nature research

Peer Review Information

Journal: Nature Microbiology
Manuscript Title: Spatial integration of transcription and splicing in a dedicated compartment sustains monogenic antigen expression in African trypanosomes
Corresponding author name(s): Professor T Siegel

Reviewer Comments & Decisions:

Decision Letter, initial version:

Dear Prof Siegel,

Thank you for submitting your Letter entitled "Antigenic variation by switching inter-chromosomal interactions with an RNA splicing locus in trypanosomes" for consideration in Nature Microbiology and please accept our apologies for the time it has taken us to contact you with a decision on your manuscript, which is due to our current high submission volume. I regret to inform you that after careful discussion within the editorial team, we have decided that we cannot consider it for publication here.

As you may know, we decline a substantial proportion of manuscripts without sending them to referees, so that they may be sent elsewhere without delay. In such cases, even if referees were to certify the manuscript as technically correct, we consider that the work does not represent the type of advance that Nature Microbiology seeks to publish. This editorial assessment is based on considerations such as the degree of conceptual advance provided, the breadth of potential interest to researchers and timeliness.

In this case, we certainly appreciate the finding that the single expressed antigen coding gene displays a specific inter-chromosomal interaction with a major mRNA splicing locus in trypanosomes, and the demonstration that this interaction is heritable and splicing dependent, as well as the finding that the two genomic loci are connected by the antigen exclusion complex, with VEX1 associating with the splicing locus and VEX2 with the antigen coding locus. We have no doubt that this will be of interest to those working closer to the field. Nevertheless, given previous links between gene expression, 3D-genome architecture and local DNA accessibility, and your previous identification and characterization of VEX1 and VEX2, together with the lack of sufficient novel mechanistic insights into the assembly/regulation of these complexes, in terms of the overall degree of advance and immediate relevance to our broad microbiological readership, we do not feel that the current work has quite met the high bar that we must unfortunately set for further consideration towards publication in Nature Microbiology. We therefore feel that the paper would find a more suitable outlet in another journal.

Although we cannot offer to publish your manuscript, I suggest that you consider transferring your manuscript to the Springer Nature journal Nature Communications, which is the Nature Research flagship Open Access journal. If you would like this work to be considered for publication there, you can easily transfer the manuscript by following the instructions below. It is not necessary to reformat your paper. Once all files are received, the editors at Nature Communications will assess your manuscript's suitability for potential publication; they aim to provide feedback quickly, with a median decision time of 8 days for first editorial decisions on suitability. The journal is also proud to offer double blind and transparent peer review options. Our open access pages contain information about article processing charges, open access funding, and advice and support from Springer Nature.

Please be assured that this editorial decision does not represent a criticism of the quality of your work, nor are we questioning its value to others working in this area. We hope that you will rapidly receive a more favourable response elsewhere.

I am sorry that we cannot respond more positively on this occasion.

Author Rebuttal to Initial comments

Dear Dr. Mayer,

Thank you very much for considering our manuscript and also for providing us a prompt decision.

We are very surprised to read that you consider our current findings not of broad enough interest to the Nature Microbiology readership. When I talked to Cláudio Nunes-Alves at the Gordon Conference in New Port, RI, in 2018, he gave me the impression that Nature Microbiology is very interested in antigenic variation. He even encouraged me to submit our previous manuscript on histone variants and genome architecture to Nature Microbiology, in case it was rejected by Nature.

Based on citations and coverage by text books, there is no other topic that is of similar general interest than antigenic variation in *Trypanosoma brucei*. This broad interest is reflected by the considerable amount of attention that our current manuscript has received on bioRxiv, including coverage by a prelights article (prelights.biologists.com/highlights/antigenic-variation-by-switching-inter-chromosomal-interactions-with-an-rna-splicing-locus-in-trypanosomes/). Understanding antigenic variation in trypanosomes has been a 50-year quest of many groups across the world. Our recent findings finally reveal the mechanism allowing the parasite to ensure monogenic antigen expression, a key requirement for its virulence. Indeed, our findings have been received with great excitement by our peers in the parasitology and chromatin communities at various scientific meetings.

Based on your feedback, we suspect that we have not stated clearly how our new findings differ from our published results. We would like to apologize for this lack of clarity and have tried to better outline the importance of our recent findings in the following paragraphs.

As you know, antigenic variation – the capacity of an infecting organism to systematically alter the identity of proteins displayed to the host immune system – plays a key role in diseases caused by some of the deadliest human parasites such as *Plasmodium falciparum* and *T. brucei*.

The two key aspects of antigenic variation are 1) the mutually exclusive expression of one antigen isoform and 2) the periodic switching of expression of one antigen isoform to another.

While our previous story on histone variants, genome architecture and antigenic variation, certainly contained important new insights with regards to the *T. brucei* genome and antigen switching, we were not able to elucidate if and how genome architecture contributes to monogenic antigen expression. It is only this current study that reveals a very elegant mechanism of sustaining monogenic antigen expression: by bringing one VSG gene into close proximity to the SL RNA locus (an RNA processing hot-spot) to ensure high mRNA processing. This is a fundamentally new concept of gene regulation. In this case, all other VSG genes are actively excluded from the interaction with the RNA processing hotspot by this VEX2 protein, which thereby ensures monogenic antigen expression. Previously, we had identified VEX2 and demonstrated its role in maintaining monogenic expression, but only now we demonstrate how - VEX2 regulates monogenic antigen expression by restricting the access to RNA maturation.

As you pointed out, we do not show how the VSG and the SL-RNA transcription compartments are assembled. However, we have elucidated a much more fundamental aspect of gene regulation: the

possibility that spatial proximity of a gene with an RNA processing hotspot can serve as a mechanism to tightly regulate the expression of antigens. Additionally, a role for stable inter-chromosomal interactions on gene expression control has long been questioned but our work clearly shows that they can indeed 1) be stably propagated, 2) impact gene expression.

In lower eukaryotes, the importance of genome architecture for gene regulation has been questioned mainly because canonical enhancers are missing. We believe that this study will change this view as our data show that close spatial proximity with RNA-hotspots can serve as enhancer-like mechanism that allow enhancement of gene expression at a post-transcriptional level. Such a novel 'strategy' of gene expression enhancement would certainly have a large impact, far beyond the *T. brucei* research community, as similar mechanisms are likely to be found in other lower eukaryotes.

Thus, we believe this manuscript is ideally suited for publication in Nature Microbiology and kindly ask you to reconsider the possibility of having the manuscript peer-reviewed.

Decision Letter, first revision:

Dear Nicolai,

Thank you for your patience while your manuscript "Antigenic variation by switching inter-chromosomal interactions with an RNA splicing locus in trypanosomes" was under peer-review at Nature Microbiology. It has now been seen by 4 referees, whose expertise and comments you will find at the end of this email. Although they find your work of some potential interest, they have raised a number of concerns that will need to be addressed before we can consider publication of the work in Nature Microbiology.

In particular, referee 3 felt that it is not clear what the consequence of loss of this interchromosomal interaction is, and that interpretation of the VEX2 knockdown phenotypes need to be done more carefully. Furthermore, referee 3 felt that "the manuscript is quite descriptive at the current stage.", and asks "what is the mechanism of VEX2's role? How does VEX2 help to locate the VSG transcription site near the RNA processing center?" Referee 4 asks if and how the VSG gene is more spliced than the ESAG genes, and also asks why the silent VSGs can't access the other SL-RNA array that also has a VEX1 foci. These concerns, together with the further points raised by the referees would need to be addressed before we would consider sending the manuscript back to the referees.

Editorially we remain very interested in your study, and should further experimental data allow you to address these criticisms, we would be happy to look at a revised manuscript.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

We strongly support public availability of data. Please place the data used in your paper into a public data repository, if one exists, or alternatively, present the data as Source Data or Supplementary Information. If data can only be shared on request, please explain why in your Data Availability Statement, and also in the correspondence with your editor. For some data types, deposition in a public repository is mandatory - more information on our data deposition policies and available repositories can be found at https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data.

Please include a data availability statement as a separate section after Methods but before

references, under the heading "Data Availability". This section should inform readers about the availability of the data used to support the conclusions of your study. This information includes accession codes to public repositories (data banks for protein, DNA or RNA sequences, microarray, proteomics data etc…), references to source data published alongside the paper, unique identifiers such as URLs to data repository entries, or data set DOIs, and any other statement about data availability. At a minimum, you should include the following statement: "The data that support the findings of this study are available from the corresponding author upon request", mentioning any restrictions on availability. If DOIs are provided, we also strongly encourage including these in the Reference list (authors, title, publisher (repository name), identifier, year). For more guidance on how to write this section please see:
http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf

If revising your manuscript:

* Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Letter format instructions at http://www.nature.com/nmicrobiol/info/final-submission. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

Please use the link below to submit a revised paper:

{REDACTED]

Note: This url links to your confidential homepage and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this e-mail to co-authors, please delete this link to your homepage first.

Nature Microbiology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision, even if a similar study has been accepted for publication at Nature Microbiology or published elsewhere (up to a maximum of 6 months). Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date would be that of the revised, not the original, version.

In the meantime we hope that you find our referees' comments helpful.

****

Reviewer Expertise:

Referee #1: Parasitology
Referee #2: Molecular cell biology of trypanosomes
Referee #3: Recombination in trypanosomes
Referee #4: Genome organization

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

The manuscript entitled " Antigenic variation by switching inter-chromosomal interactions with an RNA splicing locus in trypanosomes" reports on the discovery of a novel mechanism to ensure monogenic expression in Trypanosoma brucei. Monogenic expression is a key feature of antigenic variation, an immune evasion strategy used by pathogens such as Plasmodium falciparum or Trypanosoma brucei. Using Hi-C, ChIP-seq and super-resolution microscopy analysis, the authors described a novel mechanism involving the spatial integration of antigen transcription and mRNA splicing in a dedicated nuclear compartment to ensure monogenic expression in Trypanosoma brucei. To date, while intra-chromosomal interactions are an essential component of the regulation of gene expression in many eukaryotes, the function of inter-chromosomal interactions is still under debate. The authors provide solid here evidence of a functional and stable (even during cell division) inter-chromosomal bridge between an active VSG gene located on chromosome VI and the SL-RNA array locus on chromosome IX. These results reveal for the first time to my knowledge a selective inter-chromosomal interaction which links the transcription and splicing of the mRNA. This report adds new insights into antigenic variation mechanism and should be of a broad interest for researchers working in post-transcriptional gene regulation, and specifically those looking at monogenic expression in Plasmodium falciparum. The experiments have all been well designed with appropriate controls and biological replicates. The manuscript is extremely well written.
Below are minor suggestions for improvement:
1- lines 267-268: the authors conclude that "neither the VEX, nor the tSNAP protein levels were affected by sinefungin treatment" while describing the Extended Data Fig.3f. On the contrary, this reviewer noted significant differences in the quantity of proteins on westerns compared to the proposed loading control.

