Trends in Parasitology

CellPress
REVIEWS

Review

Mechanistic Similarities between Antigenic Variation and Antibody Diversification during *Trypanosoma brucei* Infection

Francisco Aresta-Branco,[1,2,3] Esteban Erben,[1,3] F. Nina Papavasiliou,[1,*] and C. Erec Stebbins[2,*]

***Trypanosoma brucei*, which causes African trypanosomiasis, avoids immunity by periodically switching its surface composition. The parasite is coated by 10 million identical, monoallelically expressed variant surface glycoprotein (VSG) molecules. Multiple distinct parasites (with respect to their VSG coat) coexist simultaneously during each wave of parasitemia. This substantial antigenic load is countered by B cells whose antigen receptors (antibodies or immunoglobulins) are also monoallelically expressed, and that diversify dynamically to counter each variant antigen. Here we examine parallels between the processes that generate VSGs and antibodies. We also discuss current insights into VSG mRNA regulation that may inform the emerging field of Ig mRNA biology. We conclude by extending the parallels between VSG and Ig to the protein level.**

Highlights

Identification of the VEX1 protein as a mediator of monoallelic expression has the potential to revolutionize our understanding of *T. brucei* biology but could also inform Ig allelic exclusion.

Identification of all *T. brucei* stage-specific RNA-binding proteins (RBPs) and their functions will open possibilities in the field of RNA biology in general and in the context of B cell biology in particular (where RNA biology has been lagging).

Recent structural and biochemical work on VSGs has revealed novel 3D folds, unexpected post-translational modifications, and insights into how VSGs move around the membrane aided by their GPI anchor. These recent advances have renewed a focus on VSG protein.

VSG and Ig Sequence Diversity Are Generated by Similar Processes

The Generation of the Expressed VSG and Ig Genes

Variant surface glycoproteins (**VSGs**, see Glossary) constitute the predominant antigen on the surface of *T. brucei*. The genomic repertoire (or archive) of VSGs consists of over 2000 genes, of which around 80% are annotated as incomplete or pseudogenes [1]. VSGs are transcribed by RNA polymerase I (Pol I) from one out of approximately 15 telomeric **bloodstream expression sites** (BESs) [2,3]. These are polycistronic loci encompassing (i) a variable number of **expression site-associated genes** (ESAGs) and (ii) the VSG gene which is just upstream of telomeric repeats [4] (Figure 1).

The telomeric location of the expressed VSG is mechanistically crucial for both enforcing monoallelic expression and supporting recombination (since telomeres are sites of random but frequent natural breakage and repair – [5]). Hence, VSG clonality (one VSG per trypanosome) and ease of replacement (via recombination) are essential aspects of the mechanism of antigenic variation. The establishment of monoallelic VSG expression is dependent on both chromatin conformation and subnuclear localization. The actively transcribed BES is devoid of nucleosomes and is sensitive to **DNase**, typical of an open chromatin status. In contrast, the silent BESs are enriched with regularly spaced nucleosomes in a compact conformation [6,7]. While the active BES is localized in a centrally positioned subnuclear compartment termed the **expression site body** (**ESB**) [3], a structure adjacent to the nucleolus, the inactive BESs are placed elsewhere in the nucleoplasm [8]. Whilst several factors have been associated with derepression of a silent BES and, in some cases, with disruption of monoallelic expression (Table 1), one in particular stands out. That is the VSG-exclusion-1 (VEX1) protein, which colocalizes with the active BES and whose depletion or overexpression both lead to an abrogation of monoallelic expression (marked by the

[1]Division of Immune Diversity, German Cancer Research Center, Heidelberg, Germany
[2]Division of Structural Biology of Infection and Immunity, German Cancer Research Center, Heidelberg, Germany
[3]These authors contributed equally to this work

*Correspondence:
n.papavasiliou@dkfz-heidelberg.de (F.N. Papavasiliou) and
e.stebbins@dkfz-heidelberg.de (C.E. Stebbins).

 https://doi.org/10.1016/j.pt.2019.01.011

presence of 11 or 12 silent VSGs on the surface of the parasites [9]). This identifies VEX1 as the sole simultaneous positive and negative regulator of VSG expression.

Monoallelic expression from a single BES is believed to be established early in the life of a bloodstream form parasite through a process termed ***in situ* switching** which is thought to probe BESs (together with their VSGs) in search of optimal ESAGs, ostensibly to promote host adaptation [10]. It is worth noting, however, that direct evidence for the bulk of this model is currently lacking. It is also thought that DNA recombination dominates in parasites throughout the course of infection [11]. Generally, the recombination mechanism employed is deletional gene conversion, which exchanges the active VSG gene with another, present in another BES, or replaces it with a VSG from the genomic archive (e.g., the archival VSG arrays or the VSGs that reside in minichromosomes). In all cases, translocation of the new VSG gene into the active BES occurs by RAD51-dependent homologous recombination either through gene conversion (predominantly) or break-induced replication using conserved sequences upstream (70 bp repeats) and downstream (within the 3′UTR) of the VSG-coding sequence [12,13]. Trypanosomatids appear to lack a functional **nonhomologous end-joining(NHEJ)** pathway, leaving **microhomology-mediated end-joining (MMEJ)** as a possible alternative to double-stranded break repair (though the role of MMEJ in VSG switching remains to be established) [14].

To counteract an infection with trypanosomes, the host employs mainly an **immunoglobulin**-M (IgM) centered **B cell** response, which clears the cognate parasites [15]. IgM is developmentally the first antibody on the surface of B cells. It is comprised of a heavy chain, encoded by the *IgH* locus, and a light chain, encoded by the *IgLκ* or *IgLλ* loci. The expressed heavy chain gene is somatically assembled in each B cell from variable (V), diversity (D), and joining (J) gene segments through a process termed **V(D)J recombination** [16,17]. V(D)J recombination is catalyzed by the RAG1/RAG2 complex that brings together distal gene segments, generates double-stranded breaks within specific recombination signal sequences that flank the gene segments to be recombined, and directs the post-cleavage complex to the nonhomologous end-joining machinery that repairs the DNA breaks while removing the sequence in between [18,19]. V(D)J recombination is tightly regulated in a lineage-specific but also a stage-specific fashion: D_H to J_H recombination happens early in the lymphocyte lineage; V_H is then joined to D_H–J_H to form the heavy chain segment in pro-B cells [20]. The assembly of the light chain occurs developmentally later (in pre-B cells), with V–J joining in either *IgLκ* or *IgLλ* loci. Ultimately, the heavy-chain transcript comprises the V(D)J exon, spliced together with the constant (C) region exons which are placed further downstream (Figure 1). Only one functional heavy chain gene is assembled (and one functional heavy chain transcript is produced) per B cell, a phenomenon known as 'allelic exclusion'.

The telomeric location appears to matter for the establishment of allelic exclusion. First, allelic exclusion does not operate at the level of *IgL* (which is not telomeric but rather chromosome-internal), raising the possibility that the genomic location (or the nuclear positioning of particular chromosomal regions) might enable silencing of alternate alleles. Indeed, during *IgH* locus allelic exclusion (monoallelic expression) [21], both *IgH* alleles relocalize inwards from the nuclear periphery and contract into rosette-like structures, presumably to allow for interactions between D_H–J_H with the great number of available V_H genes [22,23]. Both alleles undergo homologous pairing, and upon asynchronous DNA breakage at a single *IgH* allele, the second allele is relocalized to pericentromeric heterochromatin in an ATM-dependent fashion, presumably to avoid additional rearrangements [24]. Recombination with a given V_H at this stage requires a chromatin milieu that is characterized by open chromatin marks such as sensitivity to DNases, **CpG hypomethylation,** and **H3K9 acetylation** [25]. In contrast, the inactivated *IgH* allele presents as compact and with hypermethylated chromatin. Maintenance of Ig monoallelic

