Comparative genomics identifies male accessory gland proteins in five Glossina species [version 2; referees: 2 approved]

Muna F. Abry¹,2, Kelvin M. Kimenyi¹,2, Daniel K Masiga¹, Benard W. Kulohoma¹,2

¹International Centre for Insect Physiology and Ecology, P.O. Box 30772, Nairobi, 00100, Kenya
²Center for Biotechnology and Bioinformatics, University of Nairobi, P.O. Box 30197, Nairobi, 00100, Kenya

Abstract
Accessory gland proteins (ACPs) are important reproductive proteins produced by the male accessory glands (MAGs) of most insect species. These proteins are essential for male insect fertility, and are transferred alongside semen to females during copulation. ACPs are poorly characterized in Glossina species (tsetse fly), the principal vector of the parasite that causes life-threatening Human African Trypanosomiasis and Animal trypanosomiasis in endemic regions in Africa. The tsetse fly has a peculiar reproductive cycle because of the absence of oviposition. Females mate once and store sperm in a spermathecal, and produce a single fully developed larva at a time that pupates within minutes of exiting their uterus. This slow reproductive cycle, compared to other insects, significantly restricts reproduction to only 3 to 6 larvae per female lifespan. This unique reproductive cycle is an attractive vector control strategy entry point. We exploit comparative genomics approaches to explore the diversity of ACPs in the recently available whole genome sequence data from five tsetse fly species (Glossina morsitans, G. austeni, G. brevipalpis, G. pallidipes and G. fuscipes). We used previously described ACPs in Drosophila melanogaster and Anopheles gambiae as reference sequences. We identified 36, 27, 31, 29 and 33 diverse ACP orthologous genes in G. austeni, G. brevipalpis, G. fuscipes, G. pallidipes and G. morsitans genomes respectively, which we classified into 21 functional classes. Our findings provide genetic evidence of MAG proteins in five recently sequenced Glossina genomes. It highlights new avenues for molecular studies that evaluate potential field control strategies of these important vectors of human and animal disease.

Keywords
Accessory gland proteins, tsetse fly, reproductive cycle, vector control, trypanosomiasis
Corresponding author: Benard W. Kulohoma (bkulohoma@iscb.org)

Author roles: Abry MF: Formal Analysis, Methodology, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Kimenyi KM: Formal Analysis, Methodology, Writing – Review & Editing; Masiga DK: Conceptualization, Supervision, Writing – Review & Editing; Kulohoma BW: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Supervision, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Abry MF, Kimenyi KM, Masiga DK and Kulohoma BW. Comparative genomics identifies male accessory gland proteins in five Glossina species [version 2; referees: 2 approved] Wellcome Open Research 2017, 2:73 (doi: 10.12688/wellcomeopenres.12445.2)

Copyright: © 2017 Abry MF et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: The work was supported by the Wellcome Trust [087540], a pump-priming grant to BWK from the Training Health Researchers into Vocational Excellence in East Africa (THRiVE); UK’s Department for International Development (DFID); Swedish International Development Cooperation Agency (Sida); the Swiss Agency for Development and Cooperation (SDC); and the Kenyan Government.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Introduction

Accessory gland proteins (ACPs) are important reproductive proteins produced by the male accessory glands (MAGs) of most insect species. These proteins are essential for male insect fertility, and are transferred alongside semen to females during copulation. ACPs trigger significant physiological and behavioral changes in females after copulation, which include: egg laying, reduced sexual receptivity and refractoriness to subsequent inseminations, induce the expression of immune peptides and reduction of female lifespan. ACPs are poorly characterized in Glossina (tsetse fly), compared to Drosophila and Anopheles species. The tsetse fly is the principal vector of the parasite that causes life-threatening human (sleeping sickness) and cattle (nagana) trypanosomiasis in endemic regions in Africa. Over 60 million people and 80 million cattle are at risk of contracting disease. Female tsetse flies only mate once during their lifespan and store the male ejaculate in their spermaticae, which they subsequently use to self-fertilize. They have a peculiar reproductive cycle because of the absence of oviposition, with females producing a single fully developed larva at a time that pupates within minutes of exiting their uterus. This slow reproductive cycle, compared to other insects, significantly restricts reproduction to only 3 to 6 larvae per female lifespan. This unique reproductive cycle is an attractive vector control target. An improved understanding of tsetse fly’s reproductive biology, and specifically ACPs that are crucial determinants of successful reproduction in other insect species, may provide valuable possible vector control strategy entry points.

