

RESEARCH

Open Access

# The population structure of *Glossina fuscipes fuscipes* in the Lake Victoria basin in Uganda: implications for vector control

Chaz Hyseni<sup>1\*†</sup>, Agapitus B Kato<sup>2,3†</sup>, Loyce M Okedi<sup>4</sup>, Charles Masembe<sup>2</sup>, Johnson O Ouma<sup>5</sup>, Serap Aksoy<sup>6</sup> and Adalgisa Caccone<sup>1</sup>

## Abstract

**Background:** *Glossina fuscipes fuscipes* is the primary vector of trypanosomiasis in humans and livestock in Uganda. The Lake Victoria basin has been targeted for tsetse eradication using a rolling carpet initiative, from west to east, with four operational blocks (3 in Uganda and 1 in Kenya), under a Pan-African Tsetse and Trypanosomiasis Eradication Campaign (PATTEC). We screened tsetse flies from the three Ugandan PATTEC blocks for genetic diversity at 15 microsatellite loci from continental and offshore populations to provide empirical data to support this initiative.

**Methods:** We collected tsetse samples from 11 sites across the Lake Victoria basin in Uganda. We performed genetic analyses on 409 of the collected tsetse flies and added data collected for 278 individuals in a previous study. The flies were screened across 15 microsatellite loci and the resulting data were used to assess the temporal stability of populations, to analyze patterns of genetic exchange and structuring, to estimate dispersal rates and evaluate the sex bias in dispersal, as well as to estimate demographic parameters ( $N_E$  and  $N_C$ ).

**Results:** We found that tsetse populations in this region were stable over 4-16 generations and belong to 4 genetic clusters. Two genetic clusters (1 and 2) corresponded approximately to PATTEC blocks 1 and 2, while the other two (3 and 4) fell within PATTEC block 3. Island populations grouped into the same genetic clusters as neighboring mainland sites, suggesting presence of gene flow between these sites. There was no evidence of the stretch of water separating islands from the mainland forming a significant barrier to dispersal. Dispersal rates ranged from 2.5 km per generation in cluster 1 to 14 km per generation in clusters 3 and 4. We found evidence of male-biased dispersal. Few breeders are successfully dispersing over large distances. Effective population size estimates were low (33–310 individuals), while census size estimates ranged from 1200 (cluster 1) to 4100 (clusters 3 and 4). We present here a novel technique that adapts an existing census size estimation method to sampling without replacement, the scheme used in sampling tsetse flies.

**Conclusion:** Our study suggests that different control strategies should be implemented for the three PATTEC blocks and that, given the high potential for re-invasion from island sites, mainland and offshore sites in each block should be targeted at the same time.

**Keywords:** *Glossina fuscipes fuscipes*, Tsetse, Trypanosomiasis, Vector, Gene flow, Census size

\* Correspondence: chaz.hyseni@aya.yale.edu

†Equal contributors

<sup>1</sup>Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem Street, New Haven, CT, USA

Full list of author information is available at the end of the article

## Background

The tsetse fly (*Glossina*) is a major vector of trypanosomiasis throughout sub-Saharan Africa, causing extensive morbidity and mortality in humans and livestock [1,2]. It has been estimated that economic benefits to Africa from the eradication of tsetse could reach US\$4.5 billion per year. Currently, no vaccines exist to prevent the disease and available drugs to treat HAT are expensive, can cause severe side-effects, and are difficult to administer in remote villages [3]. As a consequence, an effective alternative for controlling the disease is to target the tsetse vector [4,5]. A variety of methods to control tsetse populations are available, including habitat modification around homesteads, trapping, insecticide-treated targets, insecticide-treated cattle, and aerial or ground spraying. The release of sterile or transgenic insects has been either used or proposed as an additional control measure [4-9]. Tsetse control is implemented using two strategies: eradication and suppression. Eradication aims at eliminating fly populations from a given area, while the objective of suppression is to greatly reduce population size.

Genetic data provide a powerful tool to help identify appropriate vector control strategies, as they can be used to estimate spatial and temporal differentiation of populations, and patterns and extent of migration. These data, together with ecological and environmental, data can be used to customize vector control efforts according to local conditions and species in order to determine whether eradication or suppression is appropriate [10,11]. For example, studies of tsetse in Burkina Faso, Guinea and Senegal have identified populations that are sufficiently isolated to warrant attempts at complete eradication [10-12]. Studies elsewhere have documented relatively high levels of gene flow, necessitating integration of barriers into eradication schemes [13,14], or warranting an area-wide control effort that encompasses populations linked by gene flow [15,16]. Regional studies such as the one on *G. palpalis palpalis* in west and central Africa [17] have provided information that is useful for control efforts at a regional scale. Other population genetic studies have pointed to specific populations at a local level where control and detection methods need improvement [18].

In 2001, the Organization for African Unity (OAU) launched a new initiative, the Pan-African Tsetse and Trypanosomiasis Eradication Campaign (PATTEC) to eradicate the tsetse flies from a vast area of sub-Saharan Africa (~10 million km<sup>2</sup>, seven *Glossina* species) by first reducing populations using area-wide approaches (odor-baited traps, insecticide-treated targets, pour-ons and ultra-low-volume aerial spraying), followed by massive release of sterile males to ensure eradication [19]. In Uganda, phase I of this initiative has been initiated in the

Lake Victoria basin, which is infested with *Glossina fuscipes fuscipes* (*Gff*). A program (Farming in Tsetse Controlled Areas, FITCA), which ended in 2004, reduced tsetse populations by 75% to 90% in the mainland sites of the target area, but the program did not include islands [20]. By 2009, a PATTEC baseline survey revealed that mainland tsetse populations had rebounded to the high levels prior to the FITCA initiative [21].

The PATTEC plan, unlike FITCA, includes islands and intends to eradicate *Gff* from the Lake Victoria basin progressively from west to east [22]. The basin has been partitioned into four operational blocks (3 in Uganda and 1 in Kenya) based on the Food and Agriculture Organization (FAO) predicted habitat suitability for *Gff*, natural barriers, major urban areas, international borders and drainage patterns [22,23]. Block 1 (Figure 1), targeted for *Gff* eradication during PATTEC phase 1, 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 has been targeted for control to create a buffer between the eradication block and the rest of the *Gff* predicted range in this area of Uganda (Figure 1). Only vector population monitoring activities are planned for the other two blocks during phase 1. Upon successful eradication in block 1, block 2 would become the eradication target and so on until the whole basin is tsetse-free. To support the PATTEC initiative, the government of Uganda and the International Atomic Energy Agency (IAEA) is planning a trial eradication of *Gff* on a remote island in the region (IAEA Project UGA5033) because of previous success of tsetse eradication on islands in Equatorial Guinea [24] and Zanzibar [25].

We used genetic variation at 15 microsatellite DNA loci to examine the genetic differentiation of *Gff* populations within and between the three Ugandan PATTEC blocks. We used these data to estimate effective population size ( $N_E$ ), to evaluate temporal stability over 4-16 generations, and to measure genetic exchange and dispersal rates within and between blocks. In view of the finding that high levels of genetic exchange occur among continental Ugandan populations of *Gff* separated by distances smaller than 100 km [16], we embarked upon an investigation of gene flow patterns at a finer scale and used the results to assess the validity of the PATTEC operational blocks with respect to vector control, to provide suggestions regarding control strategies in each block, and to evaluate the possibility of using *Gff* island populations for eradication trials.

## Methods

### Sampling

We sampled *Gff* at 3 continental localities across the Lake Victoria basin in Uganda (sites BD, EB, and MA:

