



University  
of Glasgow

Barry, J.D., Hall, J.P.J., and Plenderleith, L. (2012) Genome hyperevolution and the success of a parasite. *Annals of the New York Academy of Sciences*, 1267 (1). pp. 11-17. ISSN 0077-8923

Copyright © 2012 New York Academy of Sciences

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

The content must not be changed in any way or reproduced in any format or medium without the formal permission of the copyright holder(s)

When referring to this work, full bibliographic details must be given

<http://eprints.gla.ac.uk/73292>

Deposited on: 18 December 2012

Enlighten – Research publications by members of the University of Glasgow  
<http://eprints.gla.ac.uk>

## ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *Effects of Genome Structure and Sequence on Variation and Evolution***Genome hyperevolution and the success of a parasite**

J. David Barry, James P. J. Hall, and Lindsey Plenderleith

Wellcome Trust Centre for Molecular Parasitology, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary, &amp; Life Sciences, University of Glasgow, Glasgow, United Kingdom

Address for correspondence: J. David Barry, Wellcome Trust Centre for Molecular Parasitology, Institute of Infection, Immunity &amp; Inflammation, College of Medical, Veterinary &amp; Life Sciences, University of Glasgow, 120 University Place, Glasgow G12 8TA, United Kingdom. dave.barry@glasgow.ac.uk

The strategy of antigenic variation is to present a constantly changing population phenotype that enhances parasite transmission, through evasion of immunity arising within, or existing between, host animals. Trypanosome antigenic variation occurs through spontaneous switching among members of a silent archive of many hundreds of variant surface glycoprotein (VSG) antigen genes. As with such contingency systems in other pathogens, switching appears to be triggered through inherently unstable DNA sequences. The archive occupies subtelomeres, a genome partition that promotes hypermutagenesis and, through telomere position effects, singular expression of VSG. Trypanosome antigenic variation is augmented greatly by the formation of mosaic genes from segments of pseudo-VSG, an example of implicit genetic information. Hypermutation occurs apparently evenly across the whole archive, without direct selection on individual VSG, demonstrating second-order selection of the underlying mechanisms. Coordination of antigenic variation, and thereby transmission, occurs through networking of trypanosome traits expressed at different scales from molecules to host populations.

**Keywords:** genome hyperevolution; antigenic variation; parasite; trypanosome; subtelomere

**Introduction**

Parasites and their hosts notoriously engage in an arms race, evolving measures and countermeasures against each other in a battle for supremacy. So strong are the selective forces, and consequently the rate and extent of evolution, that we can see a variety of extreme adaptations that are informative of mutational mechanisms and of general eukaryotic biology. Highly prominent in the arms race are phenotypes associated with evasion of mammalian immunity. Escape by a parasite might seem relatively simple—hide, or perhaps block, a key step in immunity, or maybe adopt camouflage—but closer inspection often reveals a highly complex biological system. The system often includes adaptations at various scales, ranging from the genome, through cellular structures and machinery, to population behavior within and between infections. Furthermore,

the functional outcomes of such adaptations generally are coordinated, forming a network of interacting processes that underpin a long-term strategy for enhanced transmission of the parasite and hence its persistence in an ecosystem. Nowhere is this complex type of evasion system more apparent than in the antigenic variation of African trypanosomes.

**Antigenic variation: tactics for population survival**

Antigenic variation is fundamental to the success of many pathogens. As infection proceeds, the resident pathogen population is decimated by populations of highly specific antibodies directed typically against its major surface antigen. Some individuals, however, have already switched to antigenically different versions of that molecule and survive and proliferate, a process that repeats many times, producing alternating waves of antigens and antibodies.<sup>1</sup> This process reflects the classical arms race in which pathogen survival mechanisms and host immune mechanisms coevolve. With the capacity to make  $\sim 10^{10}$  distinct antibody idiotypes,

---

Re-use of this article is permitted in accordance with the Terms and Conditions set out at [http://wileyonlinelibrary.com/onlineopen#OnlineOpen\\_Terms](http://wileyonlinelibrary.com/onlineopen#OnlineOpen_Terms)

mammals seem destined to win this race. This enormous potential for immunological variation cannot be realized simply through direct encoding of antibodies in our genome, and requires a set of mechanisms that differentially shuffle and combine encoded components of antibodies.<sup>2</sup> How can a parasite, with a much smaller and less complex genome, compete?

