VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes

Glover, Lucy; Hutchinson, Sebastian; Alsford, Sam; Horn, David

Published in:
Proceedings of the National Academy of Sciences of the United States of America

DOI:
10.1073/pnas.1600344113

Publication date:
2016

Document Version
Peer reviewed version

VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes

Lucy Glover [1,3,†], Sebastian Hutchinson [1,†], Sam Alsford [2] and David Horn [1*]

[1] Division of Biological Chemistry & Drug Discovery, School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH, UK. [2] London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. [3] Current address: Institut Pasteur, 25-28 Rue du Docteur Roux 75015, Paris, France.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Allelic exclusion underpins antigenic variation and immune evasion in African trypanosomes. These bloodstream parasites employ RNA polymerase-I (pol-I) to transcribe just one telomeric variant surface glycoprotein (*VSG*) gene at a time, producing super-abundant and switchable VSG-coats. We identified trypanosome *VSG*-exclusion-1 (VEX1) using a genetic-screen for defects in telomere-exclusive expression. VEX1 was sequestered by the active *VSG* and silencing of other *VSGs* failed when VEX1 was either ectopically expressed or depleted, indicating positive and negative regulation, respectively. Positive regulation affected *VSGs* and non-telomeric pol-I transcribed genes while negative regulation primarily affected *VSGs*. Negative regulation by VEX1 also affected telomeric pol-I transcribed reporter constructs, but only when they contained blocks of sequence sharing homology with a pol-I transcribed locus. We conclude that restricted positive regulation due to VEX1 sequestration, combined with VEX1-dependent, possibly homology-dependent silencing, drives a 'winner-takes-all' mechanism of allelic exclusion.

epigenetic | monoallelic | silencing | telomere | Trypanosoma brucei

Introduction

Cells often restrict expression to a single allele of a gene or gene-family. This allelic exclusion underpins antigenic variation in pathogens, including trypanosomes that cause sleeping sickness (1) and *Plasmodium* parasites that cause malaria (2). Allelic exclusion is also essential for singular olfactory receptor expression and a sense of smell in metazoa (3). Although many factors have been identified that are required for the expression of one allele or for the silencing of other alleles in these systems, our understanding of the mechanisms by which expression and silencing are established and coordinated remains incomplete (1-3).

The African trypanosome, *Trypanosoma brucei*, is a flagellated parasitic protozoan transmitted among mammalian hosts by tsetse-flies. As well as causing trypanosomiasis in humans, a fatal and neglected tropical disease, these parasites also cause nagana in cattle. Antigenic variation is essential for persistent bloodstream infection in the face of host adaptive immune defenses and has long been a paradigm for studies on allelic exclusion (1); parasite immune evasion depends upon singular variant surface glycoprotein (VSG) gene expression and VSG-switching. While multiple subtelomeric *VSGs* are available for expression (4), only one is transcribed (5); both active and silent *VSGs* are located at the ends of polycistronic transcription units known as expression sites (ESs) (6). Notably, *VSG*-ES promoters (6) recruit RNA polymerase-I (pol-I) that typically transcribes ribosomal RNA genes (7). Indeed, the active *VSG*-ES is associated with an extranucleolar focus of pol-I known as the expression-site body (ESB) (8-10). Although the active *VSG*-ES is specifically depleted of nucleosomes (11, 12), silent *VSG*-ESs are similarly located in the extranucleolar space in bloodstream-form cells, and neither active or silent *VSG*-ESs show an appreciable association with the nuclear envelope (13). An HMG chromatin protein, that is enriched and inversely correlated with nucleosome abundance at the ESB and in the nucleolus (14), appears to maintain open chromatin at these sites (15). In addition, a highly SUMOylated focus is specific to the site of the ESB (16).

Pol-I transcription at the active *VSG*-ES, combined with attenuation at other *VSG*-ESs (17), allows trypanosomes to produce a single super-abundant VSG. Indeed, the active *VSG* generates the most abundant *T. brucei* mRNA and protein; the mRNA exceeds the next most abundant mRNA by >10-fold and approx. 10 million VSGs, constituting 10% of total cell protein (18), form a dense coat on each bloodstream-form cell (19). Antigenic variation itself occurs at low frequency and without immune selection (20), due to *VSG* rearrangement or coordinated transcription switching from one *VSG*-ES to another; the latter occurring in the absence of detectable change in the DNA sequence (1). Attempts to select for two simultaneously transcribed *VSG*-ESs indicated that double VSG-expression is highly unstable (21).

Several reports link chromatin, chromatin-associated proteins and telomere-binding proteins to *VSG* silencing (See SI Appendix, Table S1). For example, a histone H3 variant, a bloodstream stage-specific modified DNA base known as J or hydroxymethyluracil (22, 23), the chromatin remodeling ISWI complex (24), the histone H3K76 tri-methyltransferase DOT1B (25) and the telomere-associated protein RAP1 (26) all facilitate *VSG*-ES silencing. In addition, cohesin function facilitates maintenance of the active *VSG*-ES (27) and inositol phosphate signaling impacts *VSG*-ES regulation (28). Allelic exclusion though, requires the establishment of differential expression-states and coordination among members of a gene-family, which are not understood (1-3). In the case of *T. brucei*, it remains unclear how pol-I action is concentrated at one telomeric *VSG*.

Significance

Despite intense interest over a period of decades, mechanisms of allelic exclusion have remained unsolved mysteries in the field of eukaryotic gene expression control. Parasitic African trypanosomes express Variant Surface Glycoproteins (VSGs) in a monoallelic fashion and have long been a paradigm for studies in this area. We used an RNAi screen for loss of exclusion and identified and characterized VEX1 (VSG exclusion 1). VEX1 sequestration restricts expression and prevents the simultaneous establishment of more than one active *VSG* gene. VEX1 also appears to reinforce sequestration-based exclusion through homology-dependent repression. Our results indicate a 'winner-takes-all' mechanism that allows parasitic trypanosomes to express just one *VSG* gene at a time.