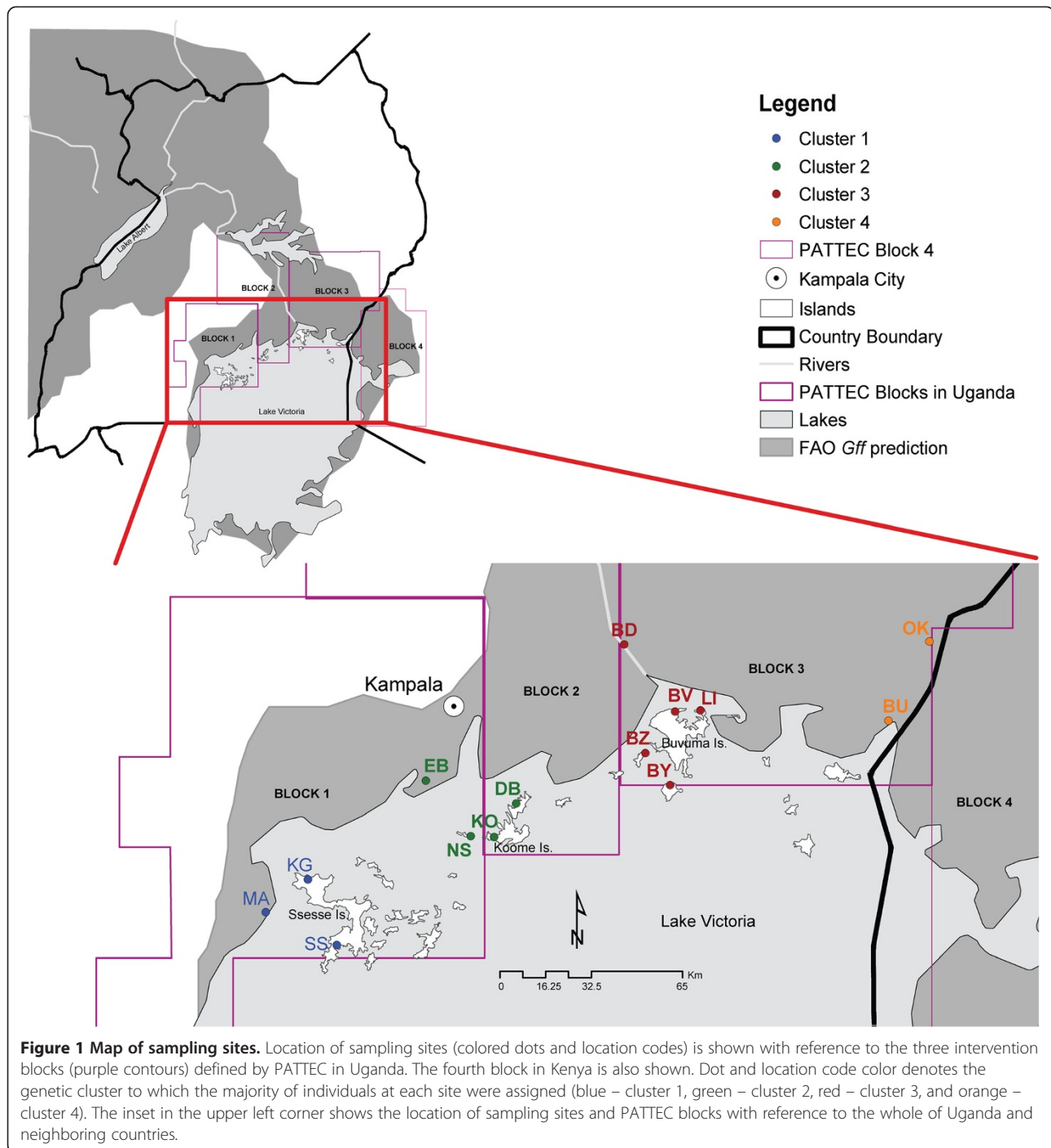


Table 1 and Figure 1) and at 8 localities from three groups of islands in Lake Victoria (sites BV, BY, BZ and LI from Buvuma islands; DB, NS and KO from Koome islands; and KG from Ssesse islands, Table 1 and Figure 1). We used 409 tsetse flies from the above 11 sampling sites for genetic analyses. Several localities were sampled over two to three months both in the wet (April-June 2010) and the dry (February-March 2010 or February-March 2011) season (Table 1a). In addition to the data

we collected from these 11 sites, we included data collected for 278 individuals sampled between the wet season of 2008 and the dry season of 2009: BU, OK [26] and SS (unpublished) (Table 1b). Assuming that *Gff* goes through approximately 8 generations per year [27,28], seasonal samples were about 4 generations apart. Tsetse flies were caught using biconical traps [29] and preserved individually in cryo-tubes containing 90% ethanol.

**Table 1 Sample collection information and genetic summary statistics**

Population	Code	N	Longitude	Latitude	Season	Collection Date	Seasonal and Total							
							Mean AR	Mean H <sub>O</sub>	Mean H <sub>E</sub>	Mean F <sub>IS</sub>				
<b>(a) New sites for this study</b>														
Budondo	BD	35	33.1209	0.5208	dry	March, 2011	4.867	4.867	0.516	0.516	0.513	0.513	-0.005	-0.005
Buvuma Is.	BV	21	33.2788	0.1368	dry	Feb, 2010	3.867	4.800	0.446	0.477	0.476	0.506	0.076	0.057
Buvuma Is.	BV	39	33.2788	0.1368	wet	June, 2010	4.400		0.493		0.513		0.035	
Bugaya Is.	BY	27	33.2684	0.0675	dry	Feb, 2010	4.067	4.867	0.499	0.486	0.496	0.497	-0.017	0.011
Bugaya Is.	BY	35	33.2684	0.0675	wet	June, 2010	4.400		0.476		0.490		0.018	
Buziri Is.	BZ	7	33.1883	0.1716	dry	Feb, 2010	2.933	3.600	0.441	0.454	0.460	0.477	0.064	0.051
Buziri Is.	BZ	11	33.1883	0.1716	wet	June, 2010	3.133		0.461		0.465		0.021	
Lingira Is.	LI	29	33.3532	0.3168	dry	Feb, 2010	4.933	5.200	0.519	0.509	0.511	0.509	-0.021	0.002
Lingira Is.	LI	17	33.3532	0.3168	wet	June, 2010	4.000		0.494		0.492		0.012	
Damba Is.	DB	32	32.7659	0.0127	wet	Apr, 2010	3.067	3.067	0.394	0.394	0.394	0.394	0.015	0.015
Koome Is.	KO	40	32.6879	-0.0911	wet	Apr, 2010	3.267	3.267	0.318	0.318	0.360	0.360	0.187	0.187
Nsazi Is.	NS	16	32.6296	-0.0956	wet	Apr, 2010	3.000	3.000	0.322	0.322	0.340	0.340	0.114	0.114
Entebbe	EB	35	32.4852	0.0823	dry	March, 2011	3.667	3.667	0.413	0.413	0.418	0.418	0.006	0.006
Kalangala Is.	KG	32	32.1083	-0.2267	wet	March, 2010	3.000	3.000	0.288	0.288	0.302	0.302	0.098	0.098
Masaka	MA	33	31.9852	-0.3564	dry	March, 2011	2.867	2.867	0.329	0.329	0.296	0.296	-0.098	-0.098
<b>(b) Sites from previous studies</b>														
Busime	BU	39	33.9711	0.2508	wet	March, 2008	4.733	5.400	0.486	0.486	0.507	0.502	0.027	0.045
Busime	BU	40	33.9711	0.2508	wet	March, 2009	4.467		0.494		0.499		0.047	
Busime	BU	40	33.9711	0.2508	dry	Oct, 2009	4.200	4.533	0.479	0.516	0.487	0.535	0.012	0.032
Okame	OK	39	33.3532	0.3168	wet	March, 2008	4.000		0.483		0.529		0.083	
Okame	OK	40	33.3532	0.3168	wet	March, 2009	3.933	4.133	0.541	0.532	0.532	0.532	-0.010	0.012
Okame	OK	39	33.3532	0.3168	dry	Oct, 2009	4.133		0.524		0.532		0.012	
Ssesse Is.	SS	40	32.1691	-0.5022	wet	Apr, 2009	3.467	3.467	0.333	0.333	0.332	0.332	-0.016	-0.016

Summary statistics are based on 15 microsatellite loci, both for (a) recently collected tsetse specimens and (b) flies collected for previous studies. N = number of individuals analyzed. Mean across all loci of AR = allelic richness, H<sub>O</sub> = observed heterozygosity, H<sub>E</sub> = expected heterozygosity and F<sub>IS</sub> = inbreeding coefficient. AR, H<sub>O</sub>, H<sub>E</sub> and F<sub>IS</sub> are reported for the dry and the wet season separately (seasonal), and for both seasons (total).

### Data collection

DNA was extracted from tsetse legs using the PrepGEM Insect DNA extraction kit (ZYGEM Corp Ltd, Hamilton, New Zealand) as per the protocol provided by the manufacturer. We collected genotypic data across 18 microsatellite loci (see Additional file 1: Supplementary Material for technical details), including the 13 loci used in Beadell *et al.* [16]. The other 5 (GmmA06, GmmB20, GmmD15, GmmL03, GmmL11) were selected among the loci described in Hyseni *et al.* [30]. Due to low amplification and scorability of large alleles, locus Pgp17 was excluded from all analyses. We also excluded locus GpC5b, because it was monomorphic for all populations, excepting BZ and NS (in both cases, F<sub>IS</sub> = 1), and locus GmmD15, because it was monomorphic in all of the 11 sites we sampled. Thus, we only used 15 loci for all subsequent genetic analyses.

### Genetic analyses

We used *Genepop 4.1* [31] to test for deviation from Hardy-Weinberg equilibrium (HWE) and also test for

linkage disequilibrium (LD). For loci with fewer than four alleles, the complete enumeration method [32] was used. All other loci were tested using the Guo and Thompson [33] Markov chain method with 100,000 dememorizations, 1,000 batches and 10,000 iterations per batch. We also used *Genepop* to carry out global tests across loci for heterozygote deficiency and heterozygote excess. Significance values were adjusted for multiple testing (HWE) and comparisons (LD) using the Benjamini-Hochberg method [34] with a false discovery rate of 0.05. Summary statistics, including allele frequencies, allelic richness, observed heterozygosity (H<sub>O</sub>), expected heterozygosity (H<sub>E</sub>) and the inbreeding coefficient (F<sub>IS</sub>) were calculated using the program *Genalex 6.41* [35]. In order to assess the statistical significance of genetic differentiation between temporal samples, we used *Genepop* to perform Fisher's exact test on both microsatellite alleles and genotypes.

We used the model-based Bayesian clustering method implemented in *Structure 2.3.3* [36] to determine the genetic structure present among *Gff* populations in the

Lake Victoria basin (details in Additional file 1: Supplementary Material). In addition to the Bayesian clustering implemented in *Structure*, we used a two-tiered multivariate ordination analysis, which makes no assumptions about deviations from Hardy-Weinberg and linkage equilibrium. This multivariate procedure, discriminant analysis of principal components (DAPC), has been shown to perform better than the Bayesian clustering approach when hierarchical and clinal structure is present in the data [37]. We used the *adegenet* package [38] in R [39] for the DAPC (details in Additional file 1: Supplementary Material).

In order to quantify the genetic heterogeneity of Lake Victoria populations, we computed pairwise  $F_{ST}$  values [40] among sampling localities and among genetic clusters.  $F_{ST}$  values were obtained using the program *Arlequin* 3.5 [41] and their significance computed via 10,000 permutations.  $F_{ST}$  describes the genetic structure produced by non-random distribution of individuals among subpopulations relative to the total population. However, our sample may contain multiple hierarchical levels of genetic differentiation, such as genetic clusters of populations with further partitioning within clusters, including potential genetic structure arising from isolation of island from mainland populations. In order to determine the contribution of different hierarchical levels to the observed genetic structure, we estimated hierarchical F-statistics using the method described in Yang [42] (details in Additional file 1: Supplementary Material) and implemented in the R package *hierfstat* [43].