Quantitative western can be done.
2- lines 229-231: the data in Fig. 1c compare cells in S phase vs G1 while the data in Fig. 3a and b oppose the total cells with those in phase G1. To support their conclusions, it would be appropriate to include the cells in S phase in Fig. 3a and b instead of Total cells.
3- This reviewer noticed potential duplication of the data in Extended data Fig. 4d and Fig.4b. The duplication itself is no problem, the distribution of data by biological condition seems different while the legends are identical. Very confusing. Must be clarified.
4- Figure 3g : an image control without SNF should be added.
5- The authors used sinefungin (SNF) to inhibit trans-splicing. Mutations in AdoMet transporters were found to be the main driver of SNF resistance in Leishmania (Bhattacharya A et al., 2019), yeast (Zheng S et al; 2007) or Toxoplasma (Behnke MS et al. 2015). Whole-genome sequencing of independent resistant mutants uncovers also secondary target for SNF in Leishmania (Bhattacharya A et al., 2019). Is it possible that sinefungin could target other pathways in Trypanosoma brucei that in turn disrupted the connection between the VSG and SL-RNA transcription compartments?

Reviewer #2 (Remarks to the Author):

In this paper, the authors investigated whether inter-chromosomal interactions could play a role in the monoallelic expression of VSG genes in African trypanosomes. By using a combination of Hi-C,

super resolution IFA and genetic tools, the authors found that the active VSG specifically interacts with one of the two Spliced Leader loci, located in a different chromosome. Transcription of Spliced Leader genes results in non-coding RNA molecules that are added by trans-splicing into every mRNA. The authors showed that this interaction is dynamic: when parasites switch from VSG2 to VSG13, there's a corresponding switch of inter-chromosomal interactions from VSG2-SL to VSG13-SL. VEX1 and VEX2 had been previously characterized as proteins associated to the active site of VSG expression. Using super resolution, the authors correct here a previous observation and conclude that: while VEX2 remains a VSG-associated component (a single site in the nucleus), VEX1 is actually associated to SL loci (two sites in the nucleus). Importantly, when VEX2 is depleted, the SL locus dissociates from the active VSG and the authors detected more frequent association of SL locus with previously silent VSGs, which become derepressed.

This is a solid work, whose combined methodology supports the conclusions taken by the authors. This work brings a new angle to our understanding of antigenic variation and of gene expression regulation, overall, in trypanosomes: the spatial proximity of a transcribed locus X to the SL locus may be important for the processing of the X transcripts. The manuscript is very clearly written and most figures are self-explanatory. Below are the points that deserve some attention:

1. Within the resolution of Hi-C, could the authors provide a more "zoomed" image of which part of the SL locus interacts with which part of VSG orf? Can the authors speculate which type of protein could bridge the two loci? The authors show that VEX2 is one of the proteins that keeps the two transcription compartments together. Are RNAs also required for this DNA:DNA interaction? For example, could the authors lightly digest parasites with RNAse and check by IFA if the two compartments are still located nearby? This would show whether the association of the two compartments (VSG and SL) depends on RNA molecules (such as VSG transcript? Spliced leader transcript?). Alternatively, the authors could consider halting transcription (Pol I and/or PolII) and ask whether this process is required to detect the two compartments closely associated by IFA.

2. In Figure 4A, the authors show that upon VEX2 depletion, the distance between the SL-transcription site and the VSG-transcription site increases, suggesting VEX2 is required to keep the two compartments close by. However, by Hi-C, the authors did not find a reduction of the interaction between the two loci. How do the authors explain this inconsistent result?

3. By IFA, did the authors observe that one of the transcription compartments showed a preferred fixed location relative to the nuclear periphery or the nucleolus? Does the SL locus move close to VSG-transcription site? Or is it the opposite? Or do they both move? Could the authors induce the expression of a second VSG (like in Battram et al., 2014 or Figueriedo et al, 2008) and test by super resolution IFA whether the two transcribed VSGs are transcribed in the same SL-transcription site or separately (one vs two VEX2/PolI sites)? According to the model that VEX2 serves as an exclusion molecule, one would expect both VSGs to be transcribed from the same SL/PolI compartment.

4. Line 110: It is not obvious from the current Figure 1a that there is "a distance-dependent decay of intra-chromosomal interactions between each viewpoint and its upstream and downstream genomic region". Can the authors show the data in a way so that this gradient is more obvious? The representations in the Extended Figure 1 are larger and easier to interpret.

5. Line 225. In the Extended Figure 3c we cannot appreciate the relative enrichment of VEX1 between SL and VSG loci. Can the figure be improved?

Reviewer #3 (Remarks to the Author):

Re: NMICROBIOL-20030686A-Z
Title: Antigenic variation by switching inter-chromosomal interactions with an RNA splicing locus in trypanosomes

Authors: Faria et al.

In this manuscript, Faria et al. reported a novel observation, where VEX1 associates with the SL RNA gene locus and the RNA splicing locus (with the SNAPc as a marker) in the nucleus, and VEX2 associates with the active VSG transcription center (with the RNA Pol I as a marker). In addition, IF followed by super-resolution microscope imaging showed that the RNA processing center and the active VSG transcription site are frequently overlapping or adjacent to each other in G1 phase, while Hi-C experiment showed that the active VSG and the SL RNA gene array are associated in chromosome conformation. The authors showed that this is true for any VSG as long as the VSG is fully active. In insect stage parasites, the procyclin genes and the SL gene array are also associated when examined by Hi-C analysis. Furthermore, highly transcribed genes, such as the tubulin gene array, is observed to associate with the RNA processing center. Upon treating cells with sinefungin, the authors found that VEX1 and VEX2 are dispersed. Upon VEX2 knockdown by RNAi (12 hr post RNAi induction), the active VSG transcription site (marked by RNAP I) moved away from the RNA processing center (marked by SNAPc). At later point (24 hr post RNAi induction), Hi-C analysis further showed that derepressed VSGs are associated with the RNA processing center at an increased frequency.

The observations described in this manuscript is novel. The partial overlapping/association between the active VSG expression site and the RNA processing center (SNAPc foci) is very interesting and suggests that localization of transcription and RNA processing machinery within close proximity is an efficient strategy for high level gene expression. The result also suggests that VEX2 is a key factor bringing the active VSG transcription to the RNA processing center, which reveals part of the key functions of VEX2 in maintaining a high level of VSG expression.

The manuscript focuses on an interesting topic: how is VSG monoallelic expression achieved. The results showed interchromosomal interaction between the active VSG locus and RNA splicing locus, which provides clues how a high level of VSG expression is achieved and how VSG transcription can be selective. However, it is not very clear what is the consequence of loss of this interchromosomal interaction, as it seems dissociation of the active VSG locus and the RNA splicing site occurs naturally every S phase. Interpretation of the VEX2 knockdown phenotypes also needs to be more careful, since the current data do not reveal a mechanistic role of VEX2 in the described interchromosome interaction between the active VSG locus and the RNA splicing locus.

Major concerns:
1. Characterization of the VSG transcription site and RNA processing center by IF-super resolution microscope only focused on G1 cells. As cells enter S phase, the active VSG transcription site is more likely to be far away from the RNA processing center. What happens in later stages of the cell cycle? What are the evidences to support the claim "the interaction is stably propagated during cell-division" and "the interaction is re-established after replication" (lines 375-376, page 16). If the VSG transcription site and the RNA processing center is more separated during S phase, what is the effect of this on VSG expression? If the different patterns of subnuclear localization of the VSG transcription site in G1 and S phase do not have any significant effect on VSG expression level, what is the significance of a closely positioned VSG transcription site and the RNA processing center? Depletion of VEX2 caused separation of the VSG transcription site and RNA processing center, this separation appears to be very comparable to what happens in the S phase. Does the separation of the two types of foci really represents a defect if this happens naturally in S phase already?
In addition, it is not clear what are the criteria to classify Pol I focus and SNAPc focus to be overlapped or adjacent. Was it all determined by eye?
2. Why treating cell with Sinefungin did not cause any dispersion of SNAPc? What is the control that Sinefungin treatment disrupted trans-splicing?
3. Depletion of VEX2 resulted in a longer distance between the active VSG transcription site and the RNA processing center by IF. However, Hi-C showed that the association of the active VSG and SL gene array is not changed. These observations are not consistent with each other. Could it because in Hi-C the cells were cross-linked, so that it catches even transient chromosome association? If Hi-C

is more prone to pick up interchromosomal association, this may present a systematic error. For example, 24 hrs after VEX2 is depleted, Hi-C detected more frequent association between derepressed VSGs and the SL gene array. Can IF-super resolution microscope imaging detect similar phenotypes, such as a close positioned derepressed VSG mRNA (RNA FISH) and the RNA processing center (SNAPc IF)?
4. VEX1 interacts with VEX2, so one possible way for the active VSG transcription site to be located near the RNA processing center is through the interaction between these two proteins. However, longer distance between the VSG transcription site and SNAPc focus was only observed in cells depleted of VEX2 but not in cells depleted of VEX1, indicating that the interaction between these two proteins are not a key for the closely positioned VSG transcription site and the RNA processing center. In this case, what is the mechanism of VEX2's role? How does VEX2 help to locate the VSG transcription site near the RNA processing center? Without the mechanistic study, the manuscript is quite descriptive at the current stage.
5. Depletion of VEX2 appears to induce a series of events, which are quite important for interpretation of its function. However, these are not carefully discussed. The earliest event occurred at 12 hr after VEX2 RNAi induction, where longer distance between the active VSG transcription site and RNA processing center was observed. Later, VSG derepression was observed at 24 hrs after VEX2 RNAi induction (but not at 12 hrs). Also, Hi-C analysis showing increased association between derepressed VSG loci and the SL gene array occurred at 24 hrs. It is therefore possible that the immediate effect of removal of VEX2 is just a loss of closely positioned active VSG transcription site and the RNA processing center. This may induce a secondary effect that indirectly caused derepression VSGs at 24 hrs. For example, not efficiently processed active VSG mRNA may cause a decrease in VSG synthesis, which can induce derepression of silent VSGs, as VSG is essential.

Reviewer #4 (Remarks to the Author):

Overall, I think this is an interesting manuscript and the conclusions are generally well supported by the data. I think it provides a good view of the assembly of the VSG transcriptional complex.

