

Published in final edited form as:

Trends Parasitol. 2013 August ; 29(8): 394–406. doi:10.1016/j.pt.2013.06.005.

***Glossina fuscipes* populations provide insights for Human African Trypanosomiasis transmission in Uganda**

Serap Aksoy¹, Adalgisa Caccone^{1,2}, Alison P. Galvani^{1,2}, and Loyce M. Okedi³

¹Yale School of Public Health, Department of Epidemiology of Microbial Diseases, Yale University, 60 College Street, New Haven, CT, USA ²Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem Street, New Haven, CT, USA ³National Livestock Resources Research Institute (NaLIRRI), P.O. Box 96, Tororo, Uganda

Abstract

Uganda has both forms of human African trypanosomiasis (HAT): the chronic *gambiense* disease in the northwest and the acute *rhodesiense* disease in the south. The recent spread of *rhodesiense* into central Uganda has raised concerns given the different control strategies the two diseases require. We present knowledge on the population genetics of the major vector species *Glossina fuscipes fuscipes* in Uganda with a focus on population structure, measures of gene flow between populations, and the occurrence of polyandry. The microbiome composition and diversity is discussed, focusing on their potential role on trypanosome infection outcomes. We discuss the implications of these findings for large-scale tsetse control programs, including suppression or eradication, being undertaken in Uganda and potential future genetic applications.

Keywords

Glossina fuscipes; *Trypanosoma brucei gambiense*; *Trypanosoma brucei rhodesiense*; Human African Trypanosomiasis; sleeping sickness; population genetics; vector control; Uganda

African trypanosomes cause devastating human and animal diseases

Diseases caused by parasitic African trypanosomes impact both human and animal health in Sub-Saharan Africa. Approximately 70 million people (1.55 million km²) are estimated to be at various levels of risk for human African trypanosomiasis (HAT) caused by two species of trypanosomes [1]; *Trypanosoma brucei gambiense* (*Tbg*) causes the chronic form of disease with long symptom-free periods lasting several years, while *Trypanosoma brucei rhodesiense* (*Tbr*) causes the acute form of disease with more than 80% mortality within the first six months if untreated [2]. Over 90% of the HAT cases reported are due to *Tbg* occurring in northwest Uganda, extending into the Central African Republic, Democratic Republic of Congo (DRC), southern Chad, along the Congo river north of Brazzaville, and by the Atlantic coast on both sides of the border between Gabon and Equatorial Guinea.

© 2013 Published by Elsevier Ltd.

Corresponding author: Aksoy, S. (serap.aksoy@yale.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Over 12 million people in eastern and southern Africa, including Uganda, Tanzania, Malawi, Zambia, and Zimbabwe are estimated to be at risk for *Tbr* [1].

Since the discovery of sleeping sickness a century ago, several waves of epidemics have plagued the continent. It is thought that during the post-independence period of the 1960s, control programs within the endemic countries gradually were run down, resulting in a steep rise in incidence during the following 40 years. It has been difficult to estimate the true burden of HAT, as the disease affects the poorest and most neglected populations living in remote and rural settings where the majority of people affected are beyond the reach of health care systems and are not captured in any of the health metrics [3]. In 2008, mortality associated with HAT ranked ninth out of 25 among the human infectious and parasitic diseases in Africa [4]. After intense international interventions, HAT cases in Africa have recently dropped below 10 000 for the first time in 50 years, signaling a possible end to the latest epidemic cycle as a major public health problem [5] Many cases, however, are likely to go undetected in areas such as Central Africa where the disease occurs in rural, hard to reach places [6].

Both forms of the disease are fatal if untreated, although recently several patients with *Tbg* disease have been reported to self-cure [7]. The drugs available for treatment are expensive [8], and drugs used to treat late stages of the disease, such as melarsoprol, can cause mortality in 5-10% of the treated patients [9,10]. Recent studies on the use of nifurtimox-eflornithine combination therapy for second-stage gambiense disease have been highly promising, and it is hoped that it will replace melarsoprol applications [11].

HAT epidemiology in Uganda

Both forms of HAT exist in Uganda

Tbg in the northwest and *Tbr* in the south, restricted to districts close to the shores of Lake Victoria until recently (Figure 1). Between 1898-1920, HAT epidemics along the shores of Lake Victoria resulted in more than 200 000 deaths in Uganda [12]. After the decline of this major outbreak, smaller outbreaks continued in the 1930s-1940s along the lake basin. In 1976 a new major outbreak occurred in the southeastern Iganga District and has since expanded into districts north of Lake Kyoga into central Uganda, including Pallisa in the Bugosa region [4], and then further north to Kaberamiado, Lira, Dokolo, and Apac districts (Figure 1) [13-15].

Since the pathology, diagnosis, and treatment varies between the two forms of disease, an overlap of the two disease belts would complicate HAT control strategy in Uganda. Identification of *Tbg* infections relies on active surveillance of infected humans and on diagnosis based on serological and parasitological examination. The identification of *Tbr* infections relies on microscopic examination of blood and central nervous system (CNS) aspirates for the presence of the parasite. The epidemiology of the two diseases also differs in that there are well-documented animal reservoirs for *Tbr* where the animal hosts are asymptomatic carriers responsible for the maintenance and spread of the parasite. By contrast, it is generally thought that the *Tbg* parasite maintains itself in human populations without a zoonotic reservoir [16], although mathematical models have recently challenged this dogma [17].

A PCR assay that amplified the *Tbr*-specific serum resistance associated (*SRA*) gene reported that 18% of cattle in the HAT focus in Soroti, Uganda, were infected with the human-infective *Tbr* parasites [18]. This was attributed to the large-scale restocking programs that may have resulted in the movement of asymptomatic carriers of human-infective *Tbr* cattle from southeastern Uganda to districts further north, thereby introducing

Tbr sleeping sickness into previously disease-free, although tsetse infested areas, such as Kaberamaido and Lira [19]. This discovery prompted the Ugandan officials to undertake an ambitious public-private partnership: Stamp Out Sleeping Sickness (SOS). The SOS campaign involved block treating cattle with a trypanocidal drug (either diminazene or isometamidium) and spraying animals with synthetic pyrethroid insecticide [20]. A recent study that investigated the presence of parasites in various domestic animals in Uganda, however, has shown that the presence of *Tbr* was highest in pigs (21.7%), followed by cattle (14.5%), dogs (12.4%), sheep (10.8%), and goats (3.2%), indicating that clearing parasite infections in the reservoir may be more challenging than previously anticipated [21].

The tsetse vector in Uganda

Based on morphological and genetic data, *G. fuscipes* is part of the subgenus *palpalis*, which also includes another riverine species from West Africa, *Glossina palpalis* [22]. *G. fuscipes* is further split into three subspecies (*G. f. fuscipes*, *G. f. martini*, and *G. f. quanzensis*), thought to have evolved within the Congo basin following forest contractions and fragmentations in the Pliocene [23]. *G. f. martini* distribution lies to the southeast of the *G. f. fuscipes* (*Gff*) range, around Lake Tanganyika, while the *G. f. quanzensis* range lies to the southwest of *Gff*. Although these three taxa differ slightly in genitalia morphology, both nuclear and mitochondrial DNA (mtDNA) data cluster (see Glossary) them into five rather than three allopatric groups, with the Ethiopian disjunct *Gff* populations as genetically distinct from other *Gff* populations as from the other two subspecies [24].

Throughout Uganda *Gff* is the main disease vector. Except for a disjunct region in Ethiopia/Sudan, Uganda, and western parts of Kenya along the Lake Victoria shore, represent the eastern edge of the *Gff* predicted range, as most of this species distribution lies in the Democratic Republic of Congo and the Central African Republic, extending as far west as Angola (Figure 1) [25]. Of the three subspecies, *Gff* is found in the most humid and forested habitats. In Uganda *Gff* is predicted to occur in high concentration in vegetation thickets along water bodies. The thickets of vegetation offer tsetse seasonal refugia and access to hosts, who are also seeking food and water and relief from heat [24]. The microclimatic conditions in these riverine and lacustrine thickets favor survival of both tsetse adults and pupae, which will die in drying soils at temperatures above ~36°C [26,27] and under low humidity conditions [26,28,29]. Ecological data suggest that *Gff* has a great capacity for dispersal and re-colonization of suitable habitats [30]. Yet, the fact that this species is found in discrete patches of riverine or lacustrine habitats suggests that the populations may be genetically discrete, unless dispersal through unsuitable habitats would homogenize tsetse flies across the spatial landscape.

Vector control programs in Uganda

While control of the infection in the mammalian host and in the large animal reservoirs is challenging, reduction of tsetse populations can be highly effective for disease control. In 2001, the African Union (AU) launched the Pan African Tsetse and Trypanosomiasis Eradication Campaign (PATTEC), a continent-wide initiative that focuses on the progressive elimination of targeted tsetse-infested areas [31]. This tsetse vector control program uses a variety of methods depending on fly species, agro-ecology, and the capacity of local and national tsetse control agencies.

The Phase I of this initiative has been initiated in the Lake Victoria basin. A previous program, Farming in Tsetse Controlled Areas (FITCA), which ended in 2004, had reduced mainland tsetse populations by 75% to 90% exclusively in southeast Uganda into districts up to Soroti and Kaberamaido, but did not include the Lake Victoria islands [32]. In 2009, a PATTEC baseline survey revealed that mainland tsetse populations had rebounded to the

high levels present prior to the FITCA initiative in most districts [33]. The PATTEC plan, unlike FITCA, includes islands and intends to eradicate *Gff* from the Lake Victoria basin progressively from west to east [34]. The basin has been partitioned into four operational blocks [three in Uganda and one in Kenya (Figure 1)] based on the Food and Agriculture Organization (FAO) predicted habitat suitability for *Gff*, natural barriers, major urban areas, international borders, and drainage patterns [34,35]. To support the PATTEC initiative, the government of Uganda and the International Atomic Energy Agency (IAEA) are planning a trial eradication of *Gff* on a remote island in the Kalangala region (IAEA Project UGA5033), building on previous success of tsetse eradication in Zanzibar island using the Sterile Insect Technique (SIT) [36].