In considering this question, we must broaden our vision. We have to adopt a systems biology approach, in which we link different functional scales of the parasite. Within infections, antigenic variation operates in concert with other intrinsic parasite parameters including growth and differentiation dynamics, to name just two.<sup>3</sup> Beyond individual infections, parasites have to be transmitted to new hosts. For a parasite that is abundant in the field, there is not an endless supply of new hosts, so a major challenge is likely to be successfully re-infecting hosts that already have antibodies corresponding to previous infections, either in circulation or stored as memory. Within an evolutionary framework, interactions at all scales, from molecules to host populations, amalgamate into a complex phenotype under complex selection. Features of antigenic variation, from molecules to parasite populations, will reflect the pressures of those selective forces.<sup>3</sup>

Trypanosomes live extracellularly in the blood, exposed to several immune mechanisms.<sup>4,5</sup> Their variant surface glycoprotein (VSG) forms a dense coat that thwarts innate immunity and prevents antibodies accessing invariant cell surface molecules and eliminating the infection.<sup>6</sup> A key event in the evolution of VSGs appears to have been dispensing with specific biochemical function, allowing their sequence to vary enormously, with merely some general structural and processing constraints. Spontaneously, at a rate of  $\sim 10^{-3}$  switch/cell/generation, trypanosomes switch from one VSG to another, using mechanisms described below. The spontaneity and high rate of switching mean constant presence of variants before the onset of specific antibodies. For example, the first wave of parasites in a cow can peak with a total in excess of  $10^{11}$  trypanosomes, and therefore  $>10^8$  will have switched to a different variant, while a mouse can support  $>10^8$  parasites with  $>10^5$  emerging variants.<sup>7</sup> Many microbes facing strong challenges in hosts display such preemptive phenotypic variation, arising specifically from highly mutable contingency loci.<sup>8</sup>

In the trypanosome system, VSG switching is coordinated with growth and density-dependent differentiation of the parasite from proliferative to the nonproliferative transmission phase, which will infect tsetse flies on uptake in a bloodmeal.<sup>9</sup> There is a fairly constant supply of the transmission stage, the level of which is critically affected by the number of easily activated VSG variants. The more VSG variants that are present at a time, the less chance each variant subpopulation has of reaching the threshold to induce an antibody response, meaning that, when the number of variants is high, overall population numbers are predicted to be controlled by differentiation.<sup>3,10</sup> Such a swing toward differentiation-based control over immunity-based control results in higher levels of the transmission stage, but complete reliance on this mechanism would lead to prolonged high parasitemia and, therefore, earlier host death. In parallel, variants appear in ordered progression,<sup>1</sup> which favors efficient use of the archive and promotes chronicity of infection. Ordered expression could be achieved by variants clustering into “blocks” of distinct activation probability, with each block corresponding to a growth peak in the host.<sup>3</sup> As infections can run for years, many VSGs are probably required over the course of a single infection. Furthermore, it has been proposed that expressing substantially different subsets of VSGs would enable successful reinfection of previously exposed, partially immune hosts.<sup>11</sup> Reinfection would be facilitated also by the known diversity in the “metacyclic” VSGs expressed by the initial population injected by the tsetse fly vector.<sup>12</sup> On top of this reinfection requirement, the wide host range of trypanosomes introduces another dimension of uncertainty with which parasites must cope. Success in the wild, therefore, requires the various components of the network to adjust, enabling success in numerous host types.

### **Multiple mechanisms contribute to the generation of VSG diversity**

The trypanosome genome appears to have evolved to provide such flexibility. To do so has required unusual adaptations, which operate at different levels.

