Understanding Tsetse Fly (*Glossina morsitans)* Behavior through its *Genome*

Kristin Talia Marbun[1], Satya Nugroho[2], Juandy Jo[1,3]

[1] *Department of Biology, Faculty of Science & Technology, Universitas Pelita Harapan, Tangerang, Banten*
[2] *Research Center for Biotechnology, National Research and Innovation Agency (BRIN), Cibinong, West Java*
[3] *Mochtar Riady Institute for Nanotechnology, Tangerang, Banten, Indonesia.*

Citation : Marbun Kristin Talia, Nugroho Satya, Jo Juandy. Understanding Tsetse Fly (Glossina morsitans) Behavior through its Genome. Medicinus. 2022 October; 10(3): 122-129.
Keywords: *Glossina morsitans*; African trypanosomiasis; Whole-genome sequence; Symbionts; Vector control
Correspondance : Juandy Jo
E-mail : juandy.jo@uph.edu
Online First : October 2022

Abstract

Glossina morsitans (G. morsitans), commonly known as tsetse fly, have caused public health concerns throughout the years. *G. morsitans* is the vector for *Trypanosoma brucei* (*T. brucei*), the parasite responsible for causing the deadly African sleeping disease (African trypanosomiasis). Researchers have searched for ways to contain this disease, but to little avail. Fortunately, new advances in sequencing methods have given researchers a new opportunity to win the war against the disease. The whole-genome sequence of *G. morsitans* provides essential data regarding involved genes that transmits *T. brucei* to humans. Information about those unique genes facilitates researchers to create new methods to prevent *G. morsitans* from becoming the vector of *T. brucei*, enabling the containment of this disease. With this, we review the unique genes in the *G. morsitans* genome, such as those that contribute to blood-feeding ability, establish a relationship with symbionts, and *G. morsitans* unique sensory genes, with an expectation that it would enhance our knowledge of *G. morsitans* as the vector for parasites causing African trypanosomiasis.

Introduction

The tsetse fly (*G. morsitans*) is the vector of *T. brucei*, a parasite that causes human African trypanosomiasis (HAT). HAT can infect humans and animals, impacting the human population, livestock and food supply. As *G. morsitans* is endemic to sub-Saharan Africa, around 70 million people in the region are at risk of contracting HAT.[1] Although treatments are available, they are relatively expensive and are not affordable for most people. Furthermore, there have been several reports of resistance to the available drugs, which calls out for a new drug development. At the moment, vaccine discovery is still an ongoing project.[2]

A way to minimize the disease's spread is by controlling the vector, *G. morsitans*. Whole-genome sequencing (WGS) of *G. morsitans* genome provides new ways to analyze its genetics and opens the door to many possibilities. Attardo and his team were ones of the pioneers who used WGS technology to sequence *G. morsitans* genome. Attardo and his team WGS project resulted in 366 megabases of *G. morsitans* genome, in which there are 12,308 predicted protein-encoding genes. The predicted protein accounts for properties such as a family of lactation-specific proteins, reduced complements of host-pathogen recognition proteins, and genes that are responsible for blood-feeding.[1] Whole-genome sequencing of *G. morsitans* enables us to analyze the unique genes that distinguish this species with other flies, and how its role as *T. brucei* vector. The most notable genes are those related with its ability as a blood feeder, the

ability to enable a symbiotic relationship with pathogens, and unique sensory genes. That genetic information subsequently could be used, controlling *G. morsitans* population to tackle the HAT.