Reserved for Publication Footnotes

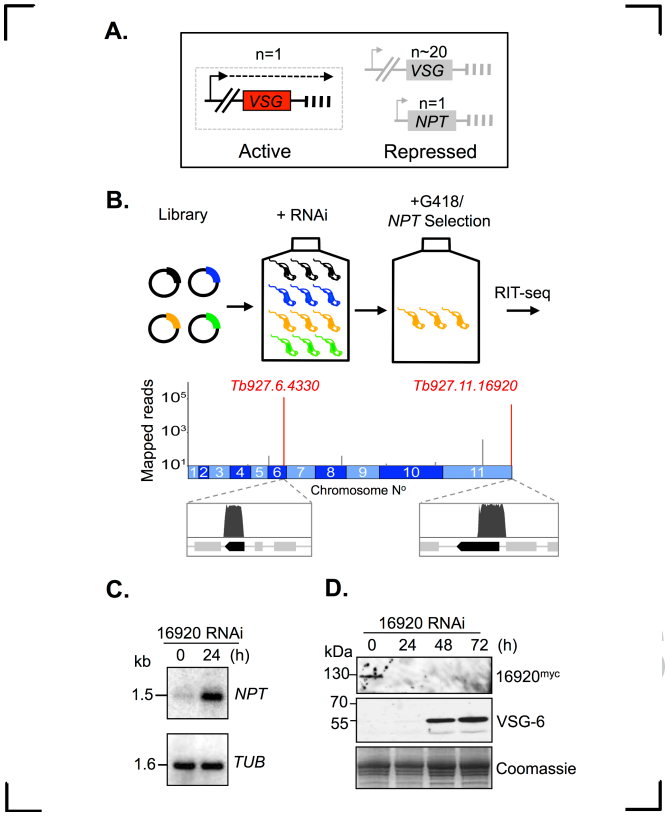

Fig. 1. A genetic screen reveals a candidate *VSG* exclusion regulator.(*A*) The bloodstream-form strain for RNAi-screening was based on a repressed *rDNA*-promoter / *Neomycin PhosphoTransferase (NPT)*-reporter cassette, with the promoter 2-kbp from the telomeric TTAGGG-repeats. Arrowheads, pol-I promoters; dashed line, transcription; vertical bars, telomeres.(*B*) The schematic shows the RNAi-screen for loss-of-exclusion (in orange cells) and the genome-map indicates hits following RIT-seq (red spikes). Mapped reads are indicated relative to gene-hits (dark bars).(*C*) Tb927.11.16920 knockdown was associated with *NPT* derepression as assessed by RNA blotting. *TUB* panel, loading-control.(*D*) Knockdown of myc-epitope-tagged Tb927.11.16920 was associated with *VSG-6* derepression as assessed by protein blotting. Coomassie-stained panel, loading-control.

Results

A genetic screen reveals Tb927.11.16920 as a candidate *VSG* exclusion regulator

To identify *T. brucei* genes that control telomere-exclusive gene expression, we assembled an RNA interference (RNAi) library in bloodstream-form trypanosomes with a pol-I transcribed telomeric reporter. The *NPT*-reporter, on a telomere-mediated chromosome-fragmentation construct (29), incorporates an *rDNA* promoter and seeds a *de-novo* telomere, comprising TTAGGG-repeats, approx. 2-kbp downstream (Fig. 1*A*). The *rDNA* promoter can be switched on and off through allelic exclusion when used to replace a native *VSG*-ES promoter (30) and is subject to repression when located close to a telomere (31). A reporter driven by an *rDNA* promoter was favored over a reporter driven by a *VSG*-ES promoter because defects in allelic exclusion were expected to result in a greater increase in *NPT*-reporter expression using the 'stronger' *rDNA* promoter (32). Since VSG expression is essential in bloodstream-form *T. brucei* (33), we also reasoned that *NPT*-activation, coupled to *VSG*-silencing during a 'telomere-switch', would fail to yield viable cells, as would knockdowns previously linked to *VSG*-silencing but associated with a severe growth-defect following RNAi (See SI Appendix, Table S1).

The population that emerged from the screen for defects in telomere-exclusive expression was subjected to RNAi target sequencing (RIT-seq), revealing two genes, Tb927.6.4330 and Tb927.11.16920, among approx. 7,400 in the genome (Fig. 1*B*). To determine their impact on *VSG* exclusion, we assembled pairs of independent RNAi knockdown strains for each gene in cells with an active *VSG-2* ES. Upon Tb927.11.16920 knockdown, we observed a moderate growth defect (See SI Appendix, Fig. S1*A*) and derepression of the telomeric *NPT*-reporter used in the screen (Fig. 1*C*), thereby validating this output from the screen. We confirmed efficient knockdown of myc-epitope-tagged Tb927.11.16920 (Fig. 1*D*) and saw derepression of silent telomeric *VSGs* using both protein blotting (Fig. 1*D*, VSG-6 panel) and microscopy (See SI Appendix, Fig. S1*B*). These findings indicated that the telomeric *NPT*-reporter was subject to the exclusion system operating in *T. brucei*. Analysis of Tb927.6.4430 supported previous reports of disrupted *VSG*-silencing when telomere structure and/or function are compromised (26, 34); knockdown of this novel telomeric TTAGGG repeat-associated factor (See SI Appendix, Fig. S2*A-B*) was associated with reporter and *VSG* derepression (See SI Appendix, Fig. S2*C-D*), but these phenotypes were substantially weaker than those observed following Tb927.11.16920 knockdown (Fig. 1*C-D*). Thus, Tb927.11.16920 emerged as the primary 'hit' in our screen for VSG exclusion regulators and we subsequently refer to this factor as VSG exclusion 1, or VEX1.

Like the majority of protein-coding genes identified in trypanosome genomes, Tb927.11.16920 / *VEX1* encodes a 'hypothetical protein' with no prior functional assignment. Analysis of the predicted peptide sequence revealed a 101.23 kDa protein incorporating a putative SWIM-type zinc-finger with a $CxCx^{17}CxH$ signature. Orthologous genes in other parasitic trypanosomatids also encode the zinc-finger motif (See SI Appendix, Fig. S3), originally found in SWI2/SNF2 family ATPase, MuDR transposase and MEK kinase (35), but these 'hypothetical proteins' also lack any prior functional assignment.