Overall Points.
1. Fig 2. And text line 182. There appears to be a major missing issue that should be discussed. The claim is that the interaction of the splicing locus is more with the VSG gene itself and not so much with the promoter. At first site this might seem reasonable. However, between the promoter and the VSG gene there are many ESAG genes – and these are known to be trans-spliced and expressed. They have to be transcribed at the same rate as the VSG gene since it is the last one in the line and there are no promoters in between (I recognise that the half-lives of the RNA or processing may be differentially regulated). So if the idea is that the association is more with the VSG gene that the promoter region is the conclusion (not in the paper I believe) also that the VSG gene is much more spliced than the ESAGs? If so, how? This result surely suggests that during or after transcription of the ESAG RNA molecules then they are not recognised, processed or matured by the splicing machinery that interacts more heavily with the VSG gene even though one passage of an RNA Pol I should produce similar amounts of RNA from each ESAG and the VSG gene? Or am I missing something? If this is so, or not I think it needs a discussion in the paper.
2. I was left a little wanting by the model. They have VEX2 as the central protein to bring the ESB to a SL-RNA array and this stops silent VSGs accessing that SL-RNA array but what about the other SL-RNA array that also has a VEX1 foci? Why can't the silent VSGs access that? Is there a natural tendency of expression sites to interact SL-RNA array due presence other unknown factors – transient interactions but locked in by VEX2? State uncovered during VEX2 RNAi.
To give this variant of their model:
VEX1 mark SL-RNA ◊ expression sites sampling SL-RNA arrays ◊engagement through VEX2 ◊ once established this site "hoovers up" the rest of the required factors so other expression sites can't establish. ie a choose and consolidate model.

Specific Points:
1. Title. In the VSG expression site world the term switching has associations with (amongst other things) chromosomal recombination. The authors might just ponder if "switching inter-chromosomal interactions" may give an initial impression that the paper is about recombination not nuclear position effects?
2. Figure 1a there is a large interaction with chm 11 as well as the spliced leader array on chm 9. What is at this point? Is there some statistical analysis to guide which are the more significant peaks than others for the HiC data representation? I think these other significant peaks need to be identified.
3. The image analysis in places left some uncertainties – the image they show for tSNAP has one large foci and one small one – is this consistent? And is the small one always close to the ESB?
4. In the extended figure 1 they include viewpoints for central core genes, which is an important control but I didn't spot this mentioned in the main text. Plus, what are the central core – which genes?
5. They say that the majority of cells 55% have 2 VEX1 foci but what about the other 45% - is there a single VEX1 foci or multiple? In the other 45% of the cells do these also have 2 tSNAP foci. For the categorisation in 3a, were any instances seen when the two sets of signals in the same cell are different? i.e. one overlapping and one adjacent. If so how was this dealt with in the counts?
6. From the image in 3a there is a large VEX1 foci associated with the small tSNAP foci and presumably the ESB – is this a consistent result?
7. The categories in 3a/b are G1 and total cells and do not look at cells specifically in S-phase like 1c. There is a drop in the number of overlapping foci in total cells and this could come from those cells in S-phase so it's difficult to be confident about whether these proteins remain in close association during S-phase.
8. Do they know if VEX1 still associates with tSNAP in PCFs interesting but not essential for their conclusions here?
9. Their previous paper showed that VEX1 RNAi had no effect on VEX2 localisation but VEX1 overexpression causes loss of VEX2 foci – what happens to VEX1 foci during VEX2 RNAi? From their data you would predict VEX1 localisation is not dependent on VEX2.

Author Rebuttal, first revision:

Reviewer #1 (Parasitology):

The manuscript entitled " Antigenic variation by switching inter-chromosomal interactions with an RNA splicing locus in trypanosomes" reports on the discovery of a novel mechanism to ensure monogenic expression in Trypanosoma brucei. Monogenic expression is a key feature of antigenic variation, an immune evasion strategy used by pathogens such as Plasmodium falciparum or Trypanosoma brucei. Using Hi-C, ChIP-seq and super-resolution microscopy analysis, the authors described a novel mechanism involving the spatial integration of antigen transcription and mRNA splicing in a dedicated nuclear compartment to ensure monogenic expression in Trypanosoma brucei. To date, while intra-chromosomal interactions are an essential component of the regulation of gene expression in many eukaryotes, the function of inter-chromosomal interactions is still under debate. The authors provide solid here evidence of a functional and stable (even during cell division) inter-chromosomal bridge between an active VSG gene located on chromosome VI and the SL-RNA array locus on chromosome

IX. These results reveal for the first time to my knowledge a selective inter-chromosomal interaction which links the transcription and splicing of the mRNA. This report adds new insights into antigenic

variation mechanism and should be of a broad interest for researchers working in post-transcriptional gene regulation, and specifically those looking at monogenic expression in Plasmodium falciparum. The experiments have all been well designed with appropriate controls and biological replicates. The manuscript is extremely well written.

Below are minor suggestions for improvement:

1.1: *lines 267-268: the authors conclude that "neither the VEX, nor the tSNAP protein levels were affected by sinefungin treatment" while describing the Extended Data Fig.3f. On the contrary, this reviewer noted significant differences in the quantity of proteins on westerns compared to the proposed loading control. Quantitative western can be done.*
R1.1: We had four independent western blots with loading controls for this experiment. The full set was quantified and fold changes +/- SD are now noted next to the example blots shown (now ED Fig. 5a). The text now states "neither…were substantially affected…".

1.2: *lines 229-231: the data in Fig. 1c compare cells in S phase vs G1 while the data in Fig. 3a and b oppose the total cells with those in phase G1. To support their conclusions, it would be appropriate to include the cells in S phase in Fig. 3a and b instead of Total cells.*
R1.2: We've adjusted Fig. 3a-b to include cells in S phase, as suggested.

1.3: *This reviewer noticed potential duplication of the data in Extended data Fig. 4d and Fig.4b. The duplication itself is no problem, the distribution of data by biological condition seems different while the legends are identical. Very confusing. Must be clarified.*
R1.3: We apologise for being unclear here. This is not duplication. We have always measured the inner AND outer distance between the compartments. We have now clarified the difference between 'inside edge' (now ED Fig. 2c and Fig. 4b) and 'outside edge' (now ED Fig. 2b and ED Fig. 6d) measurements and our reasoning: "to ensure that changes in the distance between the two protein condensates following cell cycle progression or VEX RNAi were not a consequence of changes in the diameter of the foci".

1.4: *Figure 3g: an image control without SNF should be added.*
R1.4: We've added a control image to Fig. 3g, as suggested.

1.5: *The authors used sinefungin (SNF) to inhibit trans-splicing. Mutations in AdoMet transporters were found to be the main driver of SNF resistance in Leishmania (Bhattacharya A et al., 2019), yeast*

(Zheng S et al; 2007) or Toxoplasma (Behnke MS et al. 2015). Whole-genome sequencing of independent resistant mutants uncovers also secondary target for SNF in Leishmania (Bhattacharya A et al., 2019). Is it possible that sinefungin could target other pathways in Trypanosoma brucei that in turn disrupted the connection between the VSG and SL-RNA transcription compartments?

R1.5: The cited studies suggest that sinefungin, which is an S-adenosylmethionine (AdoMet) analogue, is taken up primarily via AdoMet transporters in *Leishmania*, yeast and *Toxoplasma*. A mutation that results in receptor malfunction can therefore cause resistance to sinefungin by preventing its uptake. In *T. brucei*, sinefungin has a well-characterised mode of action; "Sinefungin inhibits the cap guanylyltransferase-methyltransferase Cgm1 (Takagi *et al*., 2007, PMID: 17416901) and subsequent methylation of the SL-RNA cap, required for *trans*-splicing (Ullu and Tschudi, 1991, PMID: 1719544; McNally and Agabian, 1992, PMID: 1406666)"; this text has now been added to the manuscript. A similar mode-of-action in both *Leishmania* (Bhattacharya *et al*., 2019, PMID: 31615876) and yeast (Zheng *et al*., 2007, PMID: 17932050) is supported by sinefungin resistance due to mRNA (guanine-N7)-methyltransferase gain-of-function, likely compensating for the sinefungin-induced defect. We're not aware of alternative targets in trypanosomes that could disrupt the VSG/SL-RNA connection.

Reviewer #2 (Molecular cell biology of trypanosomes):

In this paper, the authors investigated whether inter-chromosomal interactions could play a role in the monoallelic expression of VSG genes in African trypanosomes. By using a combination of Hi-C, super resolution IFA and genetic tools, the authors found that the active VSG specifically interacts with one of the two Spliced Leader loci, located in a different chromosome. Transcription of Spliced Leader genes results in non-coding RNA molecules that are added by trans-splicing into every mRNA. The authors showed that this interaction is dynamic: when parasites switch from VSG2 to VSG13, there's a corresponding switch of inter-chromosomal interactions from VSG2-SL to VSG13-SL. VEX1 and VEX2 had been previously characterized as proteins associated to the active site of VSG expression. Using super resolution, the authors correct here a previous observation and conclude that: while VEX2 remains a VSG-associated component (a single site in the nucleus), VEX1 is actually associated to SL loci (two sites in the nucleus). Importantly, when VEX2 is depleted, the SL locus dissociates from the active VSG and the authors detected more frequent association of SL locus with previously silent VSGs, which become derepressed.

This is a solid work, whose combined methodology supports the conclusions taken by the authors. This work brings a new angle to our understanding of antigenic variation and of gene expression regulation, overall, in trypanosomes: the spatial proximity of a transcribed locus X to the SL locus may be important for the processing of the X transcripts. The manuscript is very clearly written and most figures are self-explanatory. Below are the points that deserve some attention:

***2.1:** Within the resolution of Hi-C, could the authors provide a more "zoomed" image of which part of the SL locus interacts with which part of VSG orf? Can the authors speculate which type of protein could bridge the two loci? The authors show that VEX2 is one of the proteins that keeps the two transcription compartments together. Are RNAs also required for this DNA:DNA interaction? For example, could the authors lightly digest parasites with RNAse and check by IFA if the two compartments are still located*

nearby? This would show whether the association of the two compartments (VSG and SL) depends on RNA molecules (such as VSG transcript? Spliced leader transcript?). Alternatively, the authors could consider halting transcription (Pol I and/or PolII) and ask whether this process is required to detect the two compartments closely associated by IFA.

R2.1: Regarding the '**zoomed'** image: While our Hi-C data already provide information on DNA-DNA interactions at a very high resolution, mapping interactions to specific sites within individual genes is not possible with conventional Hi-C approaches as these rely on genome fragmentation by specific restriction enzymes. A decrease in bin size and a zoomed image of the VSG gene or SL-RNA locus is thus unlikely to reflect a true increase in resolution and may convey misleading information. Thus, we would prefer not to decrease the bin size (resolution) below the 20 kb that we have chosen for this study.

In terms of '**which type of protein could bridge the two loci**', we note that "RNAs are major actors in facilitating genomic interactions and phase transitions" and "speculate that maintenance of

VSG transcription-maturation compartment connectivity is similarly regulated by the putative RNA helicase VEX2" (see final paragraph).