To determine whether the genetic heterogeneity of *Gff* around Lake Victoria could be attributed to differences in dispersal ability between male and female flies, we performed t-tests on pairwise relatedness between individuals within genetic groups. Relatedness was computed using maximum likelihood estimation [44] implemented in *Kingroup* 2 [45]. Sex-biased dispersal was also assessed using three tests [46] implemented in *Fstat* 2.9.4 [47] (details in Additional file 1: Supplementary Material).

Isolation by distance (IBD) and dispersal were evaluated using Rousset's procedure [48] within genetic groups, using both a one-dimensional (1D) and a two-dimensional (2D) stepping-stone model. We compared the two models in order to determine differences in dispersal ability along Lake Victoria depending on whether movement happens along a line or across a surface (details in Additional file 1: Supplementary Material).

We used two methods to test for individual migrants between geographically neighboring genetic clusters (clusters 1 and 2, clusters 2 and 3, and clusters 3 and 4). In the first approach, we used the software *GeneClass* 2.0 [49] to compute the likelihood of individual assignment based on regional allele frequencies [50,51]. In the

second approach, we used *Flock 2.0* [52] to assign genetically similar individuals to  $k$  partitions (details in Additional file 1: Supplementary Material).

We estimated effective and census population sizes for each genetic group. The geographic distance between the genetic groups should reduce the bias in estimation of effective population size ( $N_E$ ) that could be introduced by migration, which influences linkage disequilibrium as well as temporal methods of  $N_E$  estimation [53].  $N_E$  was computed using Waples and Do's LD method implemented in the program *LDNe* [54]. We also used two temporal methods to estimate  $N_E$ : 1) a Bayesian algorithm based on coalescence and implemented in the program *TM3* [55] and 2) a pseudo-likelihood method [56] implemented in *MLNE*. Census size ( $N_C$ ) was computed via a sequential Bayesian method [57] adapted from Gazey and Staley [58] using an R software script [59]. This method applies to a sampling scheme with replacement (non-invasive genetic sampling). Tsetse sampling, however, was done without replacement. We designed a method to account for the difference in sampling, which allowed us to utilize the adapted Gazey-Staley method [57,58] to estimate  $N_C$  (details in Additional file 1: Supplementary Material).

## Results

We observed the lowest  $H_O$  and  $H_E$  values in KG (0.29 and 0.30, respectively) and the highest in OK (0.54 and 0.53, respectively), while mean allelic richness across the 15 loci ranged from 2.87 in MA to 5.40 in BU (Table 1). The highest  $F_{IS}$  value was observed in KO (0.187). After applying the Benjamini-Hochberg false discovery rate procedure [34] to the multiple testing of HWE, significant deviation from HWE was only observed in KO, which was due to significant heterozygote deficit. After applying the same correction for multiple comparisons of linkage between loci, no evidence of significant LD was found.

### Temporal stability

We examined the temporal stability of the samples for which temporal collections were available. Pairwise  $F_{ST}$  values between different sampling seasons revealed temporal homogeneity, i.e. temporal samples from the same sites were not significantly genetically differentiated. The smallest difference was observed between BU seasonal samples ( $F_{ST} = -0.001$ ,  $P = 0.63$ ) and highest in BY ( $F_{ST} = 0.003$ ,  $P = 0.24$ ). This was consistent with the similarity in allele frequencies between seasons observed in Additional file 2: Figure S1. We also carried out Fisher's exact test on microsatellite alleles and genotypes and found no significant differences between temporal samples (Additional file 3: Table S1).



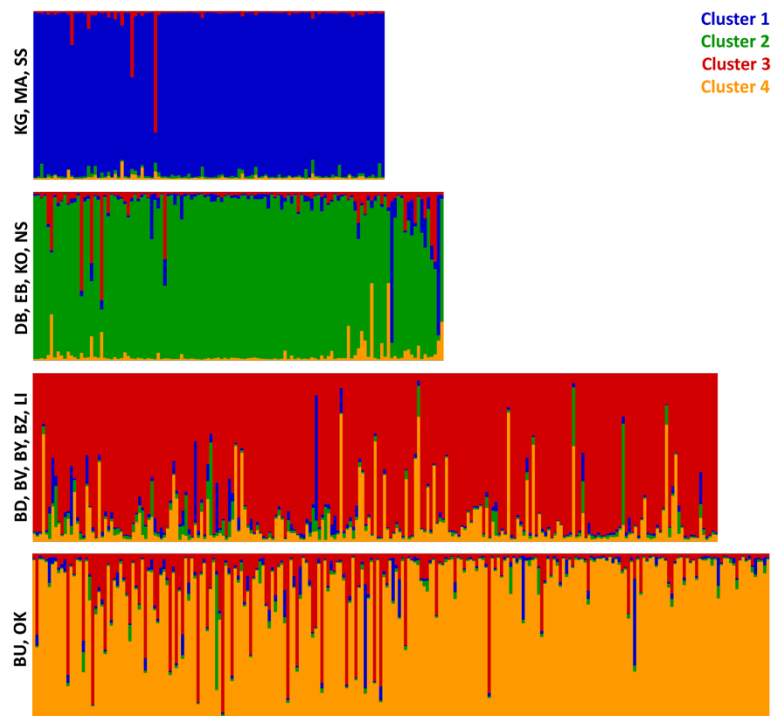
### Patterns of genetic differentiation

Using the Evanno criterion of  $\Delta K$ , the results of the Bayesian analysis in *Structure* identified four distinct genetic clusters (Figure 2). The DAPC approach (Figure 3), concordantly, detected four clusters comprising the same populations. Clusters 1 and 2 are within PATTEC blocks 1 and 2. However, one continental site and one island site within block 1 (EB and NS) are genetically closer to the continental and island sites in cluster 2 than the other samples from block 1. Cluster 3 includes the Buvuma islands and adjacent mainland sites and cluster 4 includes two inland sites to the east (BU and OK). Both clusters are within PATTEC block 3. While the two westernmost clusters (1 and 2) are genetically distinct, from each other ( $F_{ST} = 0.184$ , Additional file 4: Table S2) and from clusters 3 and 4 ( $F_{ST} = 0.124-0.191$ , Additional file 4: Table S2), there is a great amount of gene flow between clusters 3 and 4. The  $F_{ST}$  value between clusters 3 and 4 ( $F_{ST} = 0.036$ ), while significantly different, is much lower than among other clusters. The extent of gene flow between clusters 3 and 4 is also evident from posterior probabilities of assignment obtained from *Structure* (Figure 2) as well as from the DAPC method (Figure 3; posterior probabilities not shown).

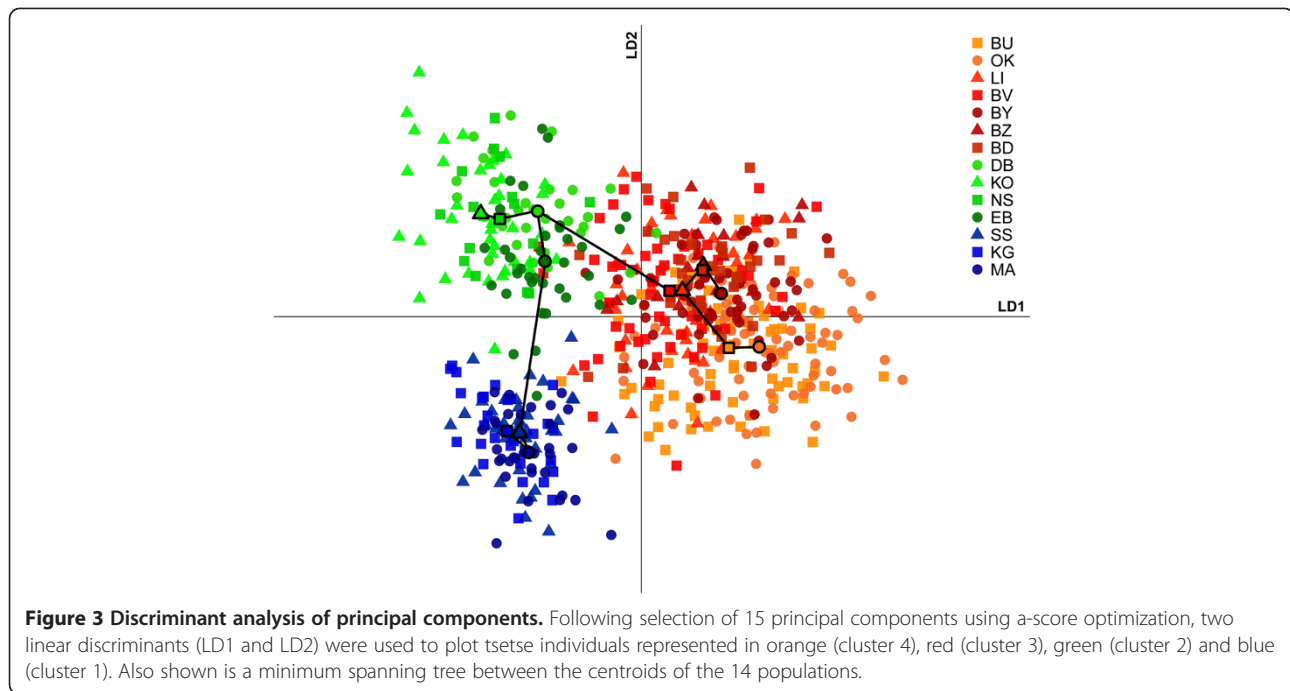
The analysis of hierarchical F-statistics (Table 2) revealed only one significant hierarchical level of genetic structure, i.e. the subdivision into three genetic groups