Glossary

Activation-induced cytidine deaminase (AID): an enzyme that converts cytidine to uracil; it is responsible for class switch recombination and somatic hypermutation.
B cells: lymphocytes that mature in the bone marrow; they belong to the adaptive immune system.
β-sandwich: a protein domain characterized by two opposing β-sheets.
β-strand: a type of secondary structure in proteins characterized by having a sheet form.
Bloodstream expression site (BES): a polycistronic subtelomeric locus from where VSGs can be transcribed.
Class-switch recombination (CSR): a maturation process of immunoglobulins which allows the transition from IgM to IgA, IgD, IgE, and IgG.
CpG hypomethylation: undermethylation of cytosine, within cytosine and guanidine dinucleotide regions in the DNA. Chromatin mark which, in gene promoters, leads to activation of gene expression.
DNase: an enzyme that cleaves and degrades DNA.
Epitope: the region of an antigen that is recognized and bound by an immunoglobulin.
Expression site-associated genes (ESAGs): genes present in BES; they are believed to be important for adaptation of the parasites in the mammalian host.
Expression site body (ESB): an extra-nucleolar compartment in the *T. brucei* nucleus which is enriched in RNA polymerase I and where the active BES is transcribed.
Germline transcript (GLT): noncoding RNA important for the CSR process.
H3K9 acetylation: acetylation of lysine 9 in histone H3. Chromatin mark characteristic of open chromatin and activation of gene expression.
Immunoglobulin (Ig): also known as antibody, a protein produced by B cells that is responsible for the recognition of antigens and the neutralization of pathogens.
***In situ* switching:** an epigenetic process used by *T. brucei* to change the expression of the VSG and,

expression in mature B cells is granted not only by the production of stable mRNA and protein, but also by compartmentalization of the active *IgH* and *IgLκ* alleles into different regions compared with the respective inactive alleles [26].

VSG and Ig genes: Similarities and Open Questions

Many similarities exist between the recombination processes that generate VSG and IgM. First, both BESs and *IgH* loci are telomeric, making them hotspots for recombination as well as locations amenable to transcriptional silencing (thus supporting monoallelic expression or allelic exclusion). Second, a substantial proportion of both gene archives is comprised of incomplete genes and pseudogenes [~80% in the case of VSG and ~30–40% in the case of the human (and mouse) *IgH*] [27]. In the context of *T. brucei* biology, pseudogenes are crucial to the generation of 'mosaics' which are new (functional) VSG recombinants [28–30] that do not pre-exist in the genome but are derived from segmental gene conversion events between the expressed VSG and other VSGs or pseudogenes. The appearance of mosaic VSG genes has been detected shortly before VSG protein expression, suggesting that the assembly of the new gene occurs directly in the active BES [31]. In addition, mosaicism could also occur through point mutations [32], although this point has been contested in the literature [33,34]. While definitive evidence that mutations contribute to VSG variability is currently lacking (and will likely need to wait for the generation of datasets of mosaics along the lines of [30]), the creation of point mutations within the surface-exposed loops has recently been shown to change antigenicity [35] and prompt immune evasion [36,37]. Similarly, while fine-tuning of Ig expression in rodent and primate B cells happens through **somatic hypermutation** (**SHM**, the accumulation of point mutations in the V(D)J exon, which is a central aspect of the adaptive immune system), birds and many farm animals (such as sheep) diversify their antibody genes by gene conversion [38,39]. Somatic hypermutation and gene conversion in the antibody system are both catalyzed by **activation-induced cytidine deaminase** (**AID**), an enzyme that converts cytidine into uracil to initiate error-prone DNA repair within the V(D)J exon [40]. AID catalyzes a third process on the Ig locus, termed **class switch recombination** (**CSR**) which, while not directly involved in the host response to VSG, still shares parallels with mechanisms that might be involved in VSG switching (see section below).

There are also gaps in knowledge in each process that could potentially be informed by progress in the other. For example, while it is clear that the RAG recombinase catalyzes V(D)J rearrangements in B cells, the switch trigger (the DNA break generator) remains a mystery in the context of VSG recombination (see Outstanding Questions). It is unclear whether an enzymatic activity exists that generates breaks (diverse models are reviewed in [41]), or whether the mechanism is simply stochastic. What is clear, however, is that sequences (repeats) surrounding the VSG are required for optimal recombination [42], such that the existence of a region- or structure-specific recombinase would not be impossible to imagine. Nonetheless, the immune system *per se* does not induce VSG switching [43], and chromatin state and transcriptional status appear to be important in the context of a VSG switching event [44,45]. An enzyme for catalyzing presumed mutations within the VSG has also not been identified, although it is possible that translesion repair polymerases, recruited to repair DNA breaks in the context of segmental gene conversion, can generate such mutations (A.Z. Zurita Leal, PhD thesis, University of Glasgow, 2017).

Overall, there are many parallels between the VSGs and Igs with regard to their regulation at the level of the DNA locus that recombines, the *cis*-acting elements required, and the reaction dynamics (Table 1). Of course, after each of these processes introduces a new gene into the

consequently, the VSG coat at the cell surface.

Microhomology-mediated end-joining (MMEJ): a pathway for the repair of double-stranded breaks; it is dependent on microhomologies flanking the break and is characterized by the appearance of deletions around the repair site.

Nonhomologous end-joining (NHEJ): a pathway for the repair of double-stranded breaks; it does not require a homologous template and typically uses the microhomologies of the overhangs created by the DNA break to ligate both ends.

Poly(A): polyadenylation, a process involved in maturation of messenger RNA characterized by the addition of adenosine monophosphates to the 3′UTR.

Poly(A) site (PAS): the location in the 3′UTR where poly(A) occurs.

Post-translational modification (PTM): modification of a peptide or protein product, following translation, by processes such as phosphorylation, acetylation, glycosylation, ubiquitination etc.

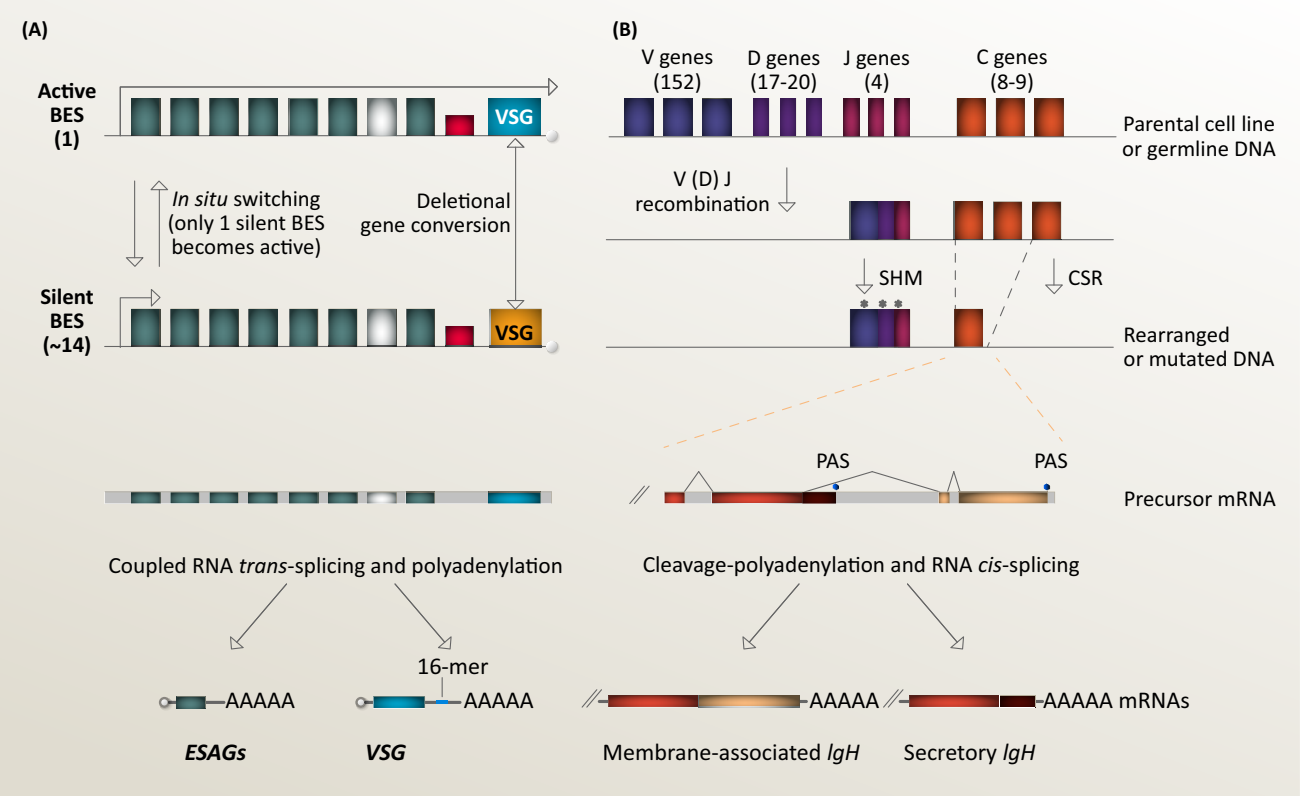
RNA-binding proteins (RBPs): proteins that bind double or single stranded RNA forming ribonucleoprotein complexes.