VEX1 is sequestered at the active *VSG*-ES

We next examined VEX1 subcellular localization. Epitope-tagged VEX1 was primarily concentrated in a single subnuclear focal compartment in bloodstream-form cells, as revealed by super-resolution microscopy (Fig. 2*A* and Video S1). The VEX1 compartment was extranucleolar (Fig. 2*B*) and closely associated, but not coincident, with the pol-I focus (Fig. 2*C* and Video S2) that is the site of the single active *VSG*-ES (10). Since all competent *VSG*-ESs are telomeric, we also used super-resolution microscopy to assess VEX1 localization relative to the telomeric TTAGGG repeat-binding factor, TRF2 (34). This revealed punctate nuclear TRF2-staining, and VEX1-staining that was coincident with a TRF2-punctum (Fig. 2*D*). VEX1 foci were observed at all cell-cycle phases, with segregated foci accumulating in G_2 in particular, following DNA replication (Fig. 2*E*). VEX1 foci were no longer detected, however, following a 30-minute exposure to the transcription inhibitor, actinomycin D.

VSG expression is inactivated during differentiation to the insect midgut-stage and reactivated in the insect salivary-gland (36). We observed a widespread punctate nuclear distribution of VEX1 in insect midgut-stage cells that substantially overlapped with telomeric TRF2-puncta (Fig. 2*F*). Thus, VEX1 is sequestered at the active *VSG*-ES, in a transcription-dependent and life-cycle stage-specific manner. VEX1 redistribution in insect-stage cells may allow *VSG*-ESs to compete for VEX1-sequestration as VSGs are reactivated in the insect salivary-gland (37).

VEX1 controls telomeric *VSG* exclusion

We next used microscopy to examine VSG-exclusion over a time-course following VEX1-knockdown in bloodstream-form cells. This analysis revealed an accumulation of cells simultaneously expressing VSG-2 and VSG-6, with no evidence for in-

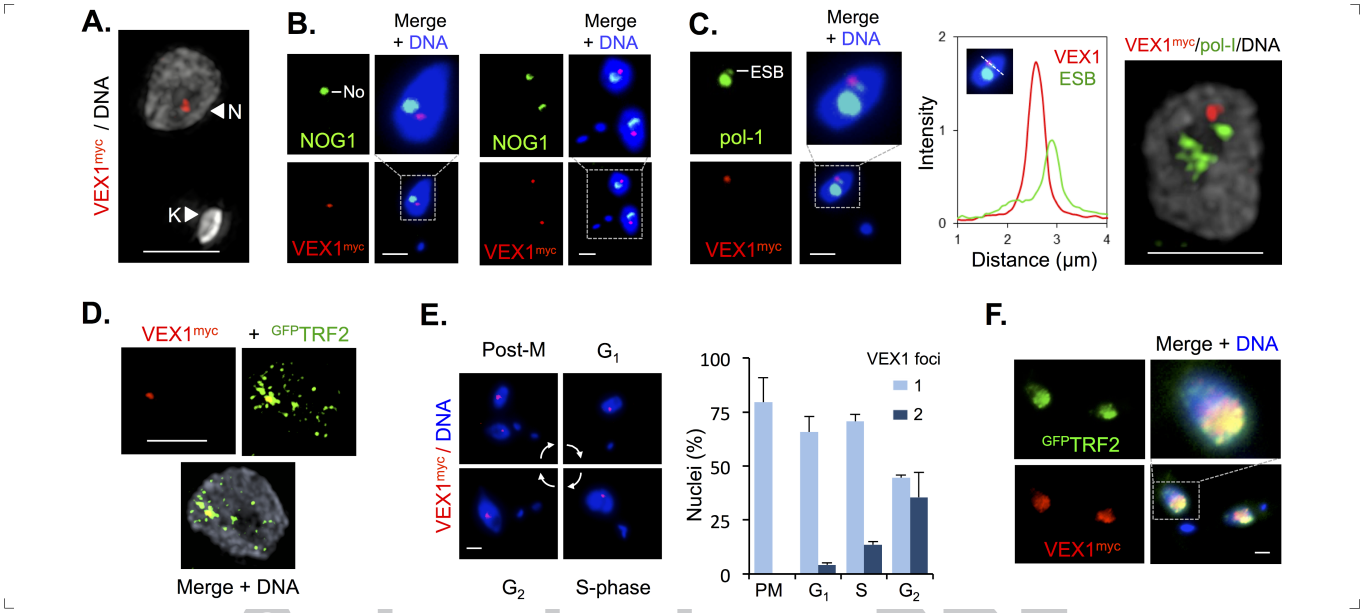

Fig. 2. VEX1 associates with the active *VSG*-ES in bloodstream-form cells. (*A*) 3D structured-illumination immunofluorescence microscopy (3D-SIM) projection of VEX1^{myc}. N, nucleus; K, kinetoplast (mitochondrial genome). (*B*) Immunofluorescence microscopy of VEX1^{myc} and a nucleolar (No) marker (NOG1). G_1 (left) and post-mitotic (right) cells are shown. (*C*) As B but with a nucleolar plus ESB marker (pol-I). The linear intensity plot shows the distance between the center of the VEX1-focus and the center of the ESB; mean distance is 0.34 +/-0.09 μm (n=7 G_1-nuclei). A 3D-SIM projection is shown on the right. (*D*) 3D-SIM projection of VEX1^{myc} and ^{GFP}TRF2. (*E*) Immunofluorescence microscopy of VEX1^{myc} during the cell cycle; phases are indicated and were inferred from DNA content. Numbers of VEX1-foci per nucleus were quantified at each cell cycle phase and are plotted on the right. (*F*) Immunofluorescence microscopy of VEX1^{myc} and ^{GFP}TRF2 in insect-stage cells. DNA was counter-stained with DAPI and all scale bars are 2 μm.

creased switching to VSG-6 (Fig. 3*A*). Flow-cytometry confirmed multi-VSG expression and again, no evidence for switching (Fig. 3*B*). To extend these findings, we carried out transcriptome analysis using pairs of wild-type sub-clones and pairs of independent knockdown strains; achieving >180 x average genome coverage in each RNA-seq experiment. Assessment of *VSG* transcript-abundance in wild-type cells revealed >1,000-fold differential between the active *VSG-2* transcript and the sum of all eighteen 'silent' pol-I promoter-associated *VSGs*; *VSG-2* transcripts represented approx. 7% of total mRNA. We also found that *VEX1* produced a low-abundance transcript in wild-type cells, within the lowest 5th percentile. Upon knockdown, *VEX1* was depleted 3.1-fold on average (Fig. 3*C* and Dataset S1), while nineteen genes (from approx. 7,400) displayed >3-fold and significantly (P<0.05) increased expression relative to wild-type. These included many bloodstream ES-linked *VSGs*, metacyclic ES-linked *VSGs* and genes immediately adjacent to *VSGs* (Fig. 3*C* and Dataset S1); metacyclic ES-linked *VSGs* are transcribed by pol-I in the insect salivary-gland (36). Indeed, expression of all eighteen 'silent' pol-I promoter-associated *VSGs* increased >26-fold overall (Dataset S1).

