Minireview

DNA breaks as triggers for antigenic variation in African trypanosomes

Sam Alsford, David Horn and Lucy Glover

Address: London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK.

Correspondence: David Horn. Email: david.horn@lshtm.ac.uk

Published: 8 June 2009

Genome ***Biology*** 2009, **10:**223 (doi:10.1186/gb-2009-10-6-223)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2009/10/6/223

© 2009 BioMed Central Ltd

Abstract

The DNA repair machinery has been co-opted for antigenic variation in African trypanosomes. New work directly demonstrates that a double-strand break initiates a switch in the expressed variant surface coat.

The African trypanosome, *Trypanosoma brucei*, is a protozoan parasite of major medical and economic importance. The phenomenon of antigenic variation in *T. brucei* and some other pathogenic microbes enables them to evade the immune system and persist in the host [1]. In *T. brucei*, antigenic variation depends on expression of a single variant surface glycoprotein (VSG) from one among multiple expression sites [2] and an archive of more than 1,600 *VSG* genes [3]. Two principal mechanisms have emerged for the activation of a new *VSG*: duplicative gene conversion into the active expression site; and *in situ* activation of a silent expression site coupled with repression of the one previously active. The former is the dominant mechanism underlying antigenic variation [4], and is the focus of new work reported recently in *Nature* by Boothroyd *et al*. [5]. As double-strand breaks (DSBs) have long been known to initiate genetic recombination processes by facilitating strand invasion, it had been proposed that a DSB in the 70 bp repeats adjacent to each *VSG* gene (Figure 1) could initiate the process that copies a new *VSG* into the active expression site [6]. Boothroyd and colleagues now provide compelling evidence in support of this idea [5].

VSG genes are copied into the active expression site

The *VSG* genes are located at subtelomeres, and immune selection has had a major impact on the evolution of these chromosomal regions in *T. brucei* [7]. The expression sites are found on the megabase (1-6 Mbp) and intermediate-sized (150-500 kbp) chromosomes. They form long polycistronic transcription units that include a variable number of expression-site-associated genes (*ESAG*s) and a *VSG* located within a few kilobases of the telomere [2]. In addition, each of the 100 or so minichromosomes contains at least one telomeric *VSG*. Surprisingly, genome sequencing revealed that the remaining 1,000-2,000 archive *VSG*s were mostly pseudogenes (more than 85%) arranged in subtelomeric tandem arrays on the megabase chromosomes [3]. In total, about 30% of the genome is dedicated to *VSG*-based immune evasion.

It has been known for some time that 70 bp repeats form the boundaries of *VSG* gene conversion events (Figure 1), suggesting that they provide a substrate for recombination between the active expression site and a silent *VSG* locus [8]. Comparison of genetic maps and DNA sequencing before and after a switch suggested that *VSG* copying initiates 5’ to the *VSG* gene, but that the termination point depends on the source of the copied gene; array *VSG*s often recombine such that their 3’ end is altered [9], whereas recombination with a silent telomeric *VSG* results in a conversion event that extends beyond this point [10,11] (Figure 1). Genome sequencing further revealed the distribution and degeneracy of the 70 bp repeats, with most (92%) full-length array *VSG*s having at least one repeat 1-2 kbp upstream (one, 74%; two, 14%; or more, 5%) [3]. Importantly, the expression-site *VSG*s are typically flanked upstream by tens to maybe more than a hundred 70 bp repeats (Figure 1).