#### *Adaptations in genome structure*

**Expanded coding capacity and complexity.** The enormous scale of antigenic variation is templated by an archive of nearly 2000 silent VSG genes

and pseudogenes—nearly one-third of the core trypanosome genome.<sup>13,14</sup> The archive is very diverse, particularly in the sequences encoding the N-terminal domain of the VSG, which carries the key epitopes seen by the immune system. Typically, amino acid identity in that domain between VSGs is less than 20%, a particularly broad range that arises presumably from the lack of specific biochemical function and because many common amino acids specify alpha-helix, which comprises most of the domain.<sup>6</sup>

**Partitioning of the genome: the importance of subtelomeres.** As with many eukaryotes, the trypanosome genome is effectively partitioned into the core, containing genes under purifying selection and where most mutation is repaired, and subtelomeres, where multigene families associated with organismal phenotypic variation diversify rapidly.<sup>15,16</sup> The entire repertoire of VSG, including the archive and expressed genes, is located in subtelomeres.<sup>14</sup> Trypanosome subtelomeres are relatively enormous, in one case being several times longer than the associated chromosome core.<sup>17</sup>

Functionally, subtelomeres contribute two special features relevant to VSG and antigenic variation. The first of these features is ectopic recombination. Unlike chromosome cores, subtelomeres recombine ectopically, permitting sequence exchange or modification between different chromosomes, introducing diversity.<sup>18</sup> A number of recombination pathways can be involved.<sup>19</sup> Other, more general, mutational mechanisms contribute to VSG diversification, but ectopic recombination plays a particularly significant role and is likely to be specific to subtelomeres.

The second feature is the telomere position effect, in which gene promoters close to telomeres are subject to reversible repression mediated by silencing proteins that bind, directly or indirectly, to the telomere tract at the chromosome end. In the trypanosome, this effect appears to have evolved further into a sophisticated regulatory system. The only loci from which VSG can be transcribed are expression sites (ES), of which there are several in the genome.<sup>20</sup> All are telomere-proximal (i.e., the last genes before the telomere). Each ES comprises a promoter, several non-VSG genes, a set of imperfect repeats (“70-bp repeats”), and the VSG adjacent to the telomere. Antigen switching involves mostly

replacement, by gene conversion, of the expressed VSG copy (or a part thereof), by all (or part) of another VSG.<sup>1,21</sup> Only one ES is active at a time, which requires occupancy of a specific nuclear niche, the ES Body (see below), and inactivity of the other ES involves a telomere position effect.<sup>22</sup> It is likely that telomere proximity also functions in antigen switching, due to the capacity of telomeres to interact with one another, promoting recombination.<sup>18</sup>

**Supernumerary chromosomes.** The trypanosome has evolved a set of ~100 nuclear minichromosomes, which carry, with little exception, only telomere-proximal VSG genes.<sup>23</sup> The two consequences are increase in archive size and availability of telomere-proximal VSG, possibly increasing the pool of telomere-proximal VSG that can interact readily with the expression site (switching) or with each other (recombination-mediated diversification).

**Adaptations in nuclear structure and function.** Trypanosome telomeric regions, including VSG genes, are located in heterochromatin at the nuclear periphery.<sup>24</sup> VSG transcription is mediated by RNA polymerase I (RNAPI), which normally transcribes only ribosomal RNA genes located in the nucleolus. The active VSG occupies the ES Body, an RNAPI-containing extranucleolar niche located within the nuclear interior rather than peripherally.<sup>25</sup> The minichromosomes behave differently from the core chromosomes, lying in the nuclear periphery, being duplicated earlier, and segregating via core microtubules rather than the spindle; the unusual behavior might be necessitated, at least partially, by overloading of the conventional machinery.<sup>25</sup>

**Recombination prone sequences.** Recombination and repair pathways serve fundamental functions in any organism, but they can be exploited to achieve specific phenotypes, such as in antigenic variation. One common mechanism is that an elevated rate of recombination causes the expressed gene to become replaced or altered. The elevated rate is specific to the contingency loci and is facilitated by the presence of recombination-prone sequences.<sup>26,27</sup> In the VSG system, the unstable, 70-bp repeats lying upstream of the expressed VSG are thought to precipitate a high level of recombination (see below). It is also apparent that the subtelomere

compartment in the trypanosome genome differs from the chromosome cores in its interaction with recombination/repair mechanisms, becoming disproportionately disrupted in null mutants of key players MRE11 and BRCA2.<sup>28,29</sup> As to which recombination/repair pathways mediate VSG interactions, some interesting novelties exist in the trypanosome general pathways, but a formal link has yet to be made.<sup>30</sup>