(group 1 = cluster 1, group 2 = cluster 2, group 3 = clusters 3 and 4;  $F_{3/T} = 0.114$ ,  $P = 0.001$ ; Table 2a). Subdivision of group 3 into clusters 3 and 4 did not contribute significantly to the observed genetic structure ( $F_{4/3} = 0.019$ ,  $P = 0.072$ ; Table 2a). Island areas with nearby mainland sites were also analyzed separately in order to look at the contribution of distance between mainland and island sites to the genetic structuring of each cluster. Tsetse flies from island sites were not significantly isolated from mainland flies within the same cluster (cluster 1:  $F_{IM/T} = -0.017$ ,  $P = 1.000$ ; cluster 2:  $F_{IM/T} = 0.042$ ,  $P = 0.259$ ; cluster 3:  $F_{IM/T} = 0.009$ ,  $P = 0.204$ ; Table 2b-d). Consistent with this finding, pairwise  $F_{ST}$  values (Additional file 5: Table S3) showed that absence of significant differentiation does not carry the proviso that the compared sites be situated on the same island or that the comparison not be between island and mainland sites. For instance, in group 1, the island site, SS, was not significantly different from the mainland site, MA ( $F_{ST} = 0.007$ ; Additional file 5: Table S3), but it was significantly different from the other island site, KG ( $F_{ST} = 0.024$ ; Additional file 5: Table S3).

Having delineated three genetic groups through multiple methods, we regressed linearized  $F_{ST}$  (i.e.,  $F_{ST}/(1-F_{ST})$ ) values against geographic distance to evaluate the occurrence of isolation by distance (IBD) within these groups (Figure 4). We detected significant IBD



**Figure 2 Bayesian clustering.** Probability of assignment of individuals from 14 sampling sites (abbreviated with two-letter codes; see Table 1) to each of the 4 identified clusters (blue – cluster 1, green – cluster 2, red – cluster 3, and orange – cluster 4) is denoted by the color composition of individual vertical bars.



within groups 2 and 3 (1D:  $P=0.021$ ,  $P=0.000$ ; 2D:  $P=0.022$ ,  $P=0.000$ ). IBD within group 1 was not significant when males were included in the model (1D:  $P=0.527$ ; 2D:  $P=0.568$ ). We did, however, find significant isolation by distance among females in group 1 (1D:  $P=0.004$ ; 2D:  $P=0.044$ ). In addition to the local patterns within groups, global linearized  $F_{ST}$  patterns between groups revealed that group 1 is as isolated from the more adjacent group 2 as it is from group 3 (Figure 4), which reiterated the DAPC results (Figure 3).

### Dispersal and migration

We tested for sex-biased dispersal within the three genetic groups using four methods (Table 3). Three of the four methods showed evidence of male-biased sex dispersal. The  $F_{ST}$ -based method did not reveal significant differences between sexes, possibly because of the reduced power of this method for low dispersal rates [46]. In group 3, male-biased dispersal was only supported by  $mAIC$  ( $P=0.012$ , Table 3). Evidence for male-biased dispersal was stronger in group 1 ( $vAIC$  and  $mPr$  were significantly different between males and females) and group 2 ( $mAIC$ ,  $vAIC$  and  $mPr$  were significant).

We calculated per generation dispersal ( $\sigma$ ), dispersal surface ( $\sigma^2$ ), Wright's neighborhood size ( $W_N$ ) and migration rate ( $m$ ), using both 1D and 2D stepping-stone models (Table 4). In groups 1 and 2, dispersal distance  $\sigma$  (group 1 (females): 1D=3.9 km, 2D=2.5 km; group 2: 1D=4.5 km, 2D=3.6 km), dispersal surface  $\sigma^2$  (2D: 6.2 km<sup>2</sup> and 12.9 km<sup>2</sup>), and neighborhood size (13 and 19 individuals) were similar to each other. These

estimates were higher for group 3 samples (2D:  $\sigma^2=200.1$  km<sup>2</sup>,  $W_N=64$  and  $\sigma=14.1$  km). Migration rate estimates per generation were also higher in group 3 (0.033) than the other two genetic groups (0.013 in group 1 (females) and 0.025 in group 2).

Figure 5 shows the scatterplots of log-likelihood of assignment of individuals to their cluster of origin and neighboring clusters using *Flock* and *GeneClass*. Migrant detection using these two methods was largely congruent with assignment of individuals to clusters using *Structure* (Figure 2) and DAPC (Figure 3; posterior probabilities not shown). Using log-likelihood ratios  $<0.5$ , we identified two migrants from cluster 1 in cluster 2 and four migrants from cluster 2 in clusters 3 and 4. The genetic exchange between clusters 3 and 4 was much higher, with 14 migrants from Buvuma islands in the BU-OK region and 12 individuals who migrated in the other direction. These two clusters shared 26 to 56 migrants for log-likelihood ratios  $<0.5$  to  $<1.0$ , respectively.

### Population size

Estimates of population size were computed for each genetic group. Table 4 shows effective population size ( $N_E$ ) estimates computed using both a linkage disequilibrium method ( $LDNe$ ) and two temporal methods (a likelihood approach implemented in *MLNE* and a Bayesian approach implemented in *TM3*). Table 4 also reports the estimated census size ( $N_C$ ).

$N_E$  and  $N_C$  were estimated for a dataset consisting of 12, 4, and 16 generations of flies from groups 1, 2 and 3, respectively. For the temporal methods only the first and

**Table 2 Hierarchical F-statistics**

<b>(a) Entire sample</b>				
	<b>3 Regions</b>	<b>4 Clusters</b>	<b>Populations</b>	<b>Individuals</b>
<b>Total</b>	0.114 ( $F_{3/T}$ )	0.130 ( $F_{4/T}$ )	0.151 ( $F_{Pop/T}$ )	0.173
<b>3 Regions</b>		0.019 ( $F_{4/3}$ )	0.043 ( $F_{Pop/3}$ )	0.067
<b>4 Clusters</b>			0.024 ( $F_{Pop/4}$ )	0.049
<b>Populations</b>				0.026
<b>(b) Cluster 1</b>				
	<b>Island/ Mainland</b>	<b>Populations</b>	<b>Individuals</b>	
<b>Total</b>	-0.017 ( $F_{IM/T}$ )	0.009 ( $F_{Pop/T}$ )	0.001	
<b>Island/Mainland</b>		0.025 ( $F_{Pop/IM}$ )	0.017	
<b>Populations</b>			-0.009	
<b>(c) Cluster 2</b>				
	<b>Island/ Mainland</b>	<b>Populations</b>	<b>Individuals</b>	
<b>Total</b>	0.042 ( $F_{IM/T}$ )	0.071 ( $F_{Pop/T}$ )	0.123	
<b>Island/Mainland</b>		0.030 ( $F_{Pop/IM}$ )	0.084	
<b>Populations</b>			0.055	
<b>(d) Cluster 3</b>				
	<b>Island/ Mainland</b>	<b>Populations</b>	<b>Individuals</b>	
<b>Total</b>	0.009 ( $F_{IM/T}$ )	0.021 ( $F_{Pop/T}$ )	0.048	
<b>Island/Mainland</b>		0.012 ( $F_{Pop/IM}$ )	0.039	
<b>Populations</b>			0.028	

**(a)** Hierarchical F-statistics computations were carried out for the entire sample based on 3 hierarchical levels: (1) populations (sampling sites), (2) populations grouped into four clusters, and (3) clusters 3 and 4 grouped together. **(b)-(d)** Hierarchical F-statistics were computed separately for clusters 1-3 (containing populations from Sese, Koome and Buvuma islands, respectively), based on 2 levels: (1) populations, and (2) populations grouped into island and mainland.

the last samplings were used for  $N_E$  estimation, as the increased sampling interval might decrease the bias caused by overlapping generations and age structure [53]. With the exception of flies in group 1, the estimates obtained with the  $LDN_e$  and temporal methods were very similar (Table 4). The lower temporal estimates for group 1 could be due to the higher substructure ( $F_{ST}$  between EB and other populations in the region was 0.07-0.09). We used  $N_E$  estimates obtained via the LD method to calculate dispersal.  $N_C$  was much higher in group 3 (4,121; Table 4) than elsewhere (1,199 in group 2 and 1,299 in group 1). The  $N_E/N_C$  ratio was lowest (0.075) in group 3; this ratio was higher in group 2 (0.101) and highest in group 1 (0.121).

