Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation

Katrina A. Lythgoe*, Liam J. Morrison*, Andrew F. Read*, and J. David Barry**

*Institutes of Evolution, Immunology, and Infection Research, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom; and †Wellcome Centre for Molecular Parasitology, Glasgow Biomedical Research Centre, University of Glasgow, 120 University Place, Glasgow G12 8TA, United Kingdom

Edited by Robert May, University of Oxford, Oxford, United Kingdom, and approved March 21, 2007 (received for review July 21, 2006)

Pathogens often persist during infection because of antigenic variation in which they evade immunity by switching between distinct surface antigen variants. A central question is how ordered appearance of variants, an important determinant of chronicity, is achieved. Theories suggest that it results directly from a complex pattern of transition connectivity between variants or indirectly from effects such as immune cross-reactivity or differential variant growth rates. Using a mathematical model based only on known infection variables, we show that order in trypanosome infections can be explained more parsimoniously by a simpler combination of two key parasite-intrinsic factors: differential activation rates of parasite variant surface glycoprotein (VSG) genes and density-dependent parasite differentiation. The model outcomes concur with empirical evidence that several variants are expressed simultaneously and that parasitaemia peaks correlate with VSG genes within distinct activation probability groups. Our findings provide a possible explanation for the enormity of the recently sequenced VSG silent archive and have important implications for field transmission.

mathematical model | Trypanosoma brucei | variant surface glycoprotein | switching | hierarchy

Trypanosomes switch antigens primarily by duplication of VSGs, in which the new gene copy replaces the expressed VSG. There is a large archive of 1,500–2,000 silent VSGs from which to duplicate (1). The VSG group activated with highest frequency is telomere-proximal, and it is duplicated by way of homologous gene flanks. The second main group, located in subtelomeric arrays, is activated with relatively low frequency. This group is mostly composed of pseudogenes, gene fragments, or genes encoding dysfunctional variant surface glycoproteins (VSGs) (1) but can be expressed as mosaic genes through donation of segments of coding sequence from different VSGs (2). A fundamental difference between these two archive groups is that VSGs in the first switch independently of each other, duplicating by way of gene flanks, whereas the second group does not switch independently, because its duplication involves coding sequence homologies.

Order in VSG expression could be imposed by the parasite or by indirect influences. The most commonly proposed molecular basis for order (3–5) is that it operates through innate differences in the activation probability for VSGs, dependent on the flanking sequence homologies of the silent gene being activated or, for mosaic genes, degrees of homology between coding sequences of the incoming and outgoing VSG genes. Indeed, the bacterium Anaplasma marginale, using an archiviral system, achieves differential switching by way of homologies in both gene flank and coding sequences (6). There is empirical support for this general trypanosome molecular mechanism (7–14), and recent quantitative analysis has revealed a higher degree of order than previously supposed (15). Three studies in particular, together involving >100 variants in >50 infections, have firmly established order in VSG expression (15–17). Most studies, however, have involved only single relapses and therefore only a few high-probability switches, and it is important now to analyze long-term infection in greater detail to help ascertain the range and patterns of switch rates that can yield order and how ordered expression interacts with trypanosome growth variables to yield the characteristic fluctuating parasitaemia. One way to do so is to use mathematical modeling on the basis of known, major variables in trypanosome growth and antigenic variation.