**Implicit information: pseudogenes and the combinatorial creation of novel intact genes.** Some two-thirds of the array VSGs are pseudogenes, with premature stop codons, frameshifts, deletions, or lack of an appropriate start codon.<sup>14</sup> At first sight, this situation appears to be a colossal waste of resources, but in fact it is probably the opposite. By what is likely to be nonreciprocal recombination (gene conversion) of wild-type fragments of pseudogenes into the expressed VSG, antigenically distinct genes can be generated.<sup>31</sup> The pseudogenes

therefore encode implicit information.<sup>32</sup> Combinatorial use of silent information has potential for greatly expanding the extent of resulting, explicit information, as is demonstrated in the case of the bacterium *Anaplasma*, which can create apparently hundreds of antigenically distinct MSP2 variants from an archive of merely five unique pseudogenes, and no intact genes.<sup>33</sup> The trypanosome pseudogene pool is more than two orders of magnitude greater in size. That short conversion fragments can introduce novel function is apparent also from the human HLA and KIR cell surface proteins of the human immune system.<sup>34–36</sup>

*How the antigenic variation phenotype is served by genome adaptations (Table 1)*

**Singular VSG expression.** The ES Body provides expression to one ES at a time, and a telomere position effect, possibly along with other epigenetic effects, provides silencing of the others, ensuring controlled expression of the archive.

**Table 1. Genome features contributing to the antigenic variation phenotype.<sup>a</sup>**

Phenotype	Genome feature
Many VSGs	Large subtelomeres populated by VSG arrays
Singular VSG expression	ES Body: specialized nuclear niche; silencing through telomere position effects
Switching VSG	Telomere proximity of expression site enables interaction with VSG in other (sub)telomeres
Ordered expression—general	Locus types within subtelomeres have different activation probabilities
Ordered expression—specific	Flanks of each VSG? For mosaic VSG expression, coding sequence identity between donor and expressed genes
Antigenically diverse growth peaks	Blocks of distinct VSG with similar activation probability
Coordinated growth, differentiation, and antigenic variation	Number of VSG constituting a block. Higher number tends toward differentiation-based control; lower number favors antibody-mediated control
Prolonged infection	Mechanisms for ordered expression Mosaic VSG formation
Reinfection, superinfection	Diversity in MVSG set expressed in infective, metacyclic population? Mosaic VSG formation Contents and organization of blocks?
Host range	Large size of archive and its organization into blocks allow it to be used flexibly in different within-host environments, providing optimum transmission phenotype in each host species?
VSG hyperevolution	Enhanced mutation/reduced mutation repair in subtelomeres

<sup>a</sup>Speculation is denoted by question marks.

**Switching arises possibly from inherent tendency for double-strand breaks.** The 70-bp repeats upstream of the transcribed *VSG* each contain a tract of the physically unstable motif  $(TAA)_n$ .<sup>37</sup> Instability of this type theoretically can cause stalling of DNA replication, leading in effect to production of a double-strand break (DSB) in the newly synthesized duplex and, consequently, induction of repair.<sup>38</sup> A common route taken is creation of a gap and filling by copying from a homologous sequence. In this case, homology at another run of 70-bp repeats, upstream of a different *VSG*, would initiate the copying of that gene into the expression site. There is indeed some evidence that artificial creation of a DSB adjacent to the 70-bp repeats prompts switching by duplication, and that DSB occur naturally in the repeats in the expression site.<sup>39</sup> As with other contingency loci systems, where runs of short repeats are thought to cause gene inactivation or reactivation through indels arising by DNA polymerase slippage on the repeat template,<sup>8</sup> this mechanism is simple—an accident waiting to happen, at reasonably high frequency, due to inherent instability in the DNA sequence.

**Ordered *VSG* expression.** Antigenic variation runs in a pattern often termed *semipredictable*, with variants tending to appear in the same general order in distinct infections.<sup>1</sup> Accomplishing order for such an enormous archive seems inordinately difficult, but it is facilitated in part by the genome environment and structure. Each gene appears to have an inherent activation probability. Those with the highest probability will activate continuously, but early onset of antibodies against their product will render all subsequent activations futile. The process will repeat as the activation hierarchy is unrolled. In general, telomere-proximal genes have the highest activation probability, followed by intact genes in the arrays, and then pseudogenes.<sup>1</sup> Presumably, the BES telomere will interact more frequently with a donor at another telomere, via standard telomere associations, than with a donor in the subtelomere, which will require more routine homology searching. When mosaic gene formation comes into play, the interaction of donor with the BES is driven by degree of identity in the *VSG* coding sequence, making switching dependent on what went before.<sup>14</sup>

This gross level of order is refined, with blocks of variants appearing within each locus category.