Trypanosoma brucei, *Glossina morsitans* and the present situation

HAT has caused unrest among sub-Saharan Africans since its discovery in 1901 due to the mortality it caused. Once transmitted by the vector *G. morsitans*, this parasite will enter the bloodstream of humans. As depicted in **Figure 1**, after two- or three weeks post-infection, the disease would progress into two clinical stages, i.e., the hemolymphatic and meningoencephalitis stages.[3] The first stage, the hemolymphatic stage, begins after two or three weeks and can be diagnosed by the occurrence of fever episodes, liver problems, and a painful lymph node's swelling. In this first stage, the parasite has successfully invaded the blood and the lymph fluid. However, this stage is often undiagnosed and untreated since people think of it as a simple fever. The second stage is the meningoencephalitis stage. The parasites had crossed the body's blood-brain barrier and invaded the central nervous system (CNS) and the cerebrospinal fluid. This second stage appears slowly, and ranges from months to years.

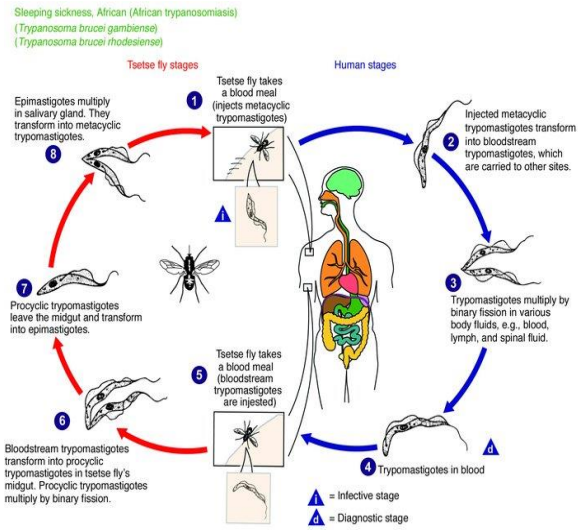

Figure 1. Life cycle of *T. brucei* and stages of infection. Human stages depict the presence of *Trypanosoma brucei* in humans, while *G. morsitans* stages depict the presence of the parasite within the fly (Centers for Disease Control and Prevention, 2022).

When *T. brucei* has successfully invaded the CNS, host immune responses will cause inflammation in the CNS. The inflammation would cause damage to the brain. In the terminal stage, patients will have dementia with incoherence and seizures due to the demyelination of the cells.[5] The most common cause of death is due to heart failure, brain inflammation (encephalitis), as well as weight and muscle loss (cachexia).[6] The term "sleeping disease" is derived from the fact that the parasite infects the hypothalamus that controls the circadian rhythm. Thus, an infection by the parasite will disturb the cycle.

Until now, HAT has been managed by using active or passive detection and treatment programs.[8] Healthcare workers conduct active detection to determine whether an individual contracts HAT by observing their symptoms and running relevant tests. Passive detection is done independently, where individuals decide whether or not they have contracted the disease by comparing their symptoms to the guidelines given by health institutions.[7] However, these methods are inefficient since the current diagnostic procedure used in active detection is costly. Moreover, the procedure used in passive detection is flawed due to the non-specific symptoms presented in the early stages of the disease.[8]

Since a specific vaccine is still unavailable, a fast and accurate diagnosis will be necessary, as poor surveillance may result in a re-emerging epidemic.[9] Scientists are currently developing new rapid serodiagnostic tests for HAT, with an expectation that those tests can be used in a particular region with a high risk of infection and thus, novel rapid serodiagnostic tests will give more countermeasures against HATin that area.[10] In parallel, controlling the vector would be crucial as *G. morsitans* is the sole vector of *T. brucei*. An ability to control *G. morsitans*

would significantly contribute to the eradication of HAT. Therefore, the following parts of this review are dedicated to better understand the vector.

Blood feeding and nutrition uptake by *G. morsitans*

G. morsitans are unique since they consume blood as their primary nutrient, despite coming from the *Diptera* order. Three genes, i.e., *tsal*, putative adenosine deaminases, and insect growth factors (*ADGF*)-related and *aquaporin* genes, play a crucial role in the blood-feeding characteristic of *G. morsitans*. When blood from another organism is transferred to another, the immune system will detect it as foreign, resulting in activation of the complement cascade and the formation of a blood clot.[11] However, in *G. morsitans* case, blood is their meal, and they need to be able to access non-coagulated blood. Thus, a non-coagulation protein is required. This is where the *tsal* gene plays the role (*tsal* stands for "tsetse salivary"), in which there are three distinct contributing genes: *tsal1* (*GMOY012071*), which encodes for tsal1 protein, and *tsal2a* (*GMOY012361*), which encodes for tsal2a protein, and *tsal2b* (*GMOY012360*) which encodes for tsal2b protein.[12] The functions of these genes were deducted through multiple sequence alignment by Caljon et al[10].