Somatic hypermutation (SHM): a maturation process of immunoglobulins characterized by gene mutations that lead to an increased binding affinity to the antigen.

Spliced leader RNA (SL RNA): a 39-nucleotide RNA sequence added to the 5′UTR of every messenger RNA in *T. brucei*.

V(D)J recombination: a process that generates a single immunoglobulin gene in B and T lymphocytes.

Variant surface glycoprotein (VSG): the major surface antigen of *T. brucei*.



Trends in Parasitology

Figure 1. Expression of the Variant Surface Glycoprotein (VSG) Antigen by *Trypanosoma brucei* and Immunoglobulins (Igs) by B Cells. (A) Expression of the VSG proteins by *T. brucei* from only one out of ~15 bloodstream expression sites (BESs) present in its genome. To continually evade the immune system, the parasites switch the expression of the VSG using one of the following mechanisms: silencing the active BES (with the blue VSG) and activating one silent BES (with yellow VSG) by *in situ* switching or replacing the VSG gene within the active BES with a VSG present in a silent BES or in the VSG archive present in internal arrays or minichromosomes (not represented) mainly by homologous recombination. Dark gray squares represent expression site-associated genes (ESAGs), red rectangle represents 70 bp repeats, white square represents pseudogenes, and the arrows represent the extent of transcription (long arrow for highly transcribed and short arrow for poorly transcribed). Genes are first transcribed by Pol I into polycistronic transcripts and then processed by coupled *trans*-splicing and polyadenylation mechanisms, giving rise to mature monocistronic mRNAs. The abundance of VSG mRNAs is much higher than all the ESAGs combined, even though the VSG and ESAGs are polycistronically transcribed. A conserved 16-mer sequence within the VSG 3′UTR is essential for mRNA stability. (B) Representation of the immunoglobulin heavy chain generation by B cells in mice. V(D)J recombination joins a single gene from each of the V (dark blue), D (purple), and J (pink) family genes, giving rise to a mature RNA that encodes a single antibody together with the first C (orange) genes that encode the antibody isotype (initially always Cμ that encodes an IgM). Upon recognition of an antigen, maturation of the antibody takes place by mutating the immunoglobulin coding sequence (asterisks) to improve binding affinity (somatic hypermutation, SHM) and by removing one or several initial C genes to change the antibody isotype (class-switch recombination, CSR). The IgLκ and IgLλ loci recombine similarly to the IgH except that these loci do not encode D gene segments; they also cannot undergo CSR. The IgH genes contain common structural features involved in alternative RNA processing to produce two mRNAs from a single primary transcript. During B cell differentiation, alternative poly(A) site (PAS) selection effects a switch of the heavy-chain expression from a membrane-bound form to the secreted form (modified after Peterson [100]). Orange rectangles: exons that are common to the two mRNAs; brown rectangles: the portion of the mRNA that remains in the transcript specific to secreted Ig; light brown boxes: exons that are found only in membrane-associated mRNA. Blue hexagons identify the two PASs. Note that the 5′end of IgH mRNA (V(D)J exon) is not shown.

genome, this new entity will have to be produced at the level of RNA and translated to generate the new VSG (or Ig) that will populate the coat of the parasite or B cell. While a fair amount has been gleaned for both of these systems at the gene level, much less is known about the stability, localization, and translation efficiency of the ensuing transcripts. This remains a large knowledge gap and a substantial opportunity for new biology to be discovered (see Outstanding Questions).

Table 1. Parallels between VSG and Ig at the DNA Level

Featurelocus	VSG	Ig (mouse)
Localization	VSGs in BES are telomeric [89]; remainder repertoire in internal megachromosomal arrays or minichromosomes [90]	*IgH* locus is telomeric; *IgLκ* locus is centromeric; *IgLλ* locus is internal [27]
Pseudogene content	80% of the entire VSG repertoire [1]	33–34% in *IgH* locus; 44–45% in *IgLκ* locus; 25–33% in *IgLλ* locus [27]
Allelic exclusion	Stringent except during a switching event [91]	Stringent in *IgH* locus [92]; allelic inclusion occurs in 1–7% in *IgLκ* locus [93,94]
Gene conversion pathway	Mainly homology recombination [12] and break-induced replication [13]	Mainly nonhomologous end-joining [95]
Role of epigenetics	Unlike silent BES, actively transcribed BES is devoid of nucleosomes [6,7] and centrally localized in the nucleus [3]. Several factors are important in maintenance of active and silent BES chromatin (reviewed in [96,97])	*IgH* alleles undergo locus contraction [22,23]; active *IgH* allele is sensitive to DNase, CpG hypomethylated and H3K9 acetylated [25]; active *IgH* and *IgLκ* alleles are differentially located to silent alleles [26]
Diversification type	Appearance of mosaic VSGs during an infection [31,37]	Class-switch recombination (CSR) to generate different antibody isotypes [98]; somatic hypermutation (SHM) to improve affinity of antibodies [99]
Initiating event for VSG switching/CSR and SHM	Unknown	Activation-induced cytidine deaminase (AID) [40]

Regulation of VSG and Ig Expression at RNA Level

The Role of RNA-binding Proteins in Trypanosomes

T. brucei protein-coding genes are arranged in long arrays that are transcribed constitutively by RNA polymerase II (Pol II) in a polycistronic manner. Individual mRNAs are resolved from their precursors by *trans*-splicing a 39 nt 5′ leader sequence, derived from an abundant capped RNA pol II-synthesized **spliced leader RNA(SL RNA)**, and by polyadenylation. As a consequence of this genomic organization, transcript levels are determined by copy number, modulated post-transcriptionally through the binding of **RNA-binding proteins(RBPs)** to *cis*-regulatory elements [46], and largely influenced by codon usage [47,48].

RBPs interact with their targets in a sequence- and secondary-structure-specific manner. In the absence of transcriptional control, the 3′UTR of developmentally regulated mRNAs are typically bound by specific RBPs, which then determine both the half-life and the translation efficiency of specific transcripts. The *T. brucei* mRNA-bound proteome (the 'RBPome') has been explored in cultured cells [49]. The potential function of many RBPs, when tethered to a transcript, has also been examined (e.g., whether it decreases/increases gene expression [49,50]). What remains unclear is the set of RBPs that is bound to (and determines the fate of) specific transcripts. This is an open question not only for *T. brucei*, but also in the field of RNA biology more generally.

his question is of particular interest when it comes to the VSG transcript, which is highly abundant and of major importance for the survival of the organism [51]. A single bloodstream form trypanosome cell contains about ~20 000 mRNA molecules [52], ~1400 of which encode the superabundant VSG [9]. This abundance (~7% of the total mRNA of the cell) is in part due to the fact that VSGs are transcribed by Pol I (which is normally reserved for pre-rRNA synthesis). Pol I transcription provides both a high rate of synthesis and enables rapid control of VSG transcripts [53,54]. These accumulate, as described above, within the ESB, a compartmentalized extra-nucleolar transcriptional factory [3]. Indeed, both ESB and Pol I transcription units are essential for functional expression of VSG [51]. However, mRNA abundance does not

directly translate to protein abundance: for example, ESAGs genes, all of which are transcribed with the same Pol I as the expressed VSG, are less abundant than VSG in terms of both mRNA [55] and protein [56]. Upon differentiation into the procyclic form, a rapid decline in the levels of VSG mRNA occurs due both to diminished synthesis capacity and loss of mRNA stability [57] – presumably due to the stage-specific loss of an RBP that stabilizes VSG mRNA or gain of an RBP that destabilizes it. It has been known for decades that the 3′UTR of VSG mRNA contains conserved elements that regulate the relative stage-specific abundance of the mRNA [58,59]. Mutation analyses of a conserved 16-mer element showed that it is essential for mRNA stability [51]; however, to date, the *trans*-acting factors (RBPs) involved in its binding and regulation remain unknown (see Outstanding Questions).