## Discussion

### Temporal stability

The genetic data we collected for 14 sampling sites (Table 1 and Figure 1), which span up to 16 generations, showed that *Gff* populations are genetically stable over

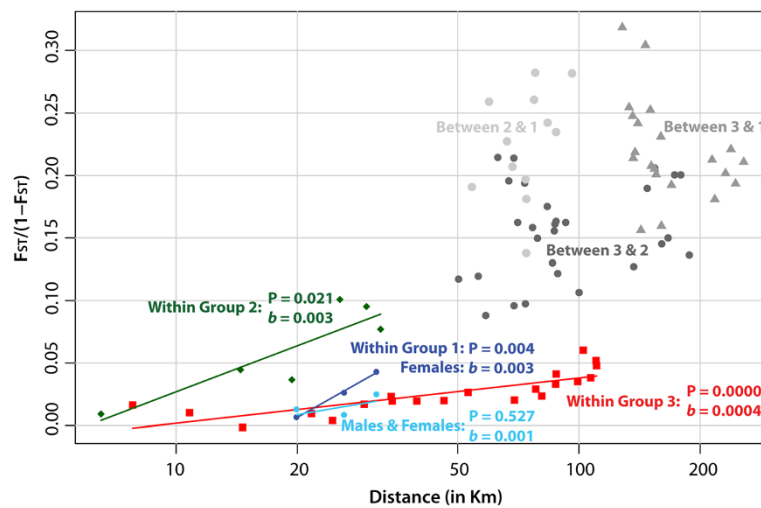
multiple seasons. Recent work on mainland sites in Uganda on the same species corroborates this result [26]. The significant genetic non-differentiation over several generations seems counterintuitive, given seasonal fluctuations in *Gff* abundance [61,62]. If this fluctuation is due to actual population reduction, however, 16 generations might not be enough time for genetic drift to produce differentiation, even at low  $N_E$  values (33-310 individuals), especially if migration alleviates the effects of drift. Seasonal fluctuations in population size could also be a reflection of the low efficiency of trapping devices used to monitor fly populations [63]. *Gff* is susceptible to high temperatures [64] and is known to inhabit the most humid habitats [65,66]. The reduced fly catches during the dry season could thus be due to refuge-seeking behavior whereby flies hide under bushes that cannot be easily accessed by traps. This behavior is shared by other riverine tsetse species [67,68]. Reduced movement during the dry season has been reported for *G. tachinoides* in northern Nigeria, which, like *Gff*, belongs to the *palpalis* group [69].

At the offshore island sites BY, BZ, and LI (Table 1), where temperatures are constantly lower and humidity is relatively high due to high rainfall and high evaporation rates throughout the year [70], dry and wet season fly catches are comparable in size. An exception to this is the low dry season tsetse abundance on the Buvuma islands (BV; Table 1). Strong dry-season winds, which are known to impact *Gff* activity [71,72] and commonly occur on Buvuma, might be responsible for the low fly densities recorded there. Windy conditions have also been reported to affect tsetse movement in Nigeria [73]. Additionally, given the genetic similarity between flies on the Buvuma islands and flies found in BY, BZ and LI, it is likely that Buvuma flies seek respite in adjacent habitats from harsh local dry season conditions.

### Patterns of genetic differentiation and dispersal within and between clusters

Clustering (Figure 2) and multivariate (Figure 3) analyses detected four genetic clusters of tsetse populations in the Ugandan Lake Victoria basin. Based on DAPC (Figure 3),  $F_{ST}$  (Additional file 4: Table S2), hierarchical F-statistics (Table 2) and IBD analyses (Figure 4), levels of genetic differentiation varied among clusters; while flies from clusters 1 and 2 were genetically distinct from each other and flies from clusters 3 or 4, the latter two clusters exchanged a large amount of genetic information. This was also confirmed by individual likelihood assignment tests, which identified numerous migrants (26 to 56) between clusters 3 and 4 (Figure 5). Clusters 1 and 2 (i.e., groups 1 and 2) correspond approximately to blocks 1 and 2, respectively, which were identified by PATTEC for eradication and suppression. However, two





**Figure 4 Isolation by distance.** Isolation by distance between sites within clusters 1, 2 and the group of clusters 3 and 4. Linearized  $F_{ST}$  (i.e.,  $F_{ST}/(1-F_{ST})$ ) values were regressed against geographic distance (in kilometers). Slopes ( $b$ ) and P-values are shown. For cluster 1, two isolation-by-distance models are pictured (light blue = both male and female individuals; dark blue = females only). In addition to populations within clusters (blue = cluster 1, green = cluster 2, red = clusters 3 and 4), populations belonging to different clusters were regressed, separately, and are depicted in three shades of grey.

sites in block 1 are genetically closer to sites from block 2 (Figure 1). Clusters 3 and 4 (i.e., group 3) fall within block 3 targeted for initial monitoring in the first phase of the rolling carpet initiative.

The high level of genetic connectivity between the two easternmost clusters (clusters 3 and 4) could be facilitated by the presence of suitable *Gff* habitat, i.e., riparian and lacustrine thickets, tree canopies with understory vegetation, or patches of banana and *Lantana camara* [74-76]. These thickets have been greatly reduced through human settlement and other human activities in the areas that make up clusters 1 and 2. The higher level of isolation of tsetse flies observed in cluster 1 could be

a function of increased habitat fragmentation and the intensive tsetse control activities dating back to the 1960s. In this area, high levels of deforestation have left only small gallery forests along the lakeshore. The growth of the city of Kampala could be an additional physical barrier to gene flow between flies from clusters 1 and 2. Increasing habitat fragmentation caused by human encroachment has been reported to have a major impact on the distribution, densities and structuring of riverine tsetse species [13,77,78].

Within clusters, island populations were not genetically differentiated from the mainland ones (Table 2 and Additional file 3: Table S1). Thus, the water body separating island and mainland sites (cluster 1: 19.9 km; cluster 2: 25.5 km; cluster 3: 34.4 km) does not act as a barrier to gene flow. Frequent human movements between islands and the mainland (fishing boats and daily ferry transport) may contribute to passive dispersal [79]. The extent of genetic connectivity of fly populations in this study is congruent with the general finding of other genetic studies on *Gff* [16,26] and other riverine species of tsetse [13,14,17,80]. In *G. p. gambiensis*, however, Solano et al. [81] report very low levels of migration between coastal sites and the Loos islands, situated 30 km off the coast of Guinea. Although this distance is not considerably greater than the distance separating adjacent island and mainland sites in our study (19.9-34.4 km), the differences in gene flow may be attributable to differences in the intensity of boat traffic, and thus more opportunity for passive dispersal of tsetse in the Lake Victoria region than off the coast of Guinea.

**Table 3 Sex-biased dispersal**

		$F_{ST}$	$mAlc$	$vAlc$	$mPr$
<b>Clusters 3 &amp; 4</b>	<b>F (202)</b>	0.028	0.476	20.05	0.031
	<b>M (179)</b>	0.028	-0.537	17.20	0.033
	<b>P-value</b>	0.991	<b>0.012</b>	0.261	0.255
<b>Cluster 2</b>	<b>F (71)</b>	0.061	0.706	11.64	0.115
	<b>M (52)</b>	0.076	-0.964	22.15	0.137
	<b>P-value</b>	0.617	<b>0.024</b>	<b>0.025</b>	<b>0.001</b>
<b>Cluster 1</b>	<b>F (61)</b>	0.020	0.541	17.19	0.099
	<b>M (44)</b>	0.013	-0.750	35.52	0.252
	<b>P-value</b>	0.931	0.144	<b>0.033</b>	<b>0.000</b>

Female and male dispersal were compared within three groups: (1) cluster 1, (2) cluster 2, and (3) clusters 3 and 4. Sex differences in dispersal were evaluated using four different statistics,  $F_{ST}$ ,  $mAlc$ ,  $vAlc$  (mean and variance, respectively, of the Favre et al. [60] assignment index) and  $mPr$  (mean pairwise relatedness (based on Konovalov and Heg [44])). P-values in bold type are significant.

**Table 4 Population size, migration and dispersal estimates**

	Linkage Diseq.		Temporal (MLNE)		Temporal (TM3)		Mean $N_C$	95% HPD
	(LDNe)		(Likelihood)		(Bayesian)			
	Mean $N_E$	(95% CI)	Mean $N_E$	(95% CI)	Mean $N_E$	(95% CI)		
Clusters 3 & 4	310	(258–380)	243	(181–327)	240	(223–250)	4121	(2422–6109)
Cluster 2	121	(88–180)	45	(35–61)	33	(22–47)	1199	(470–2191)
Cluster 1	157	(98–326)	212	(106–611)	178	(105–285)	1299	(412–2621)
	$N_E/N_C$	$\sigma$ (km)	$\sigma$ (km)	$\sigma^2$ (km <sup>2</sup> )	$m$	$W_N$		
		(1D)	(2D)	(2D)	(2D)	(2D)		
Based on LDNe estimates								
Clusters 3 & 4	0.075	15.16	14.14	200.06	0.033	64		
Cluster 2	0.101	4.53	3.59	12.92	0.025	19		
Cluster 1	0.121	3.94	2.49	6.21	0.013	13		

Estimates were computed for clusters 1, 2 and the group comprising clusters 3 and 4. Effective population size ( $N_E$ ) was computed using both a linkage disequilibrium method (in LDNe) and two temporal methods (a likelihood approach implemented in MLNE and a Bayesian approach implemented in TM3). Mean  $N_E$  and parametric 95% confidence intervals are shown. Census size ( $N_C$ ) was computed using a sequential Bayesian method [51]. Mean  $N_C$  and 95% highest probability density (HPD) intervals are shown. Dispersal distance ( $\sigma$ ) was estimated using both a one-dimensional (1D) ( $F_{ST}/(1-F_{ST}) \sim a + bGD$ ; GD = geographic distance in kilometers) and a two-dimensional (2D) ( $F_{ST}/(1-F_{ST}) \sim a + b \ln(GD)$ ;  $\ln(GD) = \log$  of geographic distance) isolation-by-distance model. Dispersal surface ( $\sigma^2$ ), migration ( $m$ ) and Wright's neighborhood size ( $W_N$ ) estimates from the 2D model are also reported.  $\sigma$ ,  $\sigma^2$ ,  $m$  and  $W_N$  in cluster 1 were only computed for female individuals (the model was not significant when males were considered) and are shown in italic type.