We next released cell-surface VSGs and used quantitative proteomics to examine relative expression (Fig. 3*D* and see SI Appendix, Materials & Methods and Table S2). VSG-2 on wild-type cells, displayed a relative abundance of 99.6%; only two significant sequences mapped to other VSGs. On VEX1-knockdown cells, the VSG-2 relative abundance was 91% and eleven 'silent' bloodstream and metacyclic ES-linked VSGs were also detected (Fig. 3*D* and see SI Appendix, Table S2). Thus, VEX1 knockdown allows for 'silent' VSGs to be transcribed, translated and delivered to the cell-surface.

Overexpressed VEX1 derepresses *VSG*-ESs and non-telomeric pol-I loci

While chromatin states may be effectively inherited, establishment of 'allele-choice' is not understood. The association between VEX1 and the active *VSG*-ES, and maintenance of the active *VSG*-ES following VEX1-knockdown, pointed to a potential role for VEX1 in establishing the active *VSG*-ES. To explore this hypothesis, we assembled a pair of independent VEX1 overexpressing bloodstream-form strains. We observed a moderate growth defect associated with VEX1 overexpression (See SI Appendix, Fig. S4*A-B*) and, as predicted, these cells failed to effectively sequester VEX1; although a VEX1-focus was often visible, we observed an additional dispersed signal through each nuclear compartment (Fig. 3*E*). Consistent with our hypothesis, when VEX1 was available to access a second telomeric *VSG*, this *VSG* was derepressed (See SI Appendix, Fig. S4*B*; VSG-6 panel). Indeed, immunofluorescence microscopy (Fig. 3*F*) and flow-cytometry (Fig. 3*G*) revealed cells simultaneously expressing both VSGs. In fact, the intensity of the cell-surface VSG-6 signal increased across the entire population of each clone and clones overexpressing VEX1 lacking a myc-tag yielded similarly increased VSG-6 expression (See SI Appendix, Fig. S4*C*).

Transcriptome analysis revealed sixty-five genes that displayed >3-fold and significantly (P<0.05) increased expression in VEX1 overexpresser strains relative to wild-type. *VEX1* mRNA was increased 16-fold on average and other up-regulated genes included all eighteen 'silent', pol-I promoter-associated *VSGs*, which increased >21-fold overall (Fig. 3*H* and Dataset S1). A striking difference compared to VEX1-knockdown was increased expression of fourteen *procyclin* and *procyclin*-associated transcripts, also >21-fold overall (Fig. 3*H*, Dataset S1 and see SI Appendix, Fig. S5*A*). These non-telomeric loci are also transcribed by pol-I, but normally produce abundant surface proteins in insect mid-gut stage cells (7). Another difference was increased expression of active (Dataset S1) and silent *VSG*-ES associated gene (*ESAG*) transcripts following VEX1 overexpression, in contrast to increased expression of the *VSGs* only following VEX1-knockdown (See SI Appendix, Fig. S5*B*). Quantitative proteomic analysis of surface-VSGs on VEX1 overexpressing cells revealed a VSG-2 abundance index reduced to 91%, with twelve additional bloodstream and metacyclic ES-linked VSGs also detected (Fig.

