathogens are known to exploit tissue tropism as a survival strategy. In malaria, differential variant antigen expression by *Plasmodium* parasites causes infected erythrocytes to avoid splenic clearance by cytoadhering to more immune-tolerant sites [73]. *Borrelia* employs an analogous strategy: spirochetes expressing alternative antigenic variants invade and persist in the brain to bypass the immune response and establish chronicity [74]. Similar strategies may be employed by trypanosomes. For instance, *T. brucei* relies on tissue invasion for the establishment of chronic infection and survival in the mammal [75], and we know that parasites in the blood and tissues express different VSGs [76]. Do certain VSGs favor tropism for specific tissues? Does *T. brucei* invade tissues as a means to amplify the antigenic diversity of the parasite population in the host and impair the host's immune response? How would this apply to fully intravascular parasites, like *T. congolense*? The potential role of VSGs in tissue preference and the importance of antigenic diversity for tissue-resident parasites has been recently revisited by us [77] and is currently under investigation.

It appears generally true that, when VSG genes gain new functions, this coincides with divergence of protein structure, including loss of domains or key amino acid residues that are essential for a variant antigen. Typically, we label these genes as 'atypical' in relation to the canonical primary structure. This is exemplified by the loss or reduction of the CTD in the VR, ESAG2, TgsGP, and SRA genes. If sequence evolution is an indicator of functional evolution, then, judging by the number of 'atypical' VSGs that remain uncharacterized in *T. brucei* (such as ESAG2, VR genes, and *T. b. evansi* RoTat1.2 [78]), there are many functions still to be discovered. Figure 3 illustrates this point by modeling the tertiary structures of *T. brucei* VR1/Fam1, VR9/Fam9, and ESAG2, after threading their amino acid sequences onto the solved b-type *T. brucei* Lister 427 VSG3 NTD crystal structure [24]. Whilst real conformations might differ from the predictions, it is clear that the diversity in genetic structure among VSGs will result in distinct proteins, perhaps distinct enough to localize to different cell compartments or adopt diverse positions in the trypanosome cell membrane. And indeed, these may be just the beginning of functional diversity within the family; the deeply conserved structural differences between *T. congolense* and *T. vivax* VSG phylotypes could also point to important functional distinctions, whether within, or in addition to, their variant antigen role.

Concluding remarks

VSGs are crucial for trypanosome survival because they provide raw material for antigenic variation, but also for innovation in host–parasite interactions. We now understand that the apparent phenotypic similarity of antigenic variation in trypanosome infections contrasts with quite distinct genetic bases in each trypanosome species, which will translate into mechanistic differences in the near future (see Outstanding questions). We can also begin to appreciate how the family has evolved. VSGs are constantly diversifying within species through gene duplication and recombination, while repertoire diverges between species through selective loss and elaboration of ancestral lineages. New evidence for functional differentiation and neofunctionalization within

Outstanding questions

Do *T. vivax* and *T. congolense* VSGs have similar 3D protein conformations to *T. brucei* VSGs? How does that affect the parasite's surface?

Have Fam25 and Fam26 originated *de novo* in *T. vivax* or have they been lost in the lineage leading to *T. congolense* and *T. brucei*?

Can *T. suis* genome sequences explain how all *T. brucei* came to have a conserved, homologous CTD when other species do not?

What are the functions of *T. brucei* VR genes (Fam1/Fam9) and ESAG2?

Why are TfRs largely confined to the expression site in *T. brucei*, but to the subtelomeres in *T. congolense*?

If the TfR subfamily is absent in *T. vivax*, does this species have an analogous receptor for sequestering host iron and, if so, is this also derived from VSGs?

Are mVSG repertoires predictable in any African trypanosome species?

Does differential VSG expression play a role in tissue invasion or vascular adhesion?

Are all VSG genes in *T. congolense* and *T. vivax* variant antigens, or do some of them carry out other functions?

Can neofunctionalized VSGs that are not variant antigens be employed as vaccine candidates?

the VSG family suggests that gene function is more nuanced than commonly assumed. In part, this explains why there are so many VSG genes. Antigenic switching evolved to ensure novel antigens are always emerging during infections, but it also encourages duplication and redundancy in the genome. Altogether, our new perspective on the VSG family indicates that its importance in host–parasite interactions, pathogenesis, and, therefore, potential for disease control, is likely to be much deeper and more complex than we commonly imagine. Everything suggests that there is much to gain by increasing efforts to dissect the functional implications of VSG diversity.