Comparisons between ACP gene orthologs in Drosophila simulans and D. melanogaster show they are rapidly evolving, relative to non-ACP genes. However, there is strong ACP peptide structural homology between closely related species, which decreases as species phylogenetic distances increase. This rapid rate of ACP genes evolution has made it challenging to reliably identify orthologs across insect species in the absence of genomic data. The recently available whole genome sequence data from five tsetse fly species (Glossina morsitans, G. austeni, G. brevipalpis, G. pallidipes and G. fuscipes) has made it possible to revisit detailed examination of ACP gene distribution and genetic diversity in tsetse flies. We exploit comparative genomics approaches to interrogate these genomes, using previously described ACPs in D. melanogaster and A. gambiae as reference sequences.

Methods

Identification of Glossina, Anopheles and Drosophila species ACP homologs

The proteomes of G. austeni, G. morsitans, G. pallidipes, G. fuscipes and G. brevipalpis were retrieved manually from VectorBase (www.vectorbase.org). The retrieved Glossina protein sequences alongside previously described ACP sequences from A. gambiae (n=57) and D. melanogaster (n=173) (Supplementary Table 1), used as references, were assigned to homologous clusters using OrthoMCL with default settings (BLASTP E-value cut-off 1e-5 and inflation index 2.5) (Supplementary File 1). Clusters with singletons were omitted from further processing. Mapped orthologs were subsequently processed using BMX.

Sequence alignment and phylogeny reconstruction

Multiple sequence alignments were performed using MUSCLE. Maximum likelihood (ML) phylogenetic analysis of the multiple aligned sequences with bootstrap values of 100 replicates was performed using PHYML version 3.5.

Determining the direction and extent of selection pressure

The magnitude and direction of selection pressure on the ACP sequences was tested based on the ratio (ω = dN/dS) of the average number of non-synonymous substitutions per non-synonymous site (dN) to the average number of synonymous substitutions per synonymous site (dS). If ω = 1, amino acid substitution is assumed to be under neutral selection. ω > 1 is indicative of positive selection whereas ω < 1 is evidence of negative or purifying selection. Sequence alignments of each of the ACP clusters...
containing A. gambiae, D. melanogaster, G. austeni, G. brevipalpis, G. fuscipes, G. moristans and G. pallidipes were generated. Each alignment was then uploaded to the SNAP program41 (www.hiv.lanl.gov), which calculates synonymous and non-synonymous substitution rates to determine the magnitude of selection pressure.

Data visualization
A list of the relative abundance of ACPs with secretory signals identified within each ortholog cluster for Glossina, Anopheles and Drosophila species was generated (Supplementary Figure 1). The 21 ACPs clusters were visually presented in a single circular ideogram using CIRCOS software22.

Results
Identification of ACP gene orthologs in Glossina species
We analyzed five recently sequenced Glossina genomes: Glossina morsitans, G. austeni, G. brevipalpis, G. pallidipes and G. fuscipes, to examine the presence of orthologs to ACP genes previously identified in Drosophila melanogaster and Anopheles gambiae (n=57)46. First, we identified 41 ACP ortholog clusters that we broadly classified into 23 groups based on the encoded protein’s functional class. These ACPs have a distinct species distribution with: 12 ortholog clusters common to Glossina, A. gambiae, and D. melanogaster; and some clusters only present in A. gambiae and Glossina species (n=7) or D. melanogaster and Glossina species (n=5) (Supplementary Table 1). The remaining 17 clusters consist of ACP orthologs exclusive to either A. gambiae (n=7) or D. melanogaster (n=10). Next, we shortlisted genes that encode proteins carrying classical secretory signals23 to distinguish the matched testes-specific polymorphism.