The *tsal1*, *tsal2a* and *tsal2b* genes from *G.morsitans* were compared to homologous genes annotated as putative salivary gland nucleases in *Culex quinquefasciatus* (southern house mosquito), *Phlebotomus argentipes* (sand flies) and *Lutzomyia longipalpis* (sand flies), as well as the *Marsupenaeus japonicus* shrimp hepatopancreatic nuclease and the *Serratia marcescens* nuclease. The analysis can be seen in **Figure 2**. The *tsal* gene family encodes high-affinity nucleic acid-binding proteins without strong endonuclease activity.[13] One of the compounds in *G. morsitans* saliva is the product of the tsetse thrombin inhibitor (*TTI*) gene, which plays a huge role in anticoagulation and antithrombotic activity during blood feeding. Not only serves as an anticoagulant agent, but this inhibitor also inhibits thrombin-induced platelet aggregation.[14] Taken together, genes encoding for *G. morsitans* saliva are really essential. Without those genes, *G. morsitans* will not be able to feed blood and survive.

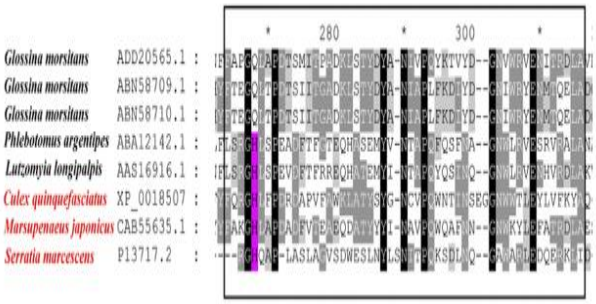

Figure 2. Multiple sequence analysis of *tsal* genes with homologous genes (Franco et al., 2018). Three tsal genes of *Glossina morsitans* were compared to homologous genes from *P. argentipes*, *L. longipalpis*, *C. quinquefasciatus* and *M. japonicus*. The species names in red have a confirmed nuclease activity in their genes. The black box represents the putative nuclease active site region. The black highlight indicates amino acid similarity in all eight species. The purple highlight represents histidine residues that supposedly are important for nuclease catalytic activity.

Other genes that enable blood-feeding are the uncharacterized adenosine deaminases insect growth factors (ADGFs)-related gene and the *aquaporin* gene. The still-yet uncharacterized abundant salivary gland's protein reduces inflammation caused by adenosine and inosine-induced mast cell activation. A mast cell is activated when an antigen is detected by IgE bound to FcεRI on the mast cell surface.[15] Blood contains a variety of antigens. There is a high possibility that one of those antigens will react with IgE and cause inflammation. Since inflammation is not desirable during feeding, the putative ADGF-related protein will suppress mast cell activation. Another essential gene during the blood-feeding by *G. morsitans* is the *aquaporin* gene. *G. morsitans* has ten *aquaporin* genes, in contrast to six and eight respective genes in mosquitoes and *Drosophila* sp., respectively. The presence of *aquaporin* genes ensures water homeostasis following the blood meal.[1] This mechanism is vital

since the ingested blood is equivalent to the *G. morsitans* weight. Hence, excess water needs to be excreted quickly to reduce body density.[1]