RBPs in B Cells

In comparison to the trypanosome system, the role of RBPs in the biology of B cells – especially as they might regulate the major transcript (the Ig transcript) – has remained almost entirely unexplored. A mature B cell contains about 100 molecules of μ heavy chain mRNA; however, upon differentiation into antibody-secreting plasma cells, the steady-state immunoglobulin mRNA level increases 10- to 100-fold (5000 to 30 000 molecules of heavy-chain message/cell) [60,61]. This apparent discrepancy is likely due to specific features of the cell type, isotype, and methodologies used in these classical experiments. Transcriptional profiling of different mouse B cell populations showed that up to 70% of total reads from bone-marrow-derived plasma cells map to immunoglobulin genes, about 30-fold higher than in peripheral B cells [62]. This high level is due to both increased transcription [63,64] and increased mRNA stability [65,66]. Upon stimulation, the half-lives of secreted mRNA forms increase, suggesting that sequence elements involved in differential mRNA stabilization are located in the 3′ end of the secreted form which are not present in the membrane counterpart [67,68]. As with the VSGs, the identity of the *trans*-acting factors is unknown (both generally, as there is no B-cell-specific RBPome described, and specifically, as very little is known about the RBPs that might bind and regulate the Ig transcript itself).

RNA Processing in Immune Cells

The little that is known regarding the regulation of Ig mRNA relates to its isoforms: membrane-bound vs secreted Ig, and IgM vs IgD. The former process is important at multiple steps of B cell differentiation: for example, during CSR (e.g., from IgM to IgG), switch isotypes can be found in the membrane-bound form (decorating the B cell surface) as well as in serum (produced by the B cells whose surface they decorate). Differentiation to plasmablasts (which are basically secreted Ig production factories) is also accompanied by a bias away from membrane-bound and toward secreted Ig. During these differentiation steps, the ratio of membrane-bound to secreted heavy-chain mRNA decreases, and that results in a shift in protein production from a membrane-bound form to a secreted form [69]. Mechanistically, this shift is due to alternative cleavage and polyadenylation [70,71]. Because mRNA 3′UTRs generally define mRNA cytoplasmic functions, including RNA stability, translatability, and localization, the use of alternative **poly(A)** sites (which lengthen or shorten the 3′UTR of a specific message, and thus likely alter its regulation) constitutes a key regulatory process in the expression of many molecules [72], including the Ig in B cells. Translational control also appears to contribute to the suppression of surface immunoglobulin expression in IgM-producing plasma cells [73,74]. Similarly, the production of IgM and IgD is subjected to transcription antitermination and alternative splicing of a precursor mRNA where Igμ heavy-chain constant region exons must be removed to produce IgD. Recent work identified the zinc finger protein Zfp318 as a repressor of transcriptional termination at the end of the μ exon series [75]. As transcripts cannot extend into the δ exons, the absence of Zfp318 results in the loss of IgD by leaving cells with precursor transcripts ending near the **poly(A) site(PAS)** for the upstream μ exons.

The Role of R-loops in Class-switch Recombination and VSG Switching

The AID-dependent CSR process is, as mentioned, not integral to the mechanism of Ig diversification with respect to the host response to trypanosomes. It is, however, a process that occurs at the Ig locus with mechanistic parallels to VSG switching. CSR is strictly dependent on noncoding transcription initiated from intronic promoters located upstream of each set of C_H exons. CSR long noncoding RNAs (lncRNAs) are termed **germline transcripts** (**GLTs**) and contain 1- to 10-kb-long sequences called switch (S) regions involved in the formation of R-loops [76]. These RNA:DNA hybrid structures result in non-template single-strand DNA (ssDNA) that acts as a substrate for AID to initiate CSR and is therefore thought to be required for optimal CSR.

R-loops may also have an important role in VSG antigenic variation [77]. As mentioned earlier, most VSGs are flanked upstream by repeats [the 70 bp repeats, a long array (3–20 kb) of a highly conserved AT-rich sequence]. Recently, have been shown that R-loops accumulate preferentially at this region of both the expressed and silent telomeric VSGs; that such structures can be resolved by the ribonuclease H1 (RNaseH1), whose downregulation enhances VSG antigenic variation, suggests a likely role for R-loops in modulating VSG switching [77]. As described for B cells [78], one of the hypotheses put forth in the context of VSG is that R-loop presence impedes DNA replication through the active VSG ES, leading to breaks that elicit a VSG switch. Interestingly, active VSG ES replicates earlier than all the silent ESs [79]; however, whether such connection between R-loops, transcription, and replication in the VSG ES relate to that described for CSR is still unknown (see Table 2 for a summary of parallels between VSGs and Igs at the RNA level).

In all, the complement of RBPs specific to B cells have not been defined, and their impact on RNA processing (e.g., via tethering as in [49,50]) has not been characterized. The very few RBPs that have been characterized in the Ig system relate to alternative splicing and polyadenylation, and their role in membrane versus secreted Ig production, or IgM versus IgD production. In addition, a detailed analysis of regulatory regions of the ~10 kB long Ig transcript has not been attempted, and regulatory elements that, in conjunction with RBPs, might define its half-life, localization, or translational efficiency remain more or less unknown. Future work that focuses on the regulation of the Ig transcript (following perhaps approaches pioneered in the *T. brucei* system) could revolutionize B cell biology but also mammalian biology more generally (see Outstanding Questions).

Parallels between Antibodies and VSGs at the Protein Level

Adaptive immunity produces antibodies, B cell, and T cell receptors that mediate recognition of foreign **epitopes**. Antigenic variation in the African trypanosome produces an ever-changing coat of epitopes to which the immune system responds. In addition to the many parallels at the genetic level for producing 'real-time' diversity in each of these systems that have been

Table 2. Parallels between VSG and Ig at the RNA Level

Transcription	Pol I	Pol II
RNA maturation	By constitutive splicing and polyadenylation of the active BES PTU	Alternative splicing and polyadenylation mechanisms regulate the synthesis of membrane-associated or secretory IgH. Many regulators characterized
mRNA stability	Requires conserved RNA elements in the 3′UTR; *trans*-acting factors unknown	Secretory IgH is generally more stable than membrane-associated. Although the 3′end regions are different, both *cis*- and *trans*-acting elements are unknown

discussed above, VSGs and antibodies possess many structural and functional similarities. All the analogies between these protein families center on an overall architectural concordance: both the VSGs and antibodies are the sum of a conserved scaffolding (core protein fold) with specific surfaces undergoing hypervariation to produce functionally distinct entities that interact with the environment for the benefit of the organism.

Antibodies are built up from multiple copies of the Ig family fold that consists of a **β-sandwich** with two Greek key topology sheets and a total of seven to nine antiparallel **β-strands**. The Ig fold typically spans 100–130 amino acids and is stabilized by an internal disulfide bond gluing the two sheets together covalently (in addition to a large hydrophobic interface). With two Ig folds in the light chain, and four in the heavy chain linked by further disulfides, this produces an elongated chain of around 75 kDa that dimerizes into a mature antibody of around 150 kDa, 10–15 nm in height depending on linker length and hinge angles between domains. The most conserved regions extend from the constant region of the heavy chain up to the penultimate Ig fold in both chains, the very tips of the last Ig domain containing the hypervariable surface patches at the distant end of the molecule (Figure 2A).