Population genetics data can offer insights into facets of ecology, reproduction, dispersal patterns, population dynamics, and genetic diversity with direct relevance for guiding control programs [30,37,38]. Information on the genetic differentiation of *Gff* populations can be used to estimate effective population size (N_e), metapopulation dynamics, and genetic exchange and dispersal rates help assess the suitability of the operational units selected for vector control. In addition, characterization of barriers to gene flow is also important in planning release of sterile or genetically modified vectors, where the extent of success will depend upon the scale at which the released vectors will mate with wild counterparts. Below we describe the current status of knowledge on *Gff* population genetics and dynamics that are fundamental to intervention strategies.

Population genetics data reveal stark divergences among *Gff* populations

A genetic screen of *Gff* from 37 localities across Uganda (Figure 1) based on mtDNA data showed two distinct genetic lineages represented by two distinct haplogroups [39]. Given the extensive genetic differentiation between them, it is likely that they may have diverged more than 100 000 years ago. The two haplogroups are allopatrically distributed to the north and south of Lake Kyoga in central Uganda with populations in these regions carrying mtDNA from either one or the other haplogroup (Figure 2A). Along Lake Kyoga there is a geographically relatively narrow contact zone, compared to the distribution of the two haplogroups, where individuals carrying either the northern or the southern mtDNA co-occur (Figure 1) [39,40].

Microsatellite data showed that *Gff* populations sampled within 1-5 km² are genetically homogeneous and distinct from other populations [40]. Bayesian clustering analyses identified northern, southern, and western clusters of genetically distinct populations (Figure 2), as also showed by the high inter-cluster F_{st} values (0.12-0.24). Estimates of genetic differentiation between sampling sites in each cluster were also quite high with an average F_{st} value of 0.2 (range 0.054 to 0.574), implying the existence of further genetic structuring within each of these three genetic clusters. Of note because of the impending disease merger in northern Uganda is the existence of two northern sub-clusters (F_{st} = 0.046, $p>0.01$). The first includes *Gff* from north of Lake Kyoga (AP, PG, PA, UGT), where the *Tbr* infections newly expanding from the south occur (Figure 3). The second genetic cluster comprises *Gff* from Northwestern Uganda (AR-MY), where *Tbg* infections occur. Interestingly *Gff* samples from these localities are also highly genetically distinct from *Gff* from sampling sites from the western cluster (F_{st} = 0.130, $p>0.01$).

Although *Gff* populations are genetically distinct, genetic admixing among populations from the same or different clusters apparently occur within a 100 km radius [40]. This, however, contradicts the ecological and physiological data on other tsetse species, which imply that tsetse cannot disperse over long distances [41]. Thus it is important to use both ecological and genetic data to understand population level traits, as the temporal scale they provide

insights for is different. Genetic data provide insights in the patterns of dispersion based on multiple generations while ecological data inform on the movement ability of single flies [30,38].

Genetic assignment tests provided information on the direction of the migration and its strength and suggest that for *Gff* in Uganda gene flow is not symmetric, being mostly from southeast to northwest (Figure 3). This asymmetry can be of relevance to HAT epidemiology as it is suggested that the HAT foci emerging in central Uganda result from parasitized cattle movement from south to north [19], and thus transmission of parasites by flies of the northern genotype. Genetic data, however, indicate the movement of flies from the southeast regions towards central Uganda [40], which in turn can imply a role for the susceptible southern genotypes for fueling the new HAT foci (Figure 3).

Generally there is strong evidence of correlation between geographic and genetic distances, as indicated by the existence of strong signals of Isolation By Distance (IBD) in the three main genetic clusters, particularly in the south (r^2 values of 0.35 –0.52) (Figure 2B) [42]. IBD patterns of isolation were also observed on a smaller geographic scale among distinct genetic clusters of sampling sites along the Lake Victoria Basin [42]. This implies that geographically close sampling sites exchange migrants more frequently than geographically distant sites. However, the finding that the strength of this correlation varies among population clusters suggests that the distance over which genetic exchange occurs can vary among clusters. From a vector control and monitoring perspective this means that it is necessary to select site-specific sizes for the operational blocks to reduce tsetse re-invasion risks from neighboring sites.

Assessing the respective contributions of each sex can contribute to moving genes across the spatial landscape, allowing us to evaluate the success of control strategies based on unequal dispersal capacities. Although it is generally thought that females disperse over longer distances than males because of their larger size, genetic analyses suggest male-biased dispersal in tsetse [42]. This finding can have significant implications, especially for Sterile Insect Technique (SIT) programs, as sterile males would likely be able to efficiently compete with resident breeders, if they tend to disperse further than females. Genetic data can also provide estimates of effective population sizes (N_e) and of their change over multiple generations (temporal stability) and help elucidate population dynamics (meta-population dynamics). N_e estimates for *Gff* were variable among Ugandan populations and generally associated with large confidence intervals, suggesting that population sizes could fluctuate seasonally, as had been reported for *Gff* [43, 44]. However, analysis of genetic variation from the same locations, separated by 8-16 generations, shows little evidence of genetic bottleneck and suggests that *Gff* populations are genetically stable over time, with little seasonal variation in abundance [40, 42, 45]. The seasonal fluctuation in trapping data observed for *Gff* is likely due to the poor efficiency of the available traps, which generally capture less than 25% of the population for *palpalis* group flies [44]. The finding of stable populations has important implications for control and suggests that re-colonization of target areas could involve both re-invasion from neighboring sources as well as re-emergence from local residual population pockets undetected by traditional monitoring methods. Moreover, it also points out the need for more sensitive sampling techniques to detect residual populations, when tsetse densities decrease due to vector control.

Given the absence of obvious alternative barriers to gene flow, and the fairly geographically narrow contact zone around Lake Kyoga, *Gff* flies from the northern and southern population clusters may have come into contact only recently, possibly in relation to changing environmental conditions. Ongoing analyses of environmental features over the past 50 years suggest that the regions north and south of Lake Kyoga differ significantly in

climate, especially in annual precipitation, which correlates significantly with genetic differentiation patterns. Understanding the environmental forces contributing to the genetic differences between geographically distinct *Gff* samples in spite of no obvious physical barriers and the potential of the flies to disperse over long distances, has important epidemiological implications, as climate changes are likely to impact *Gff* distributions and thus alter the map of the regions at risk of the disease.

Using population genetics to inform operational activities: the case of Lake Victoria basin

A genetic screen of *Gff* from island populations and the mainland over the ~200 km area along the Lake Victoria shoreline identified three genetic clusters (Figure 4) [42]. The geographic locations of the genetic clusters, however, did not match the PATTEC defined operational blocks, suggesting the need for redefinition of the boundaries of the control units. Genetically derived dispersal distances varied between clusters ranging from about 2.5 to 15 km, with the larger dispersal values found in cluster 1 [42], and matched reasonably well with dispersal rates predicted from Mark-Release-Recapture (MRR) data for *Gff* and other riverine species (14.2km/generation calculated from the movement estimate of 338 m/day [46,49]).

The differences in dispersal rates between clusters correlate well with differences in overall levels of genetic differentiation both within and among clusters, and in turn reflect different habitat fragmentation. Hierarchical F_{ST} , IBD analyses, and individual assignment tests suggest that flies in clusters 1 and 2 are genetically distinct from each other and from flies in the other two clusters, and that flies in clusters 3 and 4 are genetically less differentiated from each other and share many migrants. The difference in connectivity between these genetic clusters could reflect the impact of human activities in the region. Suitable *Gff* habitat is still present in the regions inhabited by flies in clusters 3 and 4 (i.e., riparian and lacustrine thickets, tree canopies with understory vegetation, or patches of banana and *Lantana camara* [50-52]), while it is greatly reduced where clusters 1 and 2 occur, because of human settlement and activities. The expansion of the city limits of Kampala could be one of the reasons for the genetic differentiation between clusters 1 and 2. Within cluster 1, the higher level of isolation of tsetse flies observed between sampling sites could be related to both intensive tsetse control since the 1960s and deforestation, which have left only small gallery forests along the lakeshores [42]. These results suggest that control strategies need to take into account the genetic connectivity of the tsetse populations, which varies among the different genetic clusters. The long-range dispersal capacity of *Gff* also suggests that the Lake Victoria islands should be viewed with caution as sites for eradication trials, as no significant genetic differentiation exists between offshore islands and the mainland sites only 20-30 km away. This finding, suggests that the water body does not act as a barrier to fly movements and gene flow [42]. In the Loos islands (30 km off the coast of Guinea), little genetic exchange has been found among *Glossina palpalis gambiensis* coastal and island population [53]. The difference in genetic exchange observed between these two studies may be due to the high opportunity for passive dispersal owing to active boat traffic between the mainland and the Lake Victoria islands.

Genetics of tsetse populations transmitting *Tbr* and *Tbg* parasites

Of epidemiological relevance is the analysis of *Gff* populations from northern Uganda given the expanding range of the two human parasites, *Tbr* and *Tbg*, which were reported to be less than 150 km apart [5,14]. Genetic data reveal differentiation between *Gff* from sampling sites within the two northern clusters, each one currently carrying only one of the two human-infective parasites, but that soon could be infected with both. Clustering analyses

revealed genetic differentiation within the *Gff* populations in Moyo and Apac districts bordering DRC, in southwest Uganda (MF) and in central Uganda north of Lake Kyoga (AP, BG, PD) (Figure 2). The impending contact of the two disease belts and the finding of genetic differentiation between the vector populations that carry the two parasites pose a series of epidemiologically relevant and timely questions. For instance, the vector genetic differentiation implies differential adaptation of the tsetse host to the two human parasites. Findings of ongoing gene flow in a northwestern direction also open interesting questions with regard to the impending merger of the two disease belts. A fine scale co-evolution study of host and parasites could help answer these questions. Have the two parasites coevolved with the different tsetse genome backgrounds? Will the two parasites co-exist in a single fly, or in the same location in different fly genotypes? Will the northern direction of the host gene flow mean that the *Tbr* parasites carried by the southern flies be favored? Will *Tbr* displace *Tbg*? To address these questions, it is important to understand at a local geographic scale the patterns and level of genome wide variation in both the parasite and the host, as well as the frequency and dynamics of co-infections with multiple parasite strains and the geographic and evolutionary origin of the parasite strains and the vector. Parasite and vector genome-wide SNP markers would be more informative as microsatellite loci, and polymorphism analyses of a few nuclear or mtDNA loci do not provide the statistical power to answer these questions [54,55].