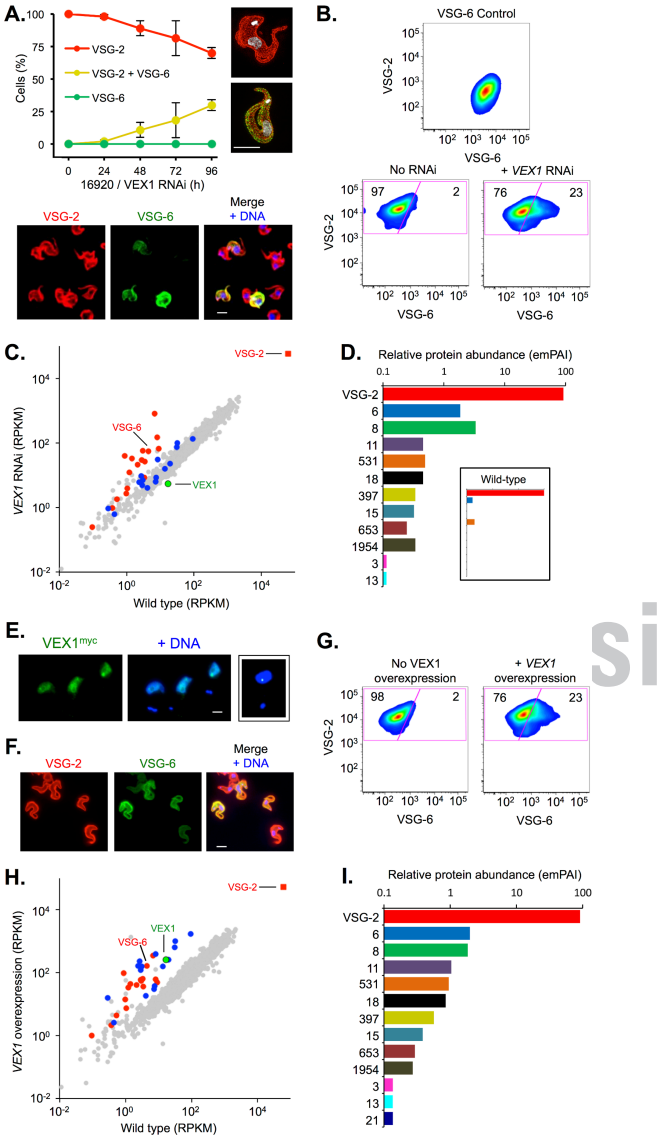

Fig. 3. VEX1 controls *VSG* allelic exclusion in bloodstream-form cells.(*A*) Immunofluorescence microscopy analysis of VSG-expression. were stained with α-VSG-2 and α-VSG-6 and counted daily during VEX1-RNAi. 3D-SIM images show a wild-type control cell and a cell expressing both VSGs. The images below the plot show cells following VEX1-RNAi (72 h). Scale bars, 5 μm.(*B*) Flow-cytometry analysis of VSG-expression following VEX1-RNAi (72 h). Numbers indicate percentage of cells in each quadrant. VSG-6 expressers serve as a control.n=10,000 cells in each case. (*C*)RNA-seq analysis following VEX1-RNAi (72 h). Values are averages for a pair of independent strains (see Dataset S1); red circles, 'silent' *VSGs*; red square, active *VSG*; blue, *procyclins* and *procyclin*-associated genes. RPKM, Reads Per Kilobase of transcript per Million mapped reads.(*D*) Quantitative mass spectrometry analysis of surface-VSGs following VEX1-RNAi (72 h, see SI Appendix, Table S2). The inset shows wild-type cells for comparison. emPAI, exponentially modified Protein Abundance Index. (*E*) Immunofluorescence microscopy of overexpressed and ectopicVEX1^{myc} (72 h). Scale bar, 2 μm. The panel to the right shows sequestered VEX1 for comparison. (*F*) Immunofluorescence microscopy analysis of VSG expression following VEX1-overexpression (72 h). Scale bar, 5 μm. (*G*)Flow-cytometry following VEX1-overexpression (72 h). Other detail as in B. (*H*)RNA-seq analysis following VEX1 overexpression (72 h). Other detail as in C. (*I*) Quantitative mass spectrometry analysis of surface-VSGs following VEX1-overexpression (72 h, see SI Appendix, Table S2).

3*I* and see SI Appendix, Table S2). We conclude that over-expressed VEX1 positively regulates both telomeric and non-telomeric pol-I transcribed loci.

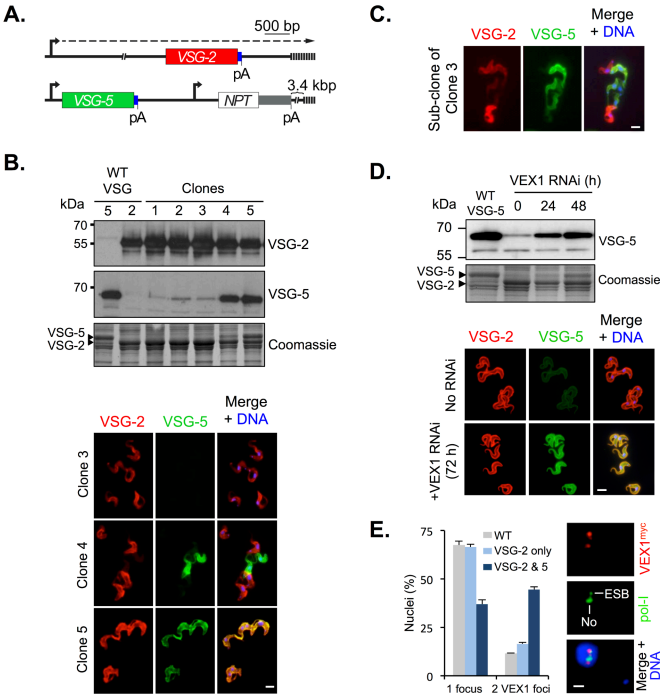

Fig. 4. VEX1-dependent communication among recombinant and native *VSGs*.(*A*) The schematic shows the *VSG-5* reporter in cells expressing native *VSG-2*. Blue boxes, common *VSG* 3´-UTRs; pA, poly-adenylation sites; other symbols as in Fig. 1A. (*B*) The protein blots show five bloodstream-form clones derived using this system. *VSG-5* or *VSG-2* expressing wild-type (WT) cells serve as controls. The Coomassie-stained panel serves as a loading-control and also reveals the major VSGs. The immunofluorescence panels show VSG expression in clones 3-5. (*C*) Immunofluorescence analysis of VSG expression in a sub-clone of clone-3 following G418/*NPT*-selection (see A). (*D*) The protein blots and immunofluorescence microscopy show *VSG-5* expression during VEX1-knockdown. Controls as in A.(*E*) Immunofluorescence microscopy analysis of VEX1^{myc} in cells equivalent to clone 5 (expressing both VSG-2 and VSG-5). Nuclei with 1 or 2 VEX1-foci were quantified relative to wild-type (WT) and cells equivalent to clone 3 above (expressing VSG-2 only). Immunofluorescence analysis of VEX1^{myc} and pol-I reveals the location of the additional VEX1-focus relative to the ESB and nucleolus (No). The mean distance between the centers of the VEX1-foci is 1.1 +/-0.3 μm (n=8 G_1-nuclei). Scale bars, 5 μm except E; scale bar, 1 μm. DNA was counter-stained with DAPI.

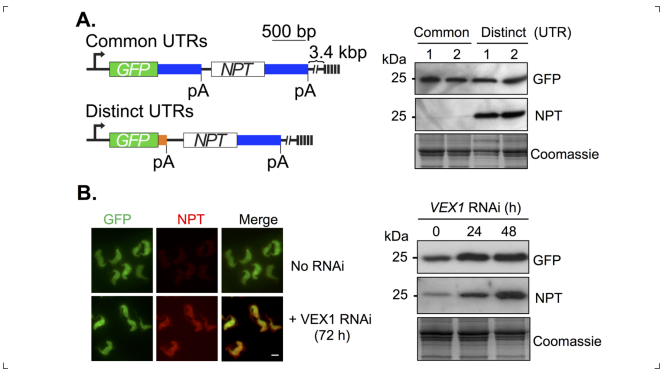

Fig. 5. VEX1-dependent communication among homologous sequences. (*A*) The schematic shows reporters with common or distinct 3'-UTRs (*aldolase*, blue; *tubulin*, orange). The protein blots show reporter expression for pairs of bloodstream-form clones derived using each construct.Coomassie-stained panel, loading-control. Symbols as in Fig. 4A.(*B*) The immunofluorescence panels and protein blots show expression of reporters with common 3'-UTRs during VEX1-knockdown. Scale bar, 5 μm.