We liked both suggested experiments to elucidate **the role of RNAs** in establishing the interaction (slight RNA digestion or halting transcription of Pol I and/or Pol II). However, we felt that the latter would provide the more specific insights. A general digestion of RNA, even at low levels, may result in global changes of the genome architecture (Khosraviani *et al*., 2019, PMID: 31921848). We therefore focused on the alternative suggested experiment; halting Pol I transcription with BMH-21 and both Pol I and Pol II transcription with actinomycin D. Either treatment disrupted VEX1 and VEX2 localisation without affecting tSNAP localisation (see now ED Fig. 5b-c and Faria *et al*., 2019, PMID: 31289266). These data suggest that "VEX protein localization and the juxtaposition of the VSG and the SL-RNA transcription compartments are dependent on Pol I transcription or VSG expression site RNAs".

2.2: *In Figure 4A, the authors show that upon VEX2 depletion, the distance between the SL-transcription site and the VSG-transcription site increases, suggesting VEX2 is required to keep the two compartments close by. However, by Hi-C, the authors did not find a reduction of the interaction between the two loci. How do the authors explain this inconsistent result?*

R2.2: We apologize for not better explaining the apparent inconsistency. While we had visualised the expression site body (Pol I) and the splicing locus (tSNAP) at the protein level in Fig. 4a, we had detected DNA-DNA interactions by Hi-C (now Fig 4e-g). Thus, the results are not inconsistent. Rather, the results are consistent with the idea that "in the absence of VEX2, Pol I initially separates from the SL-RNA compartment (12 h) and at later time points it disperses". In addition, we have further investigated DNA-DNA interactions in single cells, using DNA Fluorescence In-Situ Hybridization (see new data; Fig. 4c-d). "Following VEX2 depletion (24 h), the number of individual VSG expression site signals decreased while their size increased (Extended Data Fig. 7a), consistent with the formation of VSG expression site clusters. Indeed, the proportion of G1 nuclei displaying VSG expression site clusters overlapping or adjacent to an SL-RNA compartment increased upon VEX2 knockdown (Fig. 4d)." These results are entirely consistent with what we observed at the DNA level using Hi-C.

2.3: *By IFA, did the authors observe that one of the transcription compartments showed a preferred fixed location relative to the nuclear periphery or the nucleolus? Does the SL locus move close to VSG-transcription site? Or is it the opposite? Or do they both move? Could the authors induce the expression of a second VSG (like in Battram et al., 2014 or Figueriedo et al, 2008) and test by super resolution IFA whether the two transcribed VSGs are transcribed in the same SL-transcription site or separately (one vs two VEX2/PolI sites)? According to the model that VEX2 serves as an exclusion molecule, one would expect both VSGs to be transcribed from the same SL/PolI compartment.*

R2.3: Regarding the **location of the compartments**, we measured distances to the nucleolus for SL-RNA transcription compartments adjacent to a VSG transcription compartment and those that were not adjacent to a VSG transcription compartment. We find that "The SL-RNA transcription compartments were always extra-nucleolar (Extended Data Fig. 2d), but those adjacent to the VSG transcription compartment were significantly less intense and significantly closer to the nucleolus (Extended Data Fig. 2e-f)". In addition, the new DNA FISH data indicate that VEX2 impacts the relative localisation of both SL and VSG loci in relation to one another: As above, "Following VEX2 depletion (24 h), the number of individual VSG expression site signals decreased while their size increased (Extended Data Fig. 7a), consistent with the formation of VSG expression site clusters. Indeed, the proportion of G1 nuclei displaying VSG expression site clusters overlapping or adjacent to an SL-RNA compartment increased upon VEX2 knockdown (Fig. 4d). We also observed a decrease in the distance between SL-arrays (Extended Fig. 7b-d)." Consistent with the view that VEX2 serves as an exclusion molecule, Budzak *et al*., (2019, PMID: 31358644) recently showed that "where **two VSG genes are simultaneously active**, both co-localize at the expression site body" (see introductory text).

2.4: *Line 110: It is not obvious from the current Figure 1a that there is "a distance-dependent decay of intra-chromosomal interactions between each viewpoint and its upstream and downstream genomic region". Can the authors show the data in a way so that this gradient is more obvious? The representations in the Extended Figure 1 are larger and easier to interpret.*

R2.4: We have adjusted the scale of the graph in Fig. 1a. The distance-dependent decay is more clearly visible for expression site 3. For expression site 1, located on the left arm of chr. 6, the distance-dependent decay is not as characteristic as for other sites. The reason for this is, as we have published previously (Müller *et al*., 2018, PMID: 30333624), that expression site 1 is separated from the core region of chr. 6 by the centromere. Centromeres can serve as boundary elements, inhibiting frequent interactions between the two chromosomal arms (Müller *et al*., 2018, PMID: 30333624). Therefore, expression site 1 shows a distance-dependent decay with the core of chr. 6 that is reduced by the insulation effect of the adjacent centromere.

We have added the following explanation to the manuscript: "For the *VSG-2* gene located on chr. 6 in expression site 1 (Fig. 1a, top panel), the distance-dependent decay was not as characteristic as for other viewpoints. As we have published previously (Müller *et al*., 2018, PMID: 30333624), this expression site is separated from the core region of chr. 6 by a centromere, which can serve as a boundary element, inhibiting frequent interactions between the two chromosomal arms."

2.5: *Line 225. In the Extended Figure 3c we cannot appreciate the relative enrichment of VEX1 between SL and VSG loci. Can the figure be improved?*
R2.5: We've added new high-resolution plots showing relative VEX1 enrichment across three SL-RNA repeats and across the active VSG (see Fig. 3c).

Reviewer #3 (Recombination in trypanosomes):

In this manuscript, Faria et al. reported a novel observation, where VEX1 associates with the SL RNA gene locus and the RNA splicing locus (with the SNAPc as a marker) in the nucleus, and VEX2 associates with the active VSG transcription center (with the RNA Pol I as a marker). In addition, IF followed by super-resolution microscope imaging showed that the RNA processing center and the active VSG transcription site are frequently overlapping or adjacent to each other in G1 phase, while Hi-C experiment showed that the active VSG and the SL RNA gene array are associated in chromosome conformation. The authors showed that this is true for any VSG as long as the VSG is fully active. In insect stage parasites, the procyclin genes and the SL gene array are also associated when examined by Hi-C analysis. Furthermore, highly transcribed genes, such as the tubulin gene array, is observed to associate with the RNA processing center. Upon treating cells with sinefungin, the authors found that VEX1 and VEX2 are dispersed. Upon VEX2 knockdown by RNAi (12 hr post RNAi induction), the active VSG transcription site (marked by RNAP I) moved away from the RNA processing center (marked by SNAPc). At later point (24 hr post RNAi induction), Hi-C analysis further showed that derepressed VSGs are associated with the RNA processing center at an increased frequency.

The observations described in this manuscript is novel. The partial overlapping/association between the active VSG expression site and the RNA processing center (SNAPc foci) is very interesting and suggests that localization of transcription and RNA processing machinery within close proximity is an efficient strategy for high level gene expression. The result also suggests that VEX2 is a key factor bringing the active VSG transcription to the RNA processing center, which reveals part of the key functions of VEX2 in maintaining a high level of VSG expression.

The manuscript focuses on an interesting topic: how is VSG monoallelic expression achieved. The results showed interchromosomal interaction between the active VSG locus and RNA splicing locus, which provides clues how a high level of VSG expression is achieved and how VSG transcription can be selective. However, it is not very clear what is the consequence of loss of this interchromosomal interaction, as it seems dissociation of the active VSG locus and the RNA splicing site occurs naturally every S phase. Interpretation of the VEX2 knockdown phenotypes also needs to be more careful, since the current data do not reveal a mechanistic role of VEX2 in the described interchromosome interaction between the active VSG locus and the RNA splicing locus.

Major concerns:

3.1: *Characterization of the VSG transcription site and RNA processing center by IF-super resolution*

microscope only focused on G1 cells. As cells enter S phase, the active VSG transcription site is more likely to be far away from the RNA processing center. What happens in later stages of the cell cycle? What are the evidences to support the claim "the interaction is stably propagated during cell-division" and "the interaction is re-established after replication" (lines 375-376, page 16). If the VSG transcription site and the RNA processing center is more separated during S phase, what is the effect of this on VSG expression? If the different patterns of subnuclear localization of the VSG transcription site in G1 and S phase do not have any significant effect on VSG expression level, what is the significance of a closely positioned VSG transcription site and the RNA processing center? Depletion of VEX2 caused separation of the VSG transcription site and RNA processing center, this separation appears to be very comparable to what happens in the S phase. Does the separation of the two types of foci really represents a defect if this happens naturally in S phase already? In addition, it is not clear what are the criteria to classify Pol I focus and SNAPc focus to be overlapped or adjacent. Was it all determined by eye?

R3.1: To show what happens in **later stages of the cell cycle** and to further support the claims noted above, we have now added data for G2 cells. We find that "VSG and SL-RNA transcription compartments were once again adjacent in most G2 nuclei (Fig. 1c; Extended Data Fig. 2b-c), indicating that the interaction is resolved during S phase and successfully re-established after replication".

In terms of **VSG expression during S phase**, transcription can be an obstacle to DNA replication (see Gómez-González and Aguilera, 2019, PMID: 31123061 for example) and it has been known for some time that transcription is spatially separated from replication in S phase in mammalian nuclei (Wei *et al*., 1998, PMID: 9727975). Similarly, our observations, showing separation of the tSNAP and Pol I compartments during S phase, may reflect transient perturbation of VSG expression site transcription during DNA replication. However, given the stability of VSG mRNA (half-life ~4.5 h), a transient pause in transcription would not be expected to have a major impact on VSG mRNA levels, VSG translation or VSG protein levels (one complete cell cycle is ~6.5 h). Thus, we believe that, even with a pause of VSG transcription in S phase, "the close spatial proximity of the two compartments in a single organelle provides a sufficiently high concentration of *trans*-splicing substrate to ensure the efficient maturation of highly abundant VSG transcripts" during G1 and G2.

The **separation following VEX2 knockdown** at early time points does indeed appear comparable to what happens in S phase. "IFA data revealed a dissociation of Pol I from the SL-RNA transcription compartments in 45% of G1 nuclei" in the case of knockdown. However "at later time points it [the Pol I signal] disperses", indicating that there are differences between S phase and VEX2 depletion: during S phase we observed a transient dissociation of the two compartments. In contrast, following VEX2 depletion, we observed Pol I dispersion.