A microsatellite loci based genetic screen confirmed the co-existence of multiple strains of *Trypanosoma* from the same taxon in single flies [56]. This finding has important epidemiological consequences [57], given the impending merger of the *Tbr* and *Tbg* strains and the possibility of recombination between *Trypanosoma* strains in the tsetse salivary glands [58-64]. An analysis of microsatellite and kinetoplastid genetic variation of 142 *T. brucei* isolates across Africa confirmed that the taxonomy of the three subspecies, currently based on host range, pathogenicity, and geographic origin, does not reflect their evolutionary history [65]. While *Tbg* strains are genetically distinct from *Tbb* and *Tbr*, the last two taxa are not distinct genetic units. The population genetic analyses revealed that *Tbr* evolved multiple times from different *Tbb* genetic backgrounds [65].

The diagnostic feature of *Tbr* is the presence of the *SRA* gene, which enables *Tbr* to evade lysis by the Apolipoprotein L-1 (ApoL-1) in human serum (66). *Tbg* has evolved independent mechanism(s) from *Tbr* for evading trypanolysis, one of which involves the haptoglobin-receptor, HbHpr [67,68]. Based on a survey of a small number of strains from the three subspecies for genetic variation in *HbHpr*, five fixed, non-synonymous amino acid substitutions were identified and hypothesized to be involved in evading the immunological response [69]. A genetic screen of *HbHpr* on a geographically more comprehensive sampling than the previous study, however, confirmed only one of these substitutions as fixed [70]. A recent structural analysis study demonstrated the functional importance of the fixed substitution [71]. This result highlights the importance of carrying out screens of genetic variation that include representative samples from the geographic range of the taxa analyzed to tease apart the effect of genetic drift and selection, when trying to link genetic polymorphism to functionally relevant genomic variants.

Role of polyandry for success of control activities

Genetic data provide insights into the reproductive behaviour of tsetse, particularly on the number of times a female mates and maintains sperm from different males in the wild (polyandry). Polyandry may be fundamental to the re-infestation of cleared areas and/or of residual populations and also enhance the reproductive potential of re-invading propagules, thereby expanding their effective population size. Remating can also inhibit the success of genetic control methods, such as SIT or IIT (incompatible insect technique), or replacement

of disease susceptible natural populations with modified resistant phenotypes. Genetic analysis of the mothers and the stored sperm from two populations in Uganda indicate that remating is common (57% in Kabukanga, KB) in the West and 33% in Buvuma Island (BU) in Lake Victoria (Figure 1) and may be influenced by population age structure [72]. The temporal stability in gene frequencies of *Gff* populations over the dry and wet seasons suggests that populations do not fluctuate seasonally and that control programs based on SIT should release abundant sterile males, even within residual surviving target populations in the dry season.

Tsetse microbiota as influential factors of vector competence

Several maternally transmitted symbiotic microbes play important role(s) in tsetse physiology, including obligate *Wigglesworthia*, commensal *Sodalis*, and parasitic *Wolbachia* (Figure 5, Box 1). All *Gff* individuals analyzed in Uganda harbored *Wigglesworthia* [73,74]. Given that *Wigglesworthia* is vertically transmitted from mother to offspring, genetic variation was apparent within and among populations with two distinct symbiont lineages corresponding exactly to northern and southern *Gff* mtDNA haplogroups [74]. This congruence between host and symbiont DNA was also reported by a previous study on all three *G. fuscipes* subspecies [22]. More comprehensive geographic sampling within Uganda, however, showed incongruence within populations. This incongruence, which could potentially arise from incomplete lineage sorting, can generate novel combinations between symbiont genetic variants and host background. *Wigglesworthia* has been shown to influence host vector competence by inducing a host immune gene expression with trypanolytic activity [75,76]. It is possible that *Wigglesworthia* genotypes could differentially induce host immunity and thus provide novel targets for influencing tsetse vector competence.

Given the host genetic structuring detected in the field, populations were screened for *Wolbachia* infections, which can cause cytoplasmic incompatibility (CI) in tsetse [77], and thus can impact the reproductive success of infected and uninfected individuals, respectively. Previous surveys using standard PCR assays could not detect *Wolbachia* infections in *Gff* in Uganda or Kenya [78,79]. When more stringent PCR assay conditions and qPCR analysis were used, low-density *Wolbachia* infections were detected in different *Gff* populations [80]. The infection prevalence in different host genetic populations ranged from zero to complete fixation, with the greatest prevalence observed in the southern genotypes, followed by northern and western genotypes (55%, 42%, and 26%, respectively) [80]. To understand *Wolbachia* evolutionary dynamics and their relationship with host mtDNA, populations of known genetic background were screened for *Wolbachia* infection status, which showed that *Wolbachia* infections in *Gff* are not limited to a single host mtDNA haplotype [80]. Applying a multi-locus sequence typing approach demonstrated that *Wolbachia*-infected *Gff* individuals carry strains from at least two different lineages [73]. Very high levels of *Wolbachia* diversity are also present (evidenced by *groEL* sequence analysis) both within and between individuals (haplotype diversity = 0.945) associated with 26 host mtDNA haplotypes [73]. High *Wolbachia* sequence diversity and the association of *Wolbachia* with multiple host haplotypes suggest that different *Wolbachia* strains infected *Gff* multiple times independently. This limited analysis did not statistically support that bidirectional CI has influenced the host genetic diversity observed in Uganda [73]. Nevertheless, in general, *Wolbachia* infections in *Gff* (*wGff*) are associated with the most common mtDNA haplotypes, suggesting a potential fitness advantage of *wGff* on host physiology. Thus, future mating experiments between distinct *Gff* populations may shed light on the role of *Wolbachia* mediated CI effects on the observed host genetic patterns.

The prevalence of the entomopathogenic DNA virus (Salivary Gland Hypertrophy Virus, SGHV) differed significantly in *Gff* between the north and south mtDNA haplotypes, but not based on microsatellite genotypes [73]. This was expected given the vertical transmission of SGHV coupled with reduced female dispersal. Since SGHV can adversely affect tsetse reproduction [81], the biological significance of the associations between SGHV prevalence and densities and host genotype composition deserves further investigation. Interestingly, none of the *Gff* populations in Uganda harbored *Sodalis* infections, which had been shown to influence trypanosome infection dynamics in populations in the West Africa [82]. Absence of *Sodalis* from Kenyan *Gff* populations was also noted in a different survey, which found *Wigglesworthia* to be the dominant bacteria along with several other environmental microbes [78].

Host-symbiont-pathogen interactions

Most of our knowledge on host-pathogen interactions is based on the paradigm of one host-one pathogen. In reality multiple parasites and symbionts co-exist within a given host, but much less is known about their interactions. In addition to host genetic factors, it is becoming increasingly clear that microbiome diversity and host-microbiome interactions affect vector competence [83]. Using multivariate analysis, complex co-infection dynamics between microfauna of the tsetse and pathogen infections were uncovered, including statistically significant associations between *Gff* genetic groups and pathogen prevalence [73]. Host groups were found to be inversely correlated for *Wolbachia* and SGHV prevalence, whereas trypanosome infection prevalence co-varied with the presence of *Wolbachia* and SGHV, highlighting the importance of examining multiple pathogens simultaneously before making generalizations about infection and spatial patterns. Across Uganda, *Wolbachia* prevalence was 44% while the prevalence of SGHV and *Trypanosoma* (including all trypanosome species) was 12% and 18%, respectively. Across the 18 Uganda populations, when *Wolbachia* prevalence was below 30%, SGHV and trypanosome prevalence tended to be higher. Trypanosome prevalence was positively correlated with SGHV infection ($r=0.257$) and negatively correlated with *Wolbachia* infection ($r=-0.176$), while *Wolbachia* was negatively correlated with SGHV ($r=-0.408$) [73]. These findings imply that *Wolbachia* may prevent infection by the other two pathogens. Thus, high prevalence of SGHV in natural tsetse populations could signal an increased risk for HAT, while *Wolbachia* would indicate a decreased risk. Consistent with these observations, *Wolbachia* has been shown to confer resistance in mosquitoes to *Plasmodium*, as well as dengue and chikungunya virus infections through host immune activation [84-87]. The host resistance phenotype associated with *Wolbachia* infections is currently being exploited in the field as a novel method to interfere with disease transmission by mosquitoes [88]. Thus, further validation of *Wolbachia*-mediated trypanosome resistance in tsetse is of applied significance.

Modeling role of vector control in disease management

Global sensitivity analyses applied to mechanistic models parameterized by field and laboratory data and developed to assess the variation in the basic reproduction number, R_0 , of *Tbg* and *Tbr* rank the proportion of blood meals taken from humans by *Gff* as the most important factor for distribution of *Tbg* [89]. The second ranked parameter for *Tbg* and the highest ranked for *Tbr* was the proportion of *Gff* refractory to infection [89]. For *Tbr* the population parameters for tsetse, species composition, survival, and abundance were ranked almost as high as the proportion refractory, assuming regular treatment of livestock with trypanocides as an established practice in many areas of Uganda. These findings provide broad support for the potential effectiveness of vector control strategies, including those that are aimed at increasing refractoriness in tsetse flies for disease management.

Empirical estimates of CI in *G. m. morsitans* have been used to parameterize mathematical models to demonstrate the effectiveness of *Wolbachia* as a gene driver, which could in turn be harnessed as the basis for a population replacement approach to spread parasite resistant tsetse [70]. Specifically, modeling demonstrated that given a release corresponding to 10% of the natural population level, *Wolbachia*-infected tsetse can invade a wild-type population, with a fixation prevalence predicted at 96.9% (85.6%, 99.8%). At these fixation prevalences, the R_0 of trypanosomiasis will be driven below one and the disease eliminated potentially within a couple of years, depending on the initial release abundance [70].

Concluding remarks

Population genetics and genomics data provide powerful tools that can assist vector control activities on the ground. The availability of this information for the most important vector of HAT in Uganda can enhance the design and implementation of control activities. We describe the future studies that now need to be undertaken to ensure sustainability of the ongoing vector control activities, to enable the development of new control tools, and to predict and prevent resurgence of disease, particularly in areas where the two forms of HAT can overlap (Box 2).