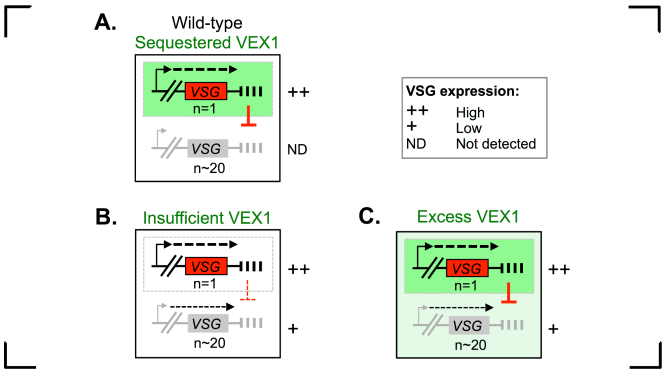

Fig. 6. A 'winner-takes-all' model for allelic exclusion by VEX1. (*A*) Wild-type bloodstream-form cells: VEX1 coordinates *VSG* positive and negative regulation. VEX1 (green) is recruited to a single *VSG*-ES through positive-feedback involving pol-I transcription. VEX1 also exerts negative regulation through homology-dependent silencing (red symbol), which enhances differential expression and ensures that the 'winner-takes-all'. (*B*) VEX1 knockdown: Insufficient VEX1 leads to relaxed homology-dependent silencing but also precludes further positive regulation. (*C*) VEX1 overexpression: Ectopic VEX1 positively regulates all *VSG*-ESs but negative regulation also persists and precludes further derepression.

A *VSG* reporter is subject to VEX1-dependent exclusion

VEX1 was identified due to its ability to repress a reporter on a telomere-mediated chromosome-fragmentation construct (Fig. 1). We reasoned that manipulation of these constructs would allow us to identify the DNA sequences involved. A new construct was assembled that contained a *VSG-5* gene with a downstream sequence (including the 3' untranslated region, 3'-UTR) that is identical to that associated with the native active *VSG-2* (Fig. 4*A*). This construct places a *VSG-5* associated *rDNA* promoter approx. 7-kbp from a *de-novo* telomere (Fig. 4*A*). Consistent with competition between VEX1-mediated positive and negative regulation, we observed variable outcomes using this construct in bloodstream-form cells; all cloned populations continued to express VSG-2, but VSG-5 expression differed. Among twenty-four independent clones, twenty displayed repressed VSG-5 and four are shown (Fig. 4*B*); three displayed uniform repression (>99% of cells, clones 1-3) while some cells in one clone displayed *VSG-5* expression that interfered with native *VSG-2* expression (clone 4). The remaining clones displayed simultaneous expression of both VSGs (>99% of cells) that was stable over many generations and yielded approximately equal quantities of each VSG (Fig. 4*B*, clone 5 and see SI Appendix, Fig. S6).

To determine whether VSG-5 repression was reversible, clone 3 cells were sub-cloned and selected for increased *NPT*-reporter expression (see Fig. 4*A*). These populations expressed VSG-5 that interfered with VSG-2 expression (Fig. 4*C*), indicating that *VSG-5* was subject to the exclusion system, being both repressible and able to repress. Also, VEX1-knockdown led to *VSG-5* derepression and produced cells simultaneously expressing both VSGs (Fig. 4*D*). This VEX1-dependent *VSG-5* reporter repression was confirmed using a second independent strain.

Simultaneous expression of two *VSGs* (Fig. 4*B*, clone 5) indicated that *VSG-5* could escape exclusion. Indeed, simultaneous *VSG* expression, apparently at distinct telomeres, has been reported before (38). Since VEX1 is concentrated at the active *VSG*-ES (Fig. 2) and can positively regulate pol-I loci (Figs. 3*E-I*), we wondered whether *VSG*-5 at a *de novo* telomere might compete for VEX1 and, in some cases, acquire sufficient material to establish a second 'privileged' site. As suspected, the proportion of nuclei with two distinct VEX1-foci was increased in cells equivalent to clone 5 (3.8-fold relative to 'wild-type' control; Fig. 4*E*) and the additional VEX1-focus associated with VSG-5 expression was typically ESB-distal and perinucleolar (Fig. 4*E*); *rDNA* loci are also perinucleolar (39). Taken together with analysis of VEX1-overexpressers (Fig. 3*H*), these findings (Fig. 4*B*, 4*E*) suggest that access to VEX1 increases pol-I transcription, that increased pol-I transcription can lead to the accumulation of VEX1, or both of the above. This is also consistent with the observation of a 'pre-active' *VSG* state when a *VSG* is in close proximity to the active *VSG*-ES (21).

VEX1-mediates a form of homology-dependent silencing

A serendipitous observation initially suggested to us that 'homologous' DNA sequences might be important for exclusion. We had assembled a bicistronic and telomeric *GFP:NPT* reporter that fortuitously contained identical sequences downstream of both the *GFP* and *NPT* genes (Fig. 5*A*, upper map). These sequences contain 3'-UTRs that are included in reporter constructs to guide mRNA polyadenylation and splicing; we did not expect them to be subject to repression or exclusion. However, we observed strong *NPT*-repression when using this construct in bloodstream-form cells (Fig. 5*A*, left-hand tracks). To determine whether this reflected homology-dependent interference, the *T. brucei aldolase* sequence downstream of the *GFP* gene was replaced with a *T. brucei tubulin* sequence (Fig. 5*A*, lower map), which is unrelated to the *aldolase* sequence but also guides efficient mRNA processing. No *NPT*-repression was observed when using this construct (Fig. 5*A*, right-hand tracks). In cells containing the construct with common *aldolase* sequences, *NPT*-repression was relieved following VEX1-knockdown (Fig. 5*B*); a result confirmed using a second independent strain. Thus, reporters with homologous sequences downstream displayed VEX1-dependent repression that no longer operated when homology was removed.