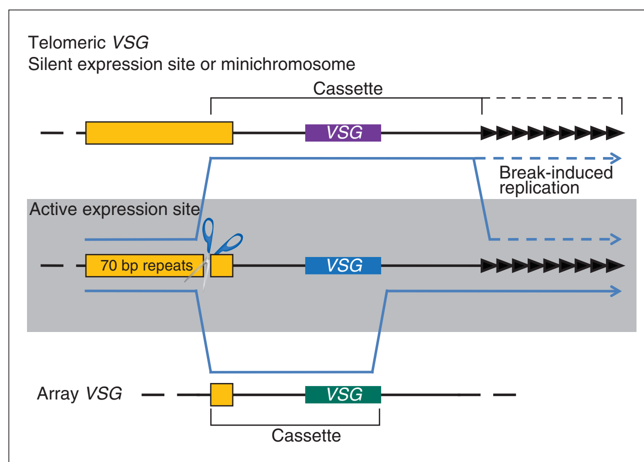

Figure 1
Model for homologous-recombination-mediated double-strand break repair and *VSG* switching. Center, a double-strand break (scissors) within the 70 bp repeats (yellow bar) at the active expression site initiates a search for a homologous repair template. Telomeric (T_2AG_3) repeats are indicated by the black arrowheads. Top, a silent telomeric *VSG* (up to 250 are available) can be copied into the active expression site via a gene-conversion cassette mechanism or by break-induced replication extending to the end of the chromosome. The longer stretches of 70 bp repeats found at these loci may increase the likelihood that these tracts are recognized and invaded by the broken strand at the active expression site. Bottom, replacement by a silent array *VSG* must use a cassette mechanism. The synthesis-dependent strand-annealing model [25] favors the type of recombination illustrated rather than crossover between silent *VSG* arrays and expression sites.

It has been challenging to determine how the variety of *VSG* templates available for recombination intersect with the host immune response and contribute to the course of a natural infection. Together with the sequence data, functional studies have begun to clarify some details. Silent *VSG*s fall into distinct activation-probability groups that correlate with the amount of flanking sequence shared with the active *VSG* [12]. As an infection progresses and immunity develops against high-probability VSGs, lower-probability VSGs emerge. Thus, *VSG* genes located adjacent to telomeres are more likely to be activated than those located in an array. The number of flanking 70 bp repeats may be particularly important if a DSB in this sequence triggers a homology search among the vast *VSG* archive. In fact, *VSG* gene conversion can occur in the absence of 70 bp repeats at the active expression site [13] and chimeric *VSG*s may be formed via *VSG*-*VSG* recombination [3], but these events are analogous to the low-probability switches in the scenario outlined above.

Double-strand breaks trigger *VSG* switching

The 70 bp repeats were first implicated as recombination 'hotspots' able to initiate the *VSG* copying process more than 25 years ago [6]. In order to test the hypothesis that a DSB adjacent to these repeats could initiate a switch, Boothroyd *et al*. [5] placed a yeast meganuclease (I-*Sce*I) recognition sequence in the active expression site immediately adjacent to the 70 bp repeats; the trypanosome strain was also modified to express an inducible copy of the meganuclease [5]. Induction of I-*Sce*I expression generated a DSB in around 1% of the cells, which increased the switching frequency by around 250-fold. In contrast, a DSB elsewhere in the expression site or after deletion of the 70 bp repeats did not increase switching frequency. These data indicate that both the location of the DSB next to the repeats and the repeats themselves are critical to increased switching.

Boothroyd *et al.* [5] then analyzed 18 of the switched progeny, revealing that they had all lost the active *VSG*, replacing it with donor *VSG*s from a silent expression site (15/18) or a mini-chromosome (3/18, top panel in Figure 1). To determine how the induced DSB was processed, the authors next sequenced across the repaired region in five of the progeny, all of which had switched by recombination with the same silent expression site. They were able to show that the DSB invaded 70 bp repeats proximal to the donor *VSG*, resulting in the replacement of some of the 70 bp repeats as well as the original *VSG*. Furthermore, they demonstrated that the subtelomeric region downstream of the original *VSG* had also been replaced. This is reminiscent of break-induced replication, where a broken chromosome invades a homologous template and initiates replication to the telomere [14] (Figure 1).

The results detailed above demonstrate that a meganuclease-generated DSB adjacent to the 70 bp repeats triggers antigenic variation, but is this what happens naturally? To investigate this the authors used ligation mediated-PCR (LM-PCR) with DNA templates from wild-type trypanosomes. More frequent, mostly staggered DSBs were detected in the 70 bp repeats at the active expression site compared with a 70 bp repeat array in a silent expression site or at a chromosomal internal locus, adding further support to the hypothesis that DSBs occur *in vivo* and trigger *VSG* switching. Interestingly, however, an I-*Sce*I-induced break appears to be a far more potent trigger for a *VSG*-switch than these natural breaks.