Acknowledgments

We are grateful for our collaborations with John Enyaru, Imna Malele, Charles Masembe, Enock Matova, Paul Mireji, Grace Murilla, Elizabeth Opiyo, and Johnson Ouma, who have contributed to the development of ideas and to the training of many MSc and PhD students employed on these projects. We are grateful to the technical staff at NaLIRRI for the meticulous temporal collections they have conducted over multiple years in Uganda that enabled this study. Some of the data we presented in this review were obtained in collaboration with the laboratories of Kostas Bourtzis, John Enyaru, and Anna Malacrida. The data presented were part of the thesis projects of the PhD students Mr Patrick Abila, Oliver Balmer, Richard Echodu, Rogers Azabo, and Agapitus Kato. We are grateful to postdoctoral fellows Drs Uzma Alam, John Beadell, Corey Brelsford, Steven Davis, Mark Sstrom, Beckie Symula, and Jan Medlock, who contributed to the development of this body of work. We are also grateful for technical support to Chaz Hyseni, Oleg Kruglov, Michelle O'Neill, and Yineng Wu. We thank Dr Katie Atkins for her critical reading of this manuscript and Dr Attardo for the art-work included in this manuscript. This work was enabled through NIAID grant RO1AI068932 and Fogarty International Center awards D43TW007391 and R03TW008755 and WHO/TDR A80132.

Glossary

Allopatric Populations	conspecific populations occurring in geographically distinct locations.
Basic Reproduction Number (R_0)	R_0 , of an infection is the number of cases one case generates on average over the course of its infectious period.
Bayesian clustering analyses	a method to infer population structures without prior knowledge about the population. Various Bayesian approaches using neutral molecular markers have been proposed, which include using Markov chain Monte Carlo (MCMC) methods to infer population structures. The most commonly used is the program STRUCTURE, which can infer the assignment of individuals to populations or the admixture proportions of individuals for a given number of populations (K). Researchers have extended Bayesian algorithms for various purposes such as to take advantage of spatial information, estimate inbreeding coefficients, and infer K values.
Bottleneck	population bottlenecks occur when a population's size is reduced for at least one generation. Undergoing a bottleneck can drastically

	reduce a population's genetic variation, even if the bottleneck is brief in duration.
Coalescence theory	population genetic theory that traces all alleles of a gene shared by all members of a population to a single ancestral copy, known as the most recent common ancestor (MRCA). The relationships between alleles are represented as a gene genealogy, called the coalescent. The point where two branches meet indicates a coalescent event. The time corresponding to how long ago the MRCA existed for a given gene genealogy is called the coalescent time.
Cluster	a genetic grouping of individuals, from a single or several nearby sampling sites, sharing a similar multi-locus genotypic profile.
Cytoplasmic incompatibility (CI)	CI is a phenomenon that results in sperm and eggs being unable to form viable offspring. The effect arises from changes in the gamete cells caused by intracellular parasites like <i>Wolbachia</i> , which infect a wide range of insect species.
Dispersal	the movement of an individual from its existing population.
Effective population size (Ne)	Ne is the size of an idealized population where all individuals contribute equally to successive generations with no changes over time. This determines how powerful drift will be relative to selection and it is almost much less than the census population size.
Fst	a measure of genetic differentiation due to genetic structure, a special case of Wright's F-statistics, and a common statistics in population genetics. It is estimates from genetic polymorphism data including microsatellite loci and SNPs. The values range from 0 to 1. A zero value implies complete panmixis (i.e., random mating among individuals), implying that the two populations are interbreeding. A value of one suggests that the two populations do not share any genetic diversity.
Genetic assignment test	test to identify immigrant individuals that exhibit a strong signal of membership in a population other than the one in which they are sampled. Immigrants in this sense could be true immigrants or recent descendants of immigrants.
Gene flow	transfer of alleles or genes from one population to another.
Genetic drift	a change in allele frequencies in a population due to random finite sampling.
Genetic drift and selection	both forces are at work in natural populations, but the degree in which they affect alleles varies according to the effective population size (Ne), as the effect of drift on allele frequencies per generation is larger in small populations.
Genome-wide association study (GWAS)	the examination of many genetic variants in multiple individuals to identify DNA regions associated with a trait.
Genotyping	the process of determining the genotype of an individual for one or more loci.

Habitat suitability map	a geographic map that relates measurable environmental variables to the suitability of a site for a species. These maps are useful for planning disease monitoring and control and for evaluating environmental changes impacts.
Haplotype	a stretch of DNA sequences that tend to be inherited together.
Haplogroup	a group of similar haplotypes that share a common ancestor.
Haplotype network	a visualization of the evolutionary relationships among DNA sequences.
Incomplete lineage sorting	a phenomenon that occurs when gene genealogies vary along the genome due to difference in coalescent times. In particular, a gene genealogy may be different from the species phylogeny, because not all gene genealogies have become unique to each derived species, some genealogies may predate speciation.
Incompatible insect technique (IIT)	a control method similar to SIT but where methods that lead to biological sterility are used instead of chemical irradiation. One such potential biological method is the CI conferred by <i>Wolbachia</i> infections.
Isolation by distance (IBD)	IBD describes a pattern of population genetic variation that derives from spatially limited gene flow. IBD is defined as a decrease in the genetic similarity among populations as the geographic distance between them increases. Statistical tests for IBD can be conducted using populations or individuals as the units of replication, although analyses at the individual level typically utilize spatial autocorrelation statistics.
Landscape genetics	the study of the influence of biotic and abiotic landscape features on microevolutionary processes within and between populations. Different from phylogeography as it is carried out in an explicit spatial framework.
Metapopulation	a group of spatially separated populations of the same species, which exchange migrants at different rates and may become locally extinct and recolonized.
Metapopulation dynamics and relevance for control	if populations are genetically diverse and stable for neutral genetic markers, this implies relatively large population sizes. If populations are not genetically stable, this might imply local extinction followed by re-colonization from neighboring locations, or periodical size contractions followed by re-emergence from residual pockets. The term sink-source is used to reflect these dynamics. In other words, the dynamics of any single population is dependent upon migration from neighboring populations as well as temporal changes within a single population, which might include local extinction and re-colonization.
Microsatellite locus	a 2 to 6 base pair repeat motif in a DNA stretch that varies in length between alleles. Their high variability, genome ubiquity, co-dominance, and bi-parental inheritance make them useful for population and individual level studies of both sexes.

Mitochondria DNA (mtDNA) haplogroups	groups of related mtDNA sequences. The faster coalescence times of mtDNA than nuclear markers make them useful for shallow divergence studies. The maternal inheritance allows studying the evolutionary processes of female lineages. The comparison of bi-parental nuclear and uni-parental mtDNA regions allows studying sex-biased phenomena.
Paratransgenic fly	a fly that carries a genetically modified symbiotic organism. The expression of products in the symbiotic bacteria can confer the paratransgenic fly beneficial traits, such as anti-parasitic properties. This mode of gene expression is an alternative approach to transgenic insects, where foreign genes are expressed in insect cells.
Phylogeography	the study of spatial distribution of genetic differences within and between populations of closely related species using a historical framework.
Polyandry	a type of breeding adaptation in which one female mates with multiple males.
Sterile insect technique (SIT)	SIT is a method of biological control, whereby overwhelming numbers of sterile male insects are released. The sterile males compete with the wild males for female insects. If a female mates with a sterile male then it will have no offspring, thus reducing the next generation's population. Repeated release of insects can eventually wipe out a population, leading to its eradication.

References

1. Simarro PP, et al. Estimating and mapping the population at risk of sleeping sickness. *PLoS Negl Trop Dis.* 2012; 6:e1859. [PubMed: 23145192]
2. Simarro PP, et al. Eliminating human African trypanosomiasis: where do we stand and what comes next? *PLoS Med.* 2008; 5:e55. [PubMed: 18303943]
3. Matemba LE, et al. Quantifying the burden of rhodesiense sleeping sickness in Urambo District, Tanzania. *PLoS Negl Trop Dis.* 2010; 4:e868. [PubMed: 21072230]
4. Fevre EM, et al. Estimating the burden of rhodesiense sleeping sickness during an outbreak in Serere, eastern Uganda. *BMC Public Health.* 2008; 8:96. [PubMed: 18366755]
5. Simarro PP, et al. The human African trypanosomiasis control and surveillance programme of the World Health Organization 2000-2009: the way forward. *PLoS Negl Trop Dis.* 2011; 5:e1007. [PubMed: 21364972]
6. Simarro PP, et al. Risk for human African trypanosomiasis, Central Africa, 2000-2009. *Emerg Infect Dis.* 2011; 17:2322. [PubMed: 22172322]
7. Jamonneau V, et al. Untreated Human Infections by *Trypanosoma brucei gambiense* Are Not 100% Fatal. *PLoS Negl Trop Dis.* 2012; 6(6):e1691. [PubMed: 22720107]
8. Simarro PP, et al. Update on field use of the available drugs for the chemotherapy of human African trypanosomiasis. *Parasitology.* 2012; 139:842. [PubMed: 22309684]
9. Jannin J, Cattand P. Treatment and control of human African trypanosomiasis. *Curr Opin Infect Dis.* 2004; 17:565. [PubMed: 15640711]
10. Nok AJ. Arsenicals (melarsoprol), pentamidine and suramin in the treatment of human African trypanosomiasis. *Parasitology Research.* 2003; 90:71. [PubMed: 12743807]
11. Alirol E, et al. Nifurtimox-eflornithine combination therapy for second-stage gambiense human African trypanosomiasis: Medecins Sans Frontieres experience in the Democratic Republic of the Congo. *Clinical infectious diseases.* 2013; 56(2):195–203. [PubMed: 23074318]