Similarly, the VSGs are also elongated, dimeric assemblies that end with a distal and divergent exposed surface. In this case the individual monomers are around 60 kDa and are centered not on the Ig fold but on an even more common and widespread core architecture, the three-helix bundle. A large protein superfamily, three-helix bundles are a stable yet adaptable scaffold on which to build diverse functions. The bundle helices themselves interdigitate to provide a tight fold with high stability and often rapid folding kinetics, especially for small polypeptides [80]. Specific, functional regions are often built into the connections between the helices, which can range from negligible in some proteins to elaborate subdomains in other proteins that serve as binding sites, docking regions, or even enzymes [81]. For the VSGs, the upper connections of the helices are the sites of extended insertions that form the 'top' of the molecule, specifically the surface exposed to the immune system in the densely packed VSG coat. This 'top' varies considerably between the VSG structures both in terms of the polypeptide fold itself and the molecular surfaces created (Figure 2B,C).

Both protein families are therefore constructed upon elongated, conserved protein cores that nevertheless accommodate a large amount of diversity, evident in hypervariable patches (Igs) or even extended hypervariable surfaces (VSGs). Indeed, when compared directly with antibodies, where diversity is established from a small set of loops, the VSGs must ratchet diversification up to encompass their entire surface or risk antibody recognition and clearance. This results in several observable traits that characterize VSGs as a whole, including (i) very low overall sequence identity (typically lower than 15%) despite the requirement for conservation in the overall architecture of the molecule for dense membrane packing, and (ii) divergent properties of the molecular surface (shape/topography, electrostatic charge, and hydrophobicity, Figure 2C).

Recent work has demonstrated that segmental gene conversion, which can increase the VSG protein 'space' by generating novel VSG genes, is most often the result of one N terminal VSG domain recombining with a new C terminal domain [30]. If we assume that antibody-exposed VSG epitopes belong exclusively to the surface-proximal N terminal domains (NTDs) of the VSGs [82], this presents a conundrum, suggesting that the same NTD epitopes are recognized differently after the acquisition of a new, antigenically shielded C terminus (a domain buried beneath the larger NTD, Figure 2B). At the same time, it has recently been shown that the presence of **post-translational modifications** (**PTMs**) within the NTD (e.g., *O*-linked

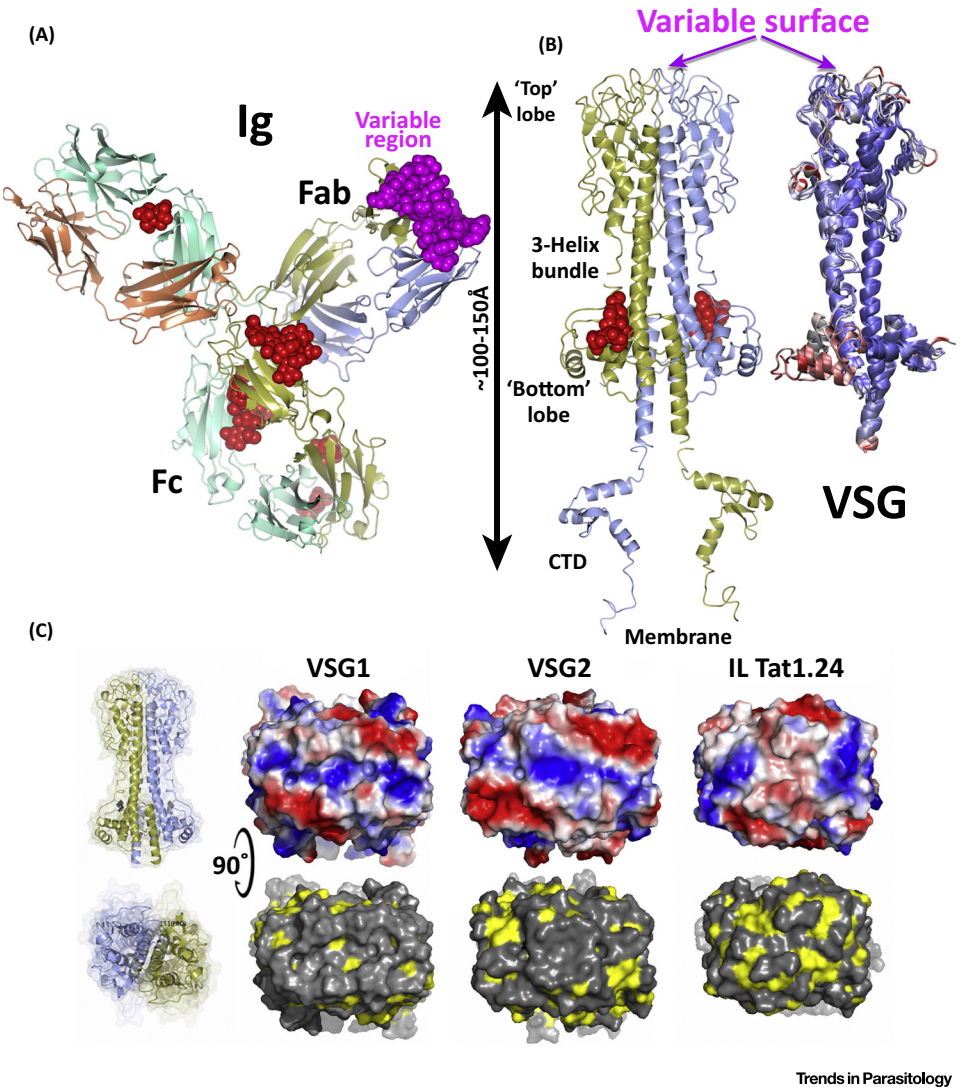


Figure 2. Structure–Function Parallels between Immunoglobulins (Igs) and Variant Surface Glycoproteins (VSGs). (A) Ribbon diagram illustrating the overall fold of an antibody (Pembrolizumab, a full length IGG4 antibody, PDB ID 5dk3 [101]). Each heavy and light chain is in a different color. The carbohydrates are shown as red space-filling spheres. The residues of the variable loops are shown as magenta spheres. The constant region (Fc) and one of the two Fab regions are labeled. (B) Left, ribbon diagram of VSG MiTat1.1 (combination of N terminal domain and C terminal domain structures as per [102], PDB IDs 5LY9 and 5M4T, respectively) with the two polypeptide chains of the dimer shown in different colors and carbohydrates shown as red space-filling spheres. Right, superposition of several VSG monomers (VSG1, VSG2, and ILTat1.24) colored by structural divergence (blue, highly similar; white, moderate differences; red, indicating large differences). (C) Membrane-distal surfaces (exposed to the immune system) of several VSG proteins shown as colored molecular surfaces (top: electrostatic, with red acidic/negative, blue basic/positive, white neutral; bottom: hydrophobicity, with yellow indicating a hydrophobic amino acid, gray indicating hydrophilic).

glycosylation) can alter the antigenicity of a sequence-identical NTD in comparison to the unmodified VSG [35]. Taken together, these discoveries suggest that antigenicity determined by the amino acid sequence of each VSG can be augmented by PTMs, and that the conundrum of N terminal and C terminal 'swaps' producing NTD antigenic differences could be related to the regulation of PTMs. It may even be possible that the substantially diverse, genomically