The relationship between *G. morsitans* and its symbionts

One of the interesting features of *G. morsitans* biology is how symbionts (i.e., an organism living in symbiosis with another organism, such as gut microbiota and humans) reside and evade *G. morsitans* immune response. Some factors enable this phenomenon, comprising external factors, genetic factors, and the capability of pathogens themselves.[16] Aside from hosting the *T. brucei* parasite, *G. morsitans* also hosts other symbionts, such as *Wigglesworthia glossinidia*, which lives intracellularly in the midgut and extracellularly in lumen milk gland, *Wolbachia* spp. which resides in gonadal tissues, *Sodalis* spp. which resides in digestive and reproductive organs, as well as the salivary gland hypertrophy virus which resides in the salivary glands.[17] *G. morsitans* competence to act as a host for these symbionts depends on multiple factors such as age, sex, and overall health condition. A study was conducted to analyze *G. morsitans* capability to eliminate pathogen invasion to find whether tsetse has an inherited defective immune system or other factors at play.[18] The experiment involved feeding *Trypanosoma*-infected blood to *G. morsitans*. In an optimal condition, less than 50% of the lab-grown tsetse flies that had fed *Trypanosoma*-infected blood became infected with *T. brucei*. This finding suggests that normal tsetse flies with adequate nutrition and *environment* can resist the parasite's infection.

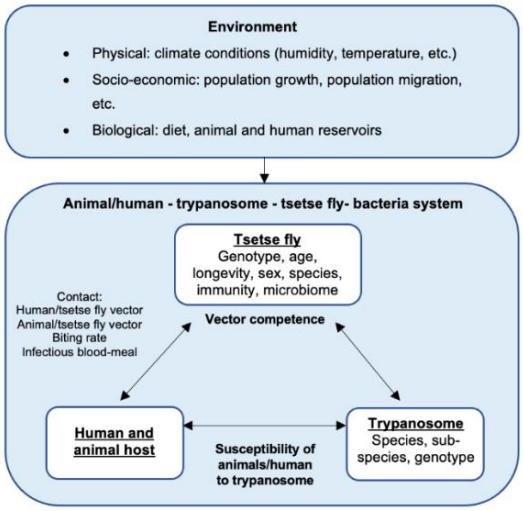

Figure 3. Factors that affect *T. brucei*, *G. morsitans*, and human infections. The interaction between the parasite, the vector and human or animal host dictate the incidence rate of African trypanosomiasis. In addition, the environmental changes will influence this three-party interaction as well.

However, the environmental condition in sub-Saharan Africa differs from the one in the lab. Global climate changes, such as changing temperature, rainfall patterns, urbanization, deforestation, grassland degradation, and overgrazing, will affect the infection rate.[16] The summary of the factors that affect *Trypanosoma* infections can be seen in **Figure 3**. The geographical distribution of *Trypanosoma* reservoirs, nutritional behavior, and development of *Trypanosoma* will be affected. Thus, *G. morsitans* and humans will also be affected. It could be that deforestation and urbanization cause *G. morsitans* to lose their original habitat and therefore move into closer contact with human populations. Furthermore, the constant climate change activity may weaken the *G. morsitans* immune system and impede the infection containment.

Glossina	*Drosophila*	Function
pgrp-la	*pgrp-la*	Toll pathway activation
pgrp-lb	*pgrp-lb*	PGN scavenging
pgrp-lc	*pgrp-lc*	Imd pathway activation
pgrp-ld	*pgrp-ld*	Unknown
	pgrp-le	Imd pathway activation
	pgrp-lf	Negative regulation of the Imd pathway
pgrp-sa	*pgrp-sa*	Toll pathway activation
pgrp-sb	*pgrp-sb1*	Bactericidal
	pgrp-sb2	Unknown
	pgrp-sc1a	PGN scavenging
	pgrp-sc1b	PGN scavenging
	pgrp-sc2	PGN scavenging
	pgrp-sd	Toll pathway activation

Figure 4. List of peptidoglycan recognition proteins (*pgrp*) genes in *Glossina* sp. and *Drosophila* sp. and their respective functions. While *Drosophila* sp. has 13 *pgrp* genes, Glossina sp. only has 6 *pgrp* genes.