12. Fevre EM, et al. Reanalyzing the 1900-1920 sleeping sickness epidemic in Uganda. *Emerg Infect Dis.* 2004; 10:567. [PubMed: 15200843]
13. Fevre EM, et al. A burgeoning epidemic of sleeping sickness in Uganda. *Lancet.* 2005; 366:745. [PubMed: 16125592]
14. Picozzi K, et al. Sleeping sickness in Uganda: a thin line between two fatal diseases. *BMJ.* 2005; 331:1238. [PubMed: 16308383]
15. Okedi, L., et al. Proceedings of the Annual East African Network Tsetse Trypanosomiasis. Whitesands Beach Hotel; Mombasa, Kenya: 2008.
16. Fevre EM, et al. Human African trypanosomiasis: Epidemiology and control. *Adv Parasitol.* 2006; 61:167. [PubMed: 16735165]
17. Funk S, et al. Identifying transmission cycles at the human-animal interface: the role of animal reservoirs in maintaining gambiense human african trypanosomiasis. *PLoS Comput Biol.* 2013; 9:e1002855. [PubMed: 23341760]
18. Welburn SC, et al. Identification of human-infective trypanosomes in animal reservoir of sleeping sickness in Uganda by means of serum-resistance-associated (SRA) gene. *Lancet.* 2001; 358:2017. [PubMed: 11755607]
19. Batchelor NA, et al. Spatial predictions of Rhodesian Human African Trypanosomiasis (sleeping sickness) prevalence in Kaberamaido and Dokolo, two newly affected districts of Uganda. *PLoS Negl Trop Dis.* 2009; 3:e563. [PubMed: 20016846]
20. Kabasa JD. Public-private partnership works to stamp out sleeping sickness in Uganda. *Trends Parasitol.* 2007; 23:191. [PubMed: 17392023]
21. Waiswa C, et al. Domestic animals as reservoirs for sleeping sickness in three endemic foci in south-eastern Uganda. *Ann Trop Med Parasitol.* 2003; 97(2):149–55. [PubMed: 12803870]
22. Dyer NA, et al. Molecular phylogenetics of tsetse flies (Diptera: Glossinidae) based on mitochondrial (COI, 16S, ND2) and nuclear ribosomal DNA sequences, with an emphasis on the palpalis group. *Mol Phylogenet Evol.* 2008; 49:227. [PubMed: 18692147]
23. Machado, A. d. B. Revision systematique des glossines du groupe palpalis (diptera). *Publ. Cult. Co. Diamates de Angola.* 1954; 22:1.
24. Dyer NA, et al. Cryptic diversity within the major trypanosomiasis vector *Glossina fuscipes* revealed by molecular markers. *PLoS Negl Trop Dis.* 2011; 5:e1266. [PubMed: 21858237]
25. Cecchi G, et al. Land cover and tsetse fly distributions in sub-Saharan Africa. *Med Vet Entomol.* 2008; 22:364. [PubMed: 18785934]
26. Hargrove JW. The effect of temperature and saturation deficit on mortality in populations of male *Glossina m. morsitans* (Diptera: Glossinidae) in Zimbabwe and Tanzania. *Bull Entomol Res.* 2001; 91:79. [PubMed: 11260721]
27. Terblanche JS, et al. Thermal tolerance in a south-east African population of the tsetse fly *Glossina pallidipes* (Diptera, Glossinidae): implications for forecasting climate change impacts. *J Insect Physiol.* 2008; 54:114. [PubMed: 17889900]
28. Nash T. A statistical analysis of the climatic factors influencing the density of Tsetse flies, *Glossina morsitans* Westwood. *Journal of Animal Ecology.* 1933; 2:197.
29. Rogers D. Tsetse Population Dynamics and Distribution: A New Analytical Approach. *Journal of Animal Ecology.* 1979; 48:825.
30. Krafur ES. Tsetse fly population genetics: an indirect approach to dispersal. *Trends Parasitol.* 2003; 19:162. [PubMed: 12689645]
31. Kabayo JP. Aiming to eliminate tsetse from Africa. *Trends Parasitol.* 2002; 18:473. [PubMed: 12473355]
32. Limited A. End of project evaluation study for FITCA Regional and the national components of five countries Kenya, Uganda, Tanzania, Rwanda and Ethiopia. Aylesbury. East Africa European Commission. 2005
33. Grant I. Creation of sustainable tsetse and trypanosomiasis free areas in East and West Africa. Environmental risk assessment report: the use of the sequential aerosol technique for the eradication of tsetse populations in Uganda. Boughton: Cybister Environmental Protection. 2009

34. Luyimbazi F. Detailed work plan/action plan for the collection of entomological baseline data. Integrated area-wide program for the creation of sustainable tsetse and trypanosomiasis free areas in the Lake Victoria basin. Entebbe: Ministry of Agriculture, Animal Industry and Fisheries; 2006. Entebbe: Ministry of Agriculture, Animal Industry and Fisheries. 2006
35. Wint W. Kilometre resolution tsetse fly distribution maps for the Lake Victoria basin and West Africa. Vienna: IAEA. Vienna: International Atomic Energy Agency. 2001
36. Vreysen MJ, et al. *Glossina austeni* (Diptera: Glossinidae) eradicated on the Island of Unguja, Zanzibar, using the sterile insect technique. *J. Econ. Entomology*. 2000; 93:123.
37. Solano P, et al. Population genetics as a tool to select tsetse control strategies: suppression or eradication of *Glossina palpalis gambiensis* in the Niayes of Senegal. *PLoS Negl Trop Dis*. 2010; 4:e692. [PubMed: 20520795]
38. Solano P, et al. How can tsetse population genetics contribute to African trypanosomiasis control? *Trends Parasitol*. 2010; 26:255. [PubMed: 20202905]
39. Abila PP, et al. High levels of genetic differentiation between Ugandan *Glossina fuscipes fuscipes* populations separated by Lake Kyoga. *PLoS Negl Trop Dis*. 2008; 2:e242. [PubMed: 18509474]
40. Beadell JS, et al. Phylogeography and population structure of *Glossina fuscipes fuscipes* in Uganda: implications for control of tsetse. *PLoS Negl Trop Dis*. 2010; 4:e636. [PubMed: 20300518]
41. Bursell E E, Taylor P P. An energy budget for *Glossina* (Diptera: Glossinidae). *Bulletin of Entomological Research*. 1980; 70:187–196.
42. Hyseni C, et al. The population structure of *Glossina fuscipes fuscipes* in the Lake Victoria basin in Uganda: implications for vector control. *Parasit Vectors*. 2012; 5:222. [PubMed: 23036153]
43. Katunguka-Rwakishaya E, Kabagambe EK. Tsetse survey in Mukono district, south-east uganda: population structure, distribution and blood meal status. *Trop. Anim. Hlth Prod*. 1996; 28:51–157.
44. Rayaisse JB, et al. Prospects for the Development of Odour Baits to Control the Tsetse Flies *Glossina tachinoides* and *G. palpalis s.l.* *PLoS Negl Trop Dis*. 2010; 4(3):e632. [PubMed: 20300513]
45. Echodu R, et al. Temporal stability of *Glossina fuscipes fuscipes* populations in Uganda. *Parasit Vectors*. 2011; 4:19. [PubMed: 21320301]
46. Cuisance D, et al. Dispersion linéaire de *Glossina palpalis gambiensis* et *G. tachinoides* dans une galerie forestière en zone soudano-guinéenne (Burkina Faso). *Rev Elev Méd vét Pays Trop*. 1985; 38:153–172.
47. Vreysen MJB, et al. Release-Recapture Studies Confirm Dispersal of *Glossina palpalis gambiensis* between River Basins in Mali. *PLoS Negl Trop Dis*. 2013; 7(4):e2022. [PubMed: 23634232]
48. Rogers D. Study of a natural population of *Glossina fuscipes fuscipes* Newstead and a model for fly movement. *J. Anim. Ecol*. 1977; 46:309–330.
49. Bouyer J, et al. Population sizes and dispersal pattern of tsetse flies: rolling on the river? *Mol Ecol*. 2009; 18:2787. [PubMed: 19457176]
50. Okoth J, Kapaata R. A study of the resting sites of *Glossina fuscipes fuscipes* (Newstead) in relation to Lantana camara thickets and coffee and banana plantations in the sleeping sickness epidemic focus, Busoga, Uganda. *International Journal of Tropical Insect Science*. 1987; 8:57–60.
51. Kuzoe F, Schofield C. Strategic review of traps and targets for tsetse and African trypanosomiasis control. UNDP/World Bank/WHO. 2004
52. Syed Z, Guerin PM. Tsetse flies are attracted to the invasive plant *Lantana camara*. *J Insect Physiol*. 2004; 50:43. [PubMed: 15037092]
53. Solano P, et al. The population structure of *Glossina palpalis gambiensis* from island and continental locations in coastal Guinea. *PLoS Negl Trop Dis*. 2009; 3:e392. [PubMed: 19290038]
54. Hide G, Tait A. Molecular epidemiology of African sleeping sickness. *Parasitology*. 2009; 1
55. Schoville SD SD, et al. Adaptive Genetic Variation on the Landscape: Methods and Cases. *Annual Review of Ecology Evolution and Systematics*. 2012; 43(1):23–43.
56. Balmer O, Caccone A. Multiple-strain infections of *Trypanosoma brucei* across Africa. *Acta Trop*. 2008; 107:275. [PubMed: 18671933]