The remaining 17 clusters consist of ACP orthologs exclusive to either A. gambiae or D. melanogaster. We identified 41 ACP ortholog clusters that we broadly classified into 23 groups based on the encoded protein’s functional class. These ACPs have a distinct species distribution with: 12 ortholog clusters common to Glossina, A. gambiae, and D. melanogaster; and some clusters only present in A. gambiae and Glossina species (n=7) or D. melanogaster and Glossina species (n=5) (Supplementary Table 1). The remaining 17 clusters consist of ACP orthologs exclusive to either A. gambiae (n=7) or D. melanogaster (n=10). Next, we shortlisted genes that encode proteins carrying classical secretory signals23 to distinguish the matched testes-specific secreted male accessory gland proteins from other insect peptides. We identified 36, 27, 31, 29 and 33 ACP orthologs with secretory signals in G. austeni, G. brevipalpis, G. fuscipes, G. pallidipes and Glossina morsitans genomes respectively (Figure 1), across 21 functional class groups. α2-macroglobulins (Group 1) and heat shock proteins (Group 17) are the most abundant ACP orthologs in Glossina species (Figure 1). Interestingly, Glossina species lack orthologs to Acp70A (Group 18) and andropin (Group 19), which has antimicrobial properties and safeguards the male ejaculate, and stimulation of long-term post mating responses in females respectively1,24,25.

Most ACP genes are under positive selection
We inferred the direction and magnitude of selection pressure on the identified ACP orthologs using dN/dS ratios. We observed positive signatures of positive selection in all genes except five are evolving under positive selection (Figure 1). We found that α2-macroglobulins, which have been shown to be important in mosquito and Drosophila immunity26,27, display signatures of purifying selection suggesting they are critical for successful reproduction and all deleterious variations are purged. Our analysis was restricted to reference genes present in A. gambiae and D. melanogaster, and future studies that integrate data from more closely related taxa will highlight evolutionary changes associated with ACPs in more detail. We reconstructed the phylogeny of ACP orthologs within each cluster (Supplementary Figure 1). We failed to identify any pattern associated with the diverse ecological niche and unique reproductive style in Glossina species in our analysis.

ACP distribution in Glossina species
Distribution of ACP orthologs varies widely between species (Figure 1). Glossina species have a disproportionately large number of α2-macroglobulin and heat shock proteins, which are important in immunity in A. gambiae and Glossina species16–29. Our analysis did not detect β-defensin orthologs, which are antimicrobial peptides involved in immune responses30, in the G. pallidipes and G. brevipalpis genomes. We also did not detect Acp29AB, Acp70A, and andropin orthologs in this comprehensive catalogue of Glossina genes. This raises the possibility that these genes were lost by tsetse flies after evolutionary radiation of insects into multiple taxa, and alternative species-specific proteins might compensate for the same roles.

Discussion
We performed comparative genomics analysis to detect the presence of male accessory gland proteins (ACPs) orthologs previously identified in A. gambiae and Drosophila1. The motivation here was to improve knowledge on the biology of Glossina species ACPs given the importance of reproductive molecules in strategic designs of vector control. We identified 21 functional classes of ACP orthologs with secretory signals in five Glossina species genomes. We observed genetic signatures of a high rate of ACP protein divergence, supporting similar findings on male reproduction-related genes in Drosophila30. ACPs exhibit high evolutionary changes, thus displaying between species divergence and within species polymorphism30. We restricted analysis to reference genes present in A. gambiae and D. melanogaster, and future studies integrating datasets from more closely related taxa will be useful to understand evolutionary changes associated with ACPs in more detail. α2-macroglobulins and heat shock proteins are the most abundant ACP orthologs in Glossina species. α2-macroglobulins are important in mosquito and Drosophila immunity26,27, and display signatures of purifying selection, suggesting they are critical for successful reproduction and all deleterious variations are purged. α2-macroglobulin over-representation, and the absence of other ACP orthologs implicated in immunity in the Glossina genomes points to their critical role in ensuring success of tsetse fly reproduction. Heat shock protein silencing in A. gambiae down-regulates up to 50% of male accessory gland proteins, half of which are male reproductive tract specific and encode the homologs of 13 known Drosophila ACPs that include Acp70A1. Interestingly, Glossina species lack orthologs to Acp70A, andropin, Acp26Ab, Acp29AB, and Acp62F, which play critical roles in successful reproduction in Anopheles and Drosophila species1,24,25.