G. morsitans genetics also influences the immune response towards its symbionts. A reduced number of genes responsible for microbial detection will simultaneously lower the immune response.[19] Genes responsible for microbial detection are crucial since they signal the body that there is a foreign invader that needs to be neutralized. Pathogen detection is a multistep process that requires contact between host (*G. morsitans*) pattern recognition receptors (PRRs) and pathogen-associated molecular patterns (PAMPs). Subsequently, a symbiont must be able to withstand the activity of various proteins, such as peptidoglycan (PGN) recognition proteins (PGRPs), antimicrobial effector peptides (AMPs) produced by immune deficiency (IMD) pathway, midgut lectins, and other proteins.[1] **Figure 4** compares PGRP components between *Drosophila* and *Glossina* spp. *Drosophila* sp. has 13 PGRPs that play a role in peptidoglycan (PGN) recognition. While, *G. morsitans* only has six identified *pgrp* genes, four in the long subfamily (*pgrp-la*, *-lb*, *-lc*, and *-ld*) and two in the short subfamily (*pgrp-sa* and *-sb*).[1,15] *Glossina morsitans* lack several PGN receptors (*pgrp-le*, *-lf*, *-sc*, and *-sd*) found in *Drosophila* sp. PGN receptors are important in microbial detection since PGN is the essential component of the cell wall of most bacteria. The absence of PGN-detecting genes causes insensitivity to pathogenic invasion. Hence, an attenuated immune response will be exhibited.[1] Therefore, this is the most probable cause of why symbionts can reside within the *G. morsitans*.

In the case of infection by *T. brucei*, this parasite has evolved. It uses two mechanisms to escape from the *G. morsitans* immune system and simultaneously uses *G. morsitans* attenuated immune system to its advantage.[9] First, *T. brucei* can evade the immune response by overcoming the complement system by recycling its variant surface glycoprotein (VSG). The classical complement pathway, activated by antibodies, can be overcome through a rapid VSG-recycling system that removes IgGs from its surface.[20]

Secondly, *T. brucei* can suppress T-cell proliferation by eliciting suppressive macrophage by activating a parasite membrane protein, Trypanosoma Suppression Immunomodulating Factor (TSIF). As a result, the induced macrophage will produce nitric oxide and prostaglandins responsible for impairing T-cell proliferation during the early infection stage. These macrophages also have a reduced ability to activate specific T cells since macrophages' ability to present peptides via their MHC class II is diminished. Therefore, parasites are free to thrive.[21]

Sensory genes of *G.morsitans*

Population growth of *G. morsitans* needs to be controlled and reduced, if possible. One way is by using traps. Therefore, it is important to determine the traps that attract *G. morsitans*. Scientists currently use one unique feature of *G. morsitans* to lure them more easily to the traps, which is through their sensitivity to color.[22] *G. morsitans* visual systems are similar to other calyptrate *Diptera* (e.g., house flies and blowflies), in which each ommatidium (a cluster of photoreceptors) consists of eight photoreceptors (R1 to R8). The *Rh5* gene encodes those photoreceptors.[1,23]

Photoreceptors R1-R6 are similar across each ommatidium, as they are sensitive to ultraviolet (UV) and blue wavelengths. On the other hand, photoreceptors R7 and R8 are located at the center of each ommatidium. The R7 and R8 have two forms: 'y' is sensitive to green-yellow wavelengths and 'p' is sensitive to blue wavelength.[19] While the R8y form is most sensitive to yellow-green wavelengths, the R7y form is most sensitive to UV wavelength since 'y' has an accessory sensitizing pigment sensitive to UV. The R8p form is most susceptible to blue wavelength, and R7p is most sensitive to lower UV lengths. Since tsetse flies are sensitive to those two types of color (green-yellow or blue-UV), a study by Santer determined which color was preferred by *G. morsitans*.[20] It was discovered that the blue wavelength was preferred over the green-yellow wavelengths. Therefore, blue-colored traps should be used since they will create a higher chance of catching *G. morsitans* in the environment.