57. Gibson W. The significance of genetic exchange in trypanosomes. *Parasitology Today*. 1995; 11:465.
58. Peacock L, et al. Identification of the meiotic life cycle stage of *Trypanosoma brucei* in the tsetse fly. *Proc Natl Acad Sci U S A*. 2011; 108:3671. [PubMed: 21321215]
59. Peacock L, et al. Intraclonal mating occurs during tsetse transmission of *Trypanosoma brucei*. *Parasit Vectors*. 2009; 2:43. [PubMed: 19772562]
60. Peacock L, et al. Fly transmission and mating of *Trypanosoma brucei brucei* strain 427. *Mol Biochem Parasitol*. 2008; 160:100. [PubMed: 18524395]
61. Gibson W, et al. The use of yellow fluorescent hybrids to indicate mating in *Trypanosoma brucei*. *Parasit Vectors*. 2008; 1:4. [PubMed: 18298832]
62. Gibson W, et al. Analysis of a cross between green and red fluorescent trypanosomes. *Biochem Soc Trans*. 2006; 34:557. [PubMed: 16856859]
63. Bingle LE, et al. A novel GFP approach for the analysis of genetic exchange in trypanosomes allowing the in situ detection of mating events. *Microbiology*. 2001; 147:3231. [PubMed: 11739755]
64. Gibson W, Stevens J. Genetic exchange in the trypanosomatidae. *Adv Parasitol*. 1999; 43:1. [PubMed: 10214689]
65. Balmer O, et al. Phylogeography and taxonomy of *Trypanosoma brucei*. *PLoS Negl Trop Dis*. 2011; 5:e961. [PubMed: 21347445]
66. Gibson WC. The SRA gene: the key to understanding the nature of *Trypanosoma brucei rhodesiense*. *Parasitology*. 2005; 131:143. [PubMed: 16145931]
67. Jackson AP, et al. Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. *Proc Natl Acad Sci U S A*. 2012; 109:3416. [PubMed: 22331916]
68. Capewell P, et al. Differences between *Trypanosoma brucei gambiense* groups 1 and 2 in their resistance to killing by trypanolytic factor 1. *PLoS Negl Trop Dis*. 2011; 5:e1287. [PubMed: 21909441]
69. Kieft R, et al. Role of expression site switching in the development of resistance to human Trypanosome Lytic Factor-1 in *Trypanosoma brucei brucei*. *Mol Biochem Parasitol*. 2012; 183:8. [PubMed: 22226682]
70. Symula RE, et al. *Trypanosoma brucei gambiense* group 1 is distinguished by a unique amino acid substitution in the HpHb receptor implicated in human serum resistance. *PLoS Negl Trop Dis*. 2012; 6:e1728. [PubMed: 22802982]
71. Higgins MK, et al. Structure of the trypanosome haptoglobin-hemoglobin receptor and implications for nutrient uptake and innate immunity. *Proc Natl Acad Sci U S A*. 2013; 110:1905. [PubMed: 23319650]
72. Bonomi A, et al. Polyandry is a common event in wild populations of the Tsetse fly *Glossina fuscipes fuscipes* and may impact population reduction measures. *PLoS Negl Trop Dis*. 2011; 5:e1190. [PubMed: 21666797]
73. Alam U, et al. Implications of microfauna-host interactions for trypanosome transmission dynamics in *Glossina fuscipes fuscipes* in Uganda. *Applied and Environmental Microbiology*. 2012; 78:4627. [PubMed: 22544247]
74. Symula RE, et al. Influence of host phylogeographic patterns and incomplete lineage sorting on within-species genetic variability in *Wigglesworthia* species, obligate symbionts of tsetse flies. *Applied and environmental microbiology*. 2011; 77:8400. [PubMed: 21948847]
75. Wang J, et al. Interactions between mutualist *Wigglesworthia* and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. *Proc Natl Acad Sci U S A*. 2009; 106:12133–12138. [PubMed: 19587241]
76. Wang J, Aksoy S. PGRP-LB is a maternally transmitted immune milk protein that influences symbiosis and parasitism in tsetse's offspring. *Proceedings of the National Academy of Sciences of the United States of America*. 2012; 109:10552. [PubMed: 22689989]
77. Alam U, et al. Wolbachia symbiont infections induce strong cytoplasmic incompatibility in the tsetse fly *Glossina morsitans*. *PLoS pathogens*. 2011; 7:e1002415. [PubMed: 22174680]

78. Lindh JM, Lehane MJ. The tsetse fly *Glossina fuscipes fuscipes* (Diptera: Glossina) harbours a surprising diversity of bacteria other than symbionts. *Antonie Van Leeuwenhoek*. 2011; 99:711. [PubMed: 21203841]
79. Doudoumis V, et al. Detection and characterization of *Wolbachia* infections in laboratory and natural populations of different species of tsetse flies (genus *Glossina*). *BMC microbiology*. 2012; 12(Suppl 1):S3. [PubMed: 22376025]
80. Symula RE, et al. *Wolbachia* association with the tsetse fly, *Glossina fuscipes fuscipes*, reveals high levels of genetic diversity and complex evolutionary dynamics. *BMC evolutionary biology*. 2013; 13:31. [PubMed: 23384159]
81. Abd-Alla AM, et al. Tsetse salivary gland hypertrophy virus: hope or hindrance for tsetse control? *PLoS neglected tropical diseases*. 2011; 5:e1220. [PubMed: 21912708]
82. Farikou O, et al. Tripartite interactions between tsetse flies, *Sodalis glossinidius* and trypanosomes--an epidemiological approach in two historical human African trypanosomiasis foci in Cameroon. *Infect Genet Evol*. 2010; 10:115. [PubMed: 19879380]
83. Weiss B, Aksoy S. Microbiome influences on insect host vector competence. *Trends Parasitol*. 2011; 27:514. [PubMed: 21697014]
84. Bian G, et al. The endosymbiotic bacterium *Wolbachia* induces resistance to dengue virus in *Aedes aegypti*. *PLoS pathogens*. 2010; 6:e1000833. [PubMed: 20368968]
85. Blagrove MS, et al. A *Wolbachia* wMel Transinfection in *Aedes albopictus* Is Not Detrimental to Host Fitness and Inhibits Chikungunya Virus. *PLoS Negl Trop Dis*. 2013; 7:e2152. [PubMed: 23556030]
86. Kambris Z, et al. *Wolbachia* stimulates immune gene expression and inhibits plasmodium development in *Anopheles gambiae*. *PLoS Pathog*. 2010; 6:e1001143. [PubMed: 20949079]
87. Blagrove MS, et al. *Wolbachia* strain wMel induces cytoplasmic incompatibility and blocks dengue transmission in *Aedes albopictus*. *Proc Natl Acad Sci U S A*. 2012; 109:255. [PubMed: 22123944]
88. McGraw EA, O'Neill SL. Beyond insecticides: new thinking on an ancient problem. *Nat Rev Microbiol*. 2013; 11:181. [PubMed: 23411863]
89. Davis S, et al. A global sensitivity analysis for African sleeping sickness. *Parasitology*. 2011; 138:516. [PubMed: 21078220]
90. Luyimbazi F. Detailed work plan/action plan for the collection of entomological baseline data. Integrated area-wide program for the creation of sustainable tsetse and trypanosomiasis free areas in the Lake Victoria basin. Ministry of Agriculture, Animal Industry and Fisheries. 2006
91. Wint, W.; Section, IPC., editors. IAEA; Vienna: 2001.
92. Chen X, et al. Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated endosymbiont, *Wigglesworthia glossinidia*. *Journal of Molecular Evolution*. 1999; 48:49. [PubMed: 9873076]
93. Pais R, et al. The obligate mutualist *Wigglesworthia glossinidia* influences reproduction, digestion, and immunity processes of its host, the tsetse fly. *Appl Environ Microbiol*. 2008; 74:5965. [PubMed: 18689507]
94. Weiss BL, et al. Tsetse immune system maturation requires the presence of obligate symbionts in larvae. *PLoS biology*. 2011; 9:e1000619. [PubMed: 21655301]
95. Weiss BL, et al. Obligate symbionts activate immune system development in the tsetse fly. *Journal of immunology*. 2012; 188:3395.
96. Wang J, et al. Interactions between mutualist *Wigglesworthia* and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:12133. [PubMed: 19587241]
97. Toh H, et al. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res*. 2006; 16:149. [PubMed: 16365377]
98. Geiger A, et al. Vector Competence of *Glossina palpalis gambiense* for *Trypanosoma brucei* s.l. and Genetic Diversity of the Symbiont *Sodalis glossinidius*. *Molecular biology and evolution*. 2006; 24:102–109. [PubMed: 17012373]

99. Jaenson TGT. Virus-Like Rods Associated with Salivary-Gland Hyperplasia in Tsetse, *Glossina Pallidipes*. *T Roy Soc Trop Med H.* 1978; 72:234.
100. Abd-Alla AMM, et al. Dynamics of the salivary gland hypertrophy virus in laboratory colonies of *Glossina pallidipes* (Diptera: Glossinidae). *Virus Res.* 2010; 150:103. [PubMed: 20214934]
101. Guerrini L, et al. Fragmentation analysis for prediction of suitable habitat for vectors: example of riverine tsetse flies in Burkina Faso. *J Med Entomol.* 2008; 45:1180. [PubMed: 19058646]
102. Boko, M.; Parry, ML.; Canziani, OF.; Palutikof, J.; Linden, P. j. v. d.; Hanson, CE., editors. Africa. Cambridge; UK: 2007. Impacts, Adaptation and Vulnerability. Contribution of working group II to the Fourth Assessment report of the Intergovernmental Panel of Climate change.
103. Cecchi G, et al. Mapping sleeping sickness in Western Africa in a context of demographic transition and climate change. *Parasite.* 2009; 16:99. [PubMed: 19585887]
104. Rio RV, et al. Strategies of the home-team: symbioses exploited for vector-borne disease control. *Trends Microbiol.* 2004; 12:325. [PubMed: 15223060]
105. Aksoy, S.; Brian, W.; Attardo, GM. Transgenesis for Vector-borne disease management. Aksoy, S., editor. Landes; NY City: 2007.

Highlights

Population genetics data offer insights with relevance for vector control programs

Genetic data can help devise control strategies to prevent reinvasion of target areas

There are three (northern, southern, and western) genetically distinct clusters of *Glossina fuscipes* in Uganda

There is apparent gene flow from southeast to northwest regions of Uganda

Movement of flies from the southeast towards central Uganda may fuel new HAT foci

Fly microbiome may influence parasite transmission traits

Low density infections with multiple *Wolbachia* strains are widely prevalent

Evidence for multiple matings are observed in females at high frequency

Box 1. Tsetse microbiome influences on host physiology

Wigglesworthia. All tsetse species rely on the obligate symbiont genus *Wigglesworthia* for supplementing their vertebrate blood specific diet with nutrients essential for fecundity [92,93], as well as maturation of the adult immune system [94,95]. In response to symbiotic peptidoglycan (PGN), tsetse express an immune molecule, peptidoglycan recognition protein, (PGRP-LB), which degrades PGN to prevent induction of host immune pathways [75], and also blocks trypanosome infections in the gut through an anti-trypanosomal activity [76]. In natural populations examined in Uganda, a positive correlation is noted between the density of *Wigglesworthia* and the level of PGRP-LB synthesized [96].

Sodalis. *Sodalis* is related to free-living enterics, but its genome indicates an active process of functional erosion [97]. Studies with *G. palpalis gambiensis* showed a statistically significant association between the presence of specific genotypes of *Sodalis* and host ability to establish midgut trypanosome infections [98]. None of the *Gff* individuals screened from Uganda harbored *Sodalis* infections [77,78].

Wolbachia. The prevalence of *Wolbachia* infections in natural populations of different tsetse species and the strains infecting different tsetse species vary [79]. In the laboratory line of *Glossina morsitans morsitans*, *Wolbachia* infections have been shown to confer strong cytoplasmic incompatibility (CI) when *Wolbachia*-infected males were mated with uninfected females [77]. Any beneficial phenotypes that accompany *Wolbachia*-infected flies are spread into natural populations as a result of CI, such as other maternally transmitted symbionts and products that they may synthesize. Mathematical modeling suggested that the time to reach fixation from a release of the size of 10% of the native population can be relatively short with the median value being 529 days [77].