Acp70A or sex peptide stimulates long-term post mating behavior, resulting in non-receptivity to mating and increased oviposition25,32. Andropin is an antimicrobial peptide transferred to the female during copulation, and defends the female reproductive tract against microbes1. Andropin also protects the male ejaculate from Gram-positive and Gram-negative bacterial infections1. Acp26Ab stimulates oviposition in Drosophila melanogaster females34, and together with Acp26Ab protects the male ejaculate from microbial infections, and displacement...
Figure 1. Identification of male accessory gland proteins (ACPs) in five Glossina species. (A) We identified 38 clusters of ACPs, which were classified into 21 functional classes. (B) The identified Glossina species ACPs are orthologous to well-characterized ACP genes in Anopheles gambiae and Drosophila melanogaster genomes. (C) The magnitude and direction of selection pressure on the ACP sequences was tested based on the ratio ($\omega = d_N/d_S$) of the average number of non-synonymous substitutions per non-synonymous site ($d_N$) to the average number of synonymous substitutions per synonymous site ($d_S$). A ratio greater than 1 indicates positive selection, and a ratio less than 1 indicates purifying selection. (D) The relative abundance of ACPs identified per ortholog cluster. We identified 36, 27, 31, 29 and 33 ACP orthologs with secretory signals in G. austeni, G. brevipalpis, G. fuscipes, G. pallidipes and Glossina morsitans genomes, respectively. (E) Schematic presentation showing the relative abundance of each ACP functional class in Glossina species, An. gambiae, and Drosophila melanogaster. $\alpha2$-macroglobulins [A] and, heat shock proteins [Q] are the most abundant ACPs in Glossina species.