Acknowledgments

S.S.P. is funded by the European Union's Horizon 2020 research and innovation program through a Marie Skłodowska-Curie Individual Standard European Fellowship, under grant agreement no. 839960. A.P.J. is funded by a Leverhulme Trust Research Fellowship (RF-2020-413). L.M.F. is an Investigator CEEC of the Fundação para a Ciência e a Tecnologia (CEECIND/03322/2018) and the laboratory is funded by the ERC (FatTryp, ref. 771714). We apologize to colleagues whose work was not cited due to space limitations. We thank the reviewers for their constructive comments.

Declaration of interests

The authors declare no competing interests.

References

1. Matthews, K.R. *et al.* (2015) The within-host dynamics of African trypanosome infections. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 370, 20140288
2. El-Sayed, N.M. *et al.* (2005) Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309, 404–409
3. Jackson, A.P. (2016) Gene family phylogeny and the evolution of parasite cell surfaces. *Mol. Biochem. Parasitol.* 209, 64–75
4. Horn, D. (2014) Antigenic variation in African trypanosomes. *Mol. Biochem. Parasitol.* 195, 123–129
5. Cross, G.A. (1984) Structure of the variant glycoproteins and surface coat of *Trypanosoma brucei*. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 307, 3–12
6. Jackson, A.P. *et al.* (2013) A cell-surface phylome for African trypanosomes. *PLoS Negl. Trop. Dis.* 7, e2121
7. Barry, J.D. *et al.* (2003) Why are parasite contingency genes often associated with telomeres? *Int. J. Parasitol.* 33, 29–45
8. Marcello, L. and Barry, J.D. (2007) Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Res.* 17, 1344–1352
9. Jackson, A.P. *et al.* (2012) Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. *Proc. Natl. Acad. Sci. U. S. A.* 109, 3416–3421
10. Cross, G. a M. *et al.* (2014) Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.* 195, 59–73
11. Hall, J.P.J. *et al.* (2013) Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. *PLoS Pathog.* 9, e1003502
12. Mugnier, M.R. *et al.* (2015) The *in vivo* dynamics of antigenic variation in *Trypanosoma brucei*. *Science* 347, 1470–1473
13. Jayaraman, S. *et al.* (2019) Application of long read sequencing to determine expressed antigen diversity in *Trypanosoma brucei* infections. *PLoS Negl. Trop. Dis.* 13, e0007262
14. Silva Pereira, S. *et al.* (2020) Variant antigen diversity in *Trypanosoma vivax* is not driven by recombination. *Nat. Commun.* 11, 844
15. Silva Pereira, S. *et al.* (2019) VAPPER: High-throughput variant antigen profiling in African trypanosomes of livestock. *Gigascience* 8, 1–8
16. Johnson, J.G. and Cross, G.A.M. (1979) Selective cleavage of variant surface glycoproteins from *Trypanosoma brucei*. *Biochem. J.* 178, 689–697
17. Ferguson, M. *et al.* (1988) Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science* 239, 753–759
18. Rice-Fitch, A. *et al.* (1981) Sequence homologies near the C-termini of the variable surface glycoproteins of *Trypanosoma brucei*. *Nature* 294, 53–57
19. Carrington, M. *et al.* (1991) Variant specific glycoprotein of *Trypanosoma brucei* consists of two domains each having an independently conserved pattern of cysteine residues. *J. Mol. Biol.* 221, 823–835
20. Blum, M.L. *et al.* (1993) A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. *Nature* 362, 603–609
21. Weirather, J.L. *et al.* (2012) Mapping of VSG similarities in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 181, 141–152
22. Schwede, A. *et al.* (2011) The VSG C-terminal domain is inaccessible to antibodies on live trypanosomes. *Mol. Biochem. Parasitol.* 175, 201–204
23. Hutchinson, O.C. *et al.* (2003) VSG structure: similar N-terminal domains can form functional VSGs with different types of C-terminal domain. *Mol. Biochem. Parasitol.* 130, 127–131
24. Pinger, J. *et al.* (2018) African trypanosomes evade immune clearance by O-glycosylation of the VSG surface coat. *Nat. Microbiol.* 3, 932–938
25. Jackson, A.P. *et al.* (2010) The genome sequence of *Trypanosoma brucei gambiense*, causative agent of chronic human African trypanosomiasis. *PLoS Negl. Trop. Dis.* 4, e658
26. Carnes, J. *et al.* (2015) Genome and phylogenetic analyses of *Trypanosoma evansi* reveal extensive similarity to *T. brucei* and multiple independent origins for dyskinetoplasty. *PLoS Negl. Trop. Dis.* 9, e3404
27. Sistrom, M. *et al.* (2016) *De novo* genome assembly shows genome wide similarity between *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense*. *PLoS One* 11, e0147660
28. Becker, M. *et al.* (2004) Isolation of the repertoire of VSG expression site containing telomeres of *Trypanosoma brucei* 427 using transformation-associated recombination in yeast. *Genome Res.* 14, 2319–2329
29. Young, R. *et al.* (2008) Isolation and analysis of the genetic diversity of repertoires of VSG expression site containing telomeres from *Trypanosoma brucei gambiense*, *T. b. brucei* and *T. equiperdum*. *BMC Genom.* 9, 1–16
30. Hertz-Fowler, C. *et al.* (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS One* 3, e3527
31. Silva Pereira, S. *et al.* (2018) Variant antigen repertoires in *Trypanosoma congolense* populations and experimental infections can be profiled from deep sequence data with a set of universal protein motifs. *Genome Res.* 28, 1383–1394