SGHV. SGHV not only causes hypertrophy of the salivary glands, but also causes reproductive organ abnormalities and reduces fecundity and fertility [100,101]. The prevalence of symptoms of SGHV in wild tsetse populations is usually very low (0.2% to 5%), but higher prevalence rates (15.2%) have been observed occasionally [77]. A *Glossina pallidipes* colony originating from Ethiopia was successfully established in 1996, but later up to 85% of adult flies displayed symptoms of SGHV, and the colony declined and became extinct by 2002.

Box 2. Future directions

1. **The relationship between *Gff* and its environment.** Habitat availability will primarily control population densities and connectivity because tsetse distribution is tightly linked to the extent of land cover. Thus, it is important to determine the correlation between genetic and environmental variation, which will allow us to: (i) improve *Gff* habitat suitability maps, which are currently based on entomological field surveys and related to vegetation or land cover [25,101]; and (ii) evaluate the impact of climate change on *Gff* geographic distribution. Based on projected climate change [102], tsetse distributions have the potential of both expanding and contracting with important epidemiological repercussions [103].
2. **Polyandry and sperm use in natural populations.** Given the high prevalence of polyandry, the sperm use in females remains to be determined. In the context of SIT programs, the efficacy of matings with sterile males can be reduced if wild-type sperm is preferred. Similarly the efficacy of *Wolbachia* CI based replacement or IIT methods would be compromised if rematings were not random and occurred more frequently following incompatible matings. Mathematical models with empirical data can help predict the success of these control methods.
3. **Microbiome diversity and influences on pathogen transmission.** The *Wolbachia* associated with *Gff* (*wGff*) infections are associated with the most common haplotypes, suggesting a potential fitness advantage of *wGff*. Thus, it is important to examine the maternal transmission efficiency of *wGff* and CI expression before excluding the possibility of bidirectional CI in the field. Additionally, given that the positive correlation found between SGHV and trypanosome infections can indicate increased HAT risk, the negative correlation found between *Wolbachia* presence and trypanosome infections needs further investigation for functional mechanisms and for potential field applications, as *Wolbachia* infections in the field can enhance trypanosome resistance in natural susceptible populations.
4. **Future applications of population and functional genomics approaches.** Analyses based on few loci fail to provide insights on the causal relationship between adaptive genetic variation and environmental variables and traits of epidemiological interest, such as refractoriness to parasite transmission. The next-generation revolution is transforming genotyping strategies for population genomic studies by providing the opportunity for large-scale Single Nucleotide Polymorphism (SNP) discovery and genotyping for Genome Wide Association Studies (GWAS). Our current understanding of the spatial genetic structuring of Ugandan *Gff* populations provides the opportunity to test for association between genomic variation and environmental or epidemiological relevant traits, while controlling for spatial dependence among samples and providing multiple independent replicates. Future GWAS investigations can help evaluate how *Gff* distributions are impacted by environmental variables and climate change scenarios and help identify genomic regions that are correlated with environmental variables. GWAS studies from multiple fly populations with differing genetic backgrounds and parasite infection status can tease apart the signal of selection from genetic drift and identify genetic variations responsible for the epidemiologically relevant traits, which can then be investigated by functional approaches.

5. Modeling. Mathematical modeling can be used as a tool for translating empirical findings related to the biology and population genetics of tsetse, *Sodalis*, and *Wolbachia* into predictions of the epidemiological impact of intervention strategies, including those of paratransgenic tsetse [104,105]. In this way, modeling acts as a bridge between field and laboratory results to an evaluation of the potential effectiveness of public health policies. Specifically, modeling can help determine whether the introduction of paratransgenic tsetse could successfully eliminate trypanosome transmission in the tsetse population, as well as significantly reduce the number of infections, and, if so, over what time frame. Additionally, uncertainty analyses can take into account empirical uncertainty in parameter values to assess the extent of confidence in quantitative predictions and the sensitivity of the system to different parameters.

In terms of translating empirical innovations into public health policy, cost-effectiveness analysis is fundamental. Work is needed to incorporate health-related costs into epidemiological models of trypanosomiasis to evaluate the cost-effectiveness of the different control strategies, including tsetse traps, insecticide, tsetse habitat reduction (land clearing), sterile tsetse male release, modification of vector populations, and treatment of humans and domestic animals.

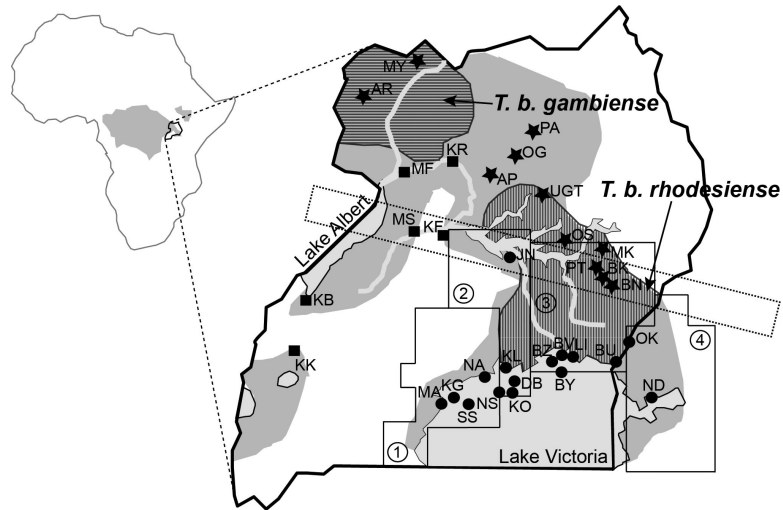


Figure 1. *Glossina fuscipes* and HAT distribution in Uganda

Left: Map of Africa showing in gray the *G. fuscipes* distribution and the location of Uganda. *Right:* Map of Uganda with *Gff* sampling sites that are referred to in the text. Different shades of gray identify main water bodies (light gray) and the current *Gff* distribution (gray). The ranges of the human infective *Trypanosoma brucei*, *T. b. rhodesiense* and *T. b. gambiense*, are superimposed on the *Gff* distribution and identified with a vertical and horizontal hatched pattern, respectively. The populations analyzed genetically in previous studies are identified with two letters and a symbol for their assignment to one of the three main genetic clusters (dot=southern; star=northern; square= western). The dashed rectangle identifies the contact zone around Lake Kyoga in central Uganda where the northern and southern mtDNA haplogroups co-occur. The four operational blocks identified by PATTEC to progressively eradicate *Gff* from the Lake Victoria basin from west to east are also identified with small dashed lines and numbers [90,91]. The blocks are based on the Food and Agriculture Organization (FAO) predicted habitat suitability for *Gff*, natural barriers, major urban areas, international borders, and drainage patterns. Block 1 in the western part of the country is the most isolated block due to the expansion of the city of Kampala and subsequent urbanization and habitat fragmentation of the surrounding area. Block 2 in the central area has been targeted for control to create a buffer zone between the eradication block and the rest of the *Gff* predicted range in Uganda, while only vector population monitoring activities are planned for the other two blocks during phase 1 of the control program. Upon successful eradication in block 1, block 2 would become the eradication target and so on until the whole basin is tsetse-free.