Key Figure

The Red Queen Paradigm: *Trypanosoma brucei* and B Cells

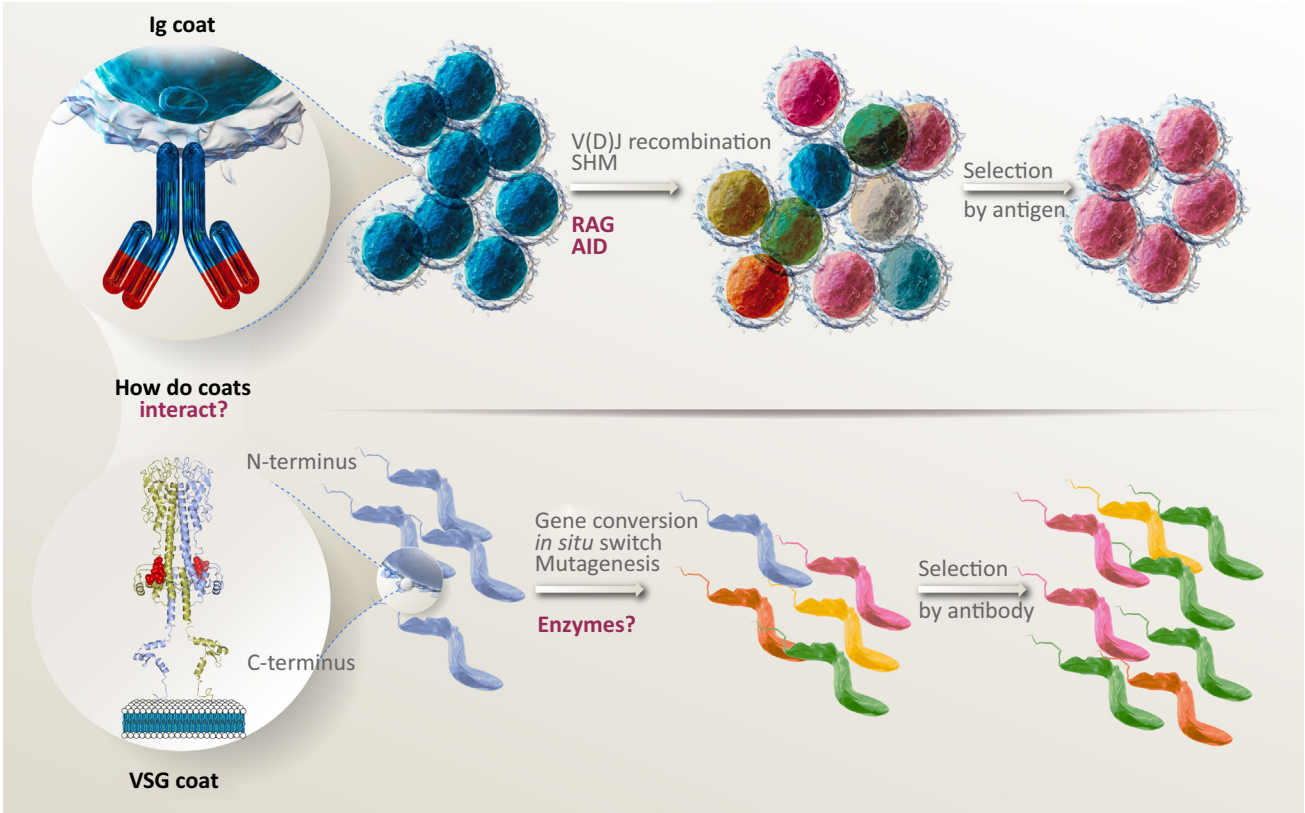


Figure 3. B cell diversification, shown here as the outcome of the combined process of V(D)J recombination and somatic hypermutation (SHM), occurs in parallel with *T. brucei* antigenic variation. Both repertoires diversify ostensibly to overtake one another. While the catalysts of B cell diversification (in the form of V(D)J recombination and SHM) are enzymatic (through RAG and activation-induced cytidine deaminase, AID), the process that catalyzes VSG switching remains unclear. Finally, both processes result in surface-bound receptors that must interact with one another – but very little is known currently as to how that interaction occurs.

encoded VSG repertoire is insufficient to explain the long-term evasion of host immunity, and that future models will need to incorporate the role of mosaics and PTMs as well as any possible relationships between them (see Outstanding Questions).

The study of antigen–immunoglobulin interactions, particularly at the structural level, has revolutionized the study of immunity and the approaches to immunotherapeutic interventions in diseases as diverse as HIV/AIDS, malaria, and cancer [83]. While there now exist growing libraries of antibody–antigen structure–function interactions in many diseases, with many of these augmented by results from B cell repertoire analyses, there is still not a single such structure of an antibody–VSG complex, nor even a firmly established antigenic epitope for any VSG coat protein. As the contact surface between antibodies and the VSG proteins can be argued to be the central host–pathogen interaction in trypanosomal infection, this paucity of

data is even more glaring. Therefore, additional structural work on VSGs, and especially VSG–antibody complexes, is necessary and timely. Such work will further allow understanding of the similarities and differences within the diverse regions of the VSG proteomic space, and would bring the study of this unique pathogen up to speed with the immunological progress in other areas of infectious disease (see Outstanding Questions).

Concluding Remarks

> Well, in our country, said Alice, still panting a little, you'd generally get to somewhere else – if you run very fast for a long time, as we've been doing.
>
> A slow sort of country! said the Queen. Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!
>
> Lewis Carroll: Through the Looking-Glass and What Alice Found There, Chapter 2

The Red Queen hypothesis states that organisms in competition suffer a relentless pressure to innovate and change, or plunge into extinction. While considerable debate has raged about the generality of this model, in certain contexts in can be better argued that it would apply. One such case is the battle between host and pathogen, and in particular, between the adaptive immune system and antigenic variation. In this context, that the 'real-time' somatic adaptability of immunity and antigenic variation would share many parallels is perhaps not surprising (Figure 3, Key Figure). 'Arms races' could lead to the employment of measures and counter-measures that harbor analogous characteristics.

Indeed, antigenic variation and antibody diversification processes share remarkable mechanistic similarities. Both VSGs and Igs coat their respective cellular surfaces at very high densities: each parasite is densely coated with $\sim 10^7$ VSG molecules [84], and each B cell can carry roughly 1.2×10^5 immunoglobulin molecules on its surface (by far the most abundant receptor present) [85]. Both genic loci diversify through DNA recombination mechanisms, and both generate abundant mRNA that must be efficiently translated to eventually populate the surface of each cell type. The end protein product of these amazing processes consists of highly stable, disulfide cross-linked, elongated molecules built around conserved protein architectures with a distal end harboring surfaces of variation that function in the host–pathogen competitive interaction. Given these similarities, advances in one area can quickly be translated to advances in the other (and there have been examples of such cross-disciplinary work in the literature; [13,86]).

Both VSG and Ig proteins have a long half-life (~32 h and ~12 h respectively). Effectively, this means that there is a substantial lag between the time that recombination at the genomic region coding for each surface protein is completed, and the time that the new VSG (or new Ig) populates the surface of the respective cell [87]. It also means that, for a significant amount of time, the old and the new versions of these surface proteins likely coexist, generating novel (chimeric) epitopes (as has been hypothesized for VSG – [88]), or generating naturally bispecific Ig receptors by combining two heavy chains that are nearly identical, save from single amino acid substitutions generated by somatic hypermutation. This could be an underappreciated mechanism of immune evasion in the context of the trypanosome, and of affinity maturation in the context of the Ig. Ultimately, a closer focus within this area of host–parasite interaction, that is defined by the evolving VSG–Ig surface, is bound to generate substantial advances for the biology of trypanosomes and also for the biology of B cells (see Outstanding Questions).

Outstanding Questions

What are the main players enforcing monoallelic expression of the VSGs in *T. brucei*? What mechanisms are involved? How is the donor sequence during switching selected? How does recombination occur?

What are the *trans*-acting factors that bind and regulate VSG and Ig RNA transcripts? For VSG, how do these factors recognize different mRNAs? For Ig, how do post-transcriptional regulators impact B cell development in general and with regard to the Ig transcript in particular? What are the main regulatory regions of Ig mRNA?

How are different VSG epitopes recognized?

Is there a pattern or code associated with VSGs post-translational modifications (PTMs)? How are antibody specificity, affinity, and avidity affected by VSG PTMs? How might the contribution of individual modifications affect immune response dynamics?