<table>
<thead>
<tr>
<th>Group</th>
<th>Functional class</th>
<th>Gene name</th>
<th>$d_S$</th>
<th>$d_N$</th>
<th>$d_{N/S}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>$\alpha2$-Macroglobulin</td>
<td>AGAP038364</td>
<td>0.496</td>
<td>0.487</td>
<td>0.940</td>
</tr>
<tr>
<td>Group 2</td>
<td>Cell adhesion</td>
<td>AGAP048438</td>
<td>0.114</td>
<td>0.203</td>
<td>1.636</td>
</tr>
<tr>
<td>Group 3</td>
<td>Carboxyditerase</td>
<td>AGAP053770</td>
<td>0.243</td>
<td>0.487</td>
<td>2.002</td>
</tr>
<tr>
<td>Group 4</td>
<td>Chaperone</td>
<td>AGAP014224</td>
<td>0.998</td>
<td>0.271</td>
<td>2.767</td>
</tr>
<tr>
<td>Group 5</td>
<td>Cytochrome</td>
<td>AGAP042121</td>
<td>0.043</td>
<td>0.952</td>
<td>2.144</td>
</tr>
<tr>
<td>Group 6</td>
<td>Fat body</td>
<td>AGAP086381</td>
<td>0.180</td>
<td>0.218</td>
<td>1.219</td>
</tr>
<tr>
<td>Group 7</td>
<td>Protease inhibitors</td>
<td>AGAP052466</td>
<td>0.172</td>
<td>0.349</td>
<td>2.033</td>
</tr>
<tr>
<td>Group 8</td>
<td>Isomerase</td>
<td>AGAP068222</td>
<td>0.075</td>
<td>0.106</td>
<td>1.426</td>
</tr>
<tr>
<td>Group 9</td>
<td>CAP</td>
<td>AGAP038881</td>
<td>0.104</td>
<td>0.079</td>
<td>0.755</td>
</tr>
<tr>
<td>Group 10</td>
<td>Redox</td>
<td>CGA170</td>
<td>0.335</td>
<td>0.240</td>
<td>1.309</td>
</tr>
<tr>
<td>Group 11</td>
<td>Transport protein</td>
<td>AGAP09966</td>
<td>0.054</td>
<td>0.166</td>
<td>2.978</td>
</tr>
<tr>
<td>Group 12</td>
<td>Ribonucleases</td>
<td>AGAP099662</td>
<td>0.116</td>
<td>0.234</td>
<td>1.721</td>
</tr>
<tr>
<td>Group 13</td>
<td>Lipase</td>
<td>CG1769</td>
<td>0.111</td>
<td>0.287</td>
<td>2.580</td>
</tr>
<tr>
<td>Group 14</td>
<td>Culecin</td>
<td>AGAP050352</td>
<td>0.066</td>
<td>0.130</td>
<td>1.979</td>
</tr>
<tr>
<td>Group 15</td>
<td>Beta-Defensin</td>
<td>AGAP070199</td>
<td>0.121</td>
<td>0.381</td>
<td>3.137</td>
</tr>
<tr>
<td>Group 16</td>
<td>Accessory gland protein</td>
<td>CG174770</td>
<td>0.064</td>
<td>0.118</td>
<td>1.865</td>
</tr>
<tr>
<td>Group 17</td>
<td>Heat shock protein</td>
<td>AGAP048132</td>
<td>0.099</td>
<td>0.293</td>
<td>2.939</td>
</tr>
<tr>
<td>Group 18</td>
<td>Sex peptide</td>
<td>AGAP0270A</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Group 19</td>
<td>Antibiotic</td>
<td>AGAP016897</td>
<td>0.105</td>
<td>0.036</td>
<td>1.032</td>
</tr>
<tr>
<td>Group 20</td>
<td>Cysteine rich venom protein</td>
<td>AGAP05683</td>
<td>0.171</td>
<td>0.340</td>
<td>1.993</td>
</tr>
<tr>
<td>Group 21</td>
<td>Embryonic pattern formation</td>
<td>AGAP068222</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Orthologous ACPs identified: 36, 27, 31, 29, 33, 34, 32, 152
by a second ejaculate\textsuperscript{1}. Drosophila Acp29AB and Acp62F up-regulate genes for egg production and muscle development, although Acp29AB or Acp62F null males do not show a reproduction impairment phenotype\textsuperscript{34}. Acp62F also protects sperm in the female reproductive tract from protease attack\textsuperscript{35}.

A limitation in our study was the absence of transcriptome data to measure ACP differential gene expression. Analysis of transcriptome data available to others using a different approach (reciprocal BLASTp hits) identified only 4 non-annotated ACPs (GMOY000024, GMOY007757, GMOY009744, GMOY012189) orthologous to D. melanogaster not detected in our analysis; and no orthologs to A. gambiae ACPs\textsuperscript{37}. Future studies focused on transcriptome datasets of male accessory glands and testis may identify rapidly evolving Glossina-specific ACPs, which perhaps do not bear a well-known protein domains, and could escape\textit{ in silico} investigation dependent on a comparative genomic approach alone.

Our analysis detected orthologous Glossina, Anopheles and Drosophila ACPs belonging to the same functional classes, suggesting a conserved role for these proteins across all three genera. However, some ACPs may represent lineage-specific ACPs that may have evolved to perform species-specific reproductive functions. Our findings support evolutionary adaptation to different reproductive styles. Tsetse fly females produce a single fully developed larva at a time that pupates within minutes of exiting their uterus, and may have lost non-essential ACP genes after adaptation.

**Conclusions**

Our findings provide genetic evidence of male accessory gland proteins in five recently sequenced Glossina genomes. It provides new avenues for molecular studies that evaluate potential field control strategies of these important vectors of human and animal disease.