Whereas the degree of coding identity will specify finer order for mosaic gene formation, we do not know how refinement is achieved for the other genes. A clue, however, lies in the *Borrelia* antigenic variation system, where fairly exquisite ordering is achieved via two sequences, one in either flank of the gene.<sup>40</sup> It is conceivable that the conserved sequences flanking *VSG* might have similar influence.

**Diverse *VSG* expression.** Each peak comprises a mixture of variants, the number of which is probably critical to the balance between antibody-mediated and density-dependent limitation of growth.<sup>3</sup> In addition, diversity within each of these blocks decreases the extent to which antibody responses might eliminate the infection. Diversity probably occurs by all variants in a block having very similar activation probabilities.

**Prolonged infection, reinfection, and superinfection.** The combination of order and enormous potential for combinatorial creation of novel, mosaic expressed *VSG* can lead, at least in theory, to infections lasting months or even years. It has been hypothesized that the combinatorial mechanism also can allow reinfection of hosts already immune to that strain.<sup>11</sup> Experimental analysis has shown that mosaic gene formation does enable reinfection of calves with the bacterium *Anaplasma*.<sup>41</sup> Comparative genomics has shown that *VSG* archives are strain specific,<sup>42</sup> a situation that enables strains to compete with each other in coinfections. In turn, competition will drive toward greater evolution of the archive.

**Hyperevolution: selection of mechanisms that generate broad diversity.** Second-order (indirect) selection<sup>43</sup> underpins evolution of antigenic variation. Every *VSG* is dispensable, and most, being very rarely expressed, have little exposure to direct selection via their encoded protein. Yet, diversity appears to be spread evenly over the archive, for example with no evident outgrouping of the high-probability activators. Typically for multigene families in subtelomeres, the archive evolves rapidly. It seems that second-order selective processes operate, in which mutational mechanisms generate diversity across the whole *VSG* archive and across the whole coding sequence within each gene, independently of the immediate phenotype;<sup>14,42,44,45</sup> it is these mechanisms that have been selected. Thus,

an individual trypanosome with a novel trait that enhances population survival will be selected, but the dominant selective pressure acts at the level of the lineage, yielding mechanisms that directly confer benefit on the population rather than the individual. The view that evolution acts on only the immediate phenotype, rather than on long-term strategy is not compatible with what is observed for trypanosomes.

It is likely that the subtelomere provides an ideal environment for hypermutation, through naturally ectopic interactions and possibly through differential positioning within the nucleus. It is more likely that standard recombination-repair activities are used differentially on subtelomeres than that exclusive mutagenic mechanisms have evolved. *VSG* genes mutate in various ways. Ectopic gene conversion events duplicate *VSG* between subtelomeres, either as blocks of genes or individually.<sup>14</sup> Recently, comparative genomics of sequential isolates of one trypanosome strain have revealed a range of mutations, including base substitutions, short indels, and conversions within the coding sequence (L.P., T. Otto, M. Berriman, and J.D.B., unpublished).

## Summary

To return to our question of how a parasite, with a smaller and less complex genome than that of its mammalian host, can compete in the face of antibody population diversity, the answer is flexibility. The trypanosome genome is adapted in somewhat extreme ways to serve the antigenic variation phenotype, with approximately 25% of the genome devoted to this phenomenon. The special features of subtelomeres, including hyperevolution and reversible gene-silencing mechanisms, have been exploited to great effect, resulting in an enormous archive of rapidly mutating silent *VSG* genes and telomere-associated mechanisms for their differential expression. Subtelomeres containing *VSG* arrays have expanded, and a large set of supernumerary chromosomes has emerged, apparently to provide even more telomeres. These structural adaptations have required changes in nuclear machinery. As with enhanced phenotypic variation systems in other pathogens, often referred to as *contingency gene systems*, the spontaneous phenotype switches that underpin antigenic variation appear to occur through inherent instability of a DNA sequence upstream of *VSG*. The silent *VSG* archive evolves at a rate sev-