To **categorize RNA Pol I and tSNAP signals as overlapping, adjacent or separate** we used "thresholded Pearson's correlation coefficients – see methods". Pearson's correlation coefficient (PCC) is one of the two widely accepted statistical measures of colocalisation (Barlow *et al*., 2010, PMID: 20946701**)** - each pixel is considered one data point, and the intensity of the 'magenta' signal and 'green' signal is measured at each pixel. The correlation coefficient is then measured across all pixels in the area of interest in an image; PCC values range from −1 to +1. Overlapping, adjacent and separate foci presented a PCC in the following ranges: ≥ 0.5 / ≤ 1, ≥ -0.5 / < 0.5, ≥ -1 / < -0.5, respectively. In the methods section we describe our approach in more detail: "For all **quantifications**, images were acquired with the same settings and equally processed. All the images were processed and scored using Fiji v. 2.0.0. (Schindelin *et al*., 2012, PMID: 22743772), using stacks of approximately 30 slices of 0.1 µm; except Fig. 4d and Extended Data Fig. 7b, where the analysis was performed using Imaris 9.5". "Pearson's correlation coefficient (PCC) was applied as a statistical measure of colocalization (Barlow *et al*., 2010, PMID: 20946701). Overlapping, adjacent and separate foci presented a PCC in the following ranges: ≥ 0.5 / ≤ 1, ≥ -0.5 / < 0.5, ≥ -1 / < -0.5, respectively." In addition, we have also

measured distances between ‘inside edges’ (ED Fig. 2c and Fig. 4b) and ‘outside edges’ (ED Fig. 2b and ED Fig. 6d) of foci (see R1.3 above) in unperturbed G1, S and G2 cells and in cells following knockdown.

3.2: *Why treating cell with Sinefungin did not cause any dispersion of SNAPc? What is the control that Sinefungin treatment disrupted trans-splicing?*

R3.2: tSNAP is "an RNA Pol II promoter-binding transcription factor" and Sinefungin is an S-adenosylmethionine (AdoMet) analogue. As mentioned in R1.5, in *T. brucei*, sinefungin has a well-characterised mode of action; "Sinefungin inhibits the cap guanylyltransferase-methyltransferase Cgm1 (Takagi *et al*., 2007, PMID: 17416901) and subsequent methylation of the SL-RNA cap, required for *trans*-splicing (Ullu and Tschudi, 1991, PMID: 1719544; McNally and Agabian, 1992, PMID: 1406666)". Thus, following sinefungin treatment to disrupt RNA processing, it was not surprising that localization of a factor required for transcription initiation, such as tSNAP, was unaffected. In contrast, the specific disruption of VEX2 and VEX1 foci following sinefungin treatment (see Fig. 3d-e and ED Fig. 5b-c) supports the view that splicing is required to maintain VEX-complex sequestration. Since sinefungin has a well-characterised mode of action and has been used extensively to inhibit splicing in *T. brucei*, we felt that further controls were not necessary.

3.3: *Depletion of VEX2 resulted in a longer distance between the active VSG transcription site and the RNA processing center by IF. However, Hi-C showed that the association of the active VSG and SL gene array is not changed. These observations are not consistent with each other. Could it because in Hi-C the cells were cross-linked, so that it catches even transient chromosome association? If Hi-C is more prone to pick up interchromosomal association, this may present a systematic error. For example, 24 hrs after VEX2 is depleted, Hi-C detected more frequent association between derepressed VSGs and the SL gene array. Can IF-super resolution microscope imaging detect similar phenotypes, such as a close positioned derepressed VSG mRNA (RNA FISH) and the RNA processing center (SNAPc IF)?*

R3.3: We apologize for not **better explaining the apparent inconsistency**. As explained in R2.2, while we had visualised the expression site body (Pol I) and the splicing locus (tSNAP) at the protein level in Fig. 4a, we had detected DNA-DNA interactions by Hi-C (now Fig 4e-g). The results are thus not inconsistent. Rather, we feel that the results are consistent with the idea that "in the absence of VEX2, Pol I initially separates from the SL-RNA compartment (12 h) and at later time points it disperses".

Both, the cells for Hi-C assays and the IFA were **cross-linked**, thus, we don't expect to observe any assay-specific cross-link-induced artefacts. Also, to control for background inter-chromosomal interactions, we performed Hi-C assays using cells expressing different VSGs. Since we only detected frequent DNA-DNA interactions between the respective actively expressed VSG gene and the SL-RNA, we expect background inter-chromosomal interactions to be negligible. Following the suggestion, to further investigate DNA-DNA interactions in single cells, we performed DNA Fluorescence In-Situ Hybridization (see new data; Fig. 4c-d). As outlined in R2.2. following VEX2 depletion for 24 h (the same level of depletion used for the Hi-C assays), DNA FISH analysis reveals a clustering of VSG expression sites with each other and with the SL-RNA loci. These results are entirely consistent with what we observed at the DNA level using Hi-C.

Together the Hi-C, IF and DNA FISH assays provide information on the position of DNA and proteins at the single cell and population level, which indicates that VEX2 is important for the localization of Pol I and VEX1 and also for exclusion of the silent expression sites.

3.4: *VEX1 interacts with VEX2, so one possible way for the active VSG transcription site to be located near the RNA processing center is through the interaction between these two proteins. However, longer distance between the VSG transcription site and SNAPc focus was only observed in cells depleted of VEX2 but not in cells depleted of VEX1, indicating that the interaction between these two proteins are not a key for the closely positioned VSG transcription site and the RNA processing center. In this case, what is the mechanism of VEX2's role? How does VEX2 help to locate the VSG transcription site near the RNA processing center? Without the mechanistic study, the manuscript is quite descriptive at the current stage.*

R3.4: Yes, while IP experiments indicated that VEX1 and VEX2 are interacting (Faria *et al*., 2019, PMID: 31289266), all previous findings and our new data suggest distinct roles for VEX1 and VEX2 beyond 'simply' connecting the two compartments. As mentioned in R3.3. our Hi-C, IF and DNA FISH

assays indicated that VEX2 is important for the localization of Pol I, VEX1 and the silent expression sites. Based on the available data (previously published and part of this manuscript) our model for the role of VEX2 is as follows:

"Our data indicate that VEX2 is not simply a tether. Instead, they led us to propose a temporal 'choose and consolidate' component to our model: VEX2 is recruited to expression sites in a stochastic

and competitive manner by Pol I transcription ('choose'), as supported by Pol I inhibition experiments that resulted in dispersion of VEX2 (Faria *et al.*, 2019, PMID: 31289266) and (Extended Data Fig. 5b). Subsequently, the presence of VEX2 at one expression site 'consolidates' expression site transcription, as supported by VEX2 depletion experiments that resulted in delocalization of the Pol I focus from the expression site (Fig. 4a-b). Once Pol I is no longer sequestered at one specific expression site, the enzyme can access and activate previously silent expression sites. Transcription of expression sites is accompanied by an increase in interactions with the SL-RNA array, as supported by Hi-C and FISH (Fig. 4c-f), leading to efficient VSG mRNA processing. According to this model, VEX2 is the key molecule excluding all but one VSG expression site from the SL-RNA transcription compartment, thereby ensuring expression of a single VSG gene per cell."

Thus, our studies have significantly advanced our understanding of VEX2 but we agree that further studies are necessary to appreciate the complex role of VEX2 as a key player in antigen expression regulation. However, we consider the primary finding of this manuscript to be "a novel mechanism to ensure monogenic expression, requiring the spatial integration of antigen transcription and mRNA splicing in a dedicated compartment. These findings suggest a new means of post-transcriptional gene regulation" (end of the abstract).

3.5: *Depletion of VEX2 appears to induce a series of events, which are quite important for interpretation of its function. However, these are not carefully discussed. The earliest event occurred at 12 hr after VEX2 RNAi induction, where longer distance between the active VSG transcription site and RNA processing center was observed. Later, VSG derepression was observed at 24 hrs after VEX2 RNAi induction (but not at 12 hrs). Also, Hi-C analysis showing increased association between derepressed VSG loci and the SL gene array occurred at 24 hrs. It is therefore possible that the immediate effect of removal of VEX2 is just a loss of closely positioned active VSG transcription site and the RNA processing center. This may induce a secondary effect that indirectly caused derepression VSGs at 24 hrs. For example, not efficiently processed active VSG mRNA may cause a decrease in VSG synthesis, which can induce derepression of silent VSGs, as VSG is essential.*

R3.5: Both the Hi-C and the new FISH analyses were carried out 24 h following VEX2 depletion, whereas "For the ESB / tSNAP localization following VEX2 or VEX1 / VEX2 RNAi (Fig. 4a-b; Extended Data Fig. 6c-d), all the imaging and analyses were performed at 12 h post-induction, a timepoint where there was sufficient VEX2 knockdown (Extended Data Fig. 6b) but both nucleolar Pol I and the ESB could be detected in > 85% of cells; the ESB is not detectable at later time points (Faria *et al.*, 2019, PMID: 31289266)". Thus, the timing of events we observe in perturbed cells, is consistent with a 'choose and consolidate' system in reverse; loss of protein condensates followed by loss of exclusion. We've now added additional discussion in relation to our model (see R3.4) to include a temporal component underpinning the exclusion mechanism.

Reviewer #4 (Genome organization):

Overall, I think this is an interesting manuscript and the conclusions are generally well supported by the data. I think it provides a good view of the assembly of the VSG transcriptional complex.

Overall Points.

4.1: *Fig 2. And text line 182. There appears to be a major missing issue that should be discussed. The claim is that the interaction of the splicing locus is more with the VSG gene itself and not so much with the promoter. At first site this might seem reasonable. However, between the promoter and the VSG*

gene there are many ESAG genes – and these are known to be trans-spliced and expressed. They have to be transcribed at the same rate as the VSG gene since it is the last one in the line and there are no promoters in between (I recognise that the half-lives of the RNA or processing may be differentially regulated). So if the idea is that the association is more with the VSG gene that the promoter region is the conclusion (not in the paper I believe) also that the VSG gene is much more spliced than the ESAGs? If so, how? This result surely suggests that during or after transcription of the ESAG RNA molecules then they are not recognised, processed or matured by the splicing machinery that interacts more heavily with the VSG gene even though one passage of an RNA Pol I should produce similar amounts

of RNA from each ESAG and the VSG gene? Or am I missing something? If this is so, or not I think it needs a discussion in the paper.

R4.1: We apologize for not being clearer on this point. The referee is correct, assuming that transcription of expression sites is only initiated at the expression site promoter, ESAGs should be transcribed at the same or higher levels than the VSG. However, from numerous RNA-seq experiments it has become clear that steady-state VSG levels are higher than those of ESAGs, pointing to more efficient VSG processing and/or higher VSG mRNA stability.

In this study our Hi-C data revealed more interactions between the active VSG and the SL-RNA array than between the ESAGs and the SL-RNA array. Previously we found that active-site ESAGs are negatively regulated by VEX2 (Faria *et al.*, 2019, PMID: 31289266). Based on these two findings, we suspect VEX2 to be important for the establishment of the juxtaposition of the SL-RNA array and the active VSG gene but not the adjacent ESAGs with the SL-RNA array.