In contrast to the molecular hypothesis, most theoretical modeling studies have assumed a uniform switching pattern among variants (Fig. 1A) and invoked indirect effects to explain order. There is, however, little empirical support for these other factors. It has been proposed that order arises from variant-specific differences in intrinsic growth rates (18), but such an effect is, at best, insignificant (18, 19) and is not supported theoretically (20, 21). Another possibility is that order arises from differential efficiencies of variant-specific immune responses, is negated by the observation that different variants are cleared in vivo by the immune system with closely similar dynamics (22). The suggestion that transient expression of two VSGs by a trypanosome, such as during switching, could enhance sensitivity to antibodies and therefore influence order (21) is not supported by the normal in vivo growth and reduced immunogenicity of artificially created double expressors (23, 24). Antigenic “cross-reactivity” on the basis of common, invariant antigens (25, 26) is unlikely to cause order, because such molecules do not elicit protective responses against trypanosomes (27–29). Order generated through VSGs sharing epitopes, corresponding to what has been proposed for PFEMP1 switching in the malaria parasite Plasmodium falciparum (30), is highly unlikely, because the degree of order observed would require multiple VSG cross-reactions, but these have been seen only for products of highly related VSG genes, and in all cases cross-protection was complete (13, 31). The cross-reactivity model proposed for malaria (30) depends on transient antibody responses, but, within the limits of available tests, they have not been detected in studies of trypanosome infections (15–17, 32).

In light of the insights into the make-up of the VSG repertoire provided by the genome project (1) and, in particular, the emerging possibility that mosaic gene formation from pseudogenes makes such substantial contribution to antigenic variation, there is a requirement to revisit and substantially revise the existing modeling approaches to trypanosome VSG switching and, in particular, to include differential VSG activation rates analogous to those seen in vivo. Moreover, it is important to involve density-dependent parasite differentiation, a key growth factor.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviation: VSG, variant surface glycoprotein.

*To whom correspondence should be addressed. E-mail: j.d.barry@bio.gla.ac.uk.

© 2007 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0606206104

PNAS | May 8, 2007 | vol. 104 | no. 19 | 8095–8100
parameter (33, 34) not previously incorporated into antigenic variation models.

Of existing mathematical models, only one has addressed differential VSG activation rates, and it successfully simulated ordered expression (35). The model was based, however, on interdependent switching between variants, in which the probability of switching on a particular variant (activation probability) depends on the currently expressed variant (Fig. 1B) and is associated with the direction of the switch, such that the probability of switching from variant i to variant j will not equal the probability of switching from variant j to variant i. This assumption is not supported by empirical evidence that early switches are governed by the VSG being activated, rather than what is currently expressed (15), and that order early in infections is independent of the infecting variant (15–17, 22, 36).

Given these data, a more plausible model is a differential-activation switching pattern during the early stage of infection (Fig. 1C), in which each variant has a distinct activation rate, independent of the currently expressed variant. This pattern changes later in infection, when mosaics of VSGs are expressed, because the involvement of coding sequence homology in switching means that the variant being switched to does depend on the variant currently being expressed (Fig. 1D).

Trypanosome survival and growth in mammalian hosts and the interplay between infection profile and antigenic variation are complex, and it is important to incorporate variables that have been demonstrated to influence parasitemia and to exclude those of little or no biological relevance. Factors known to exert significant influence on the course of infection include the density-dependent, parasite-triggered differentiation to the non-dividing short stumpy stage (33, 34), density-dependent general growth interference between trypanosomes (37) (which is likely to include any indirect immune system effects, such as reported in model host infections; ref. 38), and the VSG-specific antibody responses. All these have been quantified experimentally, permitting inclusion in mathematical modeling.

Factors thought not to impact on the generation of ordered antigenic variation include dispersed anatomical location of trypanosomes within the host (39), infection-associated immunosuppression (40), and putative, direct-host involvement in the switching mechanism (17, 41). Ordered antigenic variation occurs in the bloodstream without complication from parasites in other anatomical niches. Trypanosoma vivax is mainly vascular yet displays order (17), and direct analysis of the distribution and exchange of Trypanosoma brucei variants between vascular and extravascular sites by using VSG-specific antibodies (39) has refuted the proposal from indirect studies that extravascular colonies seed the variant pool in the blood (42). It is clear from empirical analyses of >100 variants in >50 host animals that antibody responses against individual variants generally are very persistent, usually lasting for as long as infections are followed (15–17). These antibodies are trypanolytic and, consequently, once a variant has arisen it does not reappear in long-term infection (15–17). Hence, immunosuppressive effects reported in model hosts (40) are not relevant to antigenic variation, certainly up to the terminal stages of infection. There is another VSG-triggered immune response, which activates macrophages to yield trypanocidal agents (38), but it is not known whether this response kills trypanosomes variant-specifically. Because this response is mainly extravascular (43), and order is not influenced by trypanosomes in that compartment (17, 39), it is unlikely to