### Population size and dispersal

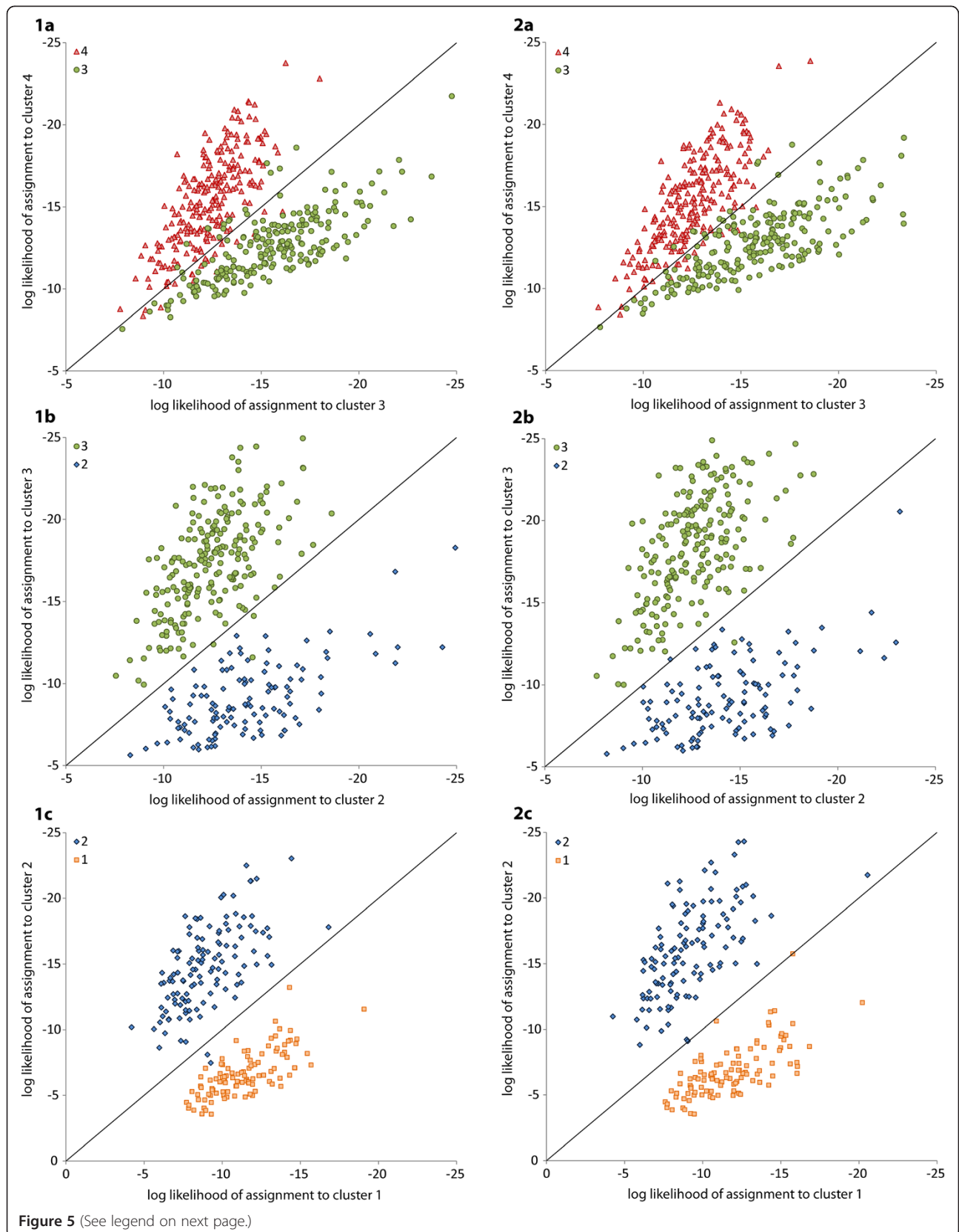
Population size estimates were calculated for three groups, which we defined as genetically distinct units using multiple analyses. The effective population size estimates ( $N_E$ ) were quite concordant among the linkage and temporal methods, ranging from 33 to 121 in group 2, 157 to 212 in group 1 and 240 to 310 in group 3. The  $N_C$  estimated for group 3 was much higher than the  $N_C$  estimates for groups 1 and 2. Despite the large difference in  $N_C$  estimates, the  $N_E$  estimates were similarly low in all three groups. The  $N_E/N_C$  ratio was, therefore, lowest in group 3 (Table 4). This ratio is affected by family-correlated survival (i.e. how families survive as a unit [82,83]) and variance in family size at high  $N_C$  values [83,84]. In terms of control, this implies that reducing population size alone without eradication would not constitute a permanent solution because  $N_E$  levels could be maintained, or even increased when  $N_C$  is reduced because of a decrease in variance of family size and increase in survival of families as units. The robustness of  $N_E$  in the face of population reduction can be seen in group 1, which has been subjected to intensive control efforts in the past; despite the low  $N_C$ ,  $N_E$  was relatively high (the  $N_E/N_C$  ratio was as high as 0.121 compared to 0.101 in group 2 and 0.075 in group 3).

Regression of linearized  $F_{ST}$  against geographic distance revealed significant IBD within groups 2 and 3, as well as females in group 1. Group 1 comprised only three sampling sites and, thus, we only had three data points for the regression. Despite the scarcity of data points, we observed a significant IBD pattern for the less mobile females by removing the better-dispersing male individuals from the model. Dispersal distances ( $\sigma$ ) were

similar between the 1D and 2D IBD models. These distances were approximately 14–15 km for flies in group 3 and lower in the other two groups, ranging from 2.5 to 4.5 km (Table 4). Based on the 2D estimates, the dispersal surfaces ( $\sigma^2$ ) within groups 1, 2 and 3 were 6.2, 12.9 and 200.1 km<sup>2</sup>, respectively. Similarly, migration rates per generation ( $m$ ) within the three units ranged from 0.01 in group 1 to 0.03 in group 3. Flies in group 3 cover greater distances and disperse over a wider surface area per generation, suggesting that conditions for movement are more favorable in this region, possibly due to environmental factors.

These genetically derived dispersal distances and migration rates are very similar to estimates of dispersal rates for *Gff* based on mark-release-recapture (MRR) studies, which are about 14.2 km per generation given the movement estimate of 338 m/day [85]. Our genetic estimates are also similar to MRR estimates for other riverine species, such as *G. palpalis gambiensis* and *G. tachinoides* [13,86]. MRR data do not necessarily correlate with genetic data, as was observed for flies from the *morsitans* group [87], suggesting that although habitat fragmentation reduces dispersal capacity, it may not impact levels of intraspecific genetic cohesiveness and that its effect is species-dependent. The relatively low dispersal rates in groups 1 and 2, as compared to group 3 and MRR estimates for *Gff* suggest that the dispersal capacity of flies is reduced in these areas, probably due to habitat loss and control efforts. The reduction in dispersal capacity has had an effect on genetic cohesiveness in the region and led to the differentiation into distinct genetic groups.

Suppression should be followed by eradication in all three groups. While eradication would be harder to



**Figure 5** (See legend on next page.)

(See figure on previous page.)

**Figure 5 Migrants.** Scatterplots of log-likelihood of assignment of individuals to the cluster of origin and neighboring clusters. Each dot represents an individual and its color denotes the individual's cluster of origin (red = cluster 4, green = cluster 3, blue = cluster 2 and orange = cluster 1; plots 1a-1c: *Flock* results, and 2a-2c: *GeneClass* results). Diagonal lines represent a log-likelihood ratio of 0. Migrants are represented by dots found across diagonals, away from their cluster of origin.

implement in group 3 because of the comparatively larger population size, larger dispersal distance and surface per generation, even the best suppression efforts would be much more difficult to maintain over time. Thus, the better option would be for an eradication campaign to follow an initial suppression phase post-haste, much more rapidly in group 3 than the other two groups.

The dispersal of flies across our study area showed evidence of sex bias. While male-biased dispersal in group 3 was only supported by mAIC, evidence for male-biased dispersal was stronger in groups 1 and 2. The finding of male-biased dispersal is very valuable for sterile insect technique (SIT) vector control efforts. The release of sterile male flies, as the better dispersing sex, holds the promise of sterile individuals being able to efficiently compete with wild breeders.

## Conclusion

The findings of this study have reaffirmed the importance of gathering genetic data prior to implementing area-wide tsetse vector control operations. The high levels of genetic mixing between islands and mainland sites suggest that these sites should be treated at the same time. The boundaries of the PATTEC blocks need to be modified to reflect the genetic composition of fly populations (i.e., three genetic groups). Fly populations from the two westernmost groups (1 and 2) are relatively isolated from flies in group 3, suggesting that suppression followed by eradication measures can be effective in these regions, but only if area-wide approaches include both island and mainland sites. For group 3, high levels of gene flow, which translate to a large dispersal surface ( $\sigma^2 = 200 \text{ km}^2$ ), as well as a very large census size and the potential for the  $N_E/N_C$  ratio to increase in the case of population size reduction, all suggest that suppression alone is not likely to produce desirable results in the long run and that it should be coupled with eradication. Additionally, following the initial suppression phase, eradication would have to ensue faster for group 3 than the other two groups in order to overcome the comparatively higher dispersal capacity of flies and prevent reinfestation. The high dispersal rate in group 3 is a strong argument against the Buvuma islands being a suitable location to evaluate eradication protocols, and that the Ssesse islands are a better target for this purpose, provided mainland sites are also included. The finding of male-biased dispersal of tsetse populations

in this region ensures that eradication efforts involving SIT are likely to be successful.

## Additional files

**Additional file 1: Supplementary Material [88,89].**

**Additional file 2: Figure S1. Temporal allele frequencies.** Frequencies of alleles (above 0.05) across 15 loci are shown for the group of BU and OK samples (wet season of 2008 – dark blue, wet season of 2009 – blue, dry season of 2009 – light blue), and BV, BY, BZ and LI samples grouped together (dry season of 2010 – dark red, wet season of 2010 – light red).