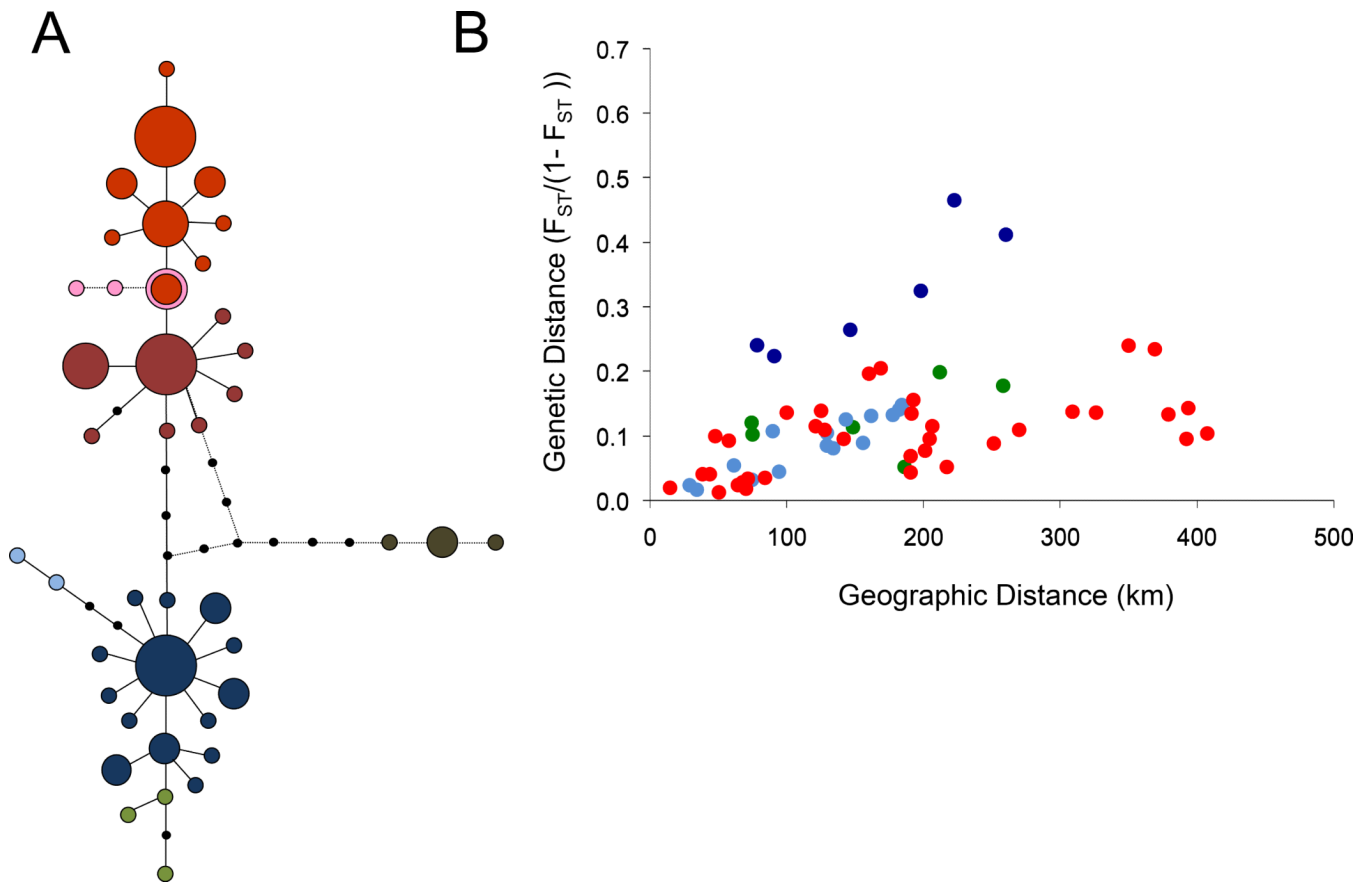


Figure 2. mtDNA based *Gff* network and isolation by distance

(A) The mtDNA sequences are grouped into genetically distinct haplogroups. The parsimony network of mtDNA sequences shows the evolutionary relationships among mtDNA haplotypes from *Gff* Ugandan populations and one population from Sudan and DRC. Each circle represents a unique haplotype, and circle size is proportional to its frequency in the *Gff* populations sampled. Black dots represent unsampled haplotypes, and each line represents one mutation. The different colors identify the geographic location of the haplogroups (blue, southern; red, northern). The pink and brown dots represent haplotypes found in the sampling sites from Sudan (pink) and DRC (brown). The geographic distribution of these haplogroups shows disjunct: one to the North of Lake Kyoga, the other to the South and West of Lake Kyoga, with the exception of a narrow geographic contact zone around Lake Kyoga, (see [40] for details). Figure modified from [39]. (B) Isolation by distance (IBD) analysis illustrates the relationship between geographic and genetic distance within each of the three population clusters for all pairwise comparisons among *Gff* populations in Uganda within the northern, western, and southern genetic clusters. Linearized F_{ST} (i.e., $F_{ST}/(1-F_{ST})$) values were regressed against geographic distance (in kilometers). Red dots refer to comparison within the northern genetic cluster. Green and blue dots refer to comparisons within the western and southern genetic clusters, respectively. Strong signals of IBD were evident in both these regions but particularly in the south (R^2 values of 0.35 to 0.52), except for comparisons involving a Kenyan population (Ndere Island, ND, dark blue dots). Figure modified from [38].

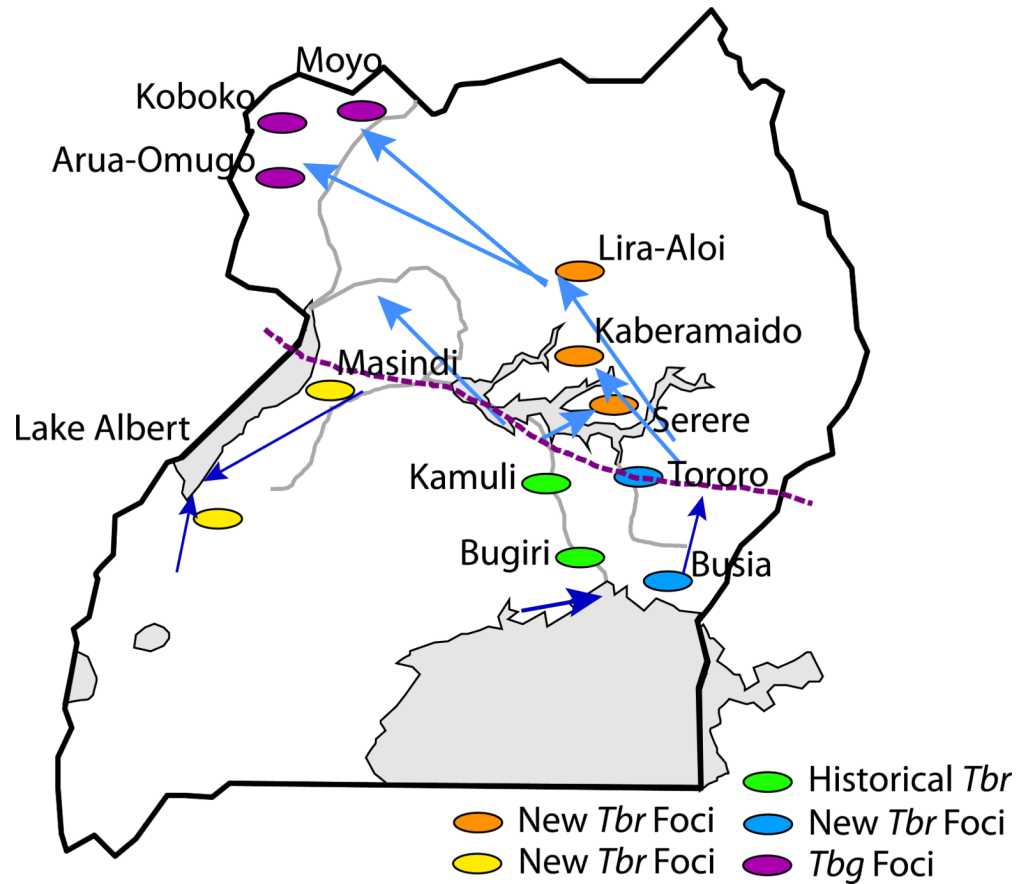


Figure 3. Inferred patterns of *Gff* migrations in Uganda based on microsatellite loci analyses
 Schematic summary of the migration patterns among *Gff* population clusters based on the results of genetic assignment test using microsatellite loci variation. The arrows indicate the direction of immigration of fly genotypes from one population cluster to the other with the thickness of each arrow illustrating the proportion of individuals in the receiving population that exhibit recent immigrant ancestry using genetic data. Names refer to historical and recent HAT foci. Grey shading indicates the water bodies. The dotted line in the central region indicates the zone of contact between northern and southern mtDNA haplotypes. The green ovals show the locations of historical *Tbr* disease foci, the blue, orange, and brown ovals show the locations of the newly emerging *Tbr* foci. The *Tbg* foci in the NW are denoted by purple dots.

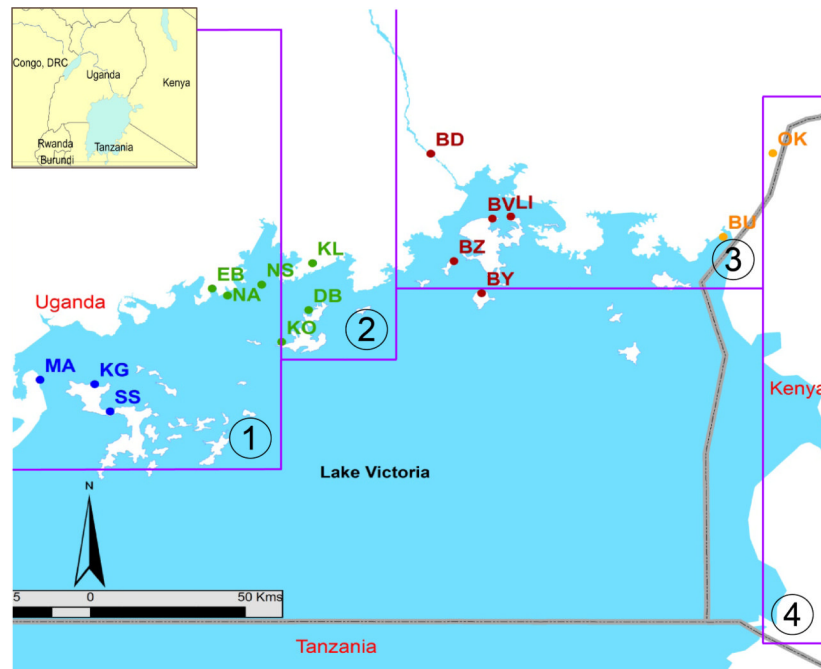


Figure 4. Genetic differentiation of *Gff* around Lake Victoria

Sixteen *Gff* populations were studied by microsatellite analysis from coastal mainland and island sampling sites. The purple lines identify the 4 PATTEC blocks. The map shows the locations of each sampling site (dots) and the genetic assignment (dot color) of each sampling site based on the Bayesian clustering program Structure: orange = cluster 1; brown = cluster 2; green= cluster 3, and blue = cluster 4. This figure shows that the operational blocks do not coincide with the genetic boundaries of the vector. Figure modified from [42].

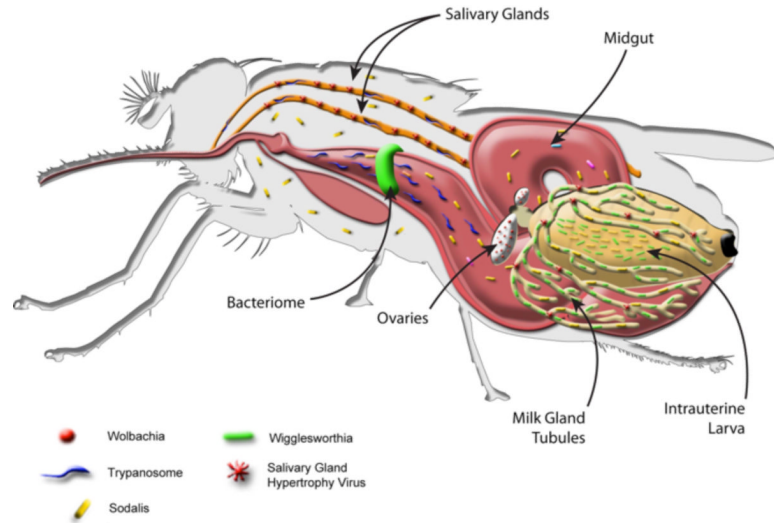


Figure 5. The tsetse

Multiple maternally transmitted symbiotic microbes have been described from laboratory lines and different field populations of tsetse, including obligate *Wigglesworthia*, commensal *Sodalis*, parasitic *Wolbachia*, and a DNA virus (Salivary Gland Hypertrophy Virus, SGHV), as schematically shown. Also shown are trypanosomes, which are acquired when tsetse feed on infected animals. The parasites reside in the gut and salivary gland organs. The midgut bacteriome organ, the intrauterine larva, and expanded milk gland tubules, which provide nutrients to the intrauterine progeny of the tsetse, are indicated. *Wigglesworthia*, *Sodalis*, and SGHV are maternally transmitted to the intrauterine larva of the tsetse in the secretions of mother's milk, while *Wolbachia* infects gonadal tissues and is vertically transmitted to developing progeny during embryogenesis.