eral times faster than genes in chromosome cores, and it appears to do so without direct selection on expressed *VSG*. Instead, in a strong example of second-order selection, mechanisms have evolved that mutate all *VSG* apparently randomly. Much of the coding information is stored in the genome implicitly, in pseudogenes, requiring formation of mosaic genes for expression of novel antigens. The trypanosome *VSG* and mammalian antibody phenotypes both derive very extensive augmentation of information through the power of combinatorial reassembly of coding sequence. Besides these adaptations at the level of the genome, management of antigenic variation requires networking of *VSG* gene expression with other trypanosome phenotypes, including growth, differentiation, and, beyond the individual infection, unpredictable transmission into a very broad range of hosts, some of which will be partially immune.

## Acknowledgment

This work was supported by the Wellcome Trust (Grant numbers 055558, 083224, and 086415). The Wellcome Trust Centre for Molecular Parasitology is supported by core funding from the Wellcome Trust (Grant number 085349).

## Conflicts of interest

The authors declare no conflicts of interest.

## References

1. Morrison, L.J., L. Marcello & R. McCulloch. 2009. Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cell Microbiol.* 11: 1724–1734.
2. Kenter, A.L., S. Feldman, R. Wuerffel, *et al.* 2012. Three-dimensional architecture of the IgH locus facilitates class switch recombination. *Ann. N.Y. Acad. Sci.* 1267: 86–94. This volume.
3. Gjini, E. *et al.* 2010. Critical interplay between parasite differentiation, host immunity, and antigenic variation in trypanosome infections. *Am. Nat.* 176: 424–439.
4. Barry, J.D. & R. McCulloch. 2001. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. *Adv. Parasitol.* 49: 1–70.
5. La Greca, F. & S. Magez. 2011. Vaccination against trypanosomiasis: can it be done or is the trypanosome truly the ultimate immune destroyer and escape artist? *Hum Vaccin.* 7: 1225–1233.
6. Schwede, A. & M. Carrington. 2010. Bloodstream form Trypanosome plasma membrane proteins: antigenic variation and invariant antigens. *Parasitology* 137: 2029–2039.