Acknowledgments

This work was supported by funds from the German Cancer Research Center (DKFZ, Heidelberg) and the Helmholtz Foundation.

References

1. Cross, G.A. *et al.* (2014) Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.* 195, 59–73
2. Gunzl, A. *et al.* (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot. Cell* 2, 542–551
3. Navarro, M. and Gull, K. (2001) A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature* 414, 759–763
4. Hertz-Fowler, C. *et al.* (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS One* 3, e3527
5. Dreesen, O. *et al.* (2007) Telomere structure and function in trypanosomes: a proposal. *Nat. Rev. Microbiol.* 5, 70–75
6. Stanne, T.M. and Rudenko, G. (2010) Active VSG expression sites in *Trypanosoma brucei* are depleted of nucleosomes. *Eukaryot. Cell* 9, 136–147
7. Figueiredo, L.M. and Cross, G.A. (2010) Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. *Eukaryot. Cell* 9, 148–154
8. Chaves, I. *et al.* (1998) Subnuclear localization of the active variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. U. S. A.* 95, 12328–12333
9. Glover, L. *et al.* (2016) VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes. *Proc. Natl. Acad. Sci. U. S. A.* 113, 7225–7230
10. Cordon-Obras, C. *et al.* (2013) *Trypanosoma brucei* gambiense adaptation to different mammalian sera is associated with VSG expression site plasticity. *PLoS One* 8, e85072
11. McCulloch, R. *et al.* (2015) DNA recombination strategies during antigenic variation in the African trypanosome. *Microbiol. Spectr.* 3, MDNA3-0016-2014
12. McCulloch, R. and Barry, J.D. (1999) A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation. *Genes Dev.* 13, 2875–2888
13. Boothroyd, C.E. *et al.* (2009) A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature* 459, 278–281
14. Glover, L. *et al.* (2008) Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes. *Nucleic Acids Res.* 36, 2608–2618
15. Reinitz, D.M. and Mansfield, J.M. (1990) 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
16. Mombaerts, P. *et al.* (1992) RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* 68, 869–877
17. Shinkai, Y. *et al.* (1992) RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 68, 855–867
18. Sakano, H. *et al.* (1981) Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. *Nature* 290, 562–565
19. Lee, G.S. *et al.* (2004) RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* 117, 171–184
20. Alt, F.W. *et al.* (1984) Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.* 3, 1209–1219
21. Jung, D. *et al.* (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* 24, 541–570
22. Kosak, S.T. *et al.* (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 296, 158–162
23. Jhunjhunwala, S. *et al.* (2008) The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell* 133, 265–279
24. Hewitt, S.L. *et al.* (2009) RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nat. Immunol.* 10, 655–664
25. Cobb, R.M. *et al.* (2006) Accessibility control of V(D)J recombination. *Adv. Immunol.* 91, 45–109
26. Skok, J.A. *et al.* (2001) Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nat. Immunol.* 2, 848–854
27. Lefranc, M.P. *et al.* (2004) IMGT-ONTOLOGY for immunogenetics and immunoinformatics. *In Silico Biol.* 4, 17–29
28. 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
29. Thon, G. *et al.* (1990) Trypanosome variable surface glycoproteins: composite genes and order of expression. *Genes Dev.* 4, 1374–1383
30. Jayaraman, S. *et al.* (2018) Application of long read sequencing to determine expressed antigen diversity in *Trypanosoma brucei* infections. *bioRxiv* Published online October 2, 2018. http://dx.doi.org/10.1101/432245
31. Mugnier, M.R. *et al.* (2015) The *in vivo* dynamics of antigenic variation in *Trypanosoma brucei*. *Science* 347, 1470–1473
32. Lu, Y. *et al.* (1993) Point mutations are associated with a gene duplication leading to the bloodstream reexpression of a trypanosome metacyclic VSG. *Cell* 72, 397–406
33. Graham, V.S. and Barry, J.D. (1996) Is point mutagenesis a mechanism for antigenic variation in *Trypanosoma brucei*? *Mol. Biochem. Parasitol.* 79, 35–45
34. Gjini, E. *et al.* (2012) The impact of mutation and gene conversion on the local diversification of antigen genes in African trypanosomes. *Mol. Biol. Evol.* 29, 3321–3331
35. Pinger, J. *et al.* (2018) African trypanosomes evade immune clearance by O-glycosylation of the VSG surface coat. *Nat. Microbiol.* 3, 932–938
36. Kamper, S.M. and Barbet, A.F. (1992) Surface epitope variation via mosaic gene formation is potential key to long-term survival of *Trypanosoma brucei*. *Mol. Biochem. Parasitol.* 53, 33–44
37. Hall, J.P. *et al.* (2013) Mosaic VSGs and the scale of *Trypanosoma brucei* antigenic variation. *PLoS Pathog.* 9, e1003502
38. Butler, J.E. (1998) Immunoglobulin diversity, B cell and antibody repertoire development in large farm animals. *Rev. Sci. Tech.* 17, 43–70
39. Arakawa, H. and Buerstedde, J.M. (2004) Immunoglobulin gene conversion: insights from bursal B cells and the DT40 cell line. *Dev. Dyn.* 229, 458–464
40. Muramatsu, M. *et al.* (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553–563
41. da Silva, M.S. *et al.* (2018) Evaluation of mechanisms that may generate DNA lesions triggering antigenic variation in African trypanosomes. *PLoS Pathog.* 14, e1007321
42. 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
43. Myler, P.J. *et al.* (1985) Antigenic variation in clones of *Trypanosoma brucei* grown in immune-deficient mice. *Infect. Immun.* 47, 684–690
44. Aresta-Branco, F. *et al.* (2016) A transcription-independent epigenetic mechanism is associated with antigenic switching in *Trypanosoma brucei*. *Nucleic Acids Res.* 44, 3131–3346