Conclusion

WGS of *G. morsitans* are able to give us a better understanding of its role as the vector of *T. brucei*. Five essential genes drive *G. morsitans* to become a prime vector for *T. brucei*. The first three are the *tsal*, *ADGF*-related, and *aquaporin* genes. These unique genes are crucial for the growth and development of *G. morsitans* as a blood-feeder and may be good candidates for vaccine targets. The next gene is the reduced *pgrp* genes which enable parasites and bacteria to evade *G. morsitans* immune response. The final gene is the *Rh5* gene that causes *G. morsitans* to be sensitive to color. This finding is used to develop a control strategy to reduce *G. morsitans* population by setting up traps with blue color. In conclusion, the whole-genome analysis of *G. morsitans* facilitates a better understanding of *G. morsitans* behavior. This information could help to control and even prevent the spread of human African trypanosomiasis

References

1. Attardo G, Abila P, Auma J, Baumann A, Benoit J, Brelsfoard C et al. Genome Sequence of the Tsetse Fly (Glossina morsitans): Vector of African Trypanosomiasis. Science. 2014; 344 (6182): 380-386. https://doi.org/10.1126/science.1249656

2. Fairlamb A, Horn D. Melarsoprol Resistance in African Trypanosomiasis. Trends in Parasitology. 2018; 34(6): 481-492. https://doi.org/10.1016/j.pt.2018.04.002

3. Bouteille B, Buguet A. The detection and treatment of human African trypanosomiasis. Research and Reports in Tropical Medicine. 2012;3: 35-45. https://doi.org/10.2147%2FRRTM.S24751

4. Centers for Disease Control and Prevention. Trypanosomiasis, African. Available on https://www.cdc.gov/dpdx/trypanosomiasisafrican/index.html cited Jan 2022.

5. Kennedy P. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). The Lancet Neurology. 2013;12(2):186-194. https://doi.org/10.1016/s1474-4422(12)70296-x

6. Ponte-Sucre A. An Overview of Trypanosoma brucei Infections: An Intense Host–Parasite Interaction. Frontiers in Microbiology. 2016;7. https://doi.org/10.3389/fmicb.2016.02126

7. Checchi F, Funk S, Chandramohan D, Chappuis F, Haydon D. The impact of passive case detection on the transmission dynamics of gambiense Human African Trypanosomiasis. PLOS Neglected Tropical Diseases. 2018;12(4): e0006276. https://doi.org/10.1371/journal.pntd.0006276

8. Snijders R, Fukinsia A, Claeys Y, Hasker E, Mpanya A, Miaka E et al. Costs and Outcomes of Integrated Human African Trypanosomiasis Surveillance System Using Rapid Diagnostic Tests, Democratic Republic of the Congo. Emerging Infectious Diseases. 2021;27(8):2144-2153. https://doi.org/10.3201%2Feid2708.202399

2. Franco J, Cecchi G, Priotto G, Paone M, Diarra A, Grout L et al. Monitoring the elimination of human African trypanosomiasis: Update to 2016. PLOS Neglected Tropical Diseases. 2020;14(5):e0008261. https://doi.org/10.1371%2Fjournal.pntd.0008261

3. Jamonneau V, Camara O, Ilboudo H, Peylhard M, Koffi M, Sakande H et al. Accuracy of Individual Rapid Tests for Serodiagnosis of Gambiense Sleeping Sickness in West Africa. PLOS Neglected Tropical Diseases. 2015; 9(2): e0003480. https://doi.org/10.1371/journal.pntd.0003480

4. Ooi C, Haines L, Southern D, Lehane M, Acosta-Serrano A. Tsetse GmmSRPN10 Has Anti-complement Activity and Is Important for Successful Establishment of Trypanosome Infections in the Fly Midgut. PLoS Neglected Tropical Diseases. 2015;9(1):e3448. https://doi.org/10.1371%2Fjournal.pntd.0003448