**Data availability**

Anopheles gambiae and Drosophila melanogaster ACPs were obtained from: DOI, 10.1073/pnas.0703904104\textsuperscript{4}.

Glossina species genome sequences were obtained from VectorBase.

---

**Competing interests**

No competing interests were disclosed.

**Grant information**

The work was supported by the Wellcome Trust [087540], a pump-priming grant to BWK from the Training Health Researchers into Vocational Excellence in East Africa (THRiVE); UK’s Department for International Development (DFID); Swedish International Development Cooperation Agency (Sida); the Swiss Agency for Development and Cooperation (SDC); and the Kenyan Government.

\textit{The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.}

**Acknowledgments**

We would like to thank Abraham Parwos, Anne Owiti, Edwin Rono, and Dr. George Obiero, and support from the University of Nairobi’s Centre for Biotechnology and Bioinformatics, and ICIPE.

This work is published with permission from the University of Nairobi and ICIPE.

---

**Supplementary material**

**Supplementary Table 1:** (A) Anopheles gambiae (n=57) and Drosophila melanogaster (n=173) ACP orthologs. (B) ACP ortholog clusters identified in Glossina, A. gambiae, and D. melanogaster species prior to signal-peptide filtering.

Click here to access the data.

**Supplementary File 1:** Amino acid sequences of the ACP ortholog clusters.

Click here to access the data.

**Supplementary Figure 1:** Reconstructed phylogenies of the ACP orthologs within each cluster.

Click here to access the data.
References


Open Peer Review

Current Referee Status: ✔ ✔

Version 2

Referee Report 06 December 2017
doi:10.21956/wellcomeopenres.14370.r28295

Kristipati Ravi Ram
Embryotoxicology Laboratory, Environmental Toxicology Group, Council of Scientific & Industrial Research (CSIR)- Indian Institute of Toxicology Research, Lucknow, Uttar Pradesh, India

Authors have adequately addressed the concerns raised on the previous version.

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Referee Report 23 October 2017
doi:10.21956/wellcomeopenres.13476.r26501

Kristipati Ravi Ram
Embryotoxicology Laboratory, Environmental Toxicology Group, Council of Scientific & Industrial Research (CSIR)- Indian Institute of Toxicology Research, Lucknow, Uttar Pradesh, India

The paper by Abry et al., is aimed at identification of accessory gland proteins (Acps) in five Glossina species through bioinformatic approaches. Given the importance of reproductive molecules in strategic designs of vector control, the reported data are of significant value. However, the study is quite preliminary, and hence it feels that authors have gone overboard with the interpretation of these preliminary results. Primarily, authors are required to be clear about the rationale of their study given that the testicular/MAG contributions towards spermatophore formation through transcriptomic and proteomic approaches in G. morsitans has already been reported earlier (Scolari et al. 2016, Scientific Reports). The inclusion of data on reciprocal hits would increase the reliability of predicted/putative Acps. Further, RT-PCR based analysis of the enrichment/expression of at least a couple of identified putative Acp orthologs in male accessory glands (MAGs) from any of the Glossina species would add excellent value to the data and support to the conclusions drawn. Alternatively, authors may compare their datasets with RNAseq/proteomic data of G. morsitans (Scolari et al., 2016) to assess if the identified genes are enriched/expressed in MAGs or form the part of male contributions to the spermatophore.
Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

**Competing Interests:** No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 20 Nov 2017

Benard Kulohoma,

We thank the reviewers for the comments. Our study is a preliminary *in silico* exploration of the MAG repertoire of *Glossina* species, and provides initial findings about genome annotations that identifies ACPs orthologs. Our aim was to establish whether there are male accessory gland (ACPs) orthologs across the Glossina species genomes, by comparing them to well characterised ACP orthologs from *Drosophila melanogaster* and *Anopheles gambiae*, given the importance of reproductive molecules in strategic designs of vector control.