45. Batram, C. *et al.* (2014) Expression site attenuation mechanistically links antigenic variation and development in *Trypanosoma brucei*. *eLife* 3, e02324
46. Clayton, C. (2013) The regulation of trypanosome gene expression by RNA-binding proteins. *PLoS Pathog.* 9, e1003680
47. de Freitas Nascimento, J. *et al.* (2018) Codon choice directs constitutive mRNA levels in trypanosomes. *eLife* 7, e32467
48. Jeacock, L. *et al.* (2018) Codon usage bias controls mRNA and protein abundance in trypanosomatids. *eLife* 7, e32496
49. Lueong, S. *et al.* (2016) Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. *Mol. Microbiol.* 100, 457–471
50. Erben, E.D. *et al.* (2014) A genome-wide tethering screen reveals novel potential post-transcriptional regulators in *Trypanosoma brucei*. *PLoS Pathog.* 10, e1004178
51. Ridewood, S. *et al.* (2017) The role of genomic location and flanking 3′UTR in the generation of functional levels of variant surface glycoprotein in *Trypanosoma brucei*. *Mol. Microbiol.* 106, 614–634
52. Haanstra, J.R. *et al.* (2008) Control and regulation of gene expression: quantitative analysis of the expression of phosphoglycerate kinase in bloodstream form *Trypanosoma brucei*. *J. Biol. Chem.* 283, 2495–2507
53. Rudenko, G. *et al.* (1991) RNA polymerase I can mediate expression of CAT and neo protein-coding genes in *Trypanosoma brucei*. *EMBO J.* 10, 3387–3397
54. Zomerdijk, J.C. *et al.* (1991) Efficient production of functional mRNA mediated by RNA polymerase I in *Trypanosoma brucei*. *Nature* 353, 772–775
55. Cully, D.F. *et al.* (1985) Coordinate transcription of variant surface glycoprotein genes and an expression site associated gene family in *Trypanosoma brucei*. *Cell* 42, 173–182
56. Gadelha, C. *et al.* (2015) Architecture of a host–parasite interface: complex targeting mechanisms revealed through proteomics. *Mol. Cell. Proteom.* 14, 1911–1926
57. Ehlers, B. *et al.* (1987) RNA turnover in *Trypanosoma brucei*. *Mol. Cell. Biol.* 7, 1242–1249
58. Majumder, H.K. *et al.* (1981) Homologous 3′-terminal regions of mRNAs for surface antigens of different antigenic variants of *Trypanosoma brucei*. *Nucleic Acids Res.* 9, 4745–4753
59. Berberof, M. *et al.* (1995) The 3′-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. *EMBO J.* 14, 2925–2934
60. Schibler, U. *et al.* (1978) The synthesis and processing of the messenger RNAs specifying heavy and light chain immunoglobulins in MPC-11 cells. *Cell* 15, 1495–1509
61. Perry, R.P. and Kelley, D.E. (1979) Immunoglobulin messenger RNAs in murine cell lines that have characteristics of immature B lymphocytes. *Cell* 18, 1333–1339
62. Shi, W. *et al.* (2015) Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. *Nat. Immunol.* 16, 663–673
63. Yuan, D. *et al.* (1992) Mechanism of lipopolysaccharide-mediated transcriptional enhancement of the mu gene. *J. Immunol.* 148, 3652–3657
64. Webb, C.F. *et al.* (1991) Novel protein-DNA interactions associated with increased immunoglobulin transcription in response to antigen plus interleukin-5. *Mol. Cell. Biol.* 11, 5197–5205
65. Cox, A. and Emtage, J.S. (1989) A 6-fold difference in the half-life of immunoglobulin mu heavy chain mRNA in cell lines representing two stages of B cell differentiation. *Nucleic Acids Res.* 17, 10439–10454
66. Genovese, C. and Milcarek, C. (1990) Increased half-life of mu immunoglobulin mRNA during mouse B cell development increases its abundancy. *Mol. Immunol.* 27, 733–743
67. Eckmann, L. *et al.* (1994) Increased transcription and coordinate stabilization of mRNAs for secreted immunoglobulin alpha heavy chain and kappa light chain following stimulation of immunoglobulin A expressing B cells. *J. Biol. Chem.* 269, 33102–33108
68. Berberich, I. and Schimpl, A. (1990) Regulation of Ig gene expression in normal lymphocytes. I. The half-life of secreted mu chain mRNA differs from that of membrane mu chain mRNA in resting and activated B cells. *Eur. J. Immunol.* 20, 445–448
69. Tyler, B.M. *et al.* (1982) mRNA for surface immunoglobulin gamma chains encodes a highly conserved transmembrane sequence and a 28-residue intracellular domain. *Proc. Natl. Acad. Sci. U. S. A.* 79, 2008–2012
70. Galli, G. *et al.* (1987) Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted mu-chain expression during B cell development. *Genes Dev.* 1, 471–481
71. Takagaki, Y. *et al.* (1996) The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* 87, 941–952
72. Singh, G. *et al.* (2015) The clothes make the mRNA: past and present trends in mRNP fashion. *Annu. Rev. Biochem.* 84, 325–354
73. Sitia, R. *et al.* (1987) Regulation of membrane IgM expression in secretory B cells: translational and post-translational events. *EMBO J.* 6, 3969–3977
74. Stockdale, A.M. *et al.* (1987) The expression of membrane and secreted immunoglobulin during the *in vitro* differentiation of the murine B cell lymphoma CH12. *J. Immunol.* 139, 3527–3535
75. Pioli, P.D. *et al.* (2014) Zfp318 regulates IgD expression by abrogating transcription termination within the Ighm/Ighd locus. *J. Immunol.* 193, 2546–2553
76. Reaban, M.E. and Griffin, J.A. (1990) Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature* 348, 342–344
77. Briggs, E. *et al.* (2018) Ribonuclease H1-targeted R-loops in surface antigen gene expression sites can direct trypanosome immune evasion. *PLoS Genet.* 14, e1007729
78. Wiedemann, E.M. *et al.* (2016) DNA replication origins in immunoglobulin switch regions regulate class switch recombination in an R-loop-dependent manner. *Cell Rep.* 17, 2927–2942
79. Devlin, R. *et al.* (2016) Mapping replication dynamics in *Trypanosoma brucei* reveals a link with telomere transcription and antigenic variation. *eLife* 5, e12765
80. Wickstrom, L. *et al.* (2006) The unfolded state of the villin headpiece helical subdomain: computational studies of the role of locally stabilized structure. *J. Mol. Biol.* 360, 1094–1107
81. Schneider, J.P. *et al.* (1998) Analysis and design of three-stranded coiled coils and three-helix bundles. *Fold. Des.* 3, R29–R40
82. Schwede, A. *et al.* (2011) The VSG C-terminal domain is inaccessible to antibodies on live trypanosomes. *Mol. Biochem. Parasitol.* 175, 201–204
83. Imkeller, K. and Wardemann, H. (2018) Assessing human B cell repertoire diversity and convergence. *Immunol. Rev.* 284, 51–66
84. Cross, G.A. (1975) Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology* 71, 393–417
85. Fiala, G.J. *et al.* (2013) Pre-clustering of the B cell antigen receptor demonstrated by mathematically extended electron microscopy. *Front. Immunol.* 4, 427
86. Zarrin, A.A. *et al.* (2007) Antibody class switching mediated by yeast endonuclease-generated DNA breaks. *Science* 315, 377–381
87. Pinger, J. *et al.* (2017) Variant surface glycoprotein density defines an immune evasion threshold for African trypanosomes undergoing antigenic variation. *Nat. Commun.* 8, 828
88. Dubois, M.E. *et al.* (2005) Trypanosomes expressing a mosaic variant surface glycoprotein coat escape early detection by the immune system. *Infect. Immun.* 73, 2690–2697

89. De Lange, T. and Borst, P. (1982) Genomic environment of the expression-linked extra copies of genes for surface antigens of *Trypanosoma brucei* resembles the end of a chromosome. *Nature* 299, 451–453
90. Williams, R.O. *et al.* (1982) Genomic environment of *T. brucei* VSG genes: presence of a minichromosome. *Nature* 299, 417–421
91. Chaves, I. *et al.* (1999) Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO J.* 18, 4846–4855
92. ten Boekel, E. *et al.* (1998) Precursor B cells showing H chain allelic inclusion display allelic exclusion at the level of pre-B cell receptor surface expression. *Immunity* 8, 199–207
93. Casellas, R. *et al.* (2007) Igkappa allelic inclusion is a consequence of receptor editing. *J. Exp. Med.* 204, 153–160
94. Velez, M.G. *et al.* (2007) Ig allotypic inclusion does not prevent B cell development or response. *J. Immunol.* 179, 1049–1057
95. Roth, D.B. (2003) Restraining the V(D)J recombinase. *Nat. Rev. Immunol.* 3, 656–666
96. Gunzl, A. *et al.* (2015) Mono-allelic VSG expression by RNA polymerase I in *Trypanosoma brucei*: expression site control from both ends? *Gene* 556, 68–73
97. Pena, A.C. *et al.* (2017) Epigenetic regulation in *T. brucei*: changing coats is a chance to survive. In *Epigenetics of Infectious Diseases* (Doerfler, W. and Casadesús, J., eds), pp. 221–241, Springer International Publishing
98. Liu, Y.J. *et al.* (1996) Within germinal centers, isotype switching of immunoglobulin genes occurs after the onset of somatic mutation. *Immunity* 4, 241–250
99. Jacob, J. *et al.* (1991) Intraclonal generation of antibody mutants in germinal centres. *Nature* 354, 389–392
100. Peterson, M.L. (2011) Immunoglobulin heavy chain gene regulation through polyadenylation and splicing competition. *Wiley Interdiscip Rev. RNA* 2, 92–105
101. Scapin, G. *et al.* (2015) Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab. *Nat. Struct. Mol. Biol.* 22, 953–958
102. Bartossek, T. *et al.* (2017) Structural basis for the shielding function of the dynamic trypanosome variant surface glycoprotein coat. *Nat. Microbiol.* 2, 1523–1532