**Additional file 3: Table S1. Fisher's exact differentiation test.** For each site or group of sites (BU-OK and BV-BY-BZ-LI), tsetse collected at different sampling times (wet and dry seasons; see Table 1) were tested for differences in allelic frequencies (genetic differentiation) and genotypic frequencies. P (genetic) and P (genotypic) – probability of non-differentiation.

**Additional file 4: Table S2. Pairwise FST values.** FST values were computed for 4 clusters, averaged across 15 loci based on Weir and Cockerham [40]. FST values are reported in the lower diagonal and significance ( $\alpha = 0.05$ ) in the upper diagonal ('-': not significant; '+': significant).

**Additional file 5: Table S3. Pairwise FST values.** FST values were computed for 14 sampling sites, averaged across 15 loci based on Weir and Cockerham [40]. FST values are reported in the lower diagonal and significance ( $\alpha = 0.05$ ) in the upper diagonal ('-': not significant; '+': significant). Non-significant FST values are represented in bold italic type.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

SA, AC, LMO, CM and JOO designed the study. ABK and LMO collected samples in Uganda. CH and ABK performed the lab work. CH carried out the statistical analyses. ABK provided background and ecological information. CH and ABK wrote the initial draft of the manuscript. CH, AC, SA, and CM revised the manuscript. SA and AC jointly supervised the work at Yale University. All authors read and approved the final manuscript.

## Acknowledgements

This study was supported by grants from NIH (R01 AI068932 and D43 TW007391) to SA, AC and LMO, and WHO-TDR (A80132) to AC and LMO. The research was accomplished while ABK was a Fogarty Research Fellow at Yale University. We are thankful for the support of Drs. Vincent Muwanika and Anne M. Akol (Makerere University, Uganda), PATTEC Uganda's STATFA Project and Mukono District Administration. We are grateful to the technical staff of NaLIRRI for excellent assistance with field sampling.

## Author details

<sup>1</sup>Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem Street, New Haven, CT, USA. <sup>2</sup>Department of Biology, Makerere University, College of Natural Sciences, School of Bio-Sciences, P.O. Box 7062, Kampala, Uganda. <sup>3</sup>Creation of Sustainable Tsetse and Trypanosomiasis Free Areas in Uganda project, Ministry of Agriculture, Animal Industry and Fisheries, P.O. Box 102, Entebbe, Uganda. <sup>4</sup>National Livestock Resources Research Institute, Tororo, Uganda, P.O. Box 96, Tororo, Uganda. <sup>5</sup>Trypanosomiasis Research Centre, Kenya Agricultural Research Institute, P.O. Box 362-00902, Kikuyu, Kenya. <sup>6</sup>Yale School of Public Health, Yale University, 60 College Street, New Haven, CT, USA.



Received: 16 May 2012 Accepted: 15 September 2012  
Published: 4 October 2012

## References

1. Dyer NA, Lawton SP, Ravel S, Choi KS, Lehane MJ, Robinson AS, Okedi LM, Hall MJR, Solano P, Donnelly MJ: **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**:227–239.
2. Leak SGA, Leak SGA: *Tsetse biology and ecology: their role in the epidemiology and control of trypanosomiasis.* Wallingford: CAB; 1999.
3. Budd L: **DFID-funded tsetse and trypanosome research and development since 1980. Vol. 2.** In *Economic analysis.* Aylesford: DFID; 1999.
4. Aksoy S, Maudlin I, Dale C, Robinson AS, O'Neill SL: **Prospects for control of African trypanosomiasis by tsetse vector manipulation.** *Trends Parasitol* 2001, **17**:29–35.
5. Aksoy S: **Control of tsetse flies and trypanosomes using molecular genetics.** *Vet Parasitol* 2003, **115**:125–145.
6. Torr SJ, Hargrove JW, Vale GA: **Towards a rational policy for dealing with tsetse.** *Trends Parasitol* 2005, **21**:537–541.
7. Vreysen MJB, Marc JB: **Prospects for area-wide integrated control of tsetse flies (Diptera: Glossinidae) and trypanosomiasis in sub-Saharan Africa.** *RSEA* 2006, **65**:1–21.
8. Schofield CJ, Kabayo JP: **Trypanosomiasis vector control in Africa and Latin America.** *Parasit Vectors* 2008, **1**:24.
9. Vreysen MJB, Robinson AS: **Ionising radiation and area-wide management of insect pests to promote sustainable agriculture. A review.** *Agron Sustain Dev* 2011, **31**:233–250.
10. Solano P, Kaba D, Ravel S, Dyer NA, Sall B, Vreysen MJ, Seck MT, Darbyshire H, Gardes L, Donnelly MJ, 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.
11. Solano P, Ravel S, de Meues T: **How can tsetse population genetics contribute to African trypanosomiasis control?** *Trends Parasitol* 2010, **26**:255–263.
12. Kagbadouno M, Camara M, Bouyer J, Hervouet JP, Courtin F, Jamonneau V, Morifaso O, Kaba D, Solano P: **Tsetse elimination: its interest and feasibility in the historical sleeping sickness focus of Loos islands, Guinea.** *Parasite* 2009, **16**:29–35.
13. Bouyer J, Balenghien T, Ravel S, Vial L, Sidibé I, Thévenon S, Solano P, Meeus T: **Population sizes and dispersal pattern of tsetse flies: rolling on the river?** *Mol Ecol* 2009, **18**:2787–2797.
14. Koné N, Bouyer J, Ravel S, Vreysen MJB, Domagni KT, Causse S, Solano P, Meeus T: **Contrasting population structures of two vectors of African trypanosomiasis in Burkina Faso: consequences for control.** *PLoS Negl Trop Dis* 2011, **5**:e1217.
15. Abila PP, Slotman MA, Parmakelis A, Dion KB, Robinson AS, Muwanika VB, Enyaru JC, Okedi LM, Aksoy S, Caccone A: **High levels of genetic differentiation between Ugandan Glossina fuscipes fuscipes populations separated by Lake Kyoga.** *PLoS Negl Trop Dis* 2008, **2**:e242.
16. Beadell JS, Hyseni C, Abila PP, Azabo R, Enyaru JC, Ouma JO, Mohammed YO, Okedi LM, Aksoy S, Caccone A: **Phylogeography and population structure of Glossina fuscipes fuscipes in Uganda: implications for control of tsetse.** *PLoS Negl Trop Dis* 2010, **4**:e636.
17. Tanekou TT, Melachio TT, Simo G, Ravel S, Meeus T, Causse S, Solano P, Lutumba P, Asonganyi T, Njikou F: **Population genetics of Glossina palpalis palpalis from central African sleeping sickness foci.** *Parasit Vectors* 2011, **4**:140.
18. Ouma JO, Beadell JS, Hyseni C, Okedi LM, Krafur ES, Aksoy S, Caccone A: **Genetic diversity and population structure of Glossina pallidipes in Uganda and western Kenya.** *Parasit Vectors* 2011, **4**:122.
19. Kabayo JP: **Aiming to eliminate tsetse from Africa.** *Trends Parasitol* 2002, **18**:473–475.
20. Agrisystems Limited: *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.
21. 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 population in Uganda.* Boughton: Cybister Environmental Protection; 2009.
22. 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.
23. Wint W: *Kilometre resolution tsetse fly distribution maps for the Lake Victoria basin and West Africa.* Vienna: IAEA; 2001.
24. da Costa BFB: *Sleeping sickness: a record of four years' war against it in Principe, Portuguese West Africa.* Lisbon: Baillière, Tindall and Cox; 1916.
25. Vreysen MJ, Saleh KM, Ali MY, Abdulla AM, Zhu ZR, Juma KG, Dyck VA, Msangi AR, Mkonyi PA, Feldmann HU: **Glossina austeni (Diptera: Glossinidae) eradicated on the island of Unguja, Zanzibar, using the sterile insect technique.** *J Econ Entomol* 2000, **93**:123–135.
26. Echodu R, Beadell JS, Okedi LM, Hyseni C, Aksoy S, Caccone A: **Temporal stability of Glossina fuscipes fuscipes populations in Uganda.** *Parasit Vectors* 2011, **4**:19.
27. Hargrove JW: **Extinction probabilities and times to extinction for populations of tsetse flies Glossina spp. (Diptera: Glossinidae) subjected to various control measures.** *Bull Entomol Res* 2005, **95**:13–21.
28. Krafur ES, Marquez JG, Ouma JO: **Structure of some East African Glossina fuscipes fuscipes populations.** *Med Vet Entomol* 2008, **22**:222–227.
29. Challier A, Laveissiere C: **A new trap for capturing Glossina flies (Diptera: Muscidae), description and field trials.** *Cah ORSTOM Entomol Med Parasitol* 1973, **11**:251–262.
30. Molecular Ecology Resources Primer Development Consortium: **Permanent genetic resources added to Molecular Ecology Resources Database 1 December 2010–31 January 2011.** *Mol Ecol Resour* 2011, **11**:586–589.
31. Rousset F: **Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux.** *Mol Ecol Resour* 2008, **8**:103–106.
32. Louis EJ, Dempster ER: **An exact test for Hardy-Weinberg and multiple alleles.** *Biometrics* 1987, **43**:805–811.
33. Guo SW, Thompson EA: **Performing the exact test of Hardy-Weinberg proportion for multiple alleles.** *Biometrics* 1992, **48**:361–372.
34. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Stat Soc Series B Stat Methodol* 1995, **57**:289–300.
35. Peakall R, Smouse PE: **GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research.** *Mol Ecol Notes* 2006, **6**:288–295.
36. Pritchard JK, Stephens M, Donnelly P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155**:945–959.
37. Jombart T, Devillard S, Baloux F: **Discriminant analysis of principal components: a new method for the analysis of genetically structured populations.** *BMC Genet* 2010, **11**:94.
38. Jombart T: **adegenet: a R package for the multivariate analysis of genetic markers.** *Bioinformatics* 2008, **24**:1403–1405.
39. R Development Core Team: *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing; 2011.
40. Weir BS, Cockerham CC: **Estimating F-Statistics for the analysis of population structure.** *Evolution* 1984, **38**:1358–1370.
41. Excoffier L, Lischer HEL: **Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows.** *Mol Ecol Resour* 2010, **10**:564–567.
42. Yang R-C: **Estimating hierarchical F-statistics.** *Evolution* 1998, **52**:950–956.
43. Goudet J: *Hierfstat: estimation and tests of hierarchical F-statistics.* Vienna: R Foundation for Statistical Computing; 2011.
44. Konovalov DA, Heg DIK: **TECHNICAL ADVANCES: a maximum-likelihood relatedness estimator allowing for negative relatedness values.** *Mol Ecol Resour* 2008, **8**:256–263.
45. Konovalov DA, Manning C, Henshaw MT: **KINGROUP: a program for pedigree relationship reconstruction and kin group assignments using genetic markers.** *Mol Ecol Notes* 2004, **4**:779–782.
46. Goudet J, Perrin N, Waser P: **Tests for sex-biased dispersal using bi-parentally inherited genetic markers.** *Mol Ecol* 2002, **11**:1103–1114.
47. Goudet J: **FSTAT (Version 1.2): a computer program to calculate F-statistics.** *J Hered* 1995, **86**:485–486.
48. Rousset F: **Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance.** *Genetics* 1997, **145**:1219–1228.
49. Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A: **GENECLASS2: a software for genetic assignment and first-generation migrant detection.** *J Hered* 2004, **95**:536–539.