Is there a rationale for why the 70 bp repeats might be more prone to DSBs, and why there should be a difference between an active and a silent expression site? The 70 bp repeats consist, in part, of a large number of TTA repeats. These have been shown to destabilize plasmid duplex DNA [15], leading Boothroyd *et al.* [5] to speculate that this region becomes destabilized during active expression site transcription, resulting in the formation of natural DSBs. Also, unusual structures formed by repetitive sequences may be prone to replication-fork stalling and collapse, or they may provide a substrate for a trypanosome endonuclease. Any combination of these factors could explain why breaks are frequently seen in the 70 bp repeats and specifically at

the active expression site. In addition, base-J, a modified thymidine implicated in suppressing homologous recombination, is found throughout silent expression sites but not in the active one nor in insect-stage trypanosomes [16]. The replacement of a proportion of the thymidines in silent expression sites may somehow restrict the occurrence of DSBs.

One intriguing feature of antigenic variation in *T. brucei* is the rapid decline in switching frequency when fresh 'natural' isolates are adapted to laboratory culture; switching drops from about 10^{-2} per population doubling and by up to 10,000-fold during growth in culture [17]. Two models that could account for this substantial but reversible decrease include changes in telomere length and accumulation of base-J. Cross and colleagues [18] proposed a model whereby increases in telomere length during growth in culture reduces the frequency of subtelomeric breaks. Alternatively, and as described above, the accumulation of base-J [16] may stabilize the 70 bp repeats in silent expression sites, making them less prone to DSBs or to strand invasion and recombination.

Questions for the future

Some additional questions are worth considering in light of the results of Boothroyd *et al.* [5]. First, if breaks in the 70 bp repeats can also initiate recombination with array *VSG*s (bottom panel in Figure 1), then what prevents replication from continuing beyond the *VSG* 3'-end? It may be that strand invasion at ectopic 70 bp repeats prompts a second homology search that favors extensive local homology (long tracts of telomeric repeats, for example), but will use more restricted homology or even microhomology if necessary. Second, is replication directional and, if so, how is this controlled? That is, can sequences be copied on either side of the 70 bp repeats and could breaks in the 70 bp repeats also be responsible for recombination events that generate upstream *ESAG* diversity? Third, subtelomeric clusters have been implicated in recombination events in *T. brucei* [19] and in other protozoa but the active expression site is sequestered in the expression-site body, an extranucleolar focus of RNA polymerase I that drives monoallelic *VSG* transcription [20]. During a recombination-mediated switch, is the DNA break repaired inside or outside of the expression-site body or does the structure disassemble transiently?

The 70 bp repeats are likely to play a key role in antigenic variation and they are probably also important for *VSG* diversification beyond the active expression site [3,21]. Although the repair pathways available to *T. brucei* [22] have yet to be dissected genetically in any detail, it is known that RAD51-dependent homologous recombination is responsible for only a subset of *VSG*-copying events [23]. This is consistent with a predominant break-induced replication model (Figure 1), as RAD51-dependent and independent break-induced replication pathways have been described in yeast [14]. Fortunately, meganucleases such as I-*Sce*I provide powerful tools to study the repair of unique DSBs *in vivo* [24]. It is exciting to see new opportunities for probing the molecular mechanisms underlying antigenic variation in trypanosomes and for genetically dissecting the pathways used to switch *VSG* gene expression.

Acknowledgements

We thank John Kelly for comments on the draft manuscript. Our research is funded by the Wellcome Trust.