- We have revised the manuscript to address the over-emphasis that ACPS have a play crucial role in the tsetse fly life reproductive cycle, and thus vector control. Lines 45, 73, 163, and 164.

- We have revised the manuscript to show that this is a preliminary study using the initial release of the Glossina genomes, which are publicly available, and future studies using RNASeq/transcriptome datasets of the male accessory glands and testis may identify rapidly evolving Glossina-specific ACPs, which perhaps do not bear a well-known protein domains, and could escape *in silico* investigation dependent on a comparative genomic approach alone. Lines 194 - 201.

- We have amended the manuscript to reflect that we identified possible signatures of selection, since a wider analysis including more closely related taxa was not performed. Lines 137, 140-143, and 165-166.
The paper by Abry et al. presents a comparative genomics analysis to identify male accessory gland proteins in *Glossina* species. The paper is an "*in silico*" exploration of the MAG repertoire of *Glossina*, and has the merit to open the road to further investigation on these important reproductive proteins and their functions. However I think that the authors should give the right relevance to their results, which are, in a sense, quite preliminary, but presented, in my opinion, with too much emphasis.

First of all the authors have to clarify the aim of the work, especially in the context of vector control. Although it is true that a deep understanding of ACPS in *Glossina* can be useful for limit fertility of *Glossina* species and thus control their spread (these proteins have important roles in female post-mating behaviour, such as single mating, storing and sustaining viability of sperms etc.), I think that ACPS do not necessarily have a realistic connection with the peculiar post-mating characteristic of the reproductive cycle of *Glossina* (e.g. absence of eggs deposition). Since this point is recurrently discussed and emphasized in the paper, I suggest the author either to better explain how the ACPS can have a role on the origin of this particular life cycle, or to eliminate this point which, otherwise, remains quite speculative.

It is important that authors stress the importance that their work should be completed in the future by RNAseq of male accessory glands and testis of *Glossina*. This because some rapidly evolving ACPS could escape *in silico* investigation, as they could be *Glossina*-specific and not corresponding to any ortholog, or simply because they do not bear a well-known protein domain. Species-specific proteins can not be obtained by a comparative genomic approach alone, so that RNA analysis is required to reveal (probably) the most important, peculiar ACPS in this important vector.

Finally, I would be very cautious in assigning positive selection to genes with dN/dS>1, since a wider analysis including more closely related taxa (and data on their intra-specific variability) should be performed. However, based on the dN/dS estimation here provided, I think that the authors should just argue about "possible signatures of positive selection" on these genes.

Is the work clearly and accurately presented and does it cite the current literature?  
Partly

Is the study design appropriate and is the work technically sound?  
Yes

Are sufficient details of methods and analysis provided to allow replication by others?  
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Partly

*Competing Interests:* No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

---

**Author Response 20 Nov 2017**

*Benard Kulohana,*

We thank the reviewers for the comments. Our study is a preliminary *in silico* exploration of the MAG repertoire of *Glossina* species, and provides initial findings about genome annotations that identifies ACPs orthologs. Our aim was to establish whether there are male accessory gland (ACPs) orthologs across the *Glossina* species genomes, by comparing them to well characterised ACP orthologs from *Drosophila melanogaster* and *Anopheles gambiae,* given the importance of reproductive molecules in strategic designs of vector control.

- We have revised the manuscript to address the over-emphasis that ACPS have a play crucial role in the tsetse fly life reproductive cycle, and thus vector control. Lines 45, 73, 163, and 164.

- We have revised the manuscript to show that this is a preliminary study using the initial release of the *Glossina* genomes, which are publicly available, and future studies using RNASeq/transcriptome datasets of the male accessory glands and testis may identify rapidly evolving *Glossina*-specific ACPs, which perhaps do not bear a well-known protein domains, and could escape *in silico* investigation dependent on a comparative genomic approach alone. Lines 194 - 201.

- We have amended the manuscript to reflect that we identified possible signatures of selection, since a wider analysis including more closely related taxa was not performed. Lines 137, 140-143, and 165-166.

*Competing Interests:* None declared