Discussion

We identified trypanosome VEX1 using a genetic screen for defects in telomere-exclusive gene expression. VEX1 is sequestered at the active *VSG*-ES and coordinates *VSG* positive and negative regulation to sustain antigenic variation. Based on our findings, we propose a 'winner-takes-all' model for the establishment of allelic exclusion by VEX1 (Fig. 6*A*). A similar model has been put forward for olfactory receptor gene choice (40) although no factor that displays similar properties to VEX1 has been identified in that system. Our results indicate that an established active *VSG*-ES is effectively inherited when VEX1-function is disrupted; this also appears to be the case when (telomeric) chromatin is disrupted by other means (22-24, 26, 34). One long-standing question, however, has been why are 'silent' *VSG*-ESs only partially derepressed when (telomeric) chromatin is disrupted? This is the case even when substantial loss-of-viability is observed following RAP1 knockdown, for example (26) and could be explained by failure to associate with sufficient transcription or RNA-processing factors. We suggest that, when VEX1-function is disrupted, directly or indirectly, negative regulation is relaxed but that silent *VSG*-ESs lack VEX1-mediated positive regulation (Fig. 6*B*). Indeed, *VSG*-ES promoters appear to be substantially 'weaker' than *rDNA* promoters (32) and may depend upon positive regulation by VEX1. In the case of excess VEX1, access to other *VSG*-ESs allows positive regulation but these sites are still subject to negative regulation exerted by the active *VSG*-ES (Fig. 6*C*). Thus, we suggest that the 'default' level for *VSG* expression is relatively low and that VEX1 drives the processes that increase expression at one locus and reduce expression elsewhere. We note that although a "winner-takes-all" mechanism may operate naturally, this can be perturbed either when VEX1 is artificially expressed in excess or, in some cases, when recombinant pol-I transcription units are introduced *de novo*.

We previously demonstrated that repression of pol-I transcribed genes spreads only a short distance from the telomere in the absence of additional *VSG*-ES sequences (31). We now

show that this repression is VEX1-dependent and also that homologous sequences can promote VEX1-dependent repression. Another recent report demonstrated that a *VSG*, transcribed at a chromosome-internal site by T7-phage polymerase, transiently silenced the native *VSG* (25). This same report demonstrated repression spreading along the *VSG*-ES in a DOT1B, histone methyltransferase-dependent manner (25). Spreading of a repressed domain may also explain why we see derepression of both bicistronic *GFP* and *NPT* reporters during VEX1-knockdown (Fig. 5*B*). In these experiments, sequences that are subject to repression also serve as repressive sequences when they are transcribed. *VSG*-transcripts (See SI Appendix, Fig. S7) and other *VSG*-associated sequences display a high degree of homology, and telomeric TTAGGG-repeat transcripts are also present in *T. brucei* (41), suggesting that transcripts could be involved. Further work will be required to delineate the mechanism, which although not involving Argonaute1-based RNAi (42), could involve alternative RNA-based repression, as reported in other cell-types (43). Thus, we tentatively suggest that VEX1-dependent *VSG*-silencing is initiated by homologous transcripts and is then propagated along the chromatin fibre in a DOT1B-dependent manner.

In summary, we report the identification of VEX1, an allelic exclusion regulator that sustains antigenic variation in trypanosomes. We describe a 'winner-takes-all' model whereby VEX1 sequestration establishes a single active *VSG*-ES that then mediates homology-dependent silencing at other *VSG*-ESs. Similar mechanisms involving positive and negative regulation, coordinated by sequestered regulators, could explain allelic exclusion in other cell-types.

Materials and Methods

For details of *T. brucei* growth and manipulation, plasmids, nucleic acid analysis, western blotting, microscopy, flow cytometry and quantitative mass spectrometry see (See SI Appendix, Materials and Methods).

T. brucei, Lister 427, MITat1.2 (VSG-2, aka VSG221), 1.5 (VSG-5, aka VSG118) and 1.6 (VSG-6, aka VSG121) cells were used for this study. RIT-seq was carried out on a MiSeq platform (Illumina) at BGI (The Beijing Genome Institute). RNA-seq was carried out on a HiSeq platform (Illumina) at the University of Dundee or at BGI. 3D structured illumination microscopy was carried out using a super-resolution OMX Blaze system (GE Healthcare). Quantitative mass spectrometry was carried out using an Ultimate 3000 RSLCnano system (Thermo Scientific) coupled to a Linear Trap Quadropole OrbiTrap Velos Pro (Thermo Scientific).

Acknowledgments:.

We would like to acknowledge R. Clark of the Flow Cytometry and Cell Sorting Facility, which is supported by the Wellcome Trust (097418/Z/11/Z); M. Posch and B. Balagopal for assistance with SIM microscopy; use of the OMX microscope was funded by an MRC Next Generation Optical Microscopy Award (MR/K015869/1); M. Febrer for assistance with Illumina sequencing; and D. Lamont & K. Beattie of the Fingerprints Proteomics Facility for assistance with quantitative proteomics. We also thank A. Mehlert, L. Guther and M. Ferguson for assistance and advice on VSG proteomics; J. Morris and P. Englund for the RNAi plasmid library; B. Dujon for the I-*Sce*I gene; and M. Ferguson, J. Wright and M. Field for comments on the manuscript. The work was funded by a project grant to D.H. (093010/B/10/Z), an Investigator Award to D.H. (100320/Z/12/Z) and a Strategic Award supporting Biological Chemistry & Drug Discovery (100476/Z/12/Z), all from The Wellcome Trust.