32. Abbas, A.H. *et al.* (2018) The structure of a conserved telomeric region associated with variant antigen loci in the blood parasite *Trypanosoma congolense*. *Genome Biol. Evol.* 10, 2458–2473
33. Jackson, A.P. *et al.* (2015) Global gene expression profiling through the complete life cycle of *Trypanosoma vivax*. *PLoS Negl. Trop. Dis.* 9, e0003975
34. Hovel-Miner, G. *et al.* (2016) A conserved DNA repeat promotes selection of a diverse repertoire of *Trypanosoma brucei* surface antigens from the genomic archive. *PLoS Genet.* 12, e1005994
35. Watt, V.M. *et al.* (1985) Homology requirements for recombination in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 82, 4768–4772
36. Rodrigues, C.M. *et al.* (2017) New insights from Gorongosa National Park and Niassa National Reserve of Mozambique increasing the genetic diversity of *Trypanosoma vivax* and *Trypanosoma vivax*-like in tsetse flies, wild ungulates and livestock from East Africa. *Parasit. Vectors* 10, 337
37. Rodrigues, C.M.F. *et al.* (2020) Expanding our knowledge on African trypanosomes of the subgenus *Pycnomonas*: a novel *Trypanosoma suis*-like in tsetse flies, livestock and wild ruminants sympatric with *Trypanosoma suis* in Mozambique. *Infect. Genet. Evol.* 78, 104143
38. Liu, A.Y.C. *et al.* (1983) The transposition unit of variant surface glycoprotein gene 118 of *Trypanosoma brucei* – presence of repeated elements at its border and absence of promoter-associated sequences. *J. Mol. Biol.* 167, 57–75
39. Campbell, D.A. *et al.* (1984) The 5′-limit of transposition and upstream barren region of a trypanosome VSG gene: tandem 76 base-pair repeats flanking (TAA)90. *Nucleic Acids Res.* 12, 2759–2774
40. Bernards, A. *et al.* (1981) Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3′ end. *Cell* 27, 497–505
41. Barbet, A.F. and Kamper, S.M. (1993) The importance of mosaic genes to trypanosome survival. *Parasitol. Today* 9, 63–66
42. Hutchinson, O.C. *et al.* (2007) Variant surface glycoprotein gene repertoires in *Trypanosoma brucei* have diverged to become strain-specific. *BMC Genom.* 8, 234
43. Stijlemans, B. *et al.* (2016) Immune evasion strategies of *Trypanosoma brucei* within the mammalian host: Progression to pathogenicity. *Front. Immunol.* 7, 233
44. Magez, S. *et al.* (1998) The glycosyl-inositol-phosphate and dimyristoylglycerol moieties of the glycosylphosphatidylinositol anchor of the trypanosome variant-specific surface glycoprotein are distinct macrophage-activating factors. *J. Immunol.* 160, 1949–1956
45. Magez, S. *et al.* (2002) VSG-GPI anchors of African trypanosomes: Their role in macrophage activation and induction of infection-associated immunopathology. *Microbes Infect.* 4, 999–1006
46. Leppert, B.J. *et al.* (2007) The soluble variant surface glycoprotein of African trypanosomes is recognized by a macrophage scavenger receptor and induces I kappa B alpha degradation independently of TRAF6-mediated TLR signaling. *J. Immunol.* 179, 548–556
47. Mansfield, J.M. and Paulnock, D.M. (2005) Regulation of innate and acquired immunity in African trypanosomiasis. *Parasite Immunol.* 27, 361–371
48. Schleifer, K.W. *et al.* (1993) Characterization of T helper cell responses to the trypanosome variant surface glycoprotein. *J. Immunol.* 150, 2910–2919
49. De Greef, C. and Hamers, R. (1994) The serum resistance-associated (SRA) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. *Mol. Biochem. Parasitol.* 68, 277–284
50. Stephens, N.A. *et al.* (2012) Trypanosome resistance to human innate immunity: Targeting Achilles' heel. *Trends Parasitol.* 28, 539–545
51. Van Xong, H. *et al.* (1998) A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma brucei rhodesiense*. *Cell* 95, 839–846
52. Campillo, N. and Carrington, M. (2003) The origin of the serum resistance associated (SRA) gene and a model of the structure of the SRA polypeptide from *Trypanosoma brucei rhodesiense*. *Mol. Biochem. Parasitol.* 127, 79–84
53. Oli, M.W. *et al.* (2006) Serum resistance-associated protein blocks lysosomal targeting of trypanosome lytic factor in *Trypanosoma brucei*. *Eukaryot. Cell* 5, 132–139