Based on this model, the processing of the VSG but not of the ESAGs would be 'enhanced'. We are aware that polyadenylation of an upstream transcript is coupled to *trans*-splicing of a downstream transcript (Matthews *et al.*, 1994, PMID: 7907303). However, this does not mean that ESAG processing has to be coupled to VSG processing. The VSG gene is separated from the ESAGs by a long 70 bp repeat region and is directly preceded by a polyadenylated ncRNA (Davies *et al.*, 1997, PMID: 9200123; Juan Jose Vasquez, 2016, urn:nbn:de:bvb:20-opus-133996). Thus, we believe that spatial organization can enhance VSG transcript processing independent of ESAG transcript processing.

The idea that ESAGs are less well processed because they interact less frequently with the SL-RNA locus under physiological conditions, is supported by our finding that ESAG genes are strongly upregulated upon VEX2 depletion, when we find them to more frequently interact with the SL-RNA array. In the manuscript we address this accordingly: "Besides the VSG genes located in previously silent expression sites, expression site associated genes were also strongly upregulated following VEX2 knockdown (Faria *et al.*, 2019, PMID: 31289266). In line with this finding, we observed the largest increase in SL-RNA interactions for the regions upstream of the VSG gene in each de-repressed expression site, where expression site associated genes are located (Fig. 4f, Extended Data Fig. 8b)."

4.2: *I was left a little wanting by the model. They have VEX2 as the central protein to bring the ESB to a SL-RNA array and this stops silent VSGs accessing that SL-RNA array but what about the other SL- RNA array that also has a VEX1 foci? Why can't the silent VSGs access that? Is there a natural tendency of expression sites to interact SL-RNA array due presence other unknown factors – transient interactions but locked in by VEX2? State uncovered during VEX2 RNAi. To give this variant of their model: VEX1 mark SL-RNA ◊ expression sites sampling SL-RNA arrays ◊engagement through VEX2 ◊ once established this site "hoovers up" the rest of the required factors so other expression sites can't establish. ie a choose and consolidate model.*

R4.2: We agree that there is likely a temporal component underpinning the exclusion mechanism and, following this suggestion, have included this very appealing concept in describing our model.
Based on all available data (previously published and in the current manuscript) our model for the role of VEX2 is as follows (see also R3.4):

"Our data indicate that VEX2 is not simply a tether. Instead, they led us to propose a temporal 'choose and consolidate' component to our model: VEX2 is recruited to expression sites in a stochastic and competitive manner by Pol I transcription ('choose'), as supported by Pol I inhibition experiments that resulted in dispersion of VEX2 (Faria *et al.*, 2019, PMID: 31289266) and (Extended Data Fig. 5b). Subsequently, the presence of VEX2 at one expression site 'consolidates' expression site transcription, as supported by VEX2 depletion experiments that resulted in delocalization of the Pol I focus from the expression site (Fig. 4a-b). Once Pol I is no longer sequestered at one specific expression site, the enzyme can access and activate previously silent expression sites. Transcription of expression sites is accompanied by an increase in interactions with the SL-RNA array, as supported by Hi-C and FISH

(Fig. 4c-f), leading to efficient VSG mRNA processing. According to this model, VEX2 is the key molecule excluding all but one VSG expression site from the SL-RNA transcription compartment, thereby ensuring expression of a single VSG gene per cell."

Specific Points.

SP4.1: *Title. In the VSG expression site world the term switching has associations with (amongst other things) chromosomal recombination. The authors might just ponder if "switching inter-chromosomal interactions" may give an initial impression that the paper is about recombination not nuclear position effects?*

R-SP4.1: We carefully considered this point and agree that our study focusses primarily on 'monogenic expression' rather than 'switching'. Accordingly, we have adjusted the title to: "An allele-specific interaction with an RNA splicing locus sustains monogenic antigen expression in trypanosomes".

SP4.2: *Figure 1a there is a large interaction with chm 11 as well as the spliced leader array on chm 9. What is at this point? Is there some statistical analysis to guide which are the more significant peaks than others for the HiC data representation? I think these other significant peaks need to be identified*. **R-SP4.2:** The referee is correct, the expression site 1 makes two strong interactions, the strongest is with the SL-array on chr. 9, the second strongest **interaction is with a site on chr. 11**. The site on chr. 11 represents a centromere. As mentioned above (R2.4), expression site 1 is located directly next to the centromere on chr. 6. As observed in other organisms, we had previously shown that centromeres cluster with each other in *T. brucei* (Müller *et al.*, 2018, PMID: 30333624). Therefore, we suspected the interaction of expression site 1 with the centromere of chr. 11 to be rather a consequence of centromere clustering than with a second specific genomic locus important for VSG processing. To validate this assumption, we confirmed that this interaction is detected in different cell lines (ED Fig. 3a), independent of the expression status of expression site 1. Therefore, we believe that this interaction is a secondary effect, caused by the location of expression site 1 next to a centromere and it does not play a role in expression site activity. We have now added information regarding this second strong interaction peak to the legend of Fig 1 "* marks the centromere on chr. 11".

Conventional **peak callers**, such as those used for ChIP-seq analyses, don't perform well in the detection of inter-chromosomal interactions, as intra-chromosomal interactions tend to be much stronger than inter-chromosomal interactions and because intra-chromosomal interactions decrease in a distant dependent manner. Thus, for this manuscript we had focused on the most frequent interactions (those with >20% of the maximum inter-chromosomal interaction frequency) that could not be explained by close proximity on the same chromosomes. In addition, we searched for peaks that changed following a switch in VSG expression (Fig. 2). Only the interaction between the SL-RNA and the active VSG appeared to meet these criteria. However, we understand that the second strong peak on chr. 11 (Fig 1a, top panel) is very intriguing and have now added information on this region to the legend and the manuscript.

"Looking for further genomic loci that made inter-chromosomal interactions with the *VSG-2* gene of at least 20% of the *VSG-2* – SL-RNA interaction frequency, we identified a second locus: The centromere of chr. 11 interacted with *VSG-2* at 37% of the frequency observed for the *VSG-2* – SL-RNA interaction (Fig. 1a, top panel). Since the *VSG-2* gene is located next to the centromere of chr. 6, we suspect these interactions to be a consequence of previously observed centromere – centromere interactions, and not to be related to the active expression of *VSG-2*."

Later we write: "As suspected, *VSG-2* – centromere (chr. 11) interactions remained unchanged after inactivation of *VSG-2* (Extended Data Fig 3a), suggesting that this is indeed a consequence of centromere – centromere interactions."

SP4.3: *The image analysis in places left some uncertainties – the image they show for tSNAP has one large foci and one small one – is this consistent? And is the small one always close to the ESB?* **R-SP4.3:** We quantified tSNAP focus signal intensity and find that those foci adjacent to an ESB are significantly less intense than those that are distal (see ED Fig. 2d-f; also see R2.3 above).

SP4.4: *In the extended figure 1 they include viewpoints for central core genes, which is an important control but I didn't spot this mentioned in the main text. Plus, what are the central core – which genes?*
R-SP4.4: The information about the control core regions was listed but in Data S1. We have now also included this information in the legend of ED Fig. 1, where it is more accessible.

"As viewpoint control regions, two actively transcribed regions from the diploid cores of chr. 2 and chr 8 and two non-transcribed regions from the subtelomeres of chr. 1 and chr. 10 were chosen. All four control regions were arbitrarily chosen. The coordinates of all viewpoints used for virtual 4C analyses are listed in Data S1." Given the bin size of 50 kb, each control region includes several genes. We have now added the information that "coordinates of all viewpoints used for virtual 4C analyses are listed in Data S1" to all legends of figures showing virtual 4C analyses.

In addition, we have added a mention of the control regions to the main text: "Conversely, just like arbitrarily chosen control region, VSG genes residing in inactive expression sites interacted less frequently or at background levels with the SL-RNA locus (Fig. 1a-b; Extended Data Fig. 1a)."

SP4.5: *They say that the majority of cells 55% have 2 VEX1 foci but what about the other 45% - is there a single VEX1 foci or multiple? In the other 45% of the cells do these also have 2 tSNAP foci. For the categorisation in 3a, were any instances seen when the two sets of signals in the same cell are different? i.e. one overlapping and one adjacent. If so how was this dealt with in the counts?*

R-SP4.5: We've added details: "one VEX1 focus was detected in the remainder". The data shown in Fig. 3a are from nuclei with two VEX1 foci and two tSNAP foci; we obtained similar results from nuclei with one VEX1 foci and two tSNAP foci (the 45% of cells with 1 VEX1 focus also have 2 tSNAP foci). In the small number of cases where nuclei had one overlapping pair of foci and one adjacent pair of foci, both pairs were included in the counts.

SP4.6: *From the image in 3a there is a large VEX1 foci associated with the small tSNAP foci and presumably the ESB – is this a consistent result?*

R-SP4.6: This is an interesting observation. We quantified tSNAP focus signal intensity (see SP4.3 above) and VEX1 focus signal intensity. "VEX1 signal intensity was significantly higher at the focus adjacent to the VSG transcription compartment (Extended Data Fig. 4a), indicating that all nuclei may have two VEX1 foci, but the second focus may be below the detection limit in some cells." We also replaced the previous example images in ED Fig. 4a with images for two new example cells to better illustrate detection of either one or two VEX1 foci.

SP4.7: *The categories in 3a/b are G1 and total cells and do not look at cells specifically in S-phase like 1c. There is a drop in the number of overlapping foci in total cells and this could come from those cells in S-phase so it's difficult to be confident about whether these proteins remain in close association during S-phase.*

R-SP4.7: We've adjusted Fig. 3a-b to include cells in S phase (also see R1.2 above).

SP4.8: *Do they know if VEX1 still associates with tSNAP in PCFs interesting but not essential for their conclusions here?*

R-SP4.8: We have added IFA data for PCFs showing that "VEX proteins also redistribute but a pool of VEX1 remains detectable at SL-RNA transcription compartments (Extended Data Fig. 4d)".

SP4.9: *Their previous paper showed that VEX1 RNAi had no effect on VEX2 localisation but VEX1 overexpression causes loss of VEX2 foci – what happens to VEX1 foci during VEX2 RNAi? From their data you would predict VEX1 localisation is not dependent on VEX2.*

R-SP4.9: VEX1 is redistributed across multiple sub-nuclear puncta following loss of VEX2 (see Fig. 6c in Faria *et al.*, 2019 (PMID: 31289266). This is consistent with the 'choose and consolidate' component of the model (see R4.2 above).

Decision Letter, second revision:

Dear Nicolai,

Thank you for your patience while your manuscript "An allele-specific interaction with an RNA splicing locus sustains monogenic antigen expression in trypanosomes" was under peer review at Nature Microbiology. It has now been seen by our referees, and in the light of their advice I am delighted to say that we can in principle offer to publish it. First, however, we would like you to revise your paper to address the points made by the reviewers, and to ensure that it is in Nature Microbiology format.