50. Paetkau D, Calvert W, Stirling I, Strobeck C: Microsatellite analysis of population structure in Canadian polar bears. *Mol Ecol* 1995, **4**:347–354.
51. Paetkau D, Slade R, Burden M, Estoup A: Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol Ecol* 2004, **13**:55–65.
52. Duchesne P, Turgeon J: FLOCK: a method for quick mapping of admixture without source samples. *Mol Ecol Resour* 2009, **9**:1333–1344.
53. Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW: Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conserv Genet* 2010, **11**:355–373.
54. Waples RS, Do C: LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* 2008, **8**:753–756.
55. Berthier P, Beaumont MA, Cornuet JM, Luikart G: Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. *Genetics* 2002, **160**:741–751.
56. Wang JL, Whitlock MC: Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* 2003, **163**:429–446.
57. Petit E, Valiere N: Estimating population size with noninvasive capture-mark-recapture data. *Conserv Biol* 2006, **20**:1062–1073.
58. Gazey WJ, Staley MJ: Population estimation from mark-recapture experiments using a sequential Bayes algorithm. *Ecology* 1986, **67**:941–951.
59. Puechmaile SJ, Petit EJ: Empirical evaluation of non-invasive capture-mark-recapture estimation of population size based on a single sampling session. *J Appl Ecol* 2007, **44**:843–852.
60. Favre L, Balloux F, Goudet J, Perrin N: Female-biased dispersal in the monogamous mammal *Crocodyrus russula*: Evidence from field data and microsatellite patterns. *Proc Biol Sci* 1997, **264**:127–132.
61. Katunguka-Rwakishaya E, Kabagambe EK: Tsetse survey in Mukono district, south-east Uganda: population structure, distribution and blood meal status. *Trop Anim Health Prod* 1996, **28**:151–157.
62. Bonomi A, Bassetti F, Gabrieli P, Beadell J, Falchetto M, Scolari F, Gomulski LM, Regazzini E, Ouma JO, Caccone 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.
63. Leak SGA, Ejigu D, Vreysen MJB: Collection of entomological baseline data for tsetse area-wide integrated pest management programmes. Rome: FAO; 2008.
64. Hargrove JW, Brady J: Activity rhythms of tsetse flies (*Glossina* spp) (Diptera: Glossinidae) at low and high temperatures in nature. *Bull Entomol Res* 1992, **82**:321–326.
65. Dyer NA, Ravel S, Choi KS, Darby AC, Causse S, Kapitano B, Hall MJR, Steen K, Lutumba P, Madinga J, others: Cryptic diversity within the major trypanosomiasis vector *Glossina fuscipes* revealed by molecular markers. *PLoS Negl Trop Dis* 2011, **5**:e1266.
66. Pollock JN: *Training manual for tsetse control personnel. Volume I. Tsetse biology, systematics and distribution; techniques.* Rome: FAO; 1982.
67. Cuisance D: Evolution sous l'action de la secheresse d'une population isolee de *Glossina tachinoides* West. (region du Bas Logone, Tchad). *Rev Elev Med Vet Pays Trop* 1974, **27**:95–107.
68. Lambrecht FL: Field studies of *Glossina morsitans* Westw. (Dipt.: Glossinidae) in relation to Rhodesian sleeping sickness in N'gamiland, Botswana. *Bull Entomol Res* 1972, **62**:183–193.
69. Dransfield RD, Brightwell R, Onah J, Okolo CJ: Population dynamics of *Glossina morsitans submorsitans* Newstead and *G. tachinoides* Westwood (Diptera: Glossinidae) in sub-Saharan savanna in northern Nigeria. I. Sampling methodology for adults and seasonal changes in numbers caught in different vegetation types. *Bull Entomol Res* 1982, **72**:175–192.
70. Kendall RL: An ecological history of the Lake Victoria basin. *Ecol Monogr* 1969, **39**:121–176.
71. Carpenter GDH: Second report on the bionomics of *Glossina fuscipes* (palpalis) of Uganda. *Rep Sleep Sicken Commn R Soc* 1913, **14**:1–37.
72. Ogwal LM, Kalyebi A, Kaddu JB: The diurnal activity, movement and trypanosome infection rates of *Glossina fuscipes fuscipes* (Diptera: Glossinidae) in Buvuma Island, Lake Victoria, Uganda. *Insect Sci* 2007, **14**:477–484.
73. Molyneux DH, Baldry DAT, Fairhurst C: Tsetse movement in wind fields - possible epidemiological and entomological implications for trypanosomiasis and its control. *Acta Trop* 1979, **36**:53–65.
74. Okoth JO, 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. *Int J Trop Insect Sci* 1987, **8**:57–60.
75. Kuzoe FAS, Schofield CJ: *Strategic review of traps and targets for tsetse and African trypanosomiasis control.* Geneva: UNICEF/UNDP/World Bank/WHO; 2004.
76. Syed Z, Guerin PM: Tsetse flies are attracted to the invasive plant *Lantana camara*. *J Insect Physiol* 2004, **50**:43–50.
77. Bouyer J, Guerrini L, Cesar J, de la Rocque S, Cuisance D: A phyto-sociological analysis of the distribution of riverine tsetse flies in Burkina Faso. *Med Vet Entomol* 2005, **19**:372–378.
78. Guerrini L, Bord JP, Ducheyne E, Bouyer J: Fragmentation analysis for prediction of suitable habitat for vectors: example of riverine tsetse flies in Burkina Faso. *J Med Entomol* 2008, **45**:1180–1186.
79. Ogwal LM, Kangwagye TN: Population dynamics of *Glossina fuscipes fuscipes* on Buvuma Island, Lake Victoria, Uganda. In *Panel Proceedings Series*. Edited by Offori ED, van der Vloedt AMV. Vienna: IAEA; 1990:67–77.
80. Bouyer J, Ravel S, Guerrini L, Dujardin JP, Sidibé I, Vreysen MJB, Solano P, Meeüs T: Population structure of *Glossina palpalis gambiensis* (Diptera: Glossinidae) between river basins in Burkina Faso: consequences for area-wide integrated pest management. *Infect Genet Evol* 2010, **10**:321–328.
81. Solano P, Ravel S, Bouyer J, Camara M, Kagbadouno MS, Dyer N, Gardes L, Herault D, Donnelly MJ, Meeüs T: The population structure of *Glossina palpalis gambiensis* from island and continental locations in coastal Guinea. *PLoS Negl Trop Dis* 2009, **3**:e392.
82. Waples RS: Evaluating the effect of stage-specific survivorship on the Ne/N ratio. *Mol Ecol* 2002, **11**:1029–1037.
83. Hedrick P: Large variance in reproductive success and the Ne/N ratio. *Evolution* 2005, **59**:1596–1599.
84. Pray LA, Goodnight CJ, Stevens L, Schwartz JM, Yan G: The effect of population size on effective population size: an empirical study in the red flour beetle *Tribolium castaneum*. *Genet Res* 1996, **68**:151–155.
85. Rogers D: Study of a natural population of *Glossina fuscipes fuscipes* Newstead and a model of fly movement. *J Anim Ecol* 1977, **46**:309–330.
86. Cuisance D, Février J, Dejardin J, Filledier J, Cuisance D, Février J, Dejardin J, Filledier J: Dispersion linéaire de *Glossina palpalis gambiensis* et *G. tachinoides* dans une galerie forestière en zone soudano-guinéenne (Burkina Faso). *Rev Elev Med Vet Pays Trop* 1985, **38**:153–172.
87. Krafurs ES, Endsley MA: Microsatellite diversities and gene flow in the tsetse fly, *Glossina morsitans* s.l. *Med Vet Entomol* 2002, **16**:292–300.
88. Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF: M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques* 2001, **31**:24–27.
89. Evanno G, Regnaut S, Goudet J: Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 2005, **14**:2611–2620.

doi:10.1186/1756-3305-5-222

Cite this article as: Hyseni et al.: The population structure of *Glossina fuscipes fuscipes* in the Lake Victoria basin in Uganda: implications for vector control. *Parasites & Vectors* 2012 **5**:222.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