54. Stephens, N.A. and Hajduk, S.L. (2011) Endosomal localization of the serum resistance-associated protein in African trypanosomes confers human infectivity. *Eukaryot. Cell* 10, 1023–1033
55. Vanhamme, L. *et al.* (2003) Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* 422, 83–87
56. Pérez-Morga, D. *et al.* (2005) Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal membranes. *Science* 309, 469–472
57. Lecordier, L. *et al.* (2009) C-Terminal mutants of apolipoprotein L-I efficiently kill both *Trypanosoma brucei brucei* and *Trypanosoma brucei rhodesiense*. *PLoS Pathog.* 5, e1000685
58. Capewell, P. *et al.* (2013) The TgsGP gene is essential for resistance to human serum in *Trypanosoma brucei gambiense*. *PLoS Pathog.* 9, e1003686
59. Uzureau, P. *et al.* (2013) Mechanism of *Trypanosoma brucei gambiense* resistance to human serum. *Nature* 501, 430–434
60. Berberof, M. *et al.* (2001) A receptor-like flagellar pocket glycoprotein specific to *Trypanosoma brucei gambiense*. *Mol. Biochem. Parasitol.* 113, 127–138
61. Gadelha, C. *et al.* (2015) Architecture of a host–parasite interface: complex targeting mechanisms revealed through proteomics. *Mol. Cell. Proteom.* 14, 1911–1926
62. Wiedemar, N. *et al.* (2018) Beyond immune escape: a variant surface glycoprotein causes suramin resistance in *Trypanosoma brucei*. *Mol. Microbiol.* 107, 57–67
63. Wiedemar, N. *et al.* (2019) Expression of a specific variant surface glycoprotein has a major impact on suramin sensitivity and endocytosis in *Trypanosoma brucei*. *FASEB BioAdv.* 1, 595–608
64. Zeelen, J. *et al.* (2021) Structure of trypanosome coat protein VSGsur and function in suramin resistance. *Nat. Microbiol.* 6, 392–400
65. Esser, K.M. *et al.* (1982) *Trypanosoma rhodesiense* blood forms express all antigen specificities relevant to protection against metacyclic (insect form) challenge. *J. Immunol.* 129, 1715–1718
66. Crowe, J.S. *et al.* (1983) All metacyclic variable antigen types of *Trypanosoma congolense* identified using monoclonal antibodies. *Nature* 306, 389–391
67. Lenardo, M.J. *et al.* (1984) Characterization of the genes specifying two metacyclic variable antigen types in *Trypanosoma brucei rhodesiense*. *Proc. Natl. Acad. Sci. U. S. A.* 81, 6642–6646
68. Lenardo, M.J. *et al.* (1986) Metacyclic variant surface glycoprotein genes of *Trypanosoma brucei* subsp. *rhodesiense* are activated *in situ*, and their expression is transcriptionally regulated. *Mol. Cell Biol.* 6, 1991–1997
69. Turner, C.M.R. *et al.* (1988) An estimate of the size of the metacyclic variable antigen repertoire of *Trypanosoma brucei rhodesiense*. *Parasitology* 97, 269–276
70. Eshita, Y. *et al.* (1992) Metacyclic form-specific variable surface glycoprotein-encoding genes of *Trypanosoma* (*Nannomonas*) *congolense*. *Gene* 113, 139–148
71. Helm, J.R. *et al.* (2009) Analysis of expressed sequence tags from the four main developmental stages of *Trypanosoma congolense*. *Mol. Biochem. Parasitol.* 168, 34–42
72. Silvester, E. *et al.* (2018) A gene expression comparison of *Trypanosoma brucei* and *Trypanosoma congolense* in the bloodstream of the mammalian host reveals species-specific adaptations to density-dependent development. *PLoS Negl. Trop. Dis.* 12, e0006863
73. Gomes, P.S. *et al.* (2016) Immune escape strategies of malaria parasites. *Front. Microbiol.* 7, 1617
74. Cadavid, D. *et al.* (2001) Isogenic serotypes of *Borrelia turicatae* show different localization in the brain and skin of mice. *Infect. Immun.* 69, 3389–3397
75. De Niz, M. *et al.* (2021) Organotypic endothelial adhesion molecules are key for *Trypanosoma brucei* tropism and virulence. *Cell Rep.* Published online April 7, 2021. http://dx.doi.org/10.2139/ssrn.3821966
76. Trindade, S. *et al.* (2016) *Trypanosoma brucei* parasites occupy and functionally adapt to the adipose tissue in mice. *Cell Host Microbe* 19, 837–848
77. Silva Pereira, S. *et al.* (2019) Tissue tropism in parasitic diseases. *Open Biol.* 9, 190036
78. Lejon, V. *et al.* (2005) Recombinant RoTat 1.2 variable surface glycoprotein as antigen for diagnosis of *Trypanosoma evansi* in dromedary camels. *Int. J. Parasitol.* 35, 455–460