Referee #2 raises three points that do not require further experiments, but which they believe should be addressed/clarified in the Discussion. The referees' remaining comments are clear, and should not be difficult to implement. Editorially, we will need you to make some changes so that the paper complies with our Guide to Authors at http://www.nature.com/nmicrobiol/info/gta.

Nature Microbiology offers a transparent peer review option for new original research manuscripts submitted from 1st December 2019. We encourage increased transparency in peer review by publishing the reviewer comments, author rebuttal letters and editorial decision letters if the authors agree. Such peer review material is made available as a supplementary peer review file. Please state in the cover letter 'I wish to participate in transparent peer review' if you want to opt in, or 'I do not wish to participate in transparent peer review' if you don't. Failure to state your preference will result in delays in accepting your manuscript for publication.

Please note: we allow redactions to authors' rebuttal and reviewer comments in the interest of confidentiality. If you are concerned about the release of confidential data, please let us know specifically what information you would like to have removed. Please note that we cannot incorporate redactions for any other reasons. Reviewer names will be published in the peer review files if the reviewer signed the comments to authors, or if reviewers explicitly agree to release their name. For more information, please refer to our FAQ page.

In recognition of the time and expertise our reviewers provide to Nature Microbiology's editorial process, we would like to formally acknowledge their contribution to the external peer review of your manuscript entitled "An allele-specific interaction with an RNA splicing locus sustains monogenic antigen expression in trypanosomes". For those reviewers who give their assent, we will be publishing their names alongside the published article.

I appreciate this email is long and recommend that you print it and use it as a checklist, reading it carefully to the end, in order to avoid delays to publication down the line.

Please note that we will be considering your paper for publication as a LETTER in our pages.

Specific points:
In particular, while checking through the manuscript and associated files, we noticed the following specific points which we will need you to address:

1. Main text display items and supplementary information. Please note that we have recently started publishing additional figures as "Extended Data". These figures appear online in the html version of the manuscript in the place they are referred to and greatly increase discoverability of the data that is presented in them.
All Supplementary Information must be submitted in accordance with the instructions in the attached Inventory of Supporting Information, and should fit into one of three categories: EXTENDED DATA (ED); SUPPLEMENTARY INFORMATION (SI); and SOURCE DATA. Below are detailed instructions on how to format each category. For your paper, we suggest that you do the following:

a. Main figures: please maintain the current 4 main figures that illustrate the main findings of the paper.

b. Extended data (ED): please maintain the current 9 ED figures. These are an integral part of the paper (presented online in the online version) and are meant to be multipanel A4 size figures. More information on file formats and how the legends should be supplied can be found below and in the attached Inventory of Supporting Information.

c. Supplementary information (SI): your study will have the 'Supplementary online data' as SI. Please submit all SI as a separate pdf file. All supplementary materials need to be assembled into a single file, including all tables (excluding those that are excessively large). In the Supplementary Information file, figure legends should be immediately below each figure and the pages should be numbered.

d. Source data: this format should be used to display source data linked to the main figures and ED figures.
We strongly encourage you include as much additional raw data underlying the graphs in the main and ED figures as possible. These data should be supplied as Excel tables, one file per main or ED figure, and should be clearly labeled and presented in a way that individual experiments are identifiable (for example, across a time course if applicable).

2. Data Availability statement. The data availability statement should clearly refer to all of the source data provided in the manuscript (more instructions on how to write this section can be found in the general formatting guidelines below).

3. Reporting checklist. Please revise this document according to the instructions found in the annotated PDF attached to this message and send in a final version with your article. The final reporting checklist will be published with your manuscript.

4. Competing interest statement. The competing interest statement needs to be included in the manuscript text (before or after the Acknowledgements).

5. New/Novel. There is an instance of the use of the word novel in the text (line 41). Please remove it, except if it is strictly necessary. It is journal policy to limit unnecessarily hyperbolic use of terms related to novelty.

6. Title. We think it may be good to modify the title and to highlight this part - spatial integration of antigen transcription and mRNA splicing in a dedicated compartment.

7. Abstract. Please make it clear in the abstract what we knew before this study and what we know now. The use of the word 'enigmatic' may not be interpretable to all, and we suggest clearly stating what was known in one line and what you wanted to find out in the next part.

8. ORCID. We now require corresponding authors to provide an ORCID identifier, and would ask that you please provide one with your final submission (please also see below). There is a step during the upload of the information to our online system in which the number can be introduced.

9. Replicates and statistics. While carefully checking the figures, we noted a few things that need to be revised so that they comply with our style guidelines and accurately report on the number of replicates, statistical testing, etc. As general rules, please note that:

General comments:

Wherever statistics have been derived (e.g., error bars, box plots, statistical significance), the legend needs to provide and define the n number (i.e., the sample size used to derive statistics) as a precise value (not a range), using the wording "n=X biologically independent samples/animals/independent experiments," etc. as applicable.

All error bars need to be defined in the legends (e.g., SD, SEM) together with a measure of centre (e.g. mean, median), and should be accompanied by their precise n number defined as noted above.

All violin plots need to be defined in the legends in terms of minima, maxima, centre, and percentiles, and should be accompanied by their precise n number defined as noted above.

The figure legends must indicate the statistical test used and if applicable, whether the test was one- or two-sided. A description of any assumptions or corrections such as tests of normality and adjustment for multiple comparisons must also be included.

For null hypothesis testing, please indicate the test statistic (e.g., F, t, r) with confidence intervals, effect sizes, degrees of freedom and P values noted.

Test results (e.g., p-values, q-values) should be given as exact values whenever possible and appropriate, and confidence intervals noted.

Please indicate how estimates of effect sizes were calculated (e.g., Cohen's d, Pearson's r).

Please state in the legends how many times each experiment was repeated independently with similar results. This is needed for all experiments but is particularly important wherever representative experiments are shown. If space in the legends is limiting, this information can be included in a section titled "Statistics and Reproducibility".

For all bar graphs, the corresponding dot plot must be overlaid.

Specific comments to address:

Please see the attached "Extended_comments" file.

General points:

Please read carefully through all of the following general formatting points when preparing the final version of your manuscript, as submitting the manuscript files in the required format will greatly speed the process to final acceptance of you work.

Titles should give an idea of the main finding of the paper and ideally not exceed 150 characters (including spaces). We discourage the use of active verbs and do not allow punctuation.

The introductory paragraph (not abstract) should be of approximately 200 words, summarizing the background, rationale, main results (introduced by "Here we show" or some equivalent phrase) and implications of the study. This paragraph should be fully referenced and should be considered part of the main text, so that any subsequent introductory material avoids too much redundancy with the introductory paragraph. Because we hope that scientists across the wider microbiology community will be interested in your work, the first paragraph should be as accessible as possible, explaining essential but specialised terms concisely. We suggest you show your summary paragraph to colleagues in other fields to uncover any problematic concepts.

Letters are not divided by headings, except for the Methods heading.

We strongly support public availability of data. Please place the data used in your paper into a public data repository, if one exists, or alternatively, present the data as Source Data or Supplementary Information. If data can only be shared on request, please explain why in your Data Availability Statement, and also in the correspondence with your editor. For some data types, deposition in a public repository is mandatory - more information on our data deposition policies and available repositories can be found at https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-data.

Please include a data availability statement as a separate section after Methods but before references, under the heading "Data Availability". This section should inform readers about the availability of the data used to support the conclusions of your study. This information includes accession codes to public repositories (data banks for protein, DNA or RNA sequences, microarray, proteomics data etc…), references to source data published alongside the paper, unique identifiers such as URLs to data repository entries, or data set DOIs, and any other statement about data availability. At a minimum, you should include the following statement: "The data that support the findings of this study are available from the corresponding author upon request", mentioning any restrictions on availability. If DOIs are provided, we also strongly encourage including these in the Reference list (authors, title, publisher (repository name), identifier, year). For more guidance on how to write this section please see:
http://www.nature.com/authors/policies/data/data-availability-statements-data-citations.pdf

Please supply the figures as vector files - EPS, PDF, AI or postscript (PS) file formats (not raster or bitmap files), preferably generated with vector-graphics software (Adobe Illustrator for example). Try to ensure that all figures are non-flattened and fully editable. All images should be at least 300 dpi resolution (when figures are scaled to approximately the size that they are to be printed at) and in RGB colour format. Please do not submit Jpeg or flattened TIFF files. Please see also 'Guidelines for Electronic Submission of Figures' at the end of this letter for further detail.
Please view http://www.nature.com/authors/editorial_policies/image.html for more detailed guidelines.

We will edit your figures/tables electronically so they conform to Nature Microbiology style. If necessary, we will re-size figures to fit single or double column width. If your figures contain several parts, the parts should be labelled lower case a, b, and so on, and form a neat rectangle when

assembled.

Please check the PDF of the whole paper and figures (on our manuscript tracking system) VERY CAREFULLY when you submit the revised manuscript. This will be used as the 'reference copy' to make sure no details (such as Greek letters or symbols) have gone missing during file-transfer/conversion and re-drawing.

All Supplementary Information must be submitted in accordance with the instructions in the attached Inventory of Supporting Information, and should fit into one of three categories:

1. EXTENDED DATA: Extended Data are an integral part of the paper and only data that directly contribute to the main message should be presented. These figures will be integrated into the full-text HTML version of your paper and will be appended to the online PDF. There is a limit of 10 Extended Data figures, and each must be referred to in the main text. Each Extended Data figure should be of the same quality as the main figures, and should be supplied at a size that will allow both the figure and legend to be presented on a single legal-sized page. Each figure should be submitted as an individual .jpg, .tif or .eps file with a maximum size of 10 MB each. All Extended Data figure legends must be provided in the attached Inventory of Accessory Information, not in the figure files themselves.

2. SUPPLEMENTARY INFORMATION: Supplementary Information is material that is essential background to the study but which is not practical to include in the printed version of the paper (for example, video files, large data sets and calculations). Each item must be referred to in the main manuscript and detailed in the attached Inventory of Accessory Information. Tables containing large data sets should be in Excel format, with the table number and title included within the body of the table. All textual information and any additional Supplementary Figures (which should be presented with the legends directly below each figure) should be provided as a single, combined PDF. Please note that we cannot accept resupplies of Supplementary Information after the paper has been formally accepted unless there has been a critical scientific error.

All Extended Data must be called you in your manuscript and cited as Extended Data 1, Extended Data 2, etc. Additional Supplementary Figures (if permitted) and other items are not required to be called out in your manuscript text, but should be numerically numbered, starting at one, as Supplementary Figure 1, not SI1, etc.

3. SOURCE DATA: We strongly encourage you to provide source data for your figures whenever possible. Full-length, unprocessed gels and blots must be provided as source data for any relevant figures, and should be provided as individual PDF files for each figure containing all supporting blots and/or gels with the linked figure noted directly in the file. Numerical source data that underlie graphs are required for in vivo experiments and strongly encouraged generally. They should be provided in Excel format, one file for each relevant figure, with the linked figure noted directly in the file. They should be clearly labelled such that individual experiments and/or animals are labelled (for example, across a time course if applicable). For imaging source data, we encourage deposition to a relevant repository, such as figshare (https://figshare.com/) or the Image Data Resource (https://idr.openmicroscopy.org).