5. Matetovici I, Caljon G, Van Den Abbeele J. Tsetse fly tolerance to T. brucei infection: transcriptome analysis of trypanosome-associated changes in the tsetse fly salivary gland. BMC Genomics. 2016;17(1). https://doi.org/10.1186%2Fs12864-016-3283-0

6. Caljon G, Ridder K, Stijlemans B, Coosemans M, Magez S, De Baetselier P et al. Tsetse Salivary Gland Proteins 1 and 2 Are High Affinity Nucleic Acid Binding Proteins with Residual Nuclease Activity. PLoS ONE. 2012;7(10):e47233. https://doi.org/10.1371%2Fjournal.pone.0047233

7. Calisto B, Ripoll-Rozada J, Dowman L, Franck C, Agten S, Parker B et al. Sulfotyrosine-Mediated Recognition of Human Thrombin by a Tsetse Fly Anticoagulant Mimics Physiological Substrates. Cell Chemical Biology. 2021; 28(1): 26-33.e8. https://doi.org/10.1016/j.chembiol.2020.10.002

8. Krystel-Whittemore M, Dileepan K, Wood J. Mast Cell: A Multi-Functional Master Cell. Frontiers in Immunology. 2016;6(6): 620. https://doi.org/10.3389/fimmu.2015.00620

9. Nnko H, Ngonyoka A, Salekwa L, Estes A, Hudson P, Gwakisa P et al. Seasonal variation of tsetse fly species abundance and prevalence of trypanosomes in the Maasai Steppe, Tanzania. Journal of Vector Ecology. 2017; 42(1): 24-33. https://doi.org/10.1111/jvec.12236

10. Wang J, Weiss B, Aksoy S. Tsetse fly microbiota: form and function. Frontiers in Cellular and Infection Microbiology. 2013;29(3): 69. https://doi.org/10.3389/fcimb.2013.00069

11. Geiger A, Ponton F, Simo G. Adult blood-feeding tsetse flies, trypanosomes, microbiota and the fluctuating environment in sub-Saharan Africa. The ISME Journal. 2014; 9(7): 1496-1507. https://doi.org/10.1038%2Fismej.2014.236

12. Myllymäki H, Valanne S, Rämet M. The Drosophila Imd Signaling Pathway. The Journal of Immunology. 2014;192(8):3455-3462. https://doi.org/10.4049/jimmunol.1303309

13. Mugnier M, Stebbins C, Papavasiliou F. Masters of Disguise: Antigenic Variation and the VSG Coat in Trypanosoma brucei. PLOS Pathogens. 2016;12(9): e1005784. https://doi.org/10.1371/journal.ppat.1005784

14. Alfituri O, Quintana J, MacLeod A, Garside P, Benson R, Brewer J et al. To the Skin and Beyond: The Immune Response to African Trypanosomes as They Enter and Exit the Vertebrate Host. Frontiers in Immunology. 2020;11. https://doi.org/10.3389%2Ffimmu.2020.01250

15. Lindh J, Goswami P, Blackburn R, Arnold S, Vale G, Lehane M et al. Optimizing the Colour and Fabric of Targets for the Control of the Tsetse Fly Glossina fuscipes fuscipes. PLoS Neglected Tropical Diseases. 2012; 6(5): e1661. https://doi.org/10.1371%2Fjournal.pntd.0001661

16. Santer R. A Colour Opponent Model That Explains Tsetse Fly Attraction to Visual Baits and Can Be Used to Investigate More Efficacious Bait Materials. PLoS Neglected Tropical Diseases. 2014; 8(12): e3360. https://doi.org/10.1371/journal.pntd.0003360

Author's Statement

The authors declared that all the images and figures in this manuscript is/are author's own work and/or has obtained necessary permission to re-use the content from the authors and publisher of respective materials.

Juandy Jo

(Juandy Jo)