79. Dagnachew, S. and Bezie, M. (2015) Review on *Trypanosoma vivax*. *Afr. J. Basic Appl. Sci.* 7, 41–64
80. Brun, R. *et al.* (1998) *Trypanosoma evansi* and *T. equiperdum*: Distribution, biology, treatment and phylogenetic relationship (a review). *Vet. Parasitol.* 79, 95–107
81. Mehlert, A. *et al.* (2002) The glycoforms of a *Trypanosoma brucei* variant surface glycoprotein and molecular modeling of a glycosylated surface coat. *Glycobiology* 12, 607–612
82. Doyle, J.J. *et al.* (1980) Antigenic variation in clones of animal-infective *Trypanosoma brucei* derived and maintained *in vitro*. *Parasitology* 80, 359–369
83. Pays, E. (1989) Pseudogenes, chimaeric genes and the timing of antigen variation in African trypanosomes. *Trends Genet.* 5, 389–391
84. Thon, G. *et al.* (1989) Antigenic diversity by the recombination of pseudogenes. *Genes Dev.* 3, 1247–1254
85. Morrison, L.J. *et al.* (2005) Probabilistic order in antigenic variation of *Trypanosoma brucei*. *Int. J. Parasitol.* 35, 961–972
86. Lythgoe, K.A. *et al.* (2007) Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 8095–8100
87. Gjini, E. *et al.* (2010) Critical Interplay between parasite differentiation, host immunity, and antigenic variation in trypanosome infections. *Am. Nat.* 176, 424–439
88. Capbern, A. *et al.* (1977) *Trypanosoma equiperdum*: Étude des variations antigéniques au cours de la trypanosomose experimentale du lapin. *Exp. Parasitol.* 42, 6–13
89. Bachmann, A. *et al.* (2016) Mosquito passage dramatically changes var gene expression in controlled human *Plasmodium falciparum* infections. *PLoS Pathog.* 12, e1005538
90. Verhey, T.B. *et al.* (2018) Analysis of recombinational switching at the antigenic variation locus of the Lyme spirochete using a novel PacBio sequencing pipeline. *Mol. Microbiol.* 107, 104–115
91. Donelson, J.E. (2003) Antigenic variation and the African trypanosome genome. *Acta Trop.* 85, 391–404
92. Majiwa, P.A.O. *et al.* (1982) Two distinct forms of surface antigen gene rearrangement in *Trypanosoma brucei*. *Nature* 297, 514–516
93. Young, J.R. *et al.* (1982) Analysis of genomic rearrangements associated with two variable antigen genes in *Trypanosoma brucei*. *Nucleic Acids Res.* 10, 1982
94. Boothroyd, C.E. *et al.* (2009) A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature* 459, 278–281
95. Glover, L. *et al.* (2013) DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes. *PLoS Pathog.* 9, e1003260
96. Majiwa, P.A.O. *et al.* (1985) Cloning and analysis of *Trypanosoma* (*Nannomonas*) *congolense* ILNat 2.1 VSG gene. *Mol. Biochem. Parasitol.* 16, 97–108
97. Hamilton, P.B. *et al.* (2004) Trypanosomes are monophyletic: Evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *Int. J. Parasitol.* 34, 1393–1404
98. Waterhouse, A. *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303
99. Gasteiger, E. *et al.* (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788