1. Horn D (2014) Antigenic variation in African trypanosomes. *Mol Biochem Parasitol* 195(2):123-129.
2. Guizetti J & Scherf A (2013) Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol* 15(5):718-726.
3. Monahan K & Lomvardas S (2015) Monoallelic Expression of Olfactory Receptors. *Annu Rev Cell Dev Biol* 31:721-740.
4. Hertz-Fowler C, *et al.* (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS One* 3(10):e3527.
5. De Lange T & 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(5882):451-453.
6. Zomerdijk JC, *et al.* (1990) The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *Embo J* 9(9):2791-2801.
7. Gunzl A, *et al.* (2003) RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryot Cell* 2(3):542-551.
8. Borst P (2002) Antigenic variation and allelic exclusion. *Cell* 109(1):5-8.
9. 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(21):12328-12333.
10. Navarro M & Gull K (2001) A pol I transcriptional body associated with *VSG* mono-allelic expression in *Trypanosoma brucei*. *Nature* 414(6865):759-763.
11. Figueiredo LM & Cross GA (2010) Nucleosomes are depleted at the *VSG* expression site transcribed by RNA polymerase I in African trypanosomes. *Eukaryot Cell* 9(1):148-154.
12. Stanne TM & Rudenko G (2010) Active *VSG* expression sites in *Trypanosoma brucei* are depleted of nucleosomes. *Eukaryot Cell* 9(1):136-147.
13. Navarro M, Penate X, & Landeira D (2007) Nuclear architecture underlying gene expression in *Trypanosoma brucei*. *Trends Microbiol* 15(6):263-270.
14. Narayanan MS & Rudenko G (2013) TDP1 is an HMG chromatin protein facilitating RNA polymerase I transcription in African trypanosomes. *Nucleic Acids Res* 41(5):2981-2992.
15. Aresta-Branco F, Pimenta S, & Figueiredo LM (2015) A transcription-independent epigenetic mechanism is associated with antigenic switching in *Trypanosoma brucei*. *Nucleic Acids Res*.
16. Lopez-Farfan D, Bart JM, Rojas-Barros DI, & Navarro M (2014) SUMOylation by the E3 ligase TbSIZ1/PIAS1 positively regulates *VSG* expression in *Trypanosoma brucei*. *PLoS Pathog* 10(12):e1004545.
17. Kassem A, Pays E, & Vanhamme L (2014) Transcription is initiated on silent variant surface glycoprotein expression sites despite monoallelic expression in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 111(24):8943-8948.
18. Cross GA (1975) Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology* 71(3):393-417.
19. Vickerman K (1969) On the surface coat and flagellar adhesion in trypanosomes. *J Cell Sci* 5(1):163-193.
20. Doyle JJ, Hirumi H, Hirumi K, Lupton EN, & Cross GA (1980) Antigenic variation in clones of animal-infective *Trypanosoma brucei* derived and maintained *in vitro*. *Parasitology* 80(2):359-369.
21. Chaves I, Rudenko G, Dirks-Mulder A, Cross M, & Borst P (1999) Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *Embo J* 18(17):4846-4855.
22. Reynolds D, *et al.* (2016) Histone H3 variant regulates RNA polymerase II transcription termination and dual strand transcription of siRNA loci in *Trypanosoma brucei*. *PLoS Genet* 12(1):e1005758.
23. Schulz D, Zaringhalam M, Papavasiliou FN, & Kim HS (2016) Base J and H3.V regulate transcriptional termination in *Trypanosoma brucei*. *PLoS Genet* 12(1):e1005762.
24. Stanne T, *et al.* (2015) Identification of the ISWI chromatin remodeling complex of the early branching eukaryote *Trypanosoma brucei*. *J Biol Chem*.
25. Batram C, Jones NG, Janzen CJ, Markert SM, & Engstler M (2014) Expression site attenuation mechanistically links antigenic variation and development in *Trypanosoma brucei*. *Elife* 3:e02324.
26. Yang X, Figueiredo LM, Espinal A, Okubo E, & Li B (2009) RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. *Cell* 137(1):99-109.
27. Landeira D, Bart JM, Van Tyne D, & Navarro M (2009) Cohesin regulates *VSG* monoallelic expression in trypanosomes. *J Cell Biol* 186(2):243-254.
28. Cestari I & Stuart K (2015) Inositol phosphate pathway controls transcription of telomeric expression sites in trypanosomes. *Proc Natl Acad Sci U S A* 112(21):E2803-2812.
29. Horn D, Spence C, & Ingram AK (2000) Telomere maintenance and length regulation in *Trypanosoma brucei*. *Embo J* 19(10):2332-2339.
30. Rudenko G, Blundell PA, Dirks-Mulder A, Kieft R, & Borst P (1995) A ribosomal DNA promoter replacing the promoter of a telomeric *VSG* gene expression site can be efficiently switched on and off in *T. brucei*. *Cell* 83(4):547-553.
31. Glover L & Horn D (2006) Repression of polymerase I-mediated gene expression at *Trypanosoma brucei* telomeres. *EMBO Rep* 7(1):93-99.
32. Horn D & Cross GA (1997) Position-dependent and promoter-specific regulation of gene expression in *Trypanosoma brucei*. *Embo J* 16(24):7422-7431.
33. Sheader K, *et al.* (2005) Variant surface glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African trypanosomes. *Proc Natl Acad Sci U S A* 102(24):8716-8721.
34. Jehi SE, *et al.* (2014) Suppression of subtelomeric *VSG* switching by *Trypanosoma brucei* TRF requires its TTAGGG repeat-binding activity. *Nucleic Acids Res* 42(20):12899-12911.
35. Makarova KS, Aravind L, & Koonin EV (2002) SWIM, a novel Zn-chelating domain present in bacteria, archaea and eukaryotes. *Trends Biochem Sci* 27(8):384-386.
36. Kolev NG, Ramey-Butler K, Cross GA, Ullu E, & Tschudi C (2012) Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. *Science* 338(6112):1352-1353.
37. Tetley L, Turner CM, Barry JD, Crowe JS, & Vickerman K (1987) Onset of expression of the variant surface glycoproteins of *Trypanosoma brucei* in the tsetse fly studied using immunoelectron microscopy. *J Cell Sci* 87 (Pt 2):363-372.
38. Baltz T, *et al.* (1986) Stable expression of two variable surface glycoproteins by cloned *Trypanosoma equiperdum*. *Nature* 319(6054):602-604.
39. Landeira D & Navarro M (2007) Nuclear repositioning of the *VSG* promoter during developmental silencing in *Trypanosoma brucei*. *J Cell Biol* 176(2):133-139.
40. Hanchate NK, *et al.* (2015) Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science* 350(6265):1251-1255.
41. Rudenko G & Van der Ploeg LH (1989) Transcription of telomere repeats in protozoa. *Embo J* 8(9):2633-2638.
42. Janzen CJ, *et al.* (2006) Expression site silencing and life-cycle progression appear normal in Argonaute1-deficient *Trypanosoma brucei*. *Mol Biochem Parasitol* 149(1):102-107.
43. Holoch D & Moazed D (2015) RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* 16(2):71-84.