![Fig. 1. Different suggested switching patterns. The green, red, and blue shapes represent trypanosomes expressing different VSG genes. The values u, v, w, x, y, and z represent switching rates. (A) Uniform switching pattern. All variants are switched to and from at the same rate. (B) Interdependent switching pattern associated with direction of switch. The rate to which a variant is switched depends on the variant from which it is switching. The rate depends on the direction of the switch so that, for example, the rate at which red switches to blue is not the same as the rate at which blue switches to red. This is the switching pattern used by Frank (35). (C) Differential activation switching pattern. Each variant has a distinct activation rate independent of the variant it is replacing so that, for example, the rate at which red switches to blue is equal to the rate at which green switches to blue. (D) Interdependent switching not associated with switch direction, as occurs in homology-dependent (mosaic) switching. The rate to which a variant is switched depends on the variant from which it was switched. The rate does not depend on the direction of switching so that, for example, the rate at which red switches to blue is equal to the rate at which blue switches to red.](https://www.pnas.org/doi/10.1073/pnas.0606206104)
be of major influence on order. Although the antitrypanosome immune response is complex (38), the dynamics of antibody production and variant killing strongly correlate, and it is empirical data regarding antibody thresholds on which we base the model. Hosts are not directly involved in switching, which is parasite-intrinsic (41) and independent of host species (17); the host acts solely to select new variants by killing old ones.

Here, we have developed a mathematical model incorporating all of the influential parameters listed above. To examine how switch rates and patterns influence order and how order influences infection profile, initially we tested whether the differential-activation switching pattern (Fig. 1C) can yield the main elements of realistic infection profiles, then we extended our model to test the homology switching pattern for mosaic genes, hitherto unincorporated in any trypanosome model (Fig. 1D), but likely to be a major influence in chronic infections. Because other parameters have been quantified empirically, the only one varied in our model was the switch rate between variants. Hence, our model has the benefit of being more parsimonious and biologically more realistic than other models.

We imagined a population of trypanosomes in which the number of replicative (slender) and nonreplicative (stumpy, cell cycle arrested) trypanosomes that are variant $i$ are given by $v_i$ and $m$, respectively. The slender cells have an intrinsic growth rate $r$, where $t$ is time after inoculation and, when a slender cell divides, there is a probability, $f$, that one daughter cell will differentiate to a stumpy cell. This rate of differentiation increases as the total number of parasites increases (33, 34). The slender cells are also able to switch VSG. We assumed that all variants switch off at the same rate, but there are differences in the activation rate of different VSGs. This switching rate we call $s_i$, which is the rate at which variant $i$ is switched to from variant $j$. As the host acquires immunity to variant $i$, the slender and stumpy cells will be killed at a rate dependent on the strength of the acquired immune response, $a_i$. The value of $a_i$ can range from 0 to 1, where 0 means no immune response and 1 means the maximum possible immune response. The maximum rates at which the immune system can kill the slender and stumpy cells are $d$ and $\delta$, respectively. It was computationally unfeasible to process the hundreds or more of possible variants, so in all simulations we included a VSG repertoire of 30. Although this clearly will abbreviate simulated infections, it will not influence what is being tested, the patterns of variants in peaks. The simulations were all initiated with a single variant, because that was the basis of almost all published empirical analyses of infections.

The dynamics of the slender and stumpy cells are given by the following:

$$\frac{dv_i}{dt} = v_i r_i (1 - f) - v_i m_i + \sum_{j=1}^{n} (v_j s_{ij} - v_i s_{ji})$$

$$\frac{dm_i}{dt} = v_i f - m_i \delta a_i,$$

where

$$f = 1 - e^{-\left(V + M\right)/K}$$

and

$$r_i = re^{-t/r}.$$ 

Here, $V$ is the total number of slender cells in the host, $M$ is the total number of stumpy cells, and $K$ is the number of cells at which maximum stumpy production occurs. At the beginning of infection, the growth rate is at a maximum, $r$, but as the infection progresses, there is generalized inhibition of the growth rate independent of the acquired immune response to each of the variants (37). For all of our simulations, we chose a $\delta$ value of 2,500 on the basis of the observation that it takes ~26 days for growth to be inhibited by 50% (37). Other parameters used in all simulations are $n = 30$; $r = 0.1$ h$^{-1}$ (44); $d = 0.5$ h$^{-1}$; $\delta = 0.1$ h$^{-1}$ (45); and $K = 1 \times 10^8$ (33).

**Fig. 3.** Time series of parasite dynamics and the immune response with a differential-activation switching pattern. The rate at which a variant is switched to varies but is independent of which variant it is replacing. (Top) The bar graphs show the distribution from which the switching rates were chosen prior to the simulation. These rates were rescaled so that the probability that a parasite switches to another or the same variant is 0.01 per population doubling. (Middle) Black lines represent the total number of parasites, and each colored line represents a variant. (Bottom) Shown is the immune response against each variant. Simulations began with 1,000 trypanosomes and a repertoire of 30 VSGs. Because this is a small proportion of the potential in vivo repertoire (~2,000 VSGs), duration of simulated infections is much briefer than occurs in vivo. Parameters are the same as in Fig. 2. (A) All variants are chosen from the same distribution. (B) Fifteen of the variants are chosen from the left-hand distribution, and 15 are chosen from the right-hand distribution. (C) Ten variants are chosen from each of the three distributions.
When modeling the impact of the immune response on infection dynamics, it is most important to capture the dynamics of antibody removal, because only 1% of the population is switching at any time, and therefore only 1% of the population is potentially vulnerable to these existing antibodies at any time. As expected, the VSG repertoire is quickly exhausted, there is no order in expression, and the infection is quickly eliminated by the host immune system (Fig. 2). To investigate the effect of differential switching, we next assumed the differential-activation pattern shown in Fig. 1C. Because the activation rate of each variant is independent of the currently expressed variant, the value of $s_{ij}$ depends only on i and not on j (so, for example, $s_{11} = s_{12} = s_{13}$ and so on). Initially, we incorporated switch rates from a continuous range, from high to low activation probability, with a range of nearly 100,000-fold in the activation rates of different variants (once chosen, the switch rates remain constant throughout the simulation). This range may or may not reach the lowest VSG switch rate, which is unknown, but a wider range of activation probabilities would simply extend the chronicity of infection. Fig. 34 reveals order in variants and in the corresponding immune responses. During the first peak of parasitemia, the inoculated variant type predominates, with different variants being generated continually, according to their inherent activation rate. The first immune response removes the infecting variant, allowing the next variant to become more predominant, and the cycle of growth and death repeats for all variants. Superimposed on this is higher-level fluctuation of the whole parasitemia, resulting from the density-dependent differentiation from long slender to short stubby, nondividing, stage (34). The general features of these initial simulations resemble what occurs in vivo (15, 16): no variants reappeared, there was positive correlation between the activation time of each variant and the appearance of a strong antibody response, many variants were present simultaneously, and the parasitemia occurred in waves. Thus, differential activation probabilities clearly can be a major factor in determining order, which does not require a complex switch matrix in which the activation probability depends on the VSG being expressed.

We then tested whether discontinuous switch rate changes, corresponding to the early and later groups in vivo (16), yield realistic patterns of antigenic variation and parasitemia. Using the differential-activation switching pattern (Fig. 1C), we included a faster and a slower group of 15 variants each for the second set of simulations (Fig. 3B). The third set of simulations included 10 variants each assigned to high, medium, and low activation probabilities (Fig. 3C). Both the second and third sets of simulations produced more distinct parasitemia peaks and more prominent variant subpeaks (Fig. 3 B and C) than the first set of simulations (Fig. 3A). As the number of switch rate distributions increases, there is a progressive clustering of the variant-specific immune responses, tending toward what has been observed to occur in vivo (15, 32, 50).

To examine the influence of mosaic-activated switching, whereby the extent of homology with the coding sequence of the active VSG determines likelihood of participation in the switch, we modeled it in isolation, with the assumptions that switching is reversible and that VSG 1 was most homologous to 2, which in turn was most and equally homologous to 1 and 3, etc. ($s_{ij} = s_{ji}$ for all i and j) (Fig. 1D). Our simulations of homology-based
switching yield, within each relapse peak, several variant sub-peaks and ordered progression (Fig. 4). Finally, to ... the same as in Fig. 2.

Discussion

It is clear from all of the discontinuous rate range simulations (Figs. 3–5) that intrinsic VSG activation rates can provide order in antigenic variation and that the switching distribution has a strong influence on antigenic variation patterns. A patent prediction is that the characteristic groupings of variants in each parasitemia peak in vivo (15, 16) arise primarily from the existence of discontinuous switch ranges. Superimposed on this primary effect are the combined effects of density-dependent differentiation to the stumpy stage and immune responses, which result in aggregation of more variants within growth peaks and greater spacing between peaks. A challenge is to search for these putative VSG groupings, which are on a finer scale than those proposed by Capbern et al. (16), and then to determine their molecular basis.

Thus, two main factors, both parasite-determined, can account for ordered antigenic variation in chronic infection. They are the simple variation in activation rates and differentiation to the nondividing transmission stage. Our model is thus more parsimonious than previous theoretical models and more consistent with experimental data (see above). It can be argued that both differential activation probabilities and density-dependent differentiation apply generally to antigenic variation in eukaryotic parasites. Stochastic, ordered switching is held to be common to bacterial and eukaryotic pathogens (51, 52). Density-dependent differentiation to a nondividing stage, although not evident in bacteria, is well documented in eukaryotic parasites. An interesting case is Giardia, which encysts with density dependence and appears to produce new antigenic variants upon excystation (53). Nevertheless, trypanosome antigenic variation differs from that of other pathogens in its likely scale, the repertoire consisting of up to 2,000 VSG genes, the majority of which are pseudogenes. In contrast, the silent MSP2 archive of the bacterium Anaplasma marginale comprises just five pseudogenes (6), whereas the P. falciparum var archive has ~60 members (54).

The trypanosome archive may have enlarged so much because, unlike the variable antigens of other pathogens, the VSG appears not to have biochemical function, presumably allowing more unrestrained, diversifying evolution. The combinatorial use of damaged genes to create novel, expressed mosaic genes can greatly enlarge the effective archive, as seen with A. marginale (6), so T. brucei theoretically has strikingly enormous potential, with corresponding capability for lengthening infection (55).

There may be two reasons for the possible silent information overload. One is that African trypanosomes live permanently extracellularly in the vasculature, provoking extensive antibody responses. The other relates to rate of archive use. In our analysis, the experimental set of 30 VSGs was expended in only two or three peaks. This is more profligate than generally has been concluded from empirical clonal studies, but those studies screened only a few selected VSGs, in unusually slowly switching strains (14, 22, 36). Nevertheless, profligacy of similar magnitude has been recorded in the first relapse after the more natural situation, tsetse fly transmission (56). It might be concluded that trypanosome antigenic variation appears effectively to be pitched in a war of attrition with the host antibody responses, the two systems attempting to match the rapid diversification of each other. By simple force of numbers, trypanosome antigenic variation apparently does not require the subtleties, such as transient cross-reactive immune responses, that have been hypothesized for Plasmodium (30). An important consequence of order being determined by the trypanosome, rather than the host, is that distinct series, or “strings,” of variants can be generated in different infections with the same strain because of the stochastic nature of mosaic gene formation. Such dissimilarity between infections might enable effective transmission in the face of herd immunity (57).

We thank Sean Nee for discussions and the Wellcome Trust for funding. J.D.B. is a Wellcome Trust Principal Research Fellow.