Please ensure that you retain unprocessed data and metadata files after publication, ideally archiving data in perpetuity, as these may be requested during the peer review and production process or after publication if any issues arise.

Figure legends must provide a brief description of the figure and the symbols used, within 350 words.

This must include definitions of any error bars employed in the figures.

Please include any references for the Methods at the end of the reference list. Any citations in the Supplemental Information will need inclusion in a separate SI reference list.

It is a condition of publication that you include a statement before the acknowledgements naming the author to whom correspondence and requests for materials should be addressed.

Finally, we require authors to include a statement of their individual contributions to the paper -- such as experimental work, project planning, data analysis, etc. -- immediately after the acknowledgements. The statement should be short, and refer to authors by their initials. For details please see the Authorship section of our joint Editorial policies at http://www.nature.com/authors/editorial_policies/authorship.html

We will not send your revised paper for further review if, in the editors' judgement, the referees' comments on the present version have been addressed. If the revised paper is in Nature Microbiology format, in accessible style and of appropriate length, we shall accept it for publication immediately.

Please resubmit electronically

* the final version of the text (not including the figures) in either Word or Latex.

* publication-quality figures. For more details, please refer to our Figure Guidelines, which is available here: https://www.nature.com/documents/NRJs-guide-to-preparing-final-artwork.pdf

* Extended Data & Supplementary Information, as instructed

* a point-by-point response to any issues raised by our referees and to any editorial suggestions.

* any suggestions for cover illustrations, which should be provided at high resolution as electronic files. Please note that such pictures should be selected more for their aesthetic appeal than for their scientific content. I am sure you will understand that we cannot make any promise as to whether any of your suggestions might be selected for the cover of Nature Microbiology.

Please use the following link to access your home page:

{REDACTED}

*This url links to your confidential homepage and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this e-mail to co-authors, please delete this link to your homepage first.

Please also send the following forms as a PDF by email to microbiology@nature.com.

* Please sign and return the Licence to Publish form.

* Or, if the corresponding author is either a Crown government employee (including Great Britain and Northern Ireland, Canada and Australia), or a US Government employee, please sign and return the Licence to Publish form for

Crown government employees, or a Licence to Publish form for US government employees.

* Should your Letter contain any items (figures, tables, images, videos or text boxes) that are the same as (or are adaptations of) items that have previously been published elsewhere and/or are owned by a third party, please note that it is your responsibility to obtain the right to use such items and to give proper attribution to the copyright holder. This includes pictures taken by professional photographers and images downloaded from the internet. If you do not hold the copyright for any such item (in whole or part) that is included in your paper, please complete and return this Third Party Rights Table, and attach any grant of rights that you have collected.

For more information on our licence policy, please consult http://npg.nature.com/authors.

AUTHORSHIP

CONSORTIA -- For papers containing one or more consortia, all members of the consortium who contributed to the paper must be listed in the paper (i.e., print/online PDF). If necessary, individual authors can be listed in both the main author list and as a member of a consortium listed at the end of the paper. When submitting your revised manuscript via the online submission system, the consortium name should be entered as an author, together with the contact details of a nominated consortium representative. See https://www.nature.com/authors/policies/authorship.html for our authorship policy and https://www.nature.com/documents/nr-consortia-formatting.pdf for further consortia formatting guidelines, which should be adhered to prior to acceptance.

ORCID

Nature Microbiology is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as ‘corresponding author’ create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS) prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. For more information please visit http://www.springernature.com/orcid

For all corresponding authors listed on the manuscript, please follow the instructions in the link below to link your ORCID to your account on our MTS before submitting the final version of the manuscript. If you do not yet have an ORCID you will be able to create one in minutes. https://www.springernature.com/gp/researchers/orcid/orcid-for-nature-research

IMPORTANT: All authors identified as ‘corresponding author’ on the manuscript must follow these instructions. Non-corresponding authors do not have to link their ORCIDs but are encouraged to do so. Please note that it will not be possible to add/modify ORCIDs at proof. Thus, if they wish to have their ORCID added to the paper they must also follow the above procedure prior to acceptance.

To support ORCID's aims, we only allow a single ORCID identifier to be attached to one account. If you have any issues attaching an ORCID identifier to your MTS account, please contact the Platform Support Helpdesk.

We hope that you will support this initiative and supply the required information. Should you have any

query or comments, please do not hesitate to contact me.

Nature Research journals encourage authors to share their step-by-step experimental protocols on a protocol sharing platform of their choice. Nature Research's Protocol Exchange is a free-to-use and open resource for protocols; protocols deposited in Protocol Exchange are citable and can be linked from the published article. More details can found at www.nature.com/protocolexchange/about.

We hope to hear from you within two weeks; please let us know if the revision process is likely to take longer.

****

Reviewer Expertise:

Referee #1: Parasitology
Referee #2: Molecular cell biology of trypanosomes
Referee #3: Recombination in trypanosomes
Referee #4: Genome organization

Reviewer Comments:

Reviewer #1 (Remarks to the Author):

To my concerns, the authors have carefully addressed all my comments and came back with a significantly better manuscript.

Reviewer #2 (Remarks to the Author):

The authors have addressed most of my comments in a convincing manner. I raise below 3 points that do not require further experiments, but I believe should be addressed/clarified in the Discussion.

1. Discussion of the Model.

I appreciate the clarifications and FISH experiments to address the changes in protein/DNA localisation and the temporal differences (12/24hr). Based on the "choose and consolidate model" (commonly referred as "establish and maintain" functions), VEX2 is the key component that limits the activation of silent VSGs. At 12 hr, depletion of VEX2 causes PolI to delocalize from BES (distance increases slightly, but Pol I still appears as a single "dot"). If I am not mistaken, there is no FISH of Hi-C data at this time point (12hr), but authors propose that BES is still interacting with SL DNA locus. Q1. Do the author means that PolI leaves the BES DNA sequence and the BES-SL interacting site, but remains nearby for a few hours? Are there examples of such dissociations in the literature?

At 24hr, PolI is dispersed (i.e., signal is fainter and detected further away from SL locus – Faria et al, 2019) and in these new locations PolI is "free" to initiate transcription from silent BESs. At this time, SL is still attached to active BES, but it also increases the frequency of interactions with silent BESs, which form clusters.

Q2. Given there is only one chromosome locus of SL genes (in each homologous chromosome), how do authors envision SL locus interacting simultaneously with active and silent BESs?
Q3. Do the authors propose that SL locus "moves" to the sites that are pre-defined by where Pol I transcription is happening? In other words, are the authors proposing that the interaction of SL - BES happens only AFTER the stochastic "choice" of VSGs (rather than being a pre-requisite for a VSG to be chosen/transcribed)?

The discussion of the model should clarify the points above.

2. In Extended Figure 5b, it is not clear to me why Pol I signal is undetectable when cells are treated for 30min with transcription inhibitors. Does Pol I signal disperse and become undetectable? While the conclusion that VEX2 localisation depends on Pol I localisation or VSG expression RNAs is valid, without detecting PolI signal, the authors cannot assess by IFA whether the interaction of VSG (PolI) and the SL-RNA transcription (tSNAPc) compartments is dependent on Pol I transcription (line 296-299). The conclusion needs to be rephrased.

3. Upon VEX2 depletion, the authors describe an increased association of SL with ESAGs, genes upstream VSG (Fig 4f). Could the authors speculate the relevance of this observation? This association was not found in isogenic parasites stably expressing two different VSG genes (Fig 2).

Reviewer #3 (Remarks to the Author):

Re: NMICROBIOL-20030686B
Title: An allele-specific interaction with an RNA splicing locus sustains monogenic antigen expression in trypanosomes
Authors: Faria et al.

The authors have added more experiment data, explained the experiment procedures and results more clearly in the revised version. The authors have also taken into consideration of the temporal effects of VEX2's action and revised their model. The authors have addressed all my previous questions.

Reviewer #4 (Remarks to the Author):

I have now had time to really go over this manuscript in detail. It is an important manuscript and I believe there is now no point in worrying about the small matters.
They have substantially added to the evidence and I believe it can now be published. I would make one specific point. I think the reviewing of this manuscript has substantially improved it (all 4 reviewers made good points). Indeed they accepted my alternative model of their results (even using my phrase for the model in the resubmitted paper). In such circumstances I think the scientific norm for courtesy would be to say in acknowledgements something like
"We thank the anonymous referees who provided insightful comments on our manuscript which added value to the work".

Final Decision Letter:

nature research

Dear Nicolai,

I am pleased to accept your Letter "Spatial integration of transcription and splicing in a dedicated compartment sustains monogenic antigen expression in African trypanosomes" for publication in Nature Microbiology. Thank you for having chosen to submit your work to us and many congratulations.

Before your manuscript is typeset, we will edit the text to ensure it is intelligible to our wide readership and conforms to house style. We look particularly carefully at the titles of all papers to ensure that they are relatively brief and understandable.

The subeditor may send you the edited text for your approval. Once your manuscript is typeset you will receive a link to your electronic proof via email within 20 working days, with a request to make any corrections within 48 hours. If you have queries at any point during the production process then please contact the production team at rjsproduction@springernature.com. Once your paper has been scheduled for online publication, the Nature press office will be in touch to confirm the details.

Acceptance of your manuscript is conditional on all authors' agreement with our publication policies (see www.nature.com/nmicrobiolate/authors/gta/content-type/index.html). In particular your manuscript must not be published elsewhere and there must be no announcement of the work to any media outlet until the publication date (the day on which it is uploaded onto our website).

The Author's Accepted Manuscript (the accepted version of the manuscript as submitted by the author) may only be posted 6 months after the paper is published, consistent with our self-archiving embargo. Please note that the Author's Accepted Manuscript may not be released under a Creative Commons license. For Nature Research Terms of Reuse of archived manuscripts please see: http://www.nature.com/authors/policies/license.html#terms
If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. All co-authors, authors' institutions and authors' funding agencies can order reprints using the form appropriate to their geographical region.

We welcome the submission of potential cover material (including a short caption of around 40 words) related to your manuscript; suggestions should be sent to Nature Microbiology as electronic files (the image should be 300 dpi at 210 x 297 mm in either TIFF or JPEG format). Please note that such pictures should be selected more for their aesthetic appeal than for their scientific content, and that colour images work better than black and white or grayscale images. Please do not try to design a cover with the Nature Microbiology logo etc., and please do not submit composites of images related to your work. I am sure you will understand that we cannot make any promise as to whether any of your suggestions might be selected for the cover of the journal.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of

your refereeing activity for the Nature journals.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

Congratulations once again and I look forward to seeing the article published.