References

1. Barry JD, Ginger ML, Burton P, McCulloch R: **Why are parasite contingency genes often associated with telomeres?** *Int J Parasitol* 2003, **33:**29-45.
2. Hertz-Fowler C, Figueiredo LM, Quail MA, Becker M, Jackson A, Bason N, Brooks K, Churcher C, Fahkro S, Goodhead I, Heath P, Kartvelishvili M, Mungall K, Harris D, Hauser H, Sanders M, Saunders D, Seeger K, Sharp S, Taylor JE, Walker D, White B, Young R, Cross GA, Rudenko G, Barry JD, Louis EJ, Berriman M: **Telomeric expression sites are highly conserved in *Trypanosoma brucei*.** *PLoS ONE* 2008, **3:**e3527.
3. Marcello L, Barry JD: **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* 2007, **17:**1344-1352.
4. Robinson NP, Burman N, Melville SE, Barry JD: **Predominance of duplicative *VSG* gene conversion in antigenic variation in African trypanosomes.** *Mol Cell Biol* 1999, **19:**5839-5846.
5. Boothroyd CE, Dreesen O, Leonova T, Ly KI, Figueiredo LM, Cross GA, Papavasiliou FN: **A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*.** *Nature* 2009, **459:**278-281.
6. Liu AY, Van der Ploeg LH, Rijsewijk FA, Borst P: **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* 1983, **167:**57-75.
7. Horn D, Barry JD: **The central roles of telomeres and subtelomeres in antigenic variation in African trypanosomes.** *Chromosome Res* 2005, **13:**525-533.
8. Campbell DA, van Bree MP, Boothroyd JC: **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* 1984, **12:**2759-2774.
9. Bernards A, Van der Ploeg LH, Frasch AC, Borst P, Boothroyd JC, Coleman S, Cross GA: **Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end.** *Cell* 1981, **27:**497-505.
10. De Lange T, Kooter JM, Michels PA, Borst P: **Telomere conversion in trypanosomes.** *Nucleic Acids Res* 1983, **11:**8149-8165.
11. Scholler JK, Myler PJ, Stuart KD: **A novel telomeric gene conversion in *Trypanosoma brucei*.** *Mol Biochem Parasitol* 1989, **35:**11-19.
12. Morrison LJ, Majiwa P, Read AF, Barry JD: **Probabilistic order in antigenic variation of *Trypanosoma brucei*.** *Int J Parasitol* 2005, **35:** 961-972.
13. McCulloch R, Rudenko G, Borst P: **Gene conversions mediating antigenic variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking 70-base-pair repeat sequences.** *Mol Cell Biol* 1997, **17:**833-843.
14. McEachern MJ, Haber JE: **Break-induced replication and recombinational telomere elongation in yeast.** *Annu Rev Biochem* 2006, **75:**111-135.
15. Ohshima K, Kang S, Larson JE, Wells RD: **TTA.TAA triplet repeats in plasmids form a non-H bonded structure.** *J Biol Chem* 1996, **271:**16784-16791.
16. Borst P, Sabatini R: **Base J: discovery, biosynthesis, and possible functions.** *Annu Rev Microbiol* 2008, **62:**235-251.
17. Turner CM: **The rate of antigenic variation in fly-transmitted and syringe-passaged infections of *Trypanosoma brucei*.** *FEMS Microbiol Lett* 1997, **153:**227-231.

18. Dreesen O, Li B, Cross GA: **Telomere structure and function in trypanosomes: a proposal.** *Nat Rev Microbiol* 2007, **5:**70-75.
19. Perez-Morga D, Amiguet-Vercher A, Vermijlen D, Pays E: **Organization of telomeres during the cell and life cycles of *Trypanosoma brucei*.** *J Eukaryot Microbiol* 2001, **48:**221-226.
20. Navarro M, Gull K: **A pol I transcriptional body associated with *VSG* mono-allelic expression in *Trypanosoma brucei*.** *Nature* 2001, **414:**759-763.
21. Aline RF, Jr., Myler PJ, Stuart KD: ***Trypanosoma brucei*: frequent loss of a telomeric variant surface glycoprotein gene.** *Exp Parasitol* 1989, **68:**8-16.
22. Glover L, McCulloch R, Horn D: **Sequence homology and microhomology dominate chromosomal double-strand break repair in African trypanosomes.** *Nucleic Acids Res* 2008, **36:**2608-2618.
23. McCulloch R, Barry JD: **A role for RAD51 and homologous recombination in *Trypanosoma brucei* antigenic variation.** *Genes Dev* 1999, **13:**2875-2888.
24. Sugawara N, Haber JE: **Repair of DNA double strand breaks: *in vivo* biochemistry.** *Methods Enzymol* 2006, **408:**416-429.
25. Paques F, Haber JE: **Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*.** *Microbiol Mol Biol Rev* 1999, **63:**349-404.