7. Barry, J.D. 1986. Antigenic variation during *Trypanosoma vivax* infections of different host species. *Parasitology* **92**: 51–65.
8. Bayliss, C.D. & M.E. Palmer. 2012. Evolution of simple sequence repeat mediated phase variation in bacterial genomes. *Ann. N.Y. Acad. Sci.* **1267**: 39–44. This volume.
9. MacGregor, P. *et al.* 2011. Transmission stages dominate trypanosome within-host dynamics during chronic infections. *Cell Host Microbe* **9**: 310–318.
10. Lythgoe, K.A. *et al.* 2007. Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation. *Proc. Natl. Acad. Sci. USA* **104**: 8095–8100.
11. Barry, J.D. *et al.* 2005. What the genome sequence is revealing about trypanosome antigenic variation. *Bioch. Soc. Trans.* **33**: 986–989.
12. Ginger, M.L. *et al.* 2002. Ex vivo and in vitro identification of a consensus promoter for VSG genes expressed by metacyclic-stage trypanosomes in the tsetse fly. *Eukaryot Cell* **1**: 1000–1009.
13. Berriman, M. *et al.* 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**: 416–422.
14. Marcello, L. & J.D. Barry. 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.
15. Barry, J.D. *et al.* 2003. Why are parasite contingency genes often associated with telomeres? *Int. J. Parasitol.* **33**: 29–45.
16. Riethman, H., A. Ambrosini & S. Paul. 2005. Human subtelomere structure and variation. *Chromosome Res.* **13**: 505–515.
17. Callejas, S. *et al.* 2006. Hemizygous subtelomeres of an African trypanosome chromosome may account for over 75% of chromosome length. *Genome Res.* **16**: 1109–1118.
18. Therizols, P. *et al.* 2010. Chromosome arm length and nuclear constraints determine the dynamic relationship of yeast subtelomeres. *Proc. Natl. Acad. Sci. USA* **107**: 2025–2030.
19. Koszul, R., B. Dujon & G. Fischer. 2006. Stability of large segmental duplications in the yeast genome. *Genetics* **172**: 2211–2222.
20. Hertz-Fowler, C. *et al.* 2008. Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS ONE* **3**: e3527.
21. Vink, C., G. Rudenko & H.S. Seifert. 2011. Microbial antigenic variation mediated by homologous DNA recombination. *FEMS Microbiol. Rev.* Dec 25. doi:10.1111/j.1574-6976.2011.00321.x. [Epub ahead of print].
22. Yang, X. *et al.* 2009. RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. *Cell* **137**: 99–109.
23. Wickstead, B., K. Ersfeld & K. Gull. 2004. The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. *Genome Res.* **14**: 1014–1024.
24. DuBois, K.N. *et al.* 2012. NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biol.* **10**: e1001287.
25. Daniels, J.P., K. Gull & B. Wickstead. 2010. Cell biology of the trypanosome genome. *Microbiol. Mol. Biol. Rev.* **74**: 552–569.
26. Fondon, J.W., 3rd *et al.* 2008. Simple sequence repeats: genetic modulators of brain function and behavior. *Trends Neurosci.* **31**: 328–334.
27. Gemayel, R. *et al.* 2010. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu. Rev. Genet.* **44**: 445–477.
28. Robinson, N.P. *et al.* 2002. Inactivation of Mre11 does not affect VSG gene duplication mediated by homologous recombination in *Trypanosoma brucei*. *J. Biol. Chem.* **277**: 26185–26193.
29. Hartley, C.L. & R. McCulloch. 2008. *Trypanosoma brucei* BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination. *Mol. Microbiol.* **68**: 1237–1251.
30. Horn, D. & R. McCulloch. 2010. Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Curr. Opin. Microbiol.* **13**: 700–705.
31. Barbet, A.F. & S.M. Kamper. 1993. The importance of mosaic genes to trypanosome survival. *Parasitol. Today* **9**: 63–66.
32. Caporale, L.H. 2006. *The Implicit Genome*. Oxford University Press, Oxford.
33. Futse, J.E. *et al.* 2005. Structural basis for segmental gene conversion in generation of *Anaplasma marginale* outer membrane protein variants. *Mol. Microbiol.* **57**: 212–221.
34. Parham, P., E.J. Adams & K.L. Arnett. 1995. The origins of HLA-A,B,C polymorphism. *Immunol. Rev.* **143**: 141–180.
35. Ohta, T. 2010. Gene conversion and evolution of gene families: an overview. *Genes* **1**: 349–356.
36. Graef, T. *et al.* 2009. KIR2DS4 is a product of gene conversion with KIR3DL2 that introduced specificity for HLA-A\*11 while diminishing avidity for HLA-C. *J. Exp. Med.* **206**: 2557–2572.
37. Ohshima, K. *et al.* 1996. TTA.TAA triplet repeats in plasmids form a non-H bonded structure. *J. Biol. Chem.* **271**: 16784–16791.
38. Labib, K. & B. Hodgson. 2007. Replication fork barriers: pausing for a break or stalling for time? *EMBO Rep.* **8**: 346–353.
39. Boothroyd, C.E. *et al.* 2009. A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature* **459**: 278–281.
40. Barbour, A.G. *et al.* 2006. Pathogen escape from host immunity by a genome program for antigenic variation. *Proc. Natl. Acad. Sci. USA* **103**: 18290–18295.
41. Futse, J.E. *et al.* 2008. Superinfection as a driver of genomic diversification in antigenically variant pathogens. *Proc. Natl. Acad. Sci. USA* **105**: 2123–2127.
42. Hutchinson, O.C. *et al.* 2007. Variant Surface Glycoprotein gene repertoires in *Trypanosoma brucei* have diverged to become strain-specific. *BMC Genomics* **8**: 234.
43. King, D.G. & Y. Kashi. 2007. Indirect selection for mutability. *Heredity* **99**: 123–124.
44. Jackson, A.P. *et al.* 2012. Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. *Proc. Natl. Acad. Sci. USA* **109**: 3416–3421.
45. Oliveira, B.M., M. Watkins, P. Bandyopadhyay, *et al.* 2012. Adaptive radiation of venomous marine snail lineages and the accelerated evolution of venom peptide genes. *Ann. N.Y. Acad. Sci.* **1267**: 